

Random Early Detection Assisted Layered Multicast

Yung-Sze Gan and Chen-Khong Tham

Department of Electrical and Computer Engineering
National University of Singapore, Singapore 119260
{engp0515, e1etck}@nus.edu.sg

Abstract. The deployment of the Random Early Detection (RED) algorithm in multicast routers can improve the performance of layered multicast congestion control schemes. In this paper, a new layered multicast protocol called RED Assisted Layered Multicast (RALM) is proposed to realise the integration of RED and layered multicast. It utilises the packet marking feature and two-level drop prioritisation capability in the Differentiated Services (DiffServ) architecture to differentiate losses in the layers of an RALM session. By marking the layers' packets appropriately and dropping them differently during congestion, the RALM protocol guides all receivers to a stable optimal subscription level that satisfies their common bandwidth requirement within a short period of time. Moreover, the DiffServ architecture permits the network administrator to control the amount of bandwidth consumed by RALM sessions through the deployment of suitable queue schedulers in the routers.

1 Introduction

The widespread employment of congestion control protocols like TCP in unicast IP networks has helped them to perform well in the face of ever increasing amount of traffic. Similarly, the success of multicast networks is dependent on the implementation of equivalent congestion control schemes. Possible candidates can be drawn from layered multicast protocols like Receiver-driven Layered Multicast (RLM) [1], Receiver-driven Layered Congestion Control (RLC) [2] and Packet Pair RLM (PLM) [3].

In layered multicast, a server encodes data into cumulative layers (i.e. a layer is decoded only if its lower layers are present) that are multicasted through separate addresses/ports. The clients adapt to their bottleneck bandwidths by subscribing to the maximum number of layers that can traverse the bottleneck links. A receiver discovers this subscription level by periodically simulating the addition of a layer to its current subscription. If the addition induces losses in its reception, it indicates that the receiver should not raise its current subscription level, and it has reached its optimal subscription level which fully utilises its bottleneck bandwidth. In current networks where droptail queues are deployed in the routers, losses occur because the queues are full and subsequent incoming packets are dropped. Thus, by the time the receivers detect and react to the

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-3-540-45812-8_28](https://doi.org/10.1007/978-3-540-45812-8_28)

K.C. Almeroth and M. Hasan (Eds.): MMNS 2002, LNCS 2496, pp. 341–353, 2002.
© IFIP International Federation for Information Processing 2002

losses, congestion would have occurred for some time. The resulting loss rate is unnecessarily high. It would be ideal if the routers can signal to the receivers that congestion is approaching by dropping some packets before their queues are full, thus permitting the receivers to lower their subscription levels to avoid the congestion. This can be accomplished by using the Random Early Detection (RED) [4] algorithm to manage the queues in the routers.

A RED queue provides early congestion indications by randomly dropping packets when a certain proportion of the queue is filled. Thus, traffic sources can reduce their transmission rates before actual congestion occurs. The packet dropping probability is calculated based on the average queue length and not on the current queue length. Hence, a RED queue can tolerate some traffic bursts due to fluctuations in the transmission rates without dropping any packets. Moreover, the random packet drops are performed across all sessions passing through the router which helps to ensure that burst losses do not occur within any session. As RED queues are implemented in routers and are transparent to all congestion control schemes, they can work with any layered multicast protocol without requiring any modification.

The focus of this paper is on exploring the possibilities of realising layered multicast congestion control with the help of intermediate routers that implement RED queues. However, simply running the RLM protocol over a RED enabled network does not improve the protocol's performance. In section 2, we will look at why the RED algorithm by itself does not help the RLM protocol. In addition, the idea of using the RED algorithm to provide loss differentiation in a layered multicast scheme to improve its performance will be introduced. In section 3, a new layered multicast protocol, RED Assisted Layered Multicast (RALM) that uses loss differentiation in its operations will be presented. Finally, some simulation results on the RALM protocol will be presented and implications raised by the simulations will be discussed before this paper is concluded.

