Addressing Heterogeneity, Scalability, and Privacy in Layered Multicast Congestion Control

Sergey Gorinsky
Dept. of Computer Science and Engineering
Washington University in St. Louis
St. Louis, MO 63130, USA
gorinsky@wustl.edu

K. K. Ramakrishnan
AT&T Labs-Research
Florham Park, NJ 07932, USA
kkrama@research.att.com

Harrick Vin
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712, USA
vin@cs.utexas.edu

***Abstract*—Multicast is attracting a resurgence of interest because it has a potential to address the explosively growing need for efficient streaming of large-volume Internet content. However, to realize the potential, large-scale multicast distribution requires effective congestion control. In this paper, we design SIM, a protocol that combines three distinct mechanisms (Selective participation, Intra-group transmission adjustment, and Menu adaptation) to provide a general solution for efficient fair scalable privacy-preserving multicast congestion control with heterogeneous receivers. Whereas the individual mechanisms have appeared in earlier multicast protocols, our main contribution lies in the cohesive integration of the techniques. SIM achieves such integration by operating the three mechanisms at different timescales and distributing the responsibility for control to different entities in the network. Besides, to instantiate and integrate the three control mechanisms, SIM employs several novel techniques of independent interest. Our extensive evaluation confirms the salient properties of SIM in diverse settings where receiving capabilities are highly heterogeneous, bottleneck capacities fluctuate, bottlenecks migrate, and session membership is dynamic.**

I. INTRODUCTION

The Internet is facing an explosive growth of content and needs efficient means for delivery of large-volume data. In particular, video streaming is becoming more common [1], [2]. Together with other large-volume traffic types that impose some real-time delivery requirements, video streaming is expected to grow at an accelerated pace fueled by the communication demands of such Internet applications as multi-party video conferencing, distributed interactive gaming, distance learning, and group collaboration.

Attracting a resurgence of interest, *network-layer multicast* provides the needed efficiency by sending data over a link only once when delivering the data to multiple destinations [3]. In comparison to end-system multicast [4], multicasting at the network layer excels in both efficiency of communications and privacy offered to session participants.

The significant presence of large-volume data makes network congestion a significant concern. Even in a well-engineered IPTV (Internet Protocol Television) distribution environment, congestion is an issue due to potential failures [5]. While multicast helps applications to meet their real-time communication requirements, it is unable to benefit from TCP [6], DCCP [7] and other more recent solutions developed to provide congestion control for unicast delivery.

Although multicast congestion control has been extensively studied [8]–[22], no single method has become dominant. One major goal that still needs to be attained is the ability of a multicast congestion control protocol to deal effectively with heterogeneity in receiving capabilities of different receivers. With large-scale content distribution, it is unrealistic to expect similar available capacities on the network paths from the sender to each of the receivers. Another important consideration is privacy of multicast services requested and provided to an individual receiver. Furthermore, efficient privacy-preserving multicast delivery should scale to a large number of receivers, allocate network resources fairly among own sessions and cross traffic, be responsive to changes in network conditions, and impose low communication and computation overhead. While existing multicast congestion control protocols achieve some of the above objectives, none of the protocols offers the full set of the desired properties.

In this paper, we present SIM, a protocol that integrates three distinct mechanisms – Selective participation, Intra-group transmission adjustment, and Menu adaptation – to provide a general solution for efficient fair scalable privacy-preserving multicast congestion control with heterogeneous receivers. Whereas the individual mechanisms have appeared in earlier multicast protocols, the main contribution of SIM lies in the cohesive integration of the techniques. SIM provides such integration by operating the three mechanisms at different timescales and ensuring that mechanisms working on longer timescales complement and take advantage of the control wielded by faster mechanisms. In SIM, selective participation works at an intermediate timescale and allows the receivers to discover their optimal subscriptions to the multi-group session. Intra-group transmission adjustment operates at the shortest timescale and bears the frontline burden of dealing with short-term changes in network conditions. Finally, menu adaptation is done at the longest timescale and improves the overall efficiency of SIM when the number of bottleneck capacities exceeds the maximum number of groups.

To instantiate and integrate the three control mechanisms cohesively, SIM employs several novel techniques of independent interest. In particular, SIM congestion detection and notification demonstrates that explicit signaling can be successfully applied in the context of multicast congestion control. Also, SIM feedback aggregation exhibits the prominent feature of providing the sender with timely accurate feedback in a scalable manner without jeopardizing privacy of the session participants.

Based on the general philosophy of our approach [23], this paper describes the specific protocol and evaluates its performance thoroughly. The evaluation shows that SIM accommodates diverse receiving capabilities efficiently and fairly. When the number of different bottleneck capacities is at most the maximum number of groups in a session, SIM aligns the group transmission rates with the receiving capabilities well and maintains stable optimal group subscriptions. For these and other settings, our experiments show that SIM has higher efficiency and stability than protocols that rely exclusively on selective participation. The salient properties of SIM are confirmed in experiments where bottleneck capacities fluctuate, bottlenecks migrate, and session membership is dynamic.

The rest of the paper is organized as follows. Section II discusses limitations of existing multicast congestion control solutions and presents a conceptual design of SIM. Section III describes our protocol in detail. Section IV reports the setup and results of the experimental study. Finally, Section V summarizes the contributions of the paper.

