Back pressure based multicast scheduling for fair bandwidth allocation

Saswati Sarkar and Leandros Tassiulas

I. Introduction

We study fair allocation of resources in multicast networks with multirate capabilities in this thesis. In multirate transmission, the session source hierarchically encodes its signal and the receivers subscribe to the appropriate number of layers. The objective of the network is to distribute the layers fairly. This can be attained either by computing the fair rates first, and then using a scheduling policy to attain the fair rates, or by using a scheduling policy which allocates the fair rates without computing them explicitly. We have studied the first approach in the earlier paper. This approach has certain disadvantages. For example, this approach requires knowledge of system parameters like link bandwidths, which are not generally known to the link schedulers. We investigate the second approach in this chapter. We would present a scheduling policy which allocates the fair rates without computing them beforehand. We have presented analytical and experimental results demonstrating the fairness of the resulting rate allocation. In addition to guaranteeing the fair rates, this policy confines the packet losses to enhancement layers, and protects the more important base layers, when there is shortage of bandwidth. Furthermore, this policy does not require any knowledge of traffic statistics, is computationally simple, and is essentially local information based.

We assume that the network uses multirate transmission to counter network heterogeneity. In multirate transmission, it is possible to serve different receivers of the same session at different rates. This is convenient as often paths for different receivers of the same session have widely different bandwidths, and serving all receivers of the same session with the same bandwidth either overwhelms the slow receivers or starves the fast ones. In multirate transmission, there is no notion of session rate as such, and the service rate of a session in a link is equal to the maximum of the session receivers downstream of the link. Thus the same session receives service at different rates in different links on its path. For this purpose, a source hierarchically encodes its signal in several layers. The lowest layer consists of the most important information and all receivers of the session should receive it. Receivers can subscribe to higher layers for successive refinement of reception quality

Parts of this paper have been presented at INFOCOM 2001, Anchorage, Alaska.

S. Sarkar is with the department of Electrical and Systems Engineering in the University of Pennsylvania, Philadelphia, PA, USA. Email: swati@ee.upenn.edu.

L. Tassiulas is with the Computer Engineering and Telecommunications department in the University of Thessaly, Volos, Greece, and the Department of Electrical and Computer Engineering, University of Maryland, College Park. Email: leandros@inf.uth.gr

at the expense of additional bandwidth. If the path to a receiver is congested, then it receives the base layer only, whereas a receiver subscribes to a large number of layers if its data path has a lot of bandwidth.

Our objective is to split bandwidth fairly among different sessions traversing a link, and also serve every receiver at a rate commensurate with the fair bandwidth share along its path. The fair share may be different for different receivers of the same session. Consider figure 2 for example. Fair rate allocation is 3.75 for u_1 and 3.25 for receivers u_2 and u_3 each. The data path of receiver u_2 consists of links e_1, e_3, e_5 and e_7. Link e_5 is the bottleneck link in this case, as it has a bandwidth of 6.5 units to be shared between sessions 1 and 2. Fair share in this bottleneck link is 3.25 for both the sessions. Link e_3 is the bottleneck for receiver u_1 of session 1 and the fair share for session 1 in this link is 3.75 after allocating 3.25 to session 2. We adopt the notion of maxmin fairness[2]. A rate allocation is maxmin fair, if no receiver can be allocated a higher rate without hurting another receiver having equal or lower rate. Maxmin fairness is a good notion of fairness and as [12] points out, maxmin fairness satisfies many intuitive fairness properties in a multirate multicast network.

The usual approach is to compute the fair shares of the receivers and then determine the service order for packets in the links so as to actually serve packets as per the fair rates. We have presented distributed algorithms for computing the fair allocations in [13]. Now, there are many scheduling policies which can attain any feasible rate allocation, once the rates are known, e.g., fair queueing strategies[3][7]. However, these scheduling policies work only when the link schedulers know the desired rates, or at least a ratio between the desired rates. But, computing the fair rates has several problems. Firstly, computation algorithms require exact knowledge of system parameters, like link bandwidths. Now, bandwidth available for real time traffic varies from time to time, depending on the service contracts between the network and the data traffic senders. In general, the schedulers at the nodes may not have exact knowledge of this available link capacity. It is also necessary to exchange messages between neighboring nodes. This adds to the overhead, and this overhead is often nonnegligible.

To overcome these limitations we present a *scheduling policy which attains the maxmin fair rates in a multirate, multicast network, without computing them beforehand.* No such scheduling policy is known for multirate multicast networks. As opposed to the fair queueing strategies, this policy can operate without any knowledge about the actual maxmin fair rates. Our scheduling policy is geared towards attaining maxmin fair rates, whereas the fair queueing strategies can attain any feasible rates *once the desired rates are known.* Our scheduling policy does not need to know the link capacities, and does not assume centralized coordination or global information at any computing processor.

This policy works in a multirate scenario where the source hierarchically encodes its signal into several layers. Information contained in a higher layer packet is meaningful only if all the lower layer packets have been successfully decoded. So, any multirate scheduling policy should first strive to attain low packet loss for the lower layers, and use the residual bandwidth to serve the higher layers. Our scheduling policy not only

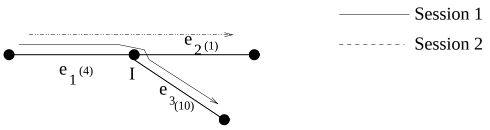

Fig. 1. An example network demonstrating the intuition behind the scheduling policy.

attains the maxmin fair rates for the individual receivers, but also offers different quality of service to different layers. Loss rates of the lower layers are negligible while those of the higher layers are somewhat higher. Let layer i emitted by a source consume b_i units of bandwidth. Let the bandwidth r allocated to a virtual session be sufficient to serve all packets of the first k layers and a portion of the packets of the $k+1$th layer, i.e., $\sum_{i=1}^{k} b_i < r < \sum_{i=1}^{k+1} b_i$. In the ideal scenario, the virtual session should receive all packets of the first k layers and at least $\frac{r-\sum_{i=1}^{k} b_i}{b_{k+1}}$ fraction of packets of the $k+1$th layer and possibly no packet from the higher layers. We will show that our scheduling policy not only delivers packets in accordance with the maxmin fair rates, but also attains the above loss rates for the layers, i.e., the first k layers suffer negligible packet loss and the loss rate for the $k+1$th layer is $1 - \frac{r-\sum_{i=1}^{k} b_i}{b_{k+1}}$ in this case. The policy attains this distinction without any knowledge of the layer bandwidths.

We would describe the scheduling policy in greater details later. The basic idea is to serve the sessions in a round robin manner in each link. When a session is sampled, it may or may not transmit a packet. The decision is based on the availability of packets for transmission and the congestion downstream. Since we consider multicast sessions, congestion of all links on the path of the session originating from the destination node of the link must be considered. For example in figure 2, the decision to transmit a session 1 packet in link e_3 during its round robin turn depends on the congestions in links e_4 and e_5. The same decision for session 2 is based on the congestion in link e_5 only. Also, a session always gives priority to a lower layer packet over a higher layer packet.

This chapter is organized as follows. Section II describes the intuition behind the scheduling policy we propose in this chapter. The scheduling policy is presented in Section III. Section IV presents analytical performance guarantees. This section also presents experimental studies. We prove the results of section IV in Section V. Section VI concludes this chapter.