2 Loss Differentiation and Layered Multicast

The RLM protocol is known to exhibit some pathological behaviours [5] and the mechanisms underlying these behaviours would also interfere with the RED algorithm. They are described in the following list:

1. A receiver probes for spare bandwidth to consume by adding a new layer to its current subscription level in a join experiment. The experiment fails when a packet loss is detected by the receiver and the newly subscribed layer is dropped immediately. Obviously, deploying a RED queue in the bottleneck router benefits the RLM protocol by providing early congestion indication before the queue is saturated. However, the router does not immediately stop forwarding the dropped layer when it receives an IGMP [6] leave request from the receiver. The router must first ensure that the layer is not subscribed by other receivers through a query/timeout mechanism before executing the leave request. The resulting multicast leave latency can be long. As a result, the benefit of providing early congestion signals is lost.

2. A layered multicast session's transmission rate is adjusted by adding or dropping a layer. Normally, a layer carries a significant portion of the total traffic generated by the server. An increase in the subscription level can result in rapid depletion of the bottleneck router's RED queue due to a rapid increase in the traversing traffic. Coupled with the problem of long multicast leave latency, the probability of a RED queue degenerating into a droptail queue during join experiments is very high. Thus, operating the RLM protocol over a RED enabled network only improves its loss rate marginally.
3. An RLM receiver is said to have reached a stable state when its join timer is sufficiently long compared to the duration of a join experiment. In this state, the receiver transits to a hysteresis state when it detects a packet drop, and waits for a detection timer's period before measuring the loss rate. This unresponsive period may be long enough that the RED queue overflows. The loss rate sampled will thus be similar to that produced by a droptail queue. Moreover, the receiver does not react to early congestion signals. As a result, a new RLM session cannot grab sufficient bandwidth from a pre-existing session by conducting join experiments to achieve max-min fairness [7] in the bandwidth share. Therefore, enhancing the bottleneck router with a RED queue does not improve the inter-session fairness of the RLM protocol.

To ensure that the RLM protocol works well with the RED algorithm, it is essential to replace the concepts of join experiments and a high loss tolerant stable state with an alternative spare bandwidth probing method that reacts fast to early congestion indications through fine-grained layer transmission rate adjustments without jeopardising the stability of the protocol.

The basic idea of the RALM protocol is to group the layers within the session's optimal subscription level together and mark their packets with a high forwarding priority. Other layers that are beyond the optimal subscription level will have their packets marked with a low forwarding priority. As the bottleneck router forwards the high priority layers, called protected layers, at the expense of the low priority layers, the layers that constitute the optimal subscription level will be protected from losses induced by oversubscription or traffic fluctuations. These losses are absorbed by the low priority layers, known also as sacrificial layers. Since low priority packets received by the RALM receiver can be decoded, the loss absorption capability of sacrificial layers can be seen as a form of fine-grained layer transmission rate adjustment.

A receiver increases its subscription level based on two different loss rates sampled in its current subscription. These rates vary as RED queues drop packets to indicate looming congestion in the bottleneck router, making the RALM protocol highly sensitive to early congestion indications. The highest current subscribed layer is used to probe for new bandwidth by comparing its loss rate against a loss threshold for the receiver to decide whether to add a higher layer. On the otherhand, the aggregated loss rate sampled in all the lower subscribed layers is used by the receiver to decide whether to reduce its subscription level.

At startup, a receiver samples both loss rates from high priority layers as it has not reached the optimal subscription level yet. Thus, the loss rates are

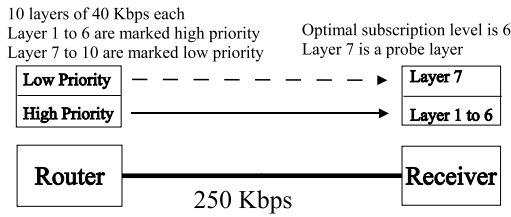


Fig. 1. The basic idea behind the RALM protocol

low and the receiver will increase its subscription level. The receiver raises its subscription level repeatedly until it subscribes a low priority layer. At this point, it has reached the optimal subscription level, excluding the low priority layer, as shown in Figure 1. The losses induced by the oversubscription of one layer will be concentrated in the low priority layer. Correspondingly, the loss rate in the layers that compose the optimal subscription level is very low, stabilising the subscription level by preventing any further layer addition or drop. If new bandwidth is available, the loss rate sampled in the subscribed low priority layer will drop below a loss threshold and the subscription level can be increased. No timer is set to probe for spare bandwidth as in join experiments. Due to the role of the highest current subscribed layer in the discovery of new bandwidth, it is called the probe layer in the RALM protocol.