II. Conceptual Design

With multicast congestion control, there is a desire to support heterogeneous receiving capabilities. While a straightforward application of feedback-driven unicast congestion control [6] matches the single transmission rate to the capability of one subset of the receivers, the resulting transmission can greatly dissatisfy other receivers in the session. To improve efficiency, the mechanism of *selective participation* composes the session from multiple multicast groups and enables a receiver to subscribe to an appropriate subset of the groups [8]–[12]. If the number of the groups and group transmission rates are fixed, selective participation can still be inefficient due to misaligning the transmission rates with dynamic receiving capabilities or using an excessively large number of groups to offer a fine-grained menu of static transmission rates. To make congestion control efficient and scalable in heterogeneous dynamic environments, we enhance selective participation with the mechanisms of *intra-group transmission adjustment* [14] and *menu adaptation* [21]. To ensure that the multiple mechanisms complement each other, we operate the mechanisms at different timescales:

- Intra-group transmission adjustment operates at the round-trip timescale and allows the sender to adjust the group transmission rates in response to short-term changes in network conditions.
- Selective participation operates at an intermediate timescale and enables each receiver to subscribe to appropriate groups.
- Menu adaptation works at the slowest timescale to improve the overall structure of the session.

While intra-group transmission adjustment and menu adaptation rely on feedback to the sender, feedback implosion is a common concern in feedback-based multicast congestion control. Our protocol addresses the problem by aggregating the feedback information. To preserve privacy of session participants, we stipulate that the feedback aggregation should be done by routers.

Packet field	Setting by a receiver
source	receiver identifier
session/group identifier	address of the top subscribed group
count of feedback reports	1
count of “congested” reports	0 or 1 depending on the congestion status
minimum received rate	measured cumulative received rate
maximum received rate	measured cumulative received rate
subtree population	1
timestamp	measured timing information

TABLE I
SIM FEEDBACK PACKET FORMAT AND FIELD SETTINGS BY A RECEIVER.

III. Protocol Details

This section discusses our layered multicast congestion control protocol SIM in more detail. Throughout the discussion, a session refers to the entire set of groups involved in the communication of layered data from one sender to many receivers. We assume that the layers are cumulative in nature. A separate multicast group carries each layer.

A. Congestion Detection and Notification

SIM uses an *integrated approach* towards congestion detection and notification for its three control mechanisms. As with RED [24], we employ the exponentially weighted averaging technique to measure the queue sizes at the routers and the received rates at the receivers. The routers detect incipient congestion by using the queue size averages similarly to RED and ECN [25] and, on detecting congestion, mark a single bit in the headers of forwarded packets. These ECN bits notify each receiver about its congestion status and are used for *both* selective participation at the receivers and intra-group transmission adjustment at the sender. Furthermore, the receivers inform the sender about their received rates, and the sender utilizes the reported rates for menu adaptation. Table I describes the format of feedback packets sent by the receivers.

1) Packet Marking for Congestion Notification: On detecting congestion, one simple approach is to mark the ECN bits of all forwarded packets and thus notify all receivers behind the congested link. However, such notification is problematic with layered multicast because some of the notified receivers might not be contributing to the congestion, e.g., when their session serves another receiver behind the congested link, and that receiver has a higher subscription level. To propagate the congestion status to the appropriate receivers only, SIM employs *priority marking*: the router marks only the packets belonging to the top subscribed group of each session. While priority marking has been proposed earlier, e.g., to improve the perceived quality of layered video [26], [27], our novel application of the technique ensures *appropriate congestion notification* within layered multicast sessions.

2) Per-Session Router State: Priority marking would require that routers know the association between groups and sessions. For scalability, it is desirable to maintain little or no per-session state in the routers. The following *naming convention* enables SIM to reduce the required per-session state dramatically:

- Allocate contiguous addresses to the groups of a session.
- Use the last $\log_2(n_g)$ bits of a group address as the rank of the group within the session, where n_g is the maximum number of groups in any session.
- Use the remaining prefix of the address to identify the session of the group.

With this convention, the router needs to maintain only $\log_2(n_g)$ bits to indicate the top subscribed group of a session. While our experiments suggest that increasing n_g beyond 8 does not improve performance significantly, the default setting of SIM enlarges the forwarding table by at most 3 per-link per-session bits. Hence, SIM priority marking adds minimal state into the routers.

3) Communication Overhead: Whereas many protocols for streaming multimedia send a feedback packet upon any packet delivery [28], the resulting communication overhead might be excessive. Because one round-trip time constitutes the lowest useful timescale for end-to-end congestion control, the sender solicits feedback once per two round-trip times. This feedback timescale enables SIM to support both *low communication overhead* and *responsive congestion control*. To enable the feedback solicitation by the sender, the data packet header includes a feedback request field. The sender requests feedback *explicitly* by incrementing this field. Upon receiving a packet with an increased value in the feedback request field, the receiver transmits a feedback packet formatted as shown in Table I.

4) Scalability and Privacy: Feedback aggregation [18] and feedback timescale adjustment [29], [30] are two common solutions to the feedback implosion problem. From the privacy perspective, feedback timescale adjustment is undesirable because it allows a receiver to infer the session size while other session participants might prefer to keep this information out of the public domain. Similarly, aggregating feedback in a receiver is problematic because receiver-based aggregation might reveal information about other receivers. To avoid feedback implosion without endangering privacy of the session participants, SIM aggregates feedback in the routers. The feedback aggregation does not have to involve all the routers. Selected routers at the branching points of the multicast tree suffice for forming an aggregation tree [31]. In SIM, each aggregation router maintains a cache of feedback information received from its children in the aggregation tree. Upon receiving a data packet with an increased value in the feedback request field, the aggregation router starts an aggregation timer. When the timer expires or upon collecting responses from all its children, the aggregation router compiles a summary of feedback information for each group and sends the summaries to the parent in the aggregation tree. The aggregation router puts its identifier in the source field of the feedback packets. Thus, in addition to the feedback aggregation, the SIM routers provide the receivers with an anonymity service [32].

To support timely accurate feedback despite potential loss or queuing of feedback packets, each aggregation router sets its aggregation timeout to two maximum round-trip times between itself and the receivers in its aggregation subtree. Such settings ensure that the sender receives representative feedback within two maximum round-trip times of the session. Moreover, the communication overhead and delay of the congestion notification do not rise when the session receiver population or aggregation tree height increases.