II. Back pressure based flow control for fairness

Intuitively, a simple round robin scheduling in every link should split link bandwidth fairly among all sessions traversing a link. However, this scheme has a deficiency, which we explain below. A session traverses multiple links, and different links offer different bandwidths to the session. The link which offers minimum bandwidth to a session is denoted as the bottleneck link of the session. For example, links e_1 and e_2 are the bottleneck links of sessions 1 and 2 respectively in figure 1. A session should not be served at a rate higher than that offered by its bottleneck link, in any link on its path. This would cause congestion and packet loss

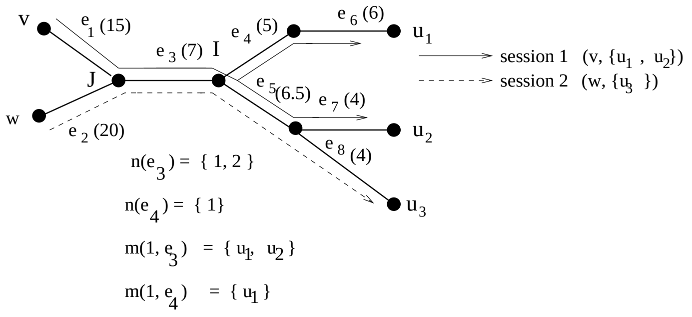

Fig. 2. Session 1 consists of receivers u_1 and u_2, and session 2 has receiver u_3. The numbers in brackets, () denote the capacities of the respective links. The capacity constraint for link e_3 is $\max(r_1, r_2) + r_3 \leq 7$ and that for link e_5 is $r_2 + r_3 \leq 6.5$. The maximum rates are 5 for session 1 and 20 for session 2. Hence, $r_1 \leq 5$, $r_2 \leq 5$ and $r_3 \leq 20$. The maxmin fair allocation is $(3.75, 3.25, 3.25)$ for receivers u_1, u_2 and u_3 respectively. Under this allocation, $\lambda_{1e_3} = 3.75$ and $\lambda_{1e_5} = 3.25$.

in the bottleneck link. Also, a significant portion of the bandwidths of nonbottleneck links will be wasted in serving packets which do not reach the destination. A simple round robin scheduling does not ensure that the service rate of a session in any link on its path is equal to that in its bottleneck link. Credit based flow control can be used for conveying the bottleneck information implicitly.

Hahne[8] used credit flow control for attaining fairness in unicast networks. A credit value (W) is decided apriori. The basic idea is to keep track of the number of session packets waiting for transmission at the destination node of a link. For example, the flow control strategy must keep track of the number of session 2 packets waiting at node I of figure 1 for controlling the flow of session 2 in link e_1. If this number is less than the credit value and the session has packets for transmission, then the session transmits a packet in the link, when it is sampled. If this number is equal to the credit value, then the session does not transmit even if it is sampled, and has packets for transmission. For example, session 2 will not transmit packets in link e_1 if there are W session 2 packets at I. Consider figure 1 to see how credit based flow control works. Simple round robin offers 2 units of bandwidths to both sessions 1 and 2 in e_1. Now, e_2 serves session 2 at a lower rate (1 per unit time). Thus there will be an accumulation of session 2 packets at node I and hence session 2 will often not transmit packet over e_1 even when it is sampled. Thus link e_1 will serve the session at a rate lower than it otherwise would. Now link e_3 can transmit session 1 packets at a higher rate than e_1. So node I will not have session 1 packets often, and this will reduce the transmission rate for session 1 packets in e_3. In fact, any link l serves any session i at the same rate as the bottleneck link of the session.

Credit flow control presents some inherent complications for multirate multicast networks. We would first explain the difficulties and then present our approach in overcoming the complications. The path of a session may consist of multiple links originating from the same node. For example, session 1 traverses through links e_4 and e_5 originating from node I in the network of figure 2. These links serve the session at different rates. As a result, the number of packets of the same session, waiting at a node for transmission in different links are different. This number may be less than the credit value for one link, but may be greater than the credit value in another link. Consider figure 2 for example. It can happen that the number of session 1 packets waiting at node I for transmission in link e_4 is 2 and the same number for link e_5 is 10, and the credit value is 3. Thus it is not clear how credit flow control can be used to determine when a link should serve a session, and when it should not. Also the choice of the flow control scheme should be such that the rate of a session in a link l is equal to the maximum of that in the links originating from the destination node of l. Rate of session 1 in link e_3 should be equal to the maximum of that in links e_4 and e_5 in figure 2. We show that this can be attained by allowing a link l to serve a session if the number of packets of the session in at least one of the links originating from the destination of l is less than the credit value W. For example in figure 2, the scheduler for e_3 keeps track of the number of session 1 packets waiting at I for transmission in e_4 and the number waiting for transmission in e_5. If at least one of these is less than W, the scheduler transmits a session 1 packet in e_3 in its round robin turn. For the time being, we assume that the queue lengths at destination node of a link are known at the link scheduler, which is normally at the origin of the link. We discuss this assumption later.

Since service rate of a session in a link is equal to the maximum of the service rates of the links downstream, the source of a link may receive packets at a rate higher than the link can serve. For example, if link e_4 serves session 1 faster than link e_5 does in figure 2, then link e_3 will serve packets at a rate equal to that of e_4 and consequently, e_5 receives packets at a rate higher than it can serve. Thus, there will be packet loss at intermediate nodes (node I in this example), since the node buffers are finite. Let the credit value be W. In the unicast case, a link l does not serve a packet if the destination node has W packets. So there is no packet loss in the intermediate nodes, if the sizes of node buffers are W, or greater. In our case there will be packet loss as long as the buffers are finite. So, our task is to attain the maxmin fair rates in presence of packet loss, and also to regulate the loss so that packets are lost from higher layers only. Again, this is because hierarchical transmission has the property that a layer yields useful information, only if packets from all lower layers have been successfully decoded. We attain this objective by using different priorities for different layers.

III. Description of the Policy

We propose a scheduling policy based on prioritized round robin with window flow control for multirate multicast networks. Here, $B_{(i,k,l)}(t)$ denotes the number of layer k packets of session i waiting for transmission in link l at time t. Packets of the same session waiting for transmission in multiple links originating from the same node, need not be stored in separate memory location. So, the quantities $B_{(i,k,l)}$s represent logical rather than physical buffers. A node needs to keep track of $B_{(i,k,l)}(t)$s for all layers k of all sessions i traversing through any link l originating from the node. Refer to figures 3, 4 and 5 for examples. We assume finite size physical buffers. Hence, memory limitations force $B_{(i,k,l)}(t)$s to be less than or equal to a quantity G, for all sessions

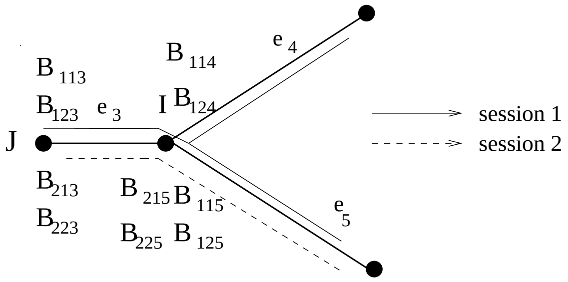

Fig. 3. We show a section of the network shown in figure 6.2(B_is). We assume that both sessions transmit two layers only. We show the different queues at the nodes. Here, B_{ijk} is the queue of session i layer j packets waiting for transmission in link e_k. The associated queues are shown in figure 6.4.

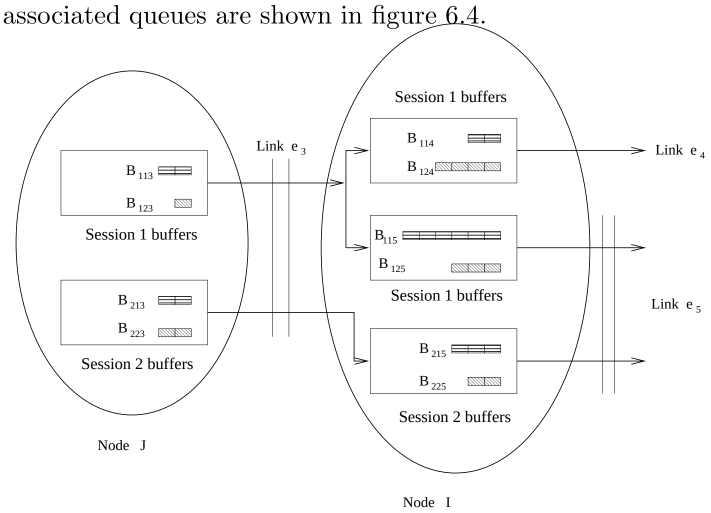

Fig. 4. We show the logical buffers of figure 6.3. Note that a session 1 packet may wait for transmission in both links e_4 and e_5. It is not necessary to have separate copies of the packet waiting at the same node, as this figure shows. This figure represents the logical queues only. The logical queues may be maintained by pointers. Physical queue corresponding to these logical queues is shown in figure 6.5.

i, layers k, link l and time t. We assume credit value W, and $G > W$.