The idea of applying two-level differentiation on layered multicast is not new. A similar concept is suggested by Gopalakrishnan et al. [8] in their Receiver-Selectable Loss Priorities (RSLP) and Receiver-driven Layered Multicast with Priorities (RLMP) protocols. In these protocols, the multicast routers are modified to support two-level drop priority in their forwarding path and the IGMP protocol is extended to permit the receiver to indicate the drop priority for each layer it subscribes. These modifications are not required in the RALM protocol. Instead, the drop priority is set by the server based on feedbacks provided by the receivers. The multicast routers simply forward the layers based on their packet markings. No intelligence is needed in the routers to determine the forwarding priority of the layers. The loss rate in the RLMP protocol is smoothed through the use of an exponential weighted moving average (EWMA). In contrast, the loss rate is used directly in the RALM protocol as it has already been smoothed by the RED algorithm operating in the routers. Finally, the loss thresholds used by receivers to decide subscription level adjustments are determined differently in the RLMP and RALM protocols.

3 RED Assisted Layered Mutlicast

In the RALM protocol, an algorithm that provides loss differentiated congestion control functionality is implemented in every receiver. This algorithm requires the cooperation of the server and routers in providing layer differentiation through the use of two-level priority packet marking and dropping mechanisms. A separate protocol that defines how the session sets the proper layer marking is

required to permit the receivers to reach a fair and stable optimal subscription level. In this paper, the focus is on the receiver driven congestion control algorithm and not on the layer marking discovery protocol, which will be investigated in a later work.

3.1 Two-Level Drop Priority Support in IP Networks

The IETF has defined the Differentiated Services (DiffServ) architecture [9] and two per-hop forwarding behaviours (PHB) to provide packet service differentiation in current IP networks. Of the two defined PHBs, the assured forwarding (AF) PHB groups [10] provide 3 drop priority levels within 4 AF PHB classes. The AF PHB classes are implemented as four separate physical queues in a DiffServ router in addition to the best-effort droptail queue. The drop priorities in each AF class are in turn implemented as virtual queues in each AF queue through the use of a 3-level RED algorithm. By using a single AF PHB class (e.g. AF1x) and restricting the number of drop priority used to just two (e.g. AF11 and AF12), two-level loss differentiation can be realised in routers.

In an RALM session, the server generates and marks the layers' packets with either an AF11 or AF12 DiffServ codepoints (DSCP) before transmitting them into the network. RFC 2597 defines AF11 to have a higher forwarding priority than AF12. In other words, packets marked with AF12 DSCP are dropped in preference to packets marked with AF11 DSCP when congestion occurs. As stated earlier, how the server decides which layer to mark with which DSCP will not be discussed here. The DiffServ routers in the multicast tree simply forward the packets in each layer based on their DSCP.

Ideally, the layers within the optimal subscription level should be marked with AF11 while other layers are marked with AF12. However, the server may occasionally mark the layers improperly due to under- or over-estimation of the session's optimal subscription level. Improper layer markings also exist when there are multiple bottlenecks in the multicast tree through which different numbers of layers can traverse. In other words, different branches of the multicast tree have different optimal subscription levels and so require different layer markings. The inability of the layer marking to match the optimal subscription level would result in instability and unfairness in the RALM protocol's performance.

In the RALM protocol, all multicast packets are marked with an AF DSCP while the best-effort traffic is not marked at all (i.e. DSCP is zero). Therefore, the multicast traffic in the DiffServ network is segregated from best-effort traffic as they are served from different physical queues in the routers. Traffic segregation ensures that TCP flows which form part of the best-effort traffic are not buffered in the RED queues. Thus, it is not necessary to consider how the TCP flows will be affected by the RED parameters which are configured solely for the RALM protocol in the AF queues (it has been shown that setting proper RED parameters to allow TCP flows to work with different bandwidth adaptive flows is difficult [11]). The DiffServ router can partition the link bandwidth between the multicast and unicast traffic fairly through the use of a fair queuing scheduler.