B. Intra-group Transmission Adjustment

As the fastest mechanism of SIM, intra-group transmission adjustment carries the frontline brunt of congestion control. The sender requests feedback from the receivers once every two round-trip times. On receiving the aggregated feedback, the sender adjusts the group transmission rates using the additive-increase multiplicative-decrease algorithm [33]. To discover the complete range of available network capacities, SIM determines the congestion status for different groups differently. The bottom group satisfies the least capable receivers and is considered congested when *at least one* receiver reports being congested. The top group satisfies the most capable receivers and is viewed as congested only when *all* reports for this group indicate congestion. Transmission adjustment in the intermediate groups depends on the menu status. If the menu is inactive, an intermediate group is considered congested when at least one of its receivers reports being congested. Once the menu is active, an intermediate group is treated as congested when a *majority* of reports for this group indicate congestion. Modern multimedia encoders enable SIM to adjust group transmission rates quickly and with fine granularity, e.g., by using layered coding [34] or modifying quantization parameters of compressed video [35]–[37].

C. Selective Participation

Unlike in purely receiver-driven protocols such as RLM [8], RLC [9], FLID [10], and WEBRC [12] where selective participation is the only congestion control mechanism, selective participation in SIM does not react to short-term changes in network conditions because intra-group transmission adjustment that operates at the shortest timescale adapts to such changes. Instead, SIM employs selective participation to improve long-term performance. In particular, a receiver changes its subscription only after intra-group transmission adjustment persistently fails to accommodate its receiving capability. We refer to the timescale for selective participation as a *selection interval*. To preclude interference with intra-group transmission adjustment without undermining the responsiveness of the control, the default setting of SIM uses twenty maximum round-trip times as the selection interval.

If a receiver is not congested over the last selection interval, and the rate of its top subscribed group decreases, then the receiver adds the group that is right above its top subscribed group. On the other hand, if the subscription contains at least

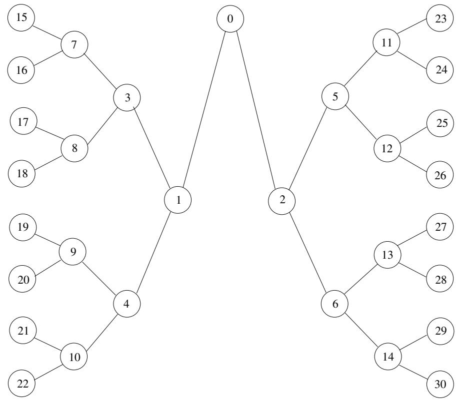


Fig. 1. Network topology.

d	Round-trip propagation delay
t	Time since the beginning of the session
t_f	Session duration
$n_f = \lfloor \frac{t_f}{d} \rfloor$	Number of complete intervals of duration d in a session
n_g	Maximum number of groups in a session
$n_b(t)$	Number of different bottleneck capacities at time t
$s(t)$	Set of receivers in a session at time t
$n(t) = \|s(t)\|$	Number of receivers in a session at time t
t_b	Interval between changes in bottleneck capacities
$p = \frac{t_b}{d}$	Number of round-trip delays between capacity changes
t_m	Interval between changes in the session membership
$m = \frac{t_m}{d}$	Number of round-trip delays between membership changes
$a_i(t)$	Available capacity for receiver i at time t
$r_i(t)$	Received rate for receiver i at time t
b	Ratio of the capacity range to the transmission range

TABLE II
DEFINITIONS OF THE EXPERIMENTAL VARIABLES.

two groups, the receiver is congested over the last selection interval, and the rate of its top subscribed group increases, the receiver drops its top subscribed group.

D. Menu Adaptation

To improve the session efficiency, SIM uses a menu that regulates the group transmission rates. Menu adaptation ensures that the regulation is sensitive to changes in network conditions. Initially, the menu is inactive, and intra-group transmission adjustment allows the sender to discover the full range of available network capacities, as discussed in Section III-B. If the top group of the session becomes congested, the number of groups is at most the number of different bottleneck capacities. On detecting this condition, menu adaptation changes the menu status to active and splits the range between the minimum and maximum received rates into $n_g - 2$ equal subranges, where n_g is the maximum number of groups in a session. The sender uses the upper boundaries of these subranges as the maximum cumulative transmission rates for the corresponding intermediate groups. Once the menu is active, menu adaptation continuously adjusts the subrange boundaries in response to the reported received rates of the least and most capable receivers. By doing so, menu adaptation improves the overall alignment of the transmission rates with dynamic receiving capabilities. To preclude interference with selective participation without undermining the responsiveness of the control, the default setting of SIM uses fifty maximum round-trip times as the menu adaptation interval.

IV. EXPERIMENTAL EVALUATION

We conduct the evaluation in ns-2 network simulator [38]. Figure 1 illustrates our network topology, a balanced tree where node 0 is the sender, the leaf nodes are the receivers, and the internal nodes aggregate feedback. All the links have such equal propagation delays that the maximum round-trip propagation time d between the sender and the receivers equals 40 ms, which is roughly the round-trip propagation delay between the east and the west coasts of the United States. For example, each link in our topology instance with 16 receivers has propagation delay of 5 ms. Table II summarizes variables in our experiments. To study the impact of changes in available network capacities, we simulate the changes directly by changing the capacities of links; this paper does not report similar results from our experiments where the available capacities change due to cross traffic. Also because of space constraints, we omit our scalability and some other experimental results and refer the reader to the earlier longer version of this work [39].