We assume than a session source is connected to the network through an access link. No other session traverses the access link. For example, link e_1 is the access link of session 1 and link e_2 is the access link of session 2 in figure 2. We first describe the scheduling for access links. Whenever the source of a session i desires to transmit a packet, the access link checks whether at least one of the logical buffers at the destination of the link has less than W session i packets of the corresponding layer. If so, the access link transmits the packet, else it does not transmit the packet. If any logical buffer at the destination of the link has G session i packets of the corresponding layer (i.e., the logical buffer is "full"), the packet is not added to the queue of packets waiting to be transmitted in the link and this packet is lost for this link. More formally, let session i source wish to transmit a layer k packet at time t over access link l. Access link l transmits the packet if $\min_{l' \in \kappa(i,l)} B_{(i,k,l')}(t) < W$, where $\kappa(i,l)$ is the set of links on session i path originating from the destination node of access link l. When access link l transmits a session i packet of layer k, it joins the queue for transmission in all links in $\kappa(i,l)$, except those which already have G session i packets of layer k. If

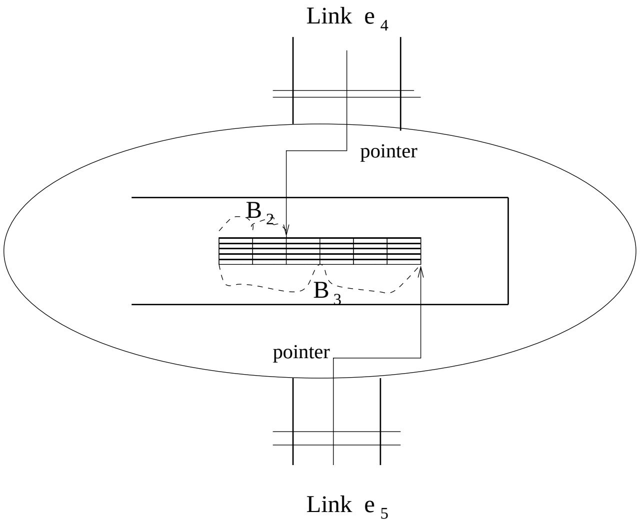

Fig. 5. We show a physical buffer at node I of figures 6.3 and 6.4. We assume that the switches are input queued. This physical buffer holds layer 1 packets of session 1 transmitted via link e_3, and corresponds to logical buffers, B_{114} and B_{115}. Packets are replicated only at the transmission epoch. So, the buffer holds 6 packets in all. All 6 need to be transmitted in link e_5 and only the last 2 need to be transmitted in link e_4. The first 4 have already been transmitted in link e_4. Hence, $B_2(B_{(114)})$ contains 2 packets and $B_3(B_{(115)})$ contains 6 packets in figure 5. Every link maintains a pointer at the first packet it needs to transmit.

$B_{(i,k,l')}(t) = G$, for some link $l' \in \kappa(i,l)$, the transmitted packet is not added to $B_{(i,k,l')}(t)$, i.e.,the packet is lost for link l' and receivers downstream of l'. If $\min_{l' \in \kappa(i,l)} B_{(i,k,l')}(t) = W$, then access link l does not transmit the packet. (Note that $\min_{l' \in \kappa(i,l)} B_{(i,k,l')}(t)$ can not exceed W at any time t.) Refer to figure 2 for an example. Here, $\kappa(1, e_1) = \{e_3\}$. So, access link e_1 transmits a layer j packet of session 1 at time t if $B_{(1,j,e_3)}(t) < W$. Here, no packet transmitted over e_3 is lost as packet transmission takes place only when $B_{(1,j,e_3)}(t) < G$ as $G > W$. Packet loss can take place only when $\kappa(i,l)$ contains more than one links. We assume that the scheduler for link l knows $\min_{l' \in \kappa(i,l)} B_{(i,k,l')}(t)$, for all sessions i traversing link l and all times t. We will argue in sections IV and VI that this assumption can be relaxed.

Now we consider the scheduling of nonaccess links. When a link l is ready to transmit a packet (which happens when it has finished transmitting the previous packet), it samples all sessions traversing the link starting from the one after the session last served. When a session i is sampled, it finds out whether it can send a lowest layer (layer 1) packet. It sends a lowest layer packet if at the sampling time t,

1. there are session i layer 1 packets waiting for transmission in link l i.e., if $B_{(i,1,l)}(t) > 0$, and

2. if the number of session i layer 1 packets waiting for transmission in link l' is less than W for at least one link l' originating from the destination of l, i.e., if $\min_{l' \in \kappa(i,l)} B_{(i,1,l')}(t) < W$.

If session i can not transmit layer 1 packet because of the violation of either of the above conditions, it tries to send a next higher layer packet. If it can not send the next higher layer packet, it tries the next and so on. In general, session i sends a layer k packet, if

1. it can not send a layer $1, \ldots, k-1$ packet,

2. $B_{(i,k,l)}(t) > 0$, and

3. $\min_{l' \in \kappa(i,l)} B_{(i,k,l')}(t) < W$.

When link l transmits a session i packet of layer k, it joins the queue for transmission in all links in $\kappa(i,l)$, except those which already have G session i packets of layer k. If $B_{(i,k,l')}(t) = G$, for some link $l' \in \kappa(i,l)$, the transmitted packet is not added to $B_{(i,k,l')}(t)$, and the packet is lost for link l' and receivers downstream of l'.

If session i can not send any packet, the link l samples the next session. If none of the sessions can send a packet, the link idles for some time, before starting the sampling once again. However, if some session transmits a packet, after transmitting the packet, the link samples the sessions, starting from the session after the one last served. We explain the policy in the following example.

Example III.1: Consider the network shown in figure 3. It is a part of the network of figure 2. The initial logical buffer contents are shown in figure 4. The quantities B_{ijk} of the figure denote $B_{(i,j,e_k)}$. We assume $W = 3$ and $G = 6$. Here $\kappa(1, e_3) = \{e_4, e_5\}$ and $\kappa(2, e_3) = e_5$. Session 2 was served last. So the link samples session 1 for packets. Note that $B_{(1,1,e_3)} > 0$ and $\min_{l' \in \kappa(1,e_3)} B_{(1,1,l')} = B_{(1,1,e_4)} = 2 < W$. So link e_3 transmits a session 1 layer 1 packet. The packet is added to the queue for transmission in link e_4, but not in link e_5 because $B_{(1,1,e_5)} = 6 = G$. This packet is lost for link e_5 and the downstream receiver u_2.* The modified buffer contents are $B_{(1,1,e_3)} = 1$, and $B_{(1,1,e_4)} = 3$. The rest of the buffer contents remain the same as before.

Next the link samples session 2 for packets. Session 2 can not transmit a layer 1 packet because $\min_{l' \in \kappa(2,e_3)} B_{(2,1,l')} = B_{(2,1,e_5)} = 3 = W$. However, it can transmit a layer 2 packet because $B_{(2,2,e_3)} > 0$, and $B_{(2,2,e_5)} = 2 < W$. So it transmits a layer 2 packet. This packet is added to the queue for transmission in link e_5. The modified buffer contents are $B_{(2,2,e_3)} = 1$ and $B_{(2,2,e_5)} = 3$. The rest of the buffer contents remain the same as before.

Next the link samples session 1 for packets. Session 1 can not send any layer 1 packet because $\min_{l' \in \kappa(1,e_3)} B_{(1,1,l')} = 3 = W$. It can not send any layer 2 packet because $\min_{l' \in \kappa(1,e_3)} B_{(1,1,l')} = 3 = W$. So the link samples session 2 for packets. Session 2 can not send any layer 1 packet because $\min_{l' \in \kappa(2,e_3)} B_{(2,1,l')} = 3 = W$. It can not send any layer 2 packet because $\min_{l' \in \kappa(2,e_3)} B_{(2,1,l')} = 3 = W$. So the link idles for some time. It starts serving only when links e_4 and e_5 serve some packets and reduce $B_{(i,k,l)}$ for $i \in \{1, 2\}$, $k \in \{1, 2\}$ and $l \in \{e_4, e_5\}$.