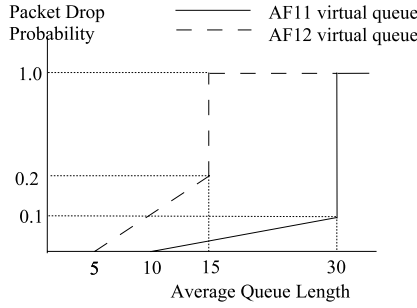


Fig. 2. The different RED parameter values used in the AF virtual queues

The RED parameters of the AF virtual queues are not configured based on any strict criterion. The only two requirements that must be kept in mind when choosing their values are that they must be close to the values recommended by the network community [4] [12] and that the AF12 queue must be penalised way before the AF11 queue when congestion occurs. In the RALM protocol, it is decided that the AF12 queue’s drop thresholds must be shorter than that of the AF11 queue and its maximum drop probability must be higher than the corresponding parameter in the AF11 queue. The net result is that AF12 packets are dropped sooner and faster than AF11 packets when the router starts to experience queue buildup due to congestion. Figure 2 shows the queue thresholds and maximum drop probability that may be configured in the AF virtual queues.

3.2 Receiver-Driven Congestion Control Algorithm

The optimal subscription level of an RALM session is reached when two conditions are satisfied: 1. There are no excessive losses in the AF11 marked protected layers such that there is a need to decrease the current subscription level; 2. There are enough losses in the AF12 marked sacrificial layers to avoid subscribing more than one sacrificial layer. For a given bottleneck link shared by multiple sessions, the optimal subscription level of each session is achieved when they share the bandwidth in a max-min fair manner. Once the layer markings in an RALM session match its optimal subscription level, the receivers will be able to reach this level by adding layers until they receive the first sacrificial layer.

Basically, the receiver joins an RALM session by subscribing the minimal two layers and increases its subscription level by using its current highest subscribed layer as a probe layer. The loss rate sampled in the probe layer is compared to the layer add loss threshold

$$T_{add} = max_{add} - \delta_{add} \times l \tag{1}$$

where max_{add} is the maximum loss rate beyond which no new layer is added, δ_{add} is the separation between adjacent subscription levels’ T_{add} , and l is the receiver’s current subscription level. If the sampled loss rate is lower than this threshold, a new layer is added unless all layers in the session have been subscribed. The

aggregated loss rate sampled in the lower layers is compared to the layer drop loss threshold

$$T_{drop} = max_{drop} - \delta_{drop} \times l \tag{2}$$

where max_{drop} is the maximum loss rate beyond which all layers are dropped, δ_{drop} is the separation between adjacent subscription levels' T_{drop} , and l is the receiver's current subscription level. If the sampled loss rate is higher than this threshold, the probe layer will be dropped unless the minimum subscription level of two has been reached. The receiver samples the loss rates in current subscribed layers using a 10 secs measurement window, which is sufficiently long to negate the effect of the long IGMP leave latency.

When a receiver's subscription level is below the optimal level, all its subscribed layers are protected from congestion due to their AF11 marking. The loss rate sampled in its probe layer will be lower than the current level's T_{add} , which is a good indication that there is sufficient bandwidth for the receiver to increase its subscription level. To avoid a contradicting layer drop indication by the aggregated loss rate in the lower layers, T_{add} is set below T_{drop} of each subscription level as shown in Figure 3. Thus, the receiver adds a new layer to its current subscription. This is continuously done until the receiver subscribes a sacrificial layer. Since the optimal subscription level is near the bottleneck bandwidth, adding the sacrificial layer causes the probe layer's loss rate to rise drastically beyond the current level's T_{add} . However, the subscribed protected layers' aggregated loss rate is still below the current level's T_{drop} . Now, the receiver has converged on the optimal subscription level, and will neither increase nor decrease its subscription level. Conversely, if the receiver's subscription level is above the optimal subscription level, part of the aggregated loss rate will be sampled from sacrificial layers. Thus, the excessive losses in these layers will cause the receiver to drop them as its current subscription level's T_{drop} is exceeded.

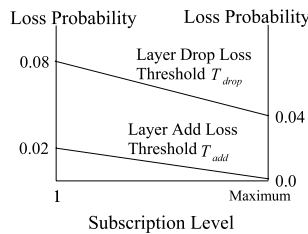


Fig. 3. The range of values that T_{add} and T_{drop} in an RALM session can take

Note that an RALM receiver converges on the optimal subscription level by oversubscribing a sacrificial layer. The condition of convergence requires the probe layer to suffer excessive losses while layers within the optimal subscription level are protected from the ongoing congestion. Thus, the loss rate performance of the RALM protocol is measured from the layers constituting the optimal subscription level and excluding the probe layer. Whether to utilise the data received in the probe layer is decided by individual receiver.