A. Performance Metrics

During each experiment, we track the following: (a) group transmission rates, (b) group subscriptions, (c) received rates, and (d) maximum cumulative group rates as specified by the menu. Besides, we evaluate the effectiveness of network utilization and the stability of subscription levels by means of the following metrics:

- *Efficiency* $E(t)$ is the average utilization at time t:

$$E(t) = \frac{1}{n(t)} \sum_{i \in s(t)} \frac{r_i(t)}{a_i(t)}; \quad 0 \le E(t) \le 1$$

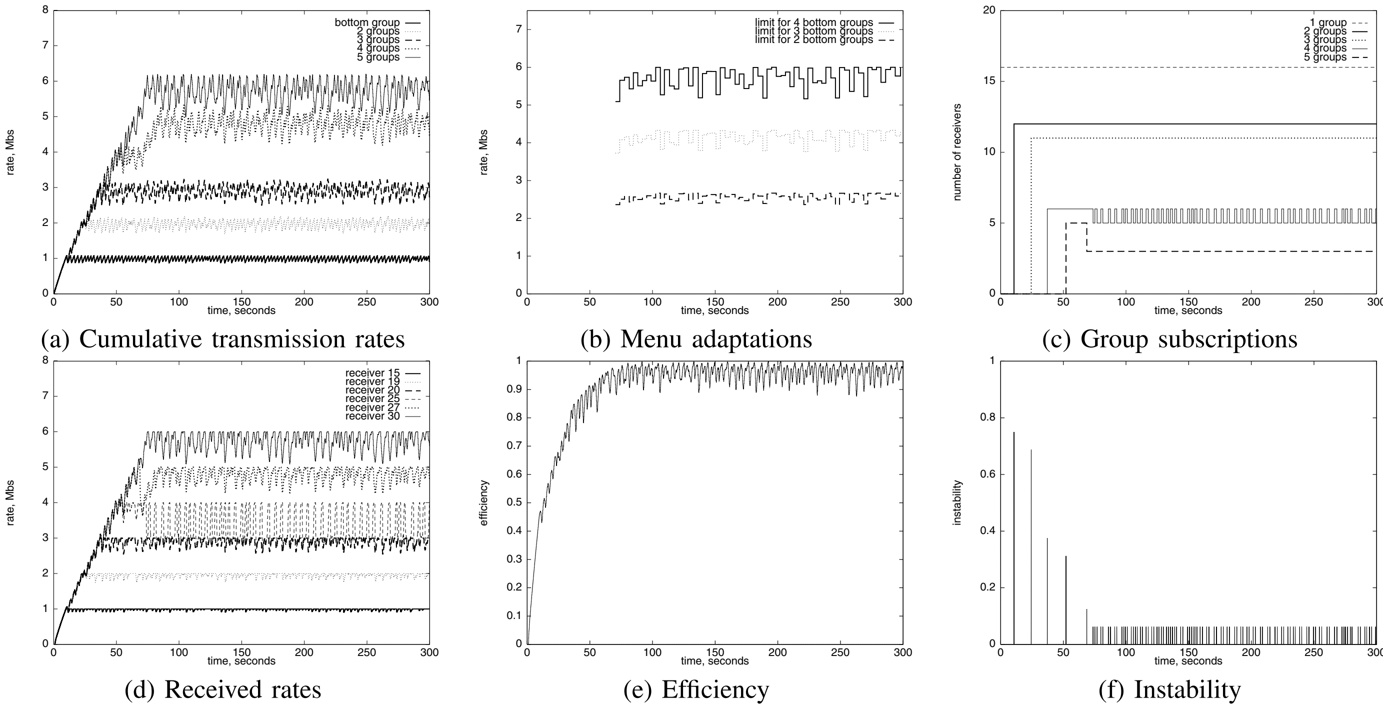
  We also compute average efficiency $\widehat{E}$ and deviation in efficiency $\tilde{E}$ as follows:

$$\widehat{E} = \frac{1}{t_f} \int_0^{t_f} E(t)dt; \quad \tilde{E} = \sqrt{\frac{1}{t_f} \int_0^{t_f} (E(t) - \widehat{E})^2 dt}$$

- *Instability*: We define instability $I(t)$ as the number of per-receiver subscription changes during interval $[t-d, t)$. We compute the average instability $\widehat{I}$ and the deviation in instability $\tilde{I}$ as follows:

$$\widehat{I} = \frac{1}{n_f} \sum_{j=1}^{n_f} I(jt); \quad \tilde{I} = \sqrt{\frac{1}{n_f} \sum_{j=1}^{n_f} (I(jt) - \widehat{I})^2}$$

  Instability is an important metric because of being directly related to the perceived quality of live video multicast and other continuous media applications. Smaller values of instability result in greater user satisfaction.



(a) Cumulative transmission rates (b) Menu adaptations (c) Group subscriptions

(d) Received rates (e) Efficiency (f) Instability

Fig. 2. Understanding the behavior of a SIM session with $n_b(t) = 6$, $n_g = 5$, $n(t) = 16$, and $t_f = 300$ seconds.

B. Session Behavior

First, we examine a SIM session with 16 receivers and a maximum of 5 groups, i.e., $n(t) = 16$ and $n_g = 5$. We configure links 1-3, 9-19, 1-4, 5-11, 12-25, and 6-13 to have capacities of 1, 2, 3, 3, 4, and 5 Mbps respectively; all other links have a capacity of 6 Mbps. This configuration with six different bottleneck capacities [i.e., $n_b(t) = 6$] has 4, 1, 5, 1, 2, and 3 receivers behind the 1-Mbps, 2-Mbps, 3-Mbps, 4-Mbps, 5-Mbps, and 6-Mbps bottleneck links respectively.

Figure 2a illustrates how intra-group transmission adjustment and selective participation enable the session to discover its highest available capacity of 6 Mbps. After the initial convergence, the session activates the menu as reflected in Figure 2b. According to the changed rules for intra-group transmission adjustment, the fourth group becomes governed not by its slowest 4-Mbps receiver but by the two receivers behind the 5-Mbps bottleneck. Figure 2a captures the respective increase in the cumulative transmission rate of the four bottom groups from 4 Mbps to 5 Mbps.