IV. Performance Evaluation

We discuss the performance of this scheduling policy. We would present analytical and simulation results. The following theorem indicates that this scheduling policy allocates maxmin fair rates to all receivers. Also, the packet loss is concentrated in the highest layer served. Refer to section V for proofs.

Let session i source transmit γ_i layers in all. Bandwidth of the kth layer of the ith session is $b_{i,k}$. Also, $\sum_{k=1}^{\gamma_i} b_{i,k} = p_i$, where p_i is the maximum rate of source i. Session i source is "well-behaved" if it attempts to transmit at most $b_{i,k}(v - u) + \xi_{i,k}$ packets and at least $b_{i,k}(v - u) - \xi_{i,k}$ packets of the kth layer in any interval

*It may happen that the new packet is added to the queue but another old packet in the queue $B_{(1,2,e_5)}$ is dropped. The bottomline is that one packet is lost. The lost packet may either be the new one or an old one.

$[u, v)$. The parameters $\xi_{i,k}$ are "transmission jitters." Many sources are well-behaved in practice. Oftentimes, a leaky bucket shaper is placed between a traffic source and a network. The output of the leaky bucket is well-behaved. Analytical results will be presented for "well behaved" sources only. However we expect these results to hold in a more general scenario.

Theorem 1: Let all sources be well behaved. Let maxmin fair rate of receiver j be r_j. Then receiver j receives at most $r_j(v-u)+\delta_j$ packets and at least $r_j(v-u)-\delta_j$ packets in any interval $[v, u)$, if $G \geq G^*$ and $W \geq W^*$. The constants $\delta_j, j = 1, \ldots, M, G^*, W^*$ depend only on the system parameters, e.g., path lengths, transmission jitters etc., and not on the length of the interval $[v, u)$ or on the time instants u, v.

Theorem 1 implies that the long term average rates of the receivers are equal to the maxmin fair rates. Also, the number of packets delivered to the receivers in any interval differs from the maxmin fair rates by at most a constant, and the value of this constant does not depend on the length of the interval. This shows that the policy is fair in short intervals as well, in some sense. The next theorem shows that packet loss is concentrated in the highest layer served. We introduce another notation. The quantity $(u)^+$ stands for $\max(u, 0)$.

Theorem 2: Let all sources be well behaved. Let $G \geq G^*$ and $W \geq W^*$. Consider any receiver j with maxmin fair rate r_j. Let receiver j belong to session i. The number of layer k packets lost in the path of receiver j in any interval $[v, u)$ is at most $\left(\min\left(\sum_{w=1}^{k} b_{i,w} - r_j, b_{i,k}\right)\right)^+ (v-u) + \iota_{j,k}$. Here, $\iota_{j,k}$ is a finite constant which depends only on system parameters and not on the length of the interval $[v, u)$, or time instants u, v.

Let receiver j belong to session i. If the sum of the transmission rates of the first k layers of session i is less than or equal to the maxmin fair rate of receiver j (i.e., $\sum_{w=1}^{k} b_{i,w} \leq r_j$), then ideally receiver j should receive all layer k packets sent by the source. In this case, $\left(\min\left(\sum_{w=1}^{k} b_{i,w} - r_j, r_j\right)\right)^+ = 0$, and hence by Theorem 2, the number of layer k packets lost in the path of receiver j in any interval $[v, u)$, is at most $\Delta_{j,k}$. Thus receiver j observes a long term loss rate of 0 for layer k packets in this case. Also, the number of layer k packets lost in any interval is upper bounded by a constant $\Delta_{j,k}$, which does not depend on the size of the interval. If the maxmin fair rate is between the sum of the transmission rates of the first $k-1$ layers and the sum of the transmission rates of the first k layers (i.e., $\sum_{w=1}^{k-1} b_{i,w} \leq r_j < \sum_{w=1}^{k} b_{i,w}$), then receiver j should receive packets of layer k partially. Theorem 2 says that in this case, the packet loss in any interval $[v, u)$, is upper bounded by $\left(\min\left(\sum_{w=1}^{k} b_{i,w} - r_j, b_{i,k}\right)\right)(v-u) + \Delta_{j,k}$. So the long term loss rate is $\min\left(1, \frac{\sum_{w=1}^{k} b_{i,w} - r_j}{b_{i,k}}\right)$. Thus, the residual bandwidth left after serving the first $k-1$ layer packets is used to serve the kth layer.

As the theorems indicate, the above results are guaranteed only when the credit value and the buffer size exceed certain lower bounds. These lower bounds depend on system parameters. In practice, it may not be possible to compute these lower bounds and decide the credit and buffer sizes accordingly. Systems have buffers of fixed size, and scheduling policies need to function within the existing memory constraints. The proposed scheduling policy shows a gradual performance improvement with increase in buffer and credit sizes.

We introduce the notation of the *rank* of a receiver. Let there be J distinct maxmin fair receiver rates. If the maxmin fair rate of receiver j is the mth smallest among the distinct maxmin fair rates, then the rank of receiver j is m.

Theorem 3: There exists a sequence of constants, $W_1, W_2, \ldots, W_J$, and G_1,
$G_2, \ldots G_J$ such that, if $W \geq W_m$, and $G \geq G_m$, then all receivers of rank m and above, receive service at a rate greater than or equal to the mth smallest maxmin fair rate. All receivers of rank smaller than m, receive service at their maxmin fair rates. Also, the entire packet loss is concentrated in the highest layer served, for all receivers.

The analytical lower bounds on window and buffer sizes, $W_1, W_2, \ldots, W_J$, $(W^* = W_J)$, $G_1, G_2, \ldots G_J$ $(G^* = G_J)$ are pessimistic bounds. Simulations indicate that, in practice, service rates converge to the maxmin fair rates, and packet loss is concentrated in the highest layer for much smaller credit and buffer sizes.

The scheduling policy we described so far assumes that a link scheduler has complete information about queue lengths at the destination node of the link. More precisely, the assumption is that at all times t the scheduler for link l knows $\min_{l' \in \tau(i,l)} B_{(i,j,l')}(t)$ for all layers j of all sessions i traversing link l. This is a reasonable assumption in networks with hop by hop congestion feedback and negligible propagation delay. Though many current day networks employs end to end congestion control only, many others use hop by hop back pressure based congestion control. In fact, in the LAN context, recent simulation results show that hop by hop back pressure can be better than TCP for dealing with short-lived congestion[15]. Though terrestrial networks have negligible propagation delay, propagation delay is significant for some other networks, e.g., those with satellite links. In these networks, the feedback will typically be delayed, i.e., at time t the scheduler will have information about destination queue lengths at some previous time t'. In this case, the policy remains the same, except that the scheduler takes decisions based on the information it has currently. The service rates of the receivers converge to the maxmin fair rates as before, even in the presence of propagation delays. Packet loss is concentrated in the highest layer served as before. Theorems 1 to 3 hold in this case as well, as long as the propagation delays are bounded (appendix V). In fact, all proofs in appendix V assume that every link l has a propagation delay Υ_l. Analytical results indicate that the effect of the propagation delay is to slower the convergence and increase the credit and buffer thresholds, W_i and G_is by multiples of the product of propagation delay and link capacity. Intuitively, the basic results hold in presence of propagation delay for the following reason. The available information can differ from the actual queue lengths at the destination node by at most a constant which depends on the propagation delay and link capacities. This constant will increase the constants $\Delta_{i,j}$s but the long term throughputs are independent of these($\Delta_{i,j}$s). Our simulation studies further confirm this observation.