The degree of congestion in a bottleneck router is matched in the loss rate curves sampled by all downstream receivers. The implicit sharing of loss rate

knowledge can help receivers in different RALM sessions to cooperate in the discovery of their optimal subscription levels that share the bottleneck bandwidth fairly. This cooperation is realised by scaling the T_{drop} and T_{add} with the subscription level. Thus, for two RALM sessions that stream the same number of layers, the session at a lower subscription level has a higher loss tolerance than the other session at a higher subscription level. Through the judicious use of layer markings, a new session can force an incumbent session to drop its subscription level by causing a lowering of the optimal subscription level that the competing session should converge on. This new optimal subscription level should represent the fair share of the network bandwidth between the RALM sessions.

4 Evaluations

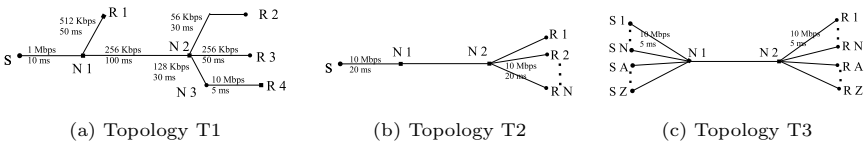


Fig. 4. Simulation topologies

Three scenarios are simulated using the three topologies shown in Figure 4 to study the performance of the RALM protocol. The sources and receivers are RALM nodes unless otherwise specified. N1, N2 and N3 are the DiffServ routers in the network. The bottleneck bandwidth is varied according to simulation scenarios but their transmission delays are fixed at 100 milliseconds and 20 milliseconds in topology T2 and T3 respectively. In all simulations, the RALM sources generate 10 equal size layers of 40 Kbps constant bit rate (CBR) traffic each. The simulation time is fixed at 1000 secs.

The congestion control algorithm in the RALM protocol is implemented in the *ns* simulator [13] that supports DiffServ mechanisms. The server generates a fixed number of layers for a multicast session and marks them with the appropriate DSCP before transmission. The optimal subscription level of the RALM session is assumed to be known before the start of the simulations. The server marks all layers within the optimal subscription level with AF11 DSCP and all higher layers with AF12 DSCP right from the start of the simulations. The layer markings are not changed for the duration of the simulations.

Protocol Independent Multicast Dense Mode (PIM-DM) [14] is used in the simulations. A two-level drop priority RED queue with its queue thresholds and max_p configured as in Figure 2 is implemented in the DiffServ routers alongside a 20 packets droptail queue for best-effort traffic. Their queue burst sensitivity weights are set to 0.02 and average packet size expected is 500 bytes. Since the layers in the RALM protocol are marked by the server and the routers forward packets based on their DSCP, the routers do not implement profile meters and policers. In the simulations, unicast traffic is not marked and thus is served through the droptail queues as best-effort traffic. A simple round robin scheduler serves the RED and droptail queues.

4.1 Accuracy and Stability

In Scenario 1, there is no single optimal subscription level that satisfies the bandwidth requirements of the receivers. Looking at the connecting bandwidth in topology T1, receiver R1 and R2 have extremely high and low bandwidth requirements respectively as R1’s optimal subscription level is 10 (maximum possible level) and R2’s optimal subscription level is 2 (minimum possible level). Receiver R3 and R4 have intermediate bandwidth requirements. As the layer markings are determined before the start of the simulations, two optimal subscription levels are picked separately (3 layers: 120 Kbps; and 6 layers: 240 Kbps) to meet the bandwidth requirement of receiver R4 (128 Kbps) and R3 (256 Kbps).

From the layer transition graphs of Figure 5, it can be seen that receiver R1 and R2 converge on their respective optimal subscription levels of 10 and 2 reasonably quickly and are able to maintain them during the session. Receiver R3 and R4 also converge on their respective optimal subscription levels of 6 and 3 reasonably quickly. But the subscription level is only stable if the source marks the layers correctly to match the optimal subscription level. Thus, receiver R4 has a stable subscription level in Figure 5(a) but finds that its subscription level fluctuates between layer 2 and 4 in Figure 5(b). Likewise, stable subscription level of receiver R3 is found only in Figure 5(b) but fluctuates between layer 6 and 7 in Figure 5(a). In both cases, the simulations illustrate the impossibility of having a session-wide optimal subscription level that satisfies the bandwidth requirements of all receivers in a heterogeneous network.