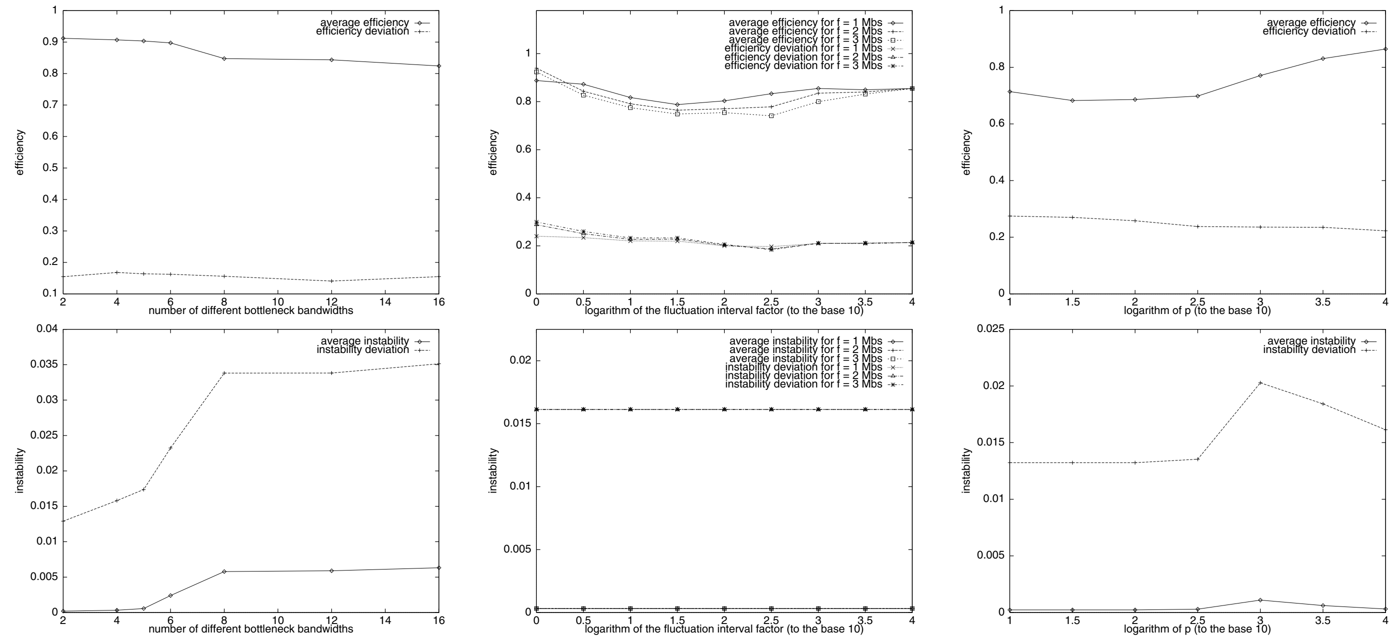
Subsequently, the session operates in a steady state where the cumulative transmission rates are 1 Mbps, 2 Mbps, 3 Mbps, 5 Mbps, and 6 Mbps, as shown in Figure 2a. Figure 2c depicts the steady-state group subscriptions: all sixteen receivers are subscribed to the bottom group; all but the four 1-Mbps receivers are subscribed to the second group; the three 6-Mbps, two 5-Mbps, one 4-Mbps, and five 3-Mbps receivers are subscribed to the third group; the 6-Mbps and 5-Mbps receivers are subscribed to the fourth group; only the 6 Mbps receivers are subscribed to the top group. As Figures 2c and 2d show, the subscription and received rate of the 4-Mbps receiver 25 fluctuate. Despite the fluctuations, Figures 2e and 2f demonstrate that the SIM session sustains high efficiency and low instability in the steady state.

C. Heterogeneity in Bottleneck Capacities

In this series of experiments, we split the range between 1 Mbps and 6 Mbps uniformly into $n_b(t)$ values and use these values as capacities for the links incident on the receivers so that each capacity gets assigned to an approximately equal fraction of such links. All other links have a capacity of 6 Mbps. Figure 3a reveals that SIM provides high efficiency and stable subscriptions after the initial convergence when $n_b(t) \leq n_g$. If $n_b(t)$ exceeds the maximum number of groups, SIM still converges to a steady state but efficiency declines and instability increases. These results indicate that SIM operates the best when different bottleneck capacities do not outnumber n_g.

D. Fluctuation of Bottleneck Capacities

To examine scenarios where the bottlenecks do not migrate but merely change their capacities, we set all link capacities to 9 Mbps except for links 2-5 and 4-9. Once per $p*d$ seconds, the capacity of link 2-5 alternates between 1 Mbps and $f+1$ Mbps, and the capacity of link 4-9 fluctuates between 5 Mbps and $f+5$ Mbps. In these experiments, f equals 1 Mbps, 2 Mbps, or 3 Mbps, and p varies from 1 to $10,000$, i.e., the period of the capacity fluctuations varies from one round-trip delay to the whole experiment duration. Figure 3b shows that SIM persistently maintains high efficiency and constant instability. Intra-group transmission adjustment successfully addresses the



(a) Capacity heterogeneity (b) Capacity fluctuation (c) Bottleneck migration

Fig. 3. Heterogeneous capacities and changing bottlenecks with $n(t) = 16$ and $t_f = 300$ seconds: (a) $n_g = 5$ and variable $n_b(t)$; (b) $n_g = 3$ and $n_b(t) = 3$; (c) $n_g = 3$ and $n_b(t) = 3$.

capacity fluctuations without triggering the selective participation mechanism. This favorably distinguishes SIM from the purely receiver-driven protocols where bottleneck capacity fluctuations can cause massive changes in group subscriptions.