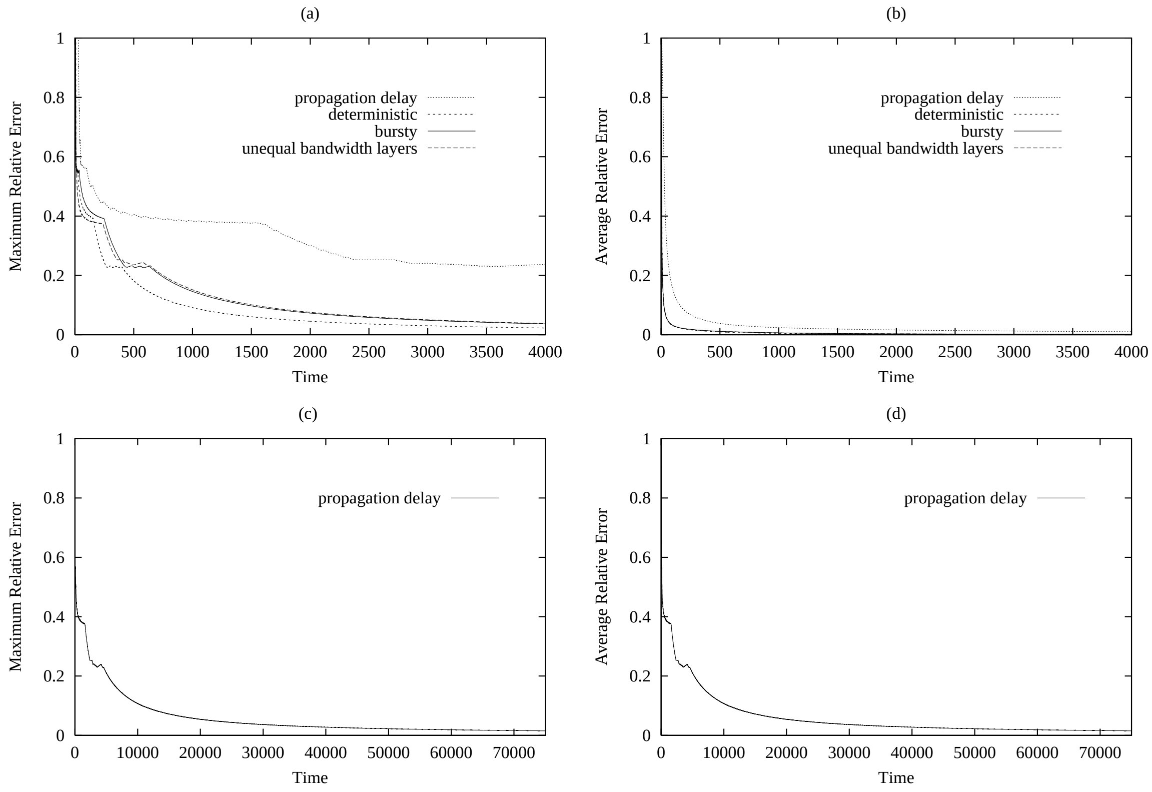

Fig. 6. These figures demonstrate the convergence of the packet delivery rates attained by the proposed scheduling policy to the maxmin fair rates. We have plotted the convergence errors as a function of time for different traffic models.

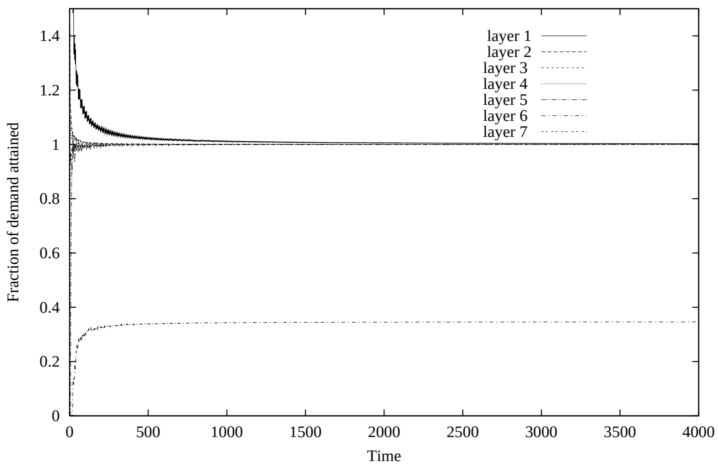

Fig. 7. The figure shows the fraction of packets of different layers delivered to one particular receiver in the random network.

Now we present the results of our experimental studies. We considered a 15 session 400 node random network for experimental evaluation. Nodes are points on a $20X20$ grid. There exists an edge between any two nodes with a probability (p) that depends on the euclidean distance between the nodes (d) ($p = \exp(\alpha(1-d))$), where α is the decay constant. We assumed $\alpha = 2$. We adopted this edge probability model because distant nodes are less likely to have an edge between them. Source and receivers of every session have been selected randomly. There are 96 receivers in all, i.e., average session size is 6.4. The session route consists of shortest paths between the source and the destinations. All sources transmit 20 layers.

Figure 6 demonstrates the convergence of the service rates attained by different receivers to the respective maxmin fair rates for different traffic patterns. We studied the difference between the actual rate of packet delivery for any receiver and its maxmin fair rate. The attained rate at time t for receiver s is the total number of packets delivered to s in the interval $[0, t)$ divided by the time t. If maxmin fair rate of a receiver s is r_s^m, and the attained rate at time t is $r_s^a(t)$, then error for receiver s is $|1 - \frac{r_s^a(t)}{r_s^m}|$ at time t. Figures $6(a)$ and $6(c)$ plot the maximum relative error, and figures $6(b)$ and $6(d)$ plot the average relative error, the maximum and average are taken over all receivers. We considered deterministic traffic patterns and traffic patterns with jitters. A traffic pattern is deterministic at rate d, if packets are generated once every $1/d$ seconds. A jittery (d, ξ) traffic pattern transmits traffic at a long term rate of d per unit time with a jitter of ξ. In this traffic model, at most $dt + \xi$ and at least $dt - \xi$ packets are generated in any interval of length t. In the deterministic case (curve labelled "deterministic" in figures 6(a) and 6(b)), packets of all layers of all sources are generated periodically at rate 1 per unit time. For the traffic model with jitters (curve labelled "bursty" in figures 6(a) and 6(b)), every source transmits packets of every layer as per a jittery$(1, 3)$ model. For unequal bandwidth layer case (curve labelled "unequal bandwidth layers" in figures 6(a) and 6(b)), source i generates layer j packets as per a jittery$(b_{i,j}, \xi_{i,j})$ model. The layer bandwidths $b_{ij}s$ and transmission jitters $\xi_{i,j}$s have been chosen randomly, and are different for different i, js. For example, the first 3 layers of session 1 have bandwidth and jitter of $(.21, 2)$, $(2.78, 3)$ and $(1.81, 0)$ respectively, whereas the first 3 layers of session 2 have bandwidths and jitters of $(3.5, 2)$, $(3.21, 0)$ and $(3.30, 1)$ respectively. All these curves ignore propagation delays. Convergence is fast for all these cases. The average error converges to 0 much faster than the maximum error, indicating that the attained rates of most of the receivers converge to the maxmin fair rates very fast, whereas convergence is relatively slow for a few others. Convergence is fastest for the deterministic traffic model. The results for the deterministic traffic model were obtained with credit value of 5 and buffer size of 10 units only. We used credit and buffer sizes of 8 and 16 for the other two traffic models with jitters. Thus, the attained rates converge to the maxmin fair rates for small credit and buffer sizes.

We also considered effect of propagation delays and subsequently delayed feedback. In our model, propagation delay of a link is equal to the euclidean distance between the end points of the link. The "propagation delay" curves in figures 6(a) to figures 6(d) study the errors for the "unequal bandwidth layer" traffic model with propagation delay. Figure 6(c) and 6(d) shows the errors for this case for a longer range of time. We used credit and buffer sizes of 100 and 200 respectively in presence of propagation delays. As expected, long term rates still converge to the maxmin fair rates, but the convergence is slower than when propagation delay is ignored. Delayed feedback increases buffer and credit size requirements. However, these requirements are still reasonable.