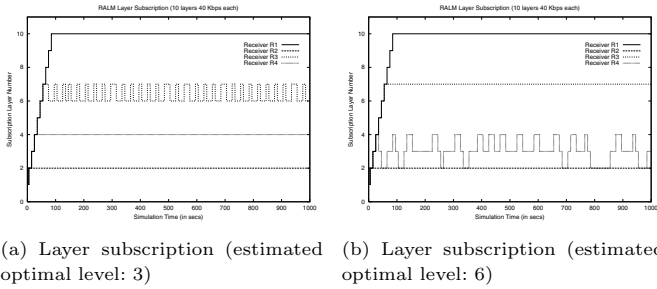


Fig. 5. Accuracy and stability of layer subscriptions in an RALM session

4.2 Intra-session Interactions

From Figure 6(c), it can be seen that 30 RALM receivers converge on the session-wide optimal subscription level of 5 (for bottleneck link bandwidth of 220 Kbps) rapidly regardless of their start time and number if the layer markings match this level accurately. The extra layer 6 that the receivers subscribed is the sacrificial layer that stabilises their subscription levels by absorbing most of the losses induced in the network as shown in Figure 6(d). This is a good example of the role of loss differentiation in guiding new receivers to the session-wide optimal subscription level. When a new receiver joins an RALM session, the low loss

rates in the protected layers within the optimal subscription level encourages it to subscribe towards the level that has already been reached by earlier receivers. Once it reaches this level, the high loss rate in the sacrificial layer discourages it from further subscription.

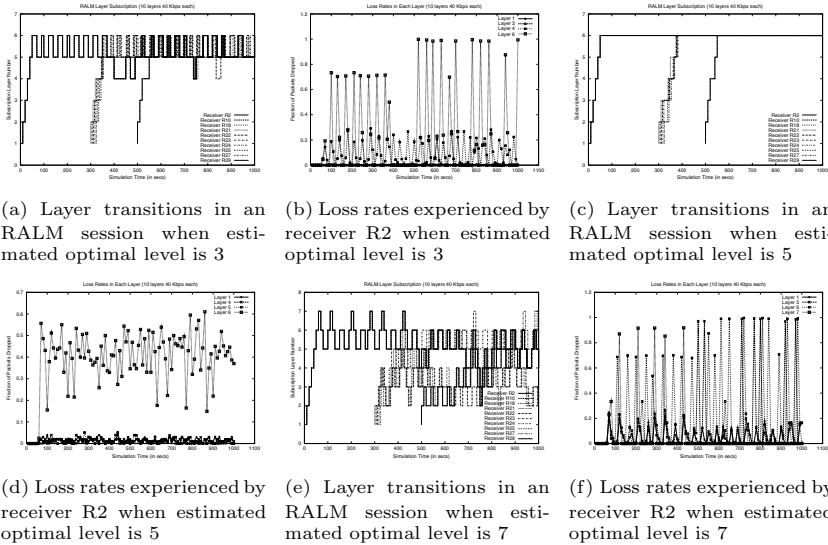


Fig. 6. Effect of layer markings on an RALM session

Figure 6(a) shows that the receivers cannot maintain stable subscription levels when the layer markings underestimate the session-wide optimal subscription level. The subscription levels fluctuate between layer 5 and 6 with occasional dips to layer 4 but never drops to layer 3. This is expected as the lower 3 layers are protected by their high priority marking. Thus, the layers higher than the estimated subscription level of 3 suffer burst losses as shown in Figure 6(b) when a receiver oversubscribes. The losses induced are shared by sacrificial layers 4, 5 and 6 with no losses in the protected layers. As a result, no receiver is forced to drop its subscription level below the level estimated by the layer markings.

In contrast, the subscription levels of the receivers in an RALM session whose layer markings overestimate the bottleneck bandwidth fluctuate wildly as shown in Figure 6(e). This is because all layers that are subscribed by the receivers have the same forwarding priority. In other words, no loss differentiation exists to protect the lower layers from the effect of oversubscribing layer 6 and 7. Therefore, losses induced by the oversubscribed layers are shared by all layers as shown in Figure 6(f).