E. Migrating Bottlenecks

We study the impact of bottleneck migration by setting all link capacities to 9 Mbps except for links 0-2, 2-5, 1-4, and 4-9. Links 2-5 and 4-9 have capacities of 3 Mbps and 7 Mbps respectively. Once every p round-trip delays, the capacity of link 0-2 fluctuates between 1 Mbps and 9 Mbps while the capacity of link 1-4 alternates between 5 Mbps and 9 Mbps. The capacity fluctuations ensure that (1) the bottleneck for receivers 19 and 20 migrates between links 1-4 and 4-9, (2) the bottleneck for receivers 21 and 22 migrates between links 1-4 and 0-1, (3) and the bottleneck for receivers 23 through 26 migrates between links 0-2 and 2-5. The overall number of different bottleneck capacities remains equal to three. Figure 3c shows that the migration at a shorter timescale than the selection interval allows the receivers to maintain high efficiency via adjustment of their subscriptions. When the migration is more frequent, SIM maintains stable but suboptimally efficient subscriptions.

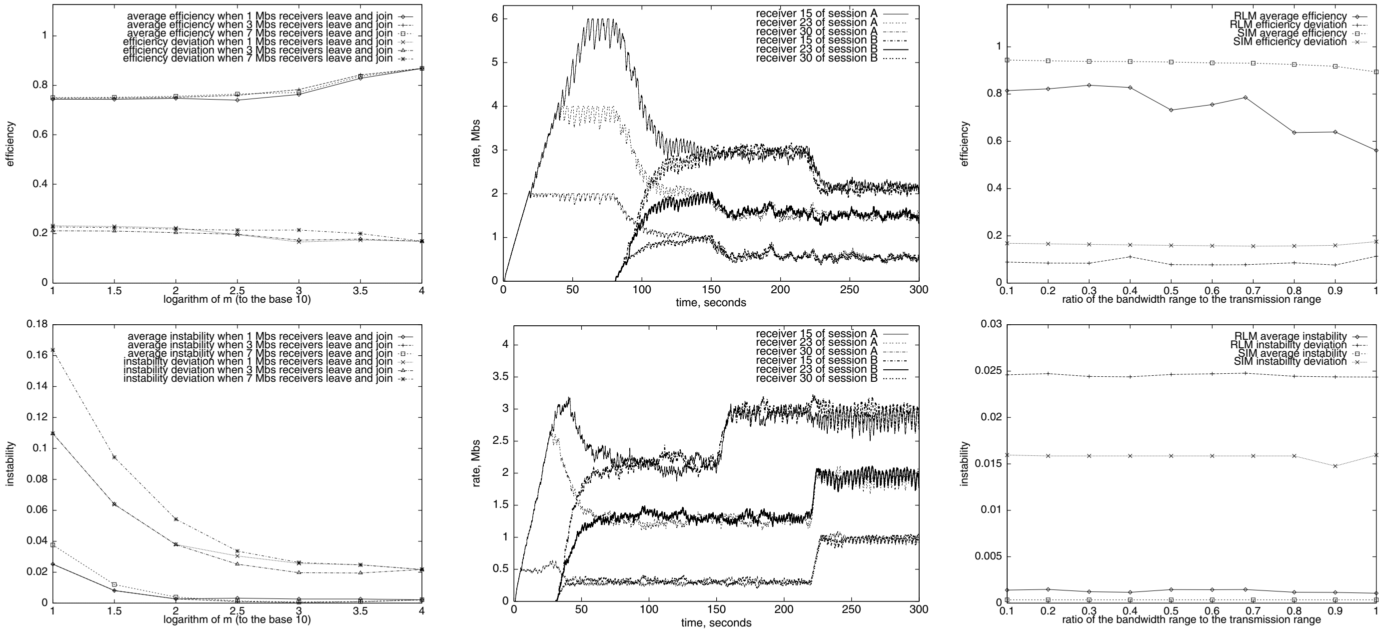
F. Changes in Session Membership

Changes in session membership impact performance the most when they modify the menu. Hence, we conduct experiments where the receivers that dictate group transmission rates join or leave the session synchronously. Links 2-5, 1-3, and 4-9 have capacities of 1 Mbps, 3 Mbps, and 5 Mbps respectively. We set all other link capacities to 9 Mbps. Once every m round-trip delays, all the receivers behind a specified bottleneck link synchronously join or leave the session. We vary m from 10 to $10,000$, i.e., the period of the membership fluctuations varies from the timescale of selective participation to the whole experiment duration. Figure 4a shows that SIM persistently maintains high efficiency: the top three and bottom three curves depict respectively average efficiencies and efficiency deviations. For small m, the dynamic session membership triggers both group subscription changes and menu adaptation, and instability is high. As the membership changes become less frequent, efficiency increases and instability drops.

G. Fairness

We examine fairness of SIM in presence of TCP [6] cross traffic and check whether the protocols provide equivalent services to the receivers behind the same bottleneck links. Links 0-1, 0-2, and 6-14 have capacities of 6, 4, and 2 Mbps respectively. All other link capacities are configured to be 10 Mbps. Two SIM sessions A and B multicast data to all the sixteen leaf nodes of the topology. Thus, receiver 15 and seven other receivers lie behind the 6 Mbps bottleneck, receiver 23 and five more receivers get the data through the 4 Mbps bottleneck while receiver 30 and the remaining receiver are served through the 2 Mbps bottleneck. Besides, two TCP sessions TCP-15 and TCP-30 transfer files from node 0 to receivers 15 and 30 respectively.