Figure 7 demonstrates the different loss rates suffered by different layers and exhibits that packet losses are confined to the highest layer served in the case we studied. It actually shows the fraction of packets delivered to one particular receiver in the random network. This fraction for layer i is the ratio between the number of packets of layer i delivered to the receiver in $[0, t)$ and the product of the layer bandwidth b_i and time t. The maxmin fair service rate of this receiver is 9 packets per unit time. The traffic model is the same as the "unequal bandwidth case" discussed before. We have ignored propagation delays in this case. The source for this receiver transmits packets of the first 6 layers at rate (b_is) 0.21, 2.78, 1.81, 3.01, 0.84 and 1 per unit time, respectively. The transmission jitters (ξ_is) for these layers are 2, 3, 0, 3 0 and 1 respectively. Ideally the receiver should receive all packets of the first 5 layers, 36% packets of layer 6 and possibly no packet of any higher layer. Analytical results guarantee delivery of packets of layer i at rate b_i, if $i \in \{1, \ldots, 5\}$ and rate $0.36 b_i$ if $i = 6$. As the figure shows, for the first 5 layers, the fraction of packets delivered to the receiver converges to 1 very fast. In fact, it is greater than 1 for some lower layers initially. This is because the source sends an initial burst of packets for every layer on account of transmission jitters. The network is able to deliver some of these bursts as well, for the lower layers. Hence the ratio between the number of packets delivered and $b_i t$ is initially greater than 1 for the lower layers. The fraction of packets of layer 6 delivered to the receiver converges to 0.36 as well. Very few packets of other layers reach the receiver. This shows that packet loss is confined to the highest layer served (layer 6 in this case). Also note that the long term rates are always higher for a lower layer as compared to that for a higher layer. Window and buffer sizes are 8 and 16 respectively for this simulation.

V. Proof of Theorems 1, 2, 3

A. Notation and Supporting Lemmas

We prove the results in Section IV in this section. As we mentioned before, the scheduling policy for attaining the maxmin fair rates in a unicast network in [8], is a special case of ours. The proofs here are similar in structure to those in [8]. However, the multirate multicast assumption makes our case more complicated. We will point out the important differences in course of the proof.

We first introduce some additional notations.

- Let E_i be the maximum number of sessions traversing a link on the path of session i. Also, E is the maximum number of sessions traversing any link in this network.

- All links on session i paths and originating from the same node are said to be session i siblings of each other.
- The set of links traversed by virtual session s is L_s.
- The session of virtual session s is $\chi(s)$.
- Maxmin fair session link rate of session i in link l is λ_{il}.
- Maxmin fair rate of virtual session s is r_s.
- The number of layer j packets of session i waiting to be served in link l at time t is $B(i,j,l,t)$.
- The number of session i packets served in link l in time interval $[s,t)$ is $P(i,l,s,t)$.
- The number of session i layer j packets served in link l in time interval $[s,t)$ is $P(i,j,l,s,t)$. Strictly speaking, $P(i,j,l,s,t)$ and $P(i,l,s,t)$ refer to the number of packets which finish service in the interval $[s,t)$.
- Number of links in the largest session i path is H_i. Every session i has at least two links on each of its paths, because every path must have at least one link in the original network and we artificially add an "access link" connecting the session i source to the rest of the network. Thus $H_i > 1$, for all sessions i.
- Let $H = \max_i H_i$.
- Let ν_i be the largest number of fanouts at any point on session i path. and let $\nu = \max_i \nu_i$.
- Let $\Gamma = \max_{l \in \mathcal{L}} C_l$ be the maximum link capacity in the network.
- Let Υ_l be the propagation delay in link l. Also, Υ is the maximum propagation delay in the network.

We assume that every link l has a propagation delay Υ_l. As a result, the minimum logical queue lengths at a node are no longer bounded by W, but the bound is $W + \Upsilon\Gamma$. This is because if this minimum queue length at the destination node of link l becomes W at time t, then the scheduler for link l will know this at time $t + \Upsilon_l$. Between t and $t + \Upsilon_l$ link l may continue to transmit packets for the particular session. So the minimum queue length can become $W + \Upsilon_l C_l$. Since $\Upsilon_l \le \Upsilon$ and $C_l \le \Gamma$, the upper bound is $W + \Upsilon\Gamma$.

We will first state and prove lemmas 1 to 3. These lemmas will be used in proving Theorem 4. The main results in Section IV will follow from Theorem 4 stated and proved in the next subsection.

Lemma 1 is similar to Lemma 2 in [8]. The important difference is that we need to assume that $B(i,j,l_2,s_k)$ is upper bounded (assumption 6) to prove result (7). The logical queue lengths, $B(i,j,l_2,t)$ are always upper bounded in [8] because of the scheduling policy. However, this does not hold in our case. So we need this additional assumption, which makes the use of Lemma 1 more restrictive. Also, since there is no congestion related packet loss in [8], the conservation equation, i.e.,

$$P(i,l_2,s,t) = P(i,l_1,s,t) - B(i,l_2,s) + B(i,l_2,t) \tag{1}$$

always holds in [8], and [8] uses this property extensively. This does not hold in our case, in general, because of congestion related packet losses.

Lemma 1: Consider layer m of session i. Let l_1 and l_2 be two links on the path of session i with l_2 originating from the sink of l_1. Let K be a positive integer and let $s_1, t_1, s_2, t_2, \ldots, s_K, t_K$ be a nondecreasing sequence of times and r, e_1, e_2, f_1, f_2 be some nonnegative real numbers. Let l_1 and l_2 satisfy the following properties:

For any $J(k)$ intervals $[u_k^1, v_k^1), \ldots, [u_k^{J(k)}, v_k^{J(k)})$ in $[s_k, t_k]$, s.t. $\min_{l \in \kappa(i,l_1)} B(i,m,l,t) < W \ \forall \ t \in \cup_{k=1}^{K} \cup_{j=1}^{J(k)} [u_k^j, v_k^j)$,

$$\sum_{k=1}^{K} \sum_{j=1}^{J(k)} P(i,m,l_1,u_k^j,v_k^j) \geq r \sum_{k=1}^{K} \sum_{j=1}^{J(k)} (v_k^j - u_k^j) - e_1 - \sum_{k=1}^{K} \sum_{j=1}^{J(k)} f_1 \tag{2}$$

For any $J(k)'$ intervals $[u_k^1, v_k^1), \ldots, [u_k^{J(k)'}, v_k^{J(k)'})$ in $[s_k, t_k]$, s.t. $B(i,m,l_2,t) > 0 \ \forall \ t \in \cup_{k=1}^{K} \cup_{j=1}^{J(k)'} [u_k^j, v_k^j)$,

$$\sum_{k=1}^{K} \sum_{j=1}^{J(k)'} P(i,m,l_2,u_k^j,v_k^j) \geq r \sum_{k=1}^{K} \sum_{j=1}^{J(k)'} (v_k^j - u_k^j) - e_2 - \sum_{k=1}^{K} \sum_{j=1}^{J(k)'} f_2 \tag{3}$$

$$W > 2 \text{ and } f_1 + f_2 + \Upsilon\Gamma + 2 \leq W < G < \infty \tag{4}$$

It follows that

$$\sum_{k=1}^{K} P(i,m,l_2,s_k,t_k) \geq r \sum_{k=1}^{K} (t_k - s_k) - (e_1 + e_2) - K(f_1 + f_2 + \Upsilon\Gamma + 1) \tag{5}$$

In addition, if

$$B(i,m,l_2,s_k) \leq W + \Upsilon\Gamma \ \forall k \in \{1, \ldots, K\} \tag{6}$$

$$\text{then } \sum_{k=1}^{K} P(i,m,l_1,s_k,t_k) \geq r \sum_{k=1}^{K} (t_k - s_k) - (e_1 + e_2) - K(f_1 + f_2 + 2\Upsilon\Gamma + 1) \tag{7}$$

Proof of Lemma 1: We first prove (5). We lower bound the throughput $P(i,m,l_2,s_k,t_k)$. We would split $[s_k, t_k)$ into subintervals such that either $B(i,m,l_2,t) < W$ or $B(i,m,l_2,t) > 0$ throughout the subinterval. In the first case, $P(i,m,l_1,s_k,t_k)$ is lower bounded by assumption (2) and $P(i,m,l_1,s_k,t_k)$ and $P(i,m,l_2,s_k,t_k)$ differ by a constant independent of the interval size. In the second case, $P(i,m,l_2,s_k,t_k)$ is lower bounded by assumption (3).