Comparing the simulation results, it can be concluded that it is preferable to underestimate than to overestimate the session-wide optimal subscription level. Underestimation does not provide the RALM receivers with a stable subscription level. But it at least protects the base layers from losses when oversubscription occurs. This is a desirable property when the base layers are significantly more important than higher layers as is the case in layered video applications.

4.3 Inter-session Interactions

In Scenario 3, 2 sets of simulations are conducted on topology T3 to investigate the RALM protocol’s behaviour in the presence of other traffic types. In the first set of simulations, 3 RALM sessions simulated across a 700 Kbps bottleneck link are squeezed by a 400 Kbps CBR source from simulation time 500 secs to 600 secs. In the second set of simulations, one RALM session shares a 400 Kbps bottleneck link with two TCP sources that generate FTP packet streams.

When multiple RALM sessions operate within a network, they share the bottleneck bandwidth fairly by adopting a network-wide optimal subscription level. However, the layer markings of an individual session might underestimate or overestimate the network-wide optimal subscription level. To investigate these possibilities, 3 layer markings (optimal: 5, underestimate: 3, overestimate: 7) are configured in the first set of simulations. Likewise, a layer marking of 3 is configured in the RALM session for the second set of simulations, assuming that the RALM session is able to discover the subscription level for fair bandwidth share with the TCP sessions.

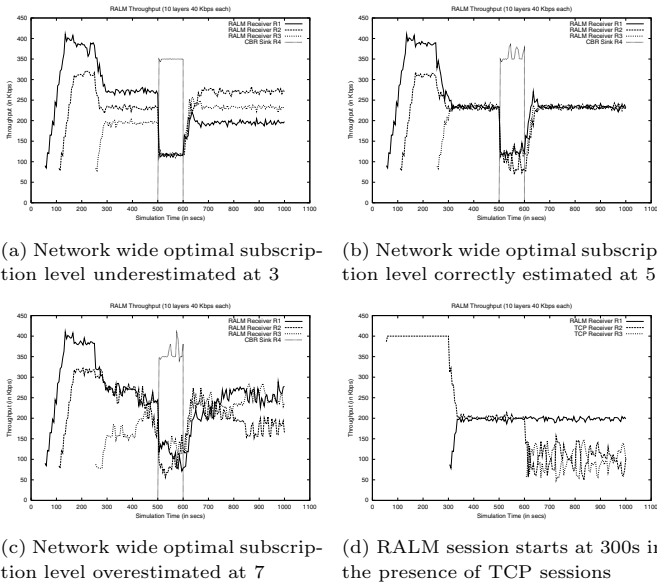


Fig. 7. Throughput in RALM sessions in the presence of CBR traffic (7(a), 7(b) and 7(c)) and TCP traffic (7(d))

The throughput curves of RALM receivers shown in Figure 7 represent the individual session-wide optimal subscription levels achieved with the indicated layer markings in the presence of CBR or TCP traffic. Figure 7(b) shows that when the layer markings match the network-wide optimal subscription level, the RALM sessions are able to share the bottleneck bandwidth fairly regardless of the presence of the CBR traffic. In contrast, Figure 7(a) and 7(c) show that

the RALM sessions do not receive their fair bandwidth share when the layer markings do not match this subscription level.

Underestimating layer markings discourage fair bandwidth share among competing RALM sessions. In fact, new sessions are disadvantaged in their ability to compete because they are prevented from increasing their subscription levels due to losses induced in their sacrificial layers by the actions of pre-existing sessions. On the other hand, overestimating layer markings cause all layers carried by the bottleneck link to be marked with the same priority. Thus, the base layers are not protected from losses induced by either the CBR traffic or an RALM session oversubscribing its fair bandwidth share. By not having a sacrificial layer in their subscriptions, the RALM sessions cannot enter a stable state.

In Figure 7, it is observed that the best-effort CBR traffic cannot reach its configured 400 Kbps rate. This is because the CBR traffic is served through a separate physical queue from the multicast traffic. The simple round robin scheduler that serves the queues partitions the bottleneck bandwidth equally between the AF marked multicast traffic and the best-effort unicast traffic. Thus, a traffic type has a minimum bandwidth of 350 Kbps that is protected from congestion caused by the other traffic type. This phenomenon is also observed in Figure 7(d) where the RALM session is able to grab its fair bandwidth share and is protected from any change in the TCP traffic.