Figure 4b tracks the received rates of representative SIM receivers in the following two scenarios: (1) session A starts at time 0 and converges to a steady state before session B starts, and sessions TCP-30 and TCP-15 start at time 150



(a) Dynamic session membership (b) Fairness (c) Comparison of SIM with RLM

Fig. 4. Dynamic session membership, fairness, and comparison with purely receiver-driven control with $t_f = 300$ seconds: (a) $n_g = 3$; (b) $n_b(t) = 3$, $n(t) = 16$, $n_g = 3$, session B starts after session A converges in the top graph, session B starts before session A converges in the bottom graph; (c) $n_b(t) = 5$, $n(t) = 16$, and $n_g = 5$.

and 220 seconds respectively; (2) sessions A, TCP-15, and TCP-30 start at time 0 but reach a steady state only after session B starts, and sessions TCP-15 and TCP-30 stop at time 150 and 220 seconds respectively. In both scenarios, the receivers of the SIM sessions discover optimal subscriptions, react appropriately to the changes in the available capacities, and their received rates converge to fair shares.

H. Comparison with Purely Receiver-driven Control

In this section, we compare SIM with RLM because the latter is a classical protocol for layered multicast. Our comparison emphasizes the fundamental reliance of purely receiver-driven control on selective participation, rather than any specific details of the RLM design. Hence, we believe that our qualitative conclusions hold for other purely receiver-driven protocols such as RLC [9], FLID [10], and WEBRC [12].

With purely receiver-driven control, the group transmission rates are predetermined, and efficiency can suffer from a mismatch between the transmission rates and dynamic receiving capabilities. To measure this effect, we quantify the mismatch with ratio b of the capacity range to the transmission range, where the capacity range refers to the difference between the highest and lowest bottleneck capacities, and the transmission range is the difference between the maximum and minimum transmission rates offered by an RLM session. In this series of experiments, the bottlenecks are links 1-3, 1-4, 2-5, 14-30, and 2-6 listed in the increasing order of their capacities. To select the capacities, we consider the uniform distribution over the interval centered at 3 Mbps and sized according to b. All the other links have a capacity of 6 Mbps. One multicast session transmits data to the whole set of sixteen receivers. Under RLM, the session uses n_g groups with cumulative transmission rates that are uniformly distributed from 1 Mbps to 5 Mbps, i.e., the bottom group transmits at a rate of 1 Mbps while each of the upper groups sends at a rate of $\frac{4}{n_g-1}$ Mbps.

Figure 4c reports experimental results for scenarios where n_g is set to 5, and b varies from 0.1 to 1. SIM is consistently superior to RLM with respect to efficiency and stability. While the cumulative transmission rates under SIM converge to the bottleneck capacities, efficiency under RLM suffers from the mismatch between the bottleneck capacities and predetermined transmission rates. When b is at least 0.8, and the capacity of link 1-3 is below 1 Mbps, receivers 15, 16, 17, and 18 are unable to sustain even the rate of the bottom group, and efficiency under RLM drops to about 60%. Moreover, instability under RLM is higher because unlike under SIM, receivers keep trying to join a higher group even after they reach the optimal subscription.

Our additional experiments show that increasing the number of groups does not allow RLM to close the performance gap with SIM. Also, the increase affects negatively RLM convergence speed and operational overhead. We omit these extra results due to space constraints.

V. Conclusion

In this paper, we showed the feasibility of combining three distinct mechanisms – Selective participation, Intra-group transmission adjustment, and Menu adaptation – into a single protocol SIM that provides a general solution for efficient fair scalable privacy-preserving multicast congestion control with heterogeneous receivers. By adjusting the group transmission rates to match the available network capacities, SIM offers higher efficiency and stability than receiver-driven protocols that rely only on selective participation. Whereas the individual mechanisms of SIM have appeared in earlier multicast protocols, our main contribution lies in the cohesive integration of the techniques. SIM achieves such integration of the control mechanisms by operating them at different timescales and ensuring that mechanisms working on longer timescales complement and benefit from the control wielded by faster mechanisms. To instantiate and integrate the three mechanisms, SIM introduces techniques of independent interest. In particular, SIM congestion detection and notification demonstrates that explicit signaling can be successfully applied in the context of multicast congestion control. Also, SIM feedback aggregation exhibits the prominent feature of providing the sender with timely accurate feedback in a scalable manner without jeopardizing privacy of the session participants. Our extensive evaluation confirms the salient properties of SIM in diverse settings where receiving capabilities are highly heterogeneous, bottleneck capacities fluctuate, bottlenecks migrate, and session membership changes.

REFERENCES

[1] X. Hei, C. Liang, J. Liang, Y. Liu, and K. Ross, "Insights into PPLive: A Measurement Study of a Large-Scale P2P IPTV System," in *Proceedings IPTV*, May 2006.

[2] S. Annapureddy, C. Gkantsidis, and P. Rodriguez, "Providing Video-on-Demand using Peer-to-Peer Networks," in *Proceedings IPTV*, May 2006.

[3] S. Deering, "Multicast Routing in a Datagram Internetwork," Ph.D. dissertation, Stanford University, December 1991.

[4] Y. Chu, S. G. Rao, S. Seshan, and H. Zhang, "A Case for End-System Multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1456 – 1471, October 2002.

[5] R. D. Doverspike, G. Li, K. Oikonomou, K. K. Ramakrishnan, and D. Wang, "IP Backbone Design for Multimedia Distribution: Architecture and Performance," in *Proceedings IEEE INFOCOM*, May 2007.

[6] M. Allman, V. Paxson, and W. Stevens, "TCP Congestion Control," RFC 2581, April 1999.

[7] E. Kohler, M. Handley, and S. Floyd, "Designing DCCP: Congestion Control Without Reliability," in *Proceedings ACM SIGCOMM*, September 2006.

[8] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven Layered Multicast," in *Proceedings ACM SIGCOMM*, August 1996.

[9] L. Vicisano, L. Rizzo, and J. Crowcroft, "TCP-like Congestion Control for Layered Multicast Data Transfer," in *Proceedings IEEE INFOCOM*, March 1998.