Initially let $B(i,m,l_2,t_k) > 0$. We break $[s_k, t_k]$ into various subintervals as follows;

1. $j = 0$, $u_k^j = t_k$.
2. $j \to j + 1$.
3. The earliest time t in $[s_k, u_k^{j-1})$ such that $B(i,m,l_2,t_1) > 0$ for all $t_1 \in [t, u_k^{j-1})$ is denoted u_k^j.
4. If $u_k^j > s_k$, $j \to j + 1$ else go to step (7).
5. The earliest time t in $[s_k, u_k^{j-1})$ such that $B(i,m,l_2,t_1) < W$ for all $t_1 \in [t, u_k^{j-1})$ is denoted u_k^j.
6. If $u_k^j > s_k$, go to step (2).
7. $J(k) = j$.

We make the following observations.

$$s_k = u_k^{J(k)} \leq \qquad \ldots \qquad \leq u_k^0 = t_k \tag{8}$$

$$B(i,m,l_2,t) > 0 \quad \forall t \in [u_k^j, u_k^{j-1}), \quad j \text{ odd} \tag{9}$$

$$
\begin{aligned}
B(i,m,l_2,u_k^j) &= 1 & & j \text{ odd}, \ j < J(k) \qquad (10)\\
B(i,m,l_2,t) &< W \quad \forall t \in [u_k^j, u_k^{j-1}), & & j \text{ even} \qquad (11)\\
B(i,m,l_2,u_k^j) &= W-1 & & j \text{ even}, \ 0 < j < J(k) \qquad (12)
\end{aligned}
$$

Let $\mathcal{B}_1^k = \{j : 1 \le j \le J(k), j \text{ odd } \}$ and $\mathcal{B}_2^k = \{j : 1 \le j \le J(k), j \text{ even } \}$ From (9) and (11),

$$
B(i,m,l_2,t) > 0, \forall t \in \cup_{j \in \mathcal{B}_1^k}[u_k^j, u_k^{j-1}) \text{ and } B(i,m,l_2,t) < W, \forall t \in \cup_{j \in \mathcal{B}_2^k}[u_k^j, u_k^{j-1}) \qquad (13)
$$

If $j \in \mathcal{B}_2^k$

$$
\begin{aligned}
P(i,m,l_2,u_k^j,u_k^{j-1}) &= P(i,m,l_1,u_k^j - \Upsilon, u_k^{j-1} - \Upsilon) - B(i,m,l_2,u_k^{j-1}) \\
&\quad + B(i,m,l_2,u_k^j) \qquad (14)\\
&\ge P(i,m,l_1,u_k^j,u_k^{j-1}) - \Upsilon\Gamma - B(i,m,l_2,u_k^{j-1}) \\
&\quad + B(i,m,l_2,u_k^j) \qquad (15)
\end{aligned}
$$

The first inequality follows from (13) since there is no loss of packets in any interval in $\mathcal{B}_2^k$ as the buffer size $G > W(4)$. The second follows since $P(i,j,l,s-\Upsilon,t-\Upsilon) \ge P(i,j,l,s,t) - \Upsilon\Gamma$ for all i,j,l,s,t. Also, $|\mathcal{B}_2^k| \le |\mathcal{B}_1^k| \le |\mathcal{B}_2^k| + 1$. Let $\mathcal{B}_2^k \ne \phi$. From (10), (12) and (14), for $j \in \mathcal{B}_2^k$,

$$
P(i,m,l_2,u_k^j,u_k^{j-1}) \ge \begin{cases} P(i,m,l_1,u_k^j,u_k^{j-1}) + W - 2 - \Upsilon\Gamma & \text{if } j < J(k) \\ P(i,m,l_2,u_k^j,u_k^{j-1}) - 1 - \Upsilon\Gamma & \text{for } j = J(k). \end{cases} \qquad (16)
$$

Note that $j = J(k)$, $j \in \mathcal{B}_2^k$ implies that $|\mathcal{B}_2^k| = |\mathcal{B}_1^k|$

$$
\begin{aligned}
P(i,m,l_2,s_k,t_k) &= \sum_{j=1}^{J(k)} P(i,m,l_2,u_k^j,u_k^{j-1}) \\
&= \sum_{j \in \mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{j \in \mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\
&\ge \sum_{j \in \mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{j \in \mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \\
&\quad + \left(|\mathcal{B}_2^k| - 1\right)(W-2) - 1 - |\mathcal{B}_2^k|\Upsilon\Gamma \\
&\quad + (W-1) I_{\mathcal{B}_1^k > \mathcal{B}_2^k} \text{ if } \mathcal{B}_2^k \ne \phi \text{ (from (16))} \qquad (17)
\end{aligned}
$$

If $\mathcal{B}_2^k = \phi$, then clearly

$$
P(i,m,l_2,s_k,t_k) = \sum_{j \in \mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{j \in \mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \qquad (18)
$$

From (17) and (18)

$$
P(i,m,l_2,s_k,t_k)
$$

$$\geq \sum_{j\in\mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1})$$
$$+I_{|\mathcal{B}_2^k|\geq 1}\left(\left(|\mathcal{B}_2^k|-1\right)(W-2)-1-|\mathcal{B}_2^k|\Upsilon\Gamma+(W-1)I_{\mathcal{B}_1^k>\mathcal{B}_2^k}\right) \tag{19}$$

If $B(i,m,l_2,t_k)=0$, then the intervals are formed somewhat differently. As before, $u_k^0=t_k$. Now, u_k^1 is the earliest time t in $[s_k,t_k]$ s.t. $B(i,m,l_2,t)<W$, $\forall\ t\in[u_k^1,t_k)$. Now, the other points are defined similarly, e.g., u_k^2 is the earliest time in the interval $[s_k,u_k^1)$ s.t. $B(i,m,l_2,t)>0$, for all t in $[u_k^2,u_k^1)$ and u_k^3 is the earliest time in the interval $[s_k,u_k^1)$ s.t. $B(i,m,l_2,t)<W$, for all t in $[u_k^3,u_k^2)$, etc.

$$s_k=u_k^{J(k)}\leq \qquad \dots \qquad \leq u_k^0=t_k \tag{20}$$
$$B(i,m,l_2,t)>0 \quad \forall t\in[u_k^j,u_k^{j-1}),\quad j \text{ even} \tag{21}$$
$$B(i,m,l_2,u_k^j)=1 \qquad j \text{ even}, 0<j<J(k) \tag{22}$$
$$B(i,m,l_2,u_k^0)=0 \tag{23}$$
$$B(i,m,l_2,t)<W \quad \forall t\in[u_k^j,u_k^{j-1}),\quad j \text{ odd} \tag{24}$$
$$B(i,m,l_2,u_k^j)=W-1 \qquad j \text{ odd}, 0<j<J(k) \tag{25}$$

Let $\mathcal{B}_1^k=\{j:1\leq j\leq J(k), j \text{ even }\}$ and $\mathcal{B}_2^k=\{j:1\leq j\leq J(k), j \text{ odd }\}$ From (21) and (24),

$$B(i,m,l_2,t)>0,\forall t\in\cup_{j\in\mathcal{B}_1^k}[u_k^j,u_k^{j-1}) \text{ and } B(i,m,l_2,t)<W,\forall t\in\cup_{j\in\mathcal{B}_2^k}[u_k^j,u_k^{j-1}) \tag{26}$$

Also $|\mathcal{B}_1^k|\leq|\mathcal{B}_2^k|$. Again (14) holds from (26) because there is no loss of packets in any interval in $\mathcal{B}_2^k$ as the buffer size $G>W(4)$. From (22), (23), (25) and (14), for $j\in\mathcal{B}_2^k$,

$$P(i,m,l_2,u_k^j,u_k^{j-1})\geq\begin{cases} P(i,m,l_1,u_k^j,u_k^{j-1})+W-2-\Upsilon\Gamma & \text{for } 1\leq j<J(k),\\ P(i,m,l_1,u_k^j,u_k^{j-1})-1-\Upsilon\Gamma & \text{for } j=J(k).\end{cases} \tag{27}$$

Note that $\mathcal{B}_2^k\neq\phi$. If $|\mathcal{B}_2^k|=1$, then $\mathcal{B}_1^k=\{2\}$ or $\mathcal{B}_1^k=\phi$ and $\mathcal{B}_2^k=\{1\}$. In this case,

$$P(i,m,l_2,s_k,t_k)\geq\sum_{j\in\mathcal{B}_1^k}P(i,m,l_2,u_k^j,u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}P(i,m,l_1,u_k^j,u_k^{j-1})-\Upsilon\Gamma \tag{28}$$