The segregation of RALM traffic from other traffic types offers the network administrator the ability to limit multicast bandwidth consumption by configuring a weighted queue scheduler appropriately. Thus, a rogue multicast session that does not implement any congestion control scheme will not be able to congest competing best-effort traffic sharing the same router. An additional advantage of traffic segregation is that the network administrator can configure the AF queue's RED parameters without affecting TCP performance. This should make it easier to deploy the RALM protocol in current IP networks.

5 Conclusion

The RALM protocol that provides layered multicast congestion control through the use of loss differentiation is presented in this paper. A series of simulations have shown that receivers in an RALM session converge accurately, rapidly and stably on the session-wide optimal subscription level when the layer markings match this level in a homogeneous network. And the RALM protocol scales with the session size and provides excellent inter-session bandwidth fairness when all sessions mark their layers to match the network-wide subscription level. Unfortunately, in a heterogeneous network where a single optimal subscription level cannot be found to satisfy varied receivers' bandwidth requirements, some receivers will not be able to maintain stable subscription levels even though they may receive a throughput close to their bandwidth requirements. In such cases, it is preferable for the layer markings to underestimate the bandwidth requirements so that the base layers received by the mismatched receivers are protected from any losses induced by their frequent oversubscriptions. Finally, the proportion

of network bandwidth consumed by RALM traffic with respect to best-effort traffic can be adjusted by deploying weighted schedulers to serve the queues. Since the RALM traffic is segregated from the best-effort traffic, the AF queues' RED parameters in the RALM protocol can be configured without affecting the TCP performance of a best-effort session.

To actually deploy the RALM protocol in a DiffServ network, the servers in the sessions must discover the layer markings that will provide the network-wide optimal subscription level. Therefore, a layer marking discovery protocol that receivers can use to estimate their optimal subscription levels and communicate these estimates back to the servers for actual layer markings is required. The actual scheme to be used in this protocol will be investigated in the future.

References

1. S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven Layered Multicast", in Proc. ACM SIGCOMM '96, Stanford, CA, Aug. 1996, pp. 117-130 [341](#)
2. L. Vicisano, L. Rizzo, and J. Crowcroft, "TCP-like Congestion Control for Layered Multicast Data Transfer", in Proc. IEEE INFOCOM '98, San Francisco, CA, Mar. 1998, pp. 996-1003 [341](#)
3. A. Legout and E. W. Biersack, "PLM: Fast Convergence for Cumulative Layered Multicast Transmission Schemes", in Proc. ACM SIGMETRICS '00, Santa Clara, CA, June 2000 [341](#)
4. S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, V.1 N.4, Aug. 1993, pp. 397-413 [342](#), [346](#)
5. A. Legout and E. W. Biersack, "Pathological Behaviors for RLM and RLC", in Proc. NOSSDAV '00, Chapel Hill, NC, June 2000 [342](#)
6. W. Fenner, "Internet Group Management Protocol, Version 2", RFC 2236, IETF, Nov. 1997 [342](#)
7. D. Rubenstein, J. Kurose, and D. Towsley, "The Impact of Multicast Layering on Network Fairness", in Proc. ACM SIGCOMM '99, Sept. 1999, pp 27-38 [343](#)
8. R. Gopalakrishnan, J. Griffioen, G. Hjalmytsson, C. Sreenan and S. Wen, "A Simple Loss Differentiation Approach to Layered Multicast", in Proc. IEEE INFOCOM '00, Tel Aviv, Israel, Mar. 2000 [344](#)
9. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, IETF, Dec. 1998 [345](#)
10. J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, IETF, June 1999 [345](#)
11. D. Lin and R. Morris, "Dynamics of Random Early Detection", in Proc. ACM SIGCOMM '97, Cannes, France, Oct. 1997, pp. 127-137 [345](#)
12. S. Floyd, various notes on RED, <http://www.aciri.org/floyd/red.html> [346](#)
13. NS2, Network Simulator - ns-allinone-2.1b8a, LBL / Xerox PARC / UCB / ISI, <http://www.isi.edu/nsnam/ns> [348](#)
14. S. Deering, D. Estrin, D. Farinacci, V. Jacobson, A. Helmy, D. Meyer, and L. Wei, "Protocol Independent Multicast Version 2 Dense Mode Specification (PIM-DM)", work in progress, IETF, Nov. 1998 [348](#)