[10] J. Byers, G. Horn, M. Luby, M. Mitzenmacher, and W. Shaver, "FLID-DL: Congestion Control for Layered Multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1558 – 1570, October 2002.

[11] J. Byers, G. Kwon, M. Luby, and M. Mitzenmacher, "Fine-Grained Layered Multicast with STAIR," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 81 – 93, February 2006.

[12] M. Luby, V. Goyal, S. Skaria, and G. Horn, "Wave and Equation Based Rate Control Using Multicast Round Trip Time," in *Proceedings ACM SIGCOMM*, August 2002.

[13] S. Gorinsky, S. Jain, H. Vin, and Y. Zhang, "Design of Multicast Protocols Robust against Inflated Subscription," *IEEE/ACM Transactions on Networking*, vol. 14, no. 2, pp. 249–262, April 2006.

[14] L. Rizzo, "pgmcc: A TCP-friendly Single-Rate Multicast Congestion Control Scheme," in *Proceedings ACM SIGCOMM*, August 2000.

[15] G. Iannaccone and L. Rizzo, "Fairness of a Single-Rate Multicast Congestion Control Scheme," in *Proceedings IWDC*, September 2001.

[16] A. Legout and E. W. Biersack, "PLM: Fast Convergence for Cumulative Layered Multicast Transmission Schemes," in *Proceedings ACM SIGMETRICS*, June 2000.

[17] D. Sisalem and A. Wolisz, "MLDA: A TCP-friendly Congestion Control Framework for Heterogeneous Multicast Environments," in *Proceedings IWQoS*, June 2000.

[18] S. Paul, K. Sabnani, J. Lin, and S. Bhattacharyya, "Reliable Multicast Transport Protocol (RMTP)," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 3, April 1997.

[19] S. Gorinsky and H. Vin, "The Utility of Feedback in Layered Multicast Congestion Control," in *Proceedings Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2001)*, June 2001.

[20] J. Widmer and M. Handley, "Extending Equation-Based Congestion Control to Multicast Applications," in *Proceedings ACM SIGCOMM*, August 2001.

[21] S. Y. Cheung and M. H. Ammar, "Using Destination Set Grouping to Improve the Performance of Window-controlled Multipoint Connections," *Computer Communications Journal*, vol. 19, no. 8, pp. 723–736, July 1996.

[22] S. Gorinsky, S. Jain, and H. Vin, "Multicast Congestion Control with Distrusted Receivers," in *Proceedings Networked Group Communication (NGC 2002)*, October 2002.

[23] N. Duffield, M. Grossglauser, and K. K. Ramakrishnan, "Distrust and Privacy: Axioms for Multicast Congestion Control," in *Proceedings NOSSDAV*, June 1999.

[24] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, August 1993.

[25] K. K. Ramakrishnan and S. Floyd, "A Proposal to Add Explicit Congestion Notification (ECN) to IP," RFC 2481, January 1999.

[26] S. Bajaj, L. Breslau, and S. Shenker, "Uniform versus Priority Dropping for Layered Video," in *Proceedings ACM SIGCOMM*, August 1998.

[27] R. Gopalakrishnan, J. Griffioen, G. Hjalmtysson, C. Sreenan, and S. Wen, "A Simple Loss Differentiation Approach to Layered Multicast," in *Proceedings IEEE INFOCOM*, March 2000.

[28] R. Rejaie, M. Handley, and D. Estrin, "RAP: An End-to-end Rate-based Congestion Control Mechanism for Realtime Streams in the Internet," in *Proceedings IEEE INFOCOM*, March 1999.

[29] J.-C. Bolot, T. Turletti, and I. Wakeman, "Scalable Feedback Control for Multicast Video Distribution in the Internet," in *Proceedings ACM SIGCOMM*, October 1994.

[30] H. Schulzrinne and J. Rosenberg, "Timer Reconsideration for Enhanced RTP Scalability," in *Proceedings IEEE INFOCOM*, March 1998.

[31] T. Speakman, J. Crowcroft, J. Gemmell, D. Farinacci, S. Lin, D. Leshchiner, M. Luby, T. Montgomery, L. Rizzo, A. Tweedly, N. Bhaskar, R. Edmonstone, R. Sumanasekera, and L. Vicisano, "PGM Reliable Transport Protocol Specification," RFC 3208, December 2001.

[32] M. Reiter and A. Rubin, "Crowds: Anonymity for Web Transactions," *ACM Transactions on Information and System Security*, vol. 1, no. 1, pp. 66–92, November 1998.

[33] D. Chiu and R. Jain, "Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks," *Journal of Computer Networks and ISDN*, vol. 17, no. 1, pp. 1–14, June 1989.

[34] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple Description Coding Using Pairwise Correlating Transforms," *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 351–366, March 2001.

[35] N. Duffield, K. K. Ramakrishnan, and A. Reibman, "SAVE: an Algorithm for Smoothed Adaptive Video over Explicit Rate Networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 6, pp. 717–728, December 1998.

[36] T. Lakshman, P. Mishra, and K. K. Ramakrishnan, "Transporting Compressed Video over ATM Networks with Explicit-Rate Feedback Control," *IEEE/ACM Transactions on Networking*, vol. 7, no. 5, pp. 710–723, October 1999.

[37] B. Vandalore, W. Feng, R. Jain, and S. Fahmy, "A Survey of Application Layer Techniques for Adaptive Streaming of Multimedia," *Real-Time Imaging*, vol. 7, no. 2, pp. 221–235, June 2001.

[38] S. McCanne and S. Floyd, *ns Network Simulator. http://www.isi.edu/nsnam/ns/.*

[39] S. Gorinsky, K. K. Ramakrishnan, and H. Vin, "Addressing Heterogeneity and Scalability in Layered Multicast Congestion Control," Department of Computer Sciences, University of Texas at Austin, Tech. Rep. TR2000-31, November 2000.