If $|\mathcal{B}_2^k|\geq 2$, proceeding as in the proof of (19),

$$\text{Thus } P(i,m,l_2,s_k,t_k) \geq \sum_{j\in\mathcal{B}_1^k}P(i,m,l_2,u_k^j,u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}P(i,m,l_1,u_k^j,u_k^{j-1})$$
$$+(|\mathcal{B}_2^k|-1)(W-2)-1-|\mathcal{B}_2^k|\Upsilon\Gamma \tag{29}$$

Combining (28) and (29),

$$P(i,m,l_2,s_k,t_k) \geq \sum_{j\in\mathcal{B}_1^k}P(i,m,l_2,u_k^j,u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}P(i,m,l_1,u_k^j,u_k^{j-1})$$
$$+I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)(W-2)-1-|\mathcal{B}_2^k|\Upsilon\Gamma \tag{30}$$

From (19), (30) and the fact that $|\mathcal{B}_1^k| \leq |\mathcal{B}_2^k|$, for the case considered in (30), we have

$$\begin{aligned}
& P(i,m,l_2,s_k,t_k) \\
\geq\ & \sum_{j\in\mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \\
+\ & I_{|\mathcal{B}_2^k|\geq 1}\Big((|\mathcal{B}_2^k|-1)(W-2) - 1 - |\mathcal{B}_2^k|\Upsilon\Gamma + I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|}(W-1)\Big)
\end{aligned} \tag{31}$$

$$\begin{aligned}
\min_{l\in\kappa(i,l_1)} B(i,m,l,t) &\leq B(i,m,l_2,t) \text{ (since } l_2 \in \kappa(i,l_1)) \\
&< W, \forall t \in \cup_{k=1}^K \cup_{j\in\mathcal{B}_k^2} [u_k^j, u_k^{j-1})
\end{aligned} \tag{32}$$

The last inequality follows from (13) and (26).

$$\begin{aligned}
& \sum_{k=1}^K P(i,m,l_2,s_k,t_k) \\
\geq\ & \sum_{k=1}^K \sum_{j\in\mathcal{B}_1^k} P(i,m,l_2,u_k^j,u_k^{j-1}) + \sum_{k=1}^K \sum_{j\in\mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \\
& +(W-2)\sum_{k=1}^K \Big(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\Big) - K - \sum_{k=1}^K |\mathcal{B}_2^k|\Upsilon\Gamma \\
& +\sum_{k=1}^K I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}(W-1) \text{ (from (31))} \\
\geq\ & r\sum_{k=1}^K \sum_{j\in\mathcal{B}_1^k} (u_k^j - u_k^{j-1}) + \sum_{k=1}^K \sum_{j\in\mathcal{B}_2^k} (u_k^j - u_k^{j-1}) - e_1 - e_2 \\
& -\sum_{k=1}^K |\mathcal{B}_1^k| f_2 - \sum_{k=1}^K |\mathcal{B}_2^k|(f_1 + \Upsilon\Gamma) + (W-2)\sum_{k=1}^K \Big(I_{|\mathcal{B}_2^k|\geq 2}(|\mathcal{B}_2^k|-1)\Big) \\
& -K + \sum_{k=1}^K I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}(W-1) \\
& \text{ (from (13), (26), (32) and assumptions (2) and (3))} \\
\geq\ & r\left(\sum_{k=1}^K \left(\sum_{j\in\mathcal{B}_1^k} (u_k^j - u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} (u_k^j - u_k^{j-1})\right)\right) - e_1 \\
& -e_2 - \sum_{k=1}^K \Big(|\mathcal{B}_2^k| + I_{\mathcal{B}_2^k=\phi} I_{|\mathcal{B}_1^k|=1} + I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}\Big) f_2 \\
& -\sum_{k=1}^K |\mathcal{B}_2^k|(f_1 + \Upsilon\Gamma) + (W-2)\sum_{k=1}^K \Big(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\Big) \\
& -K + \sum_{k=1}^K I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}(W-1) \text{ (since } |\mathcal{B}_1^k| \leq |\mathcal{B}_2^k| + 1\ \forall\ k) \\
\geq\ & r\left(\sum_{k=1}^K \left(\sum_{j\in\mathcal{B}_1^k} (u_k^j - u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} (u_k^j - u_k^{j-1})\right)\right) - e_1 - e_2
\end{aligned}$$

$$
\begin{aligned}
& -\sum_{k=1}^{K}|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}f_2-\sum_{k=1}^{K}f_2I_{\mathcal{B}_2^k=\phi}-\sum_{k=1}^{K}|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}(f_1+\Upsilon\Gamma)\\
& -\sum_{k=1}^{K}(f_1+\Upsilon\Gamma)I_{\mathcal{B}_2^k=\phi}+(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right)-K\\
& +(W-1-f_2)\sum_{k=1}^{K}I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}\\
\geq\ & r\sum_{k=1}^{K}(t_k-s_k)+(W-2-f_1-\Upsilon\Gamma-f_2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 2}(|\mathcal{B}_2^k|-1)\right)\\
& -\sum_{k=1}^{K}\Bigg(\left(f_2I_{|\mathcal{B}_2^k|\geq 1}+f_2I_{\mathcal{B}_2^k=\phi}\right)\\
& +\left((f_1+\Upsilon\Gamma)I_{|\mathcal{B}_2^k|\geq 1}+(f_1+\Upsilon\Gamma)I_{\mathcal{B}_2^k=\phi}\right)\Bigg)-K\\
& \qquad\qquad (\text{assumption (4)})\\
\geq\ & r\sum_{k=1}^{K}(t_k-s_k)-K(f_1+f_2+1+\Upsilon\Gamma)\ (\text{assumption (4)})
\end{aligned}
$$

Thus (5) follows.

Now we prove (7) under the additional assumption that (6) holds. We lower bound the throughput $P(i,m,l_1,s_k,t_k)$. Like before, we would split $[s_k,t_k)$ into subintervals such that either $B(i,m,l_2,t) < W$ or $B(i,m,l_2,t) > 0$ throughout the subinterval. In the first case, $P(i,m,l_1,s_k,t_k)$ is lower bounded by assumption (2). In the second case, $P(i,m,l_2,s_k,t_k)$ is lower bounded by assumption (3) and $P(i,m,l_1,s_k,t_k)$ and $P(i,m,l_2,s_k,t_k)$ differ by a constant independent of the interval size. The splitting order is different in this case.

Initially let $B(i,m,l_2,t_k) < W$. We break $[s_k,t_k]$ into various subintervals as follows;

1. $j=0$, $u_k^j=t_k$.
2. $j\rightarrow j+1$.
3. The earliest time t in $[s_k,u_k^{j-1})$ such that $B(i,m,l_2,t_1) < W$ for all $t_1\in[t,u_k^{j-1})$ is denoted u_k^j.
4. If $u_k^j > s_k$, $j\rightarrow j+1$ else go to step (7).
5. The earliest time t in $[s_k,u_k^{j-1})$ such that $B(i,m,l_2,t_1) > 0$ for all $t_1\in[t,u_k^{j-1})$ is denoted u_k^j.
6. If $u_k^j > s_k$, go to step (2).
7. $J(k)=j$.

We make the following observations.

$$
\begin{aligned}
s_k=u_k^{J(k)}\leq & \qquad \dots & \leq u_k^0=t_k && (33)\\
B(i,m,l_2,t)<W & \quad \forall t\in[u_k^j,u_k^{j-1}), & j \text{ odd} && (34)\\
B(i,m,l_2,u_k^j)=W-1 & & j\text{ odd},\ j<J(k) && (35)\\
B(i,m,l_2,t)>0 & \quad \forall t\in[u_k^j,u_k^{j-1}), & j\text{ even} && (36)
\end{aligned}
$$

$$B(i,m,l_2,u_k^j) = 1 \qquad\qquad j \text{ even, } 0 < j < J(k) \tag{37}$$

$$P(i,m,l_1,u_k^j,u_k^{j-1}) \geq P(i,m,l_2,u_k^j,u_k^{j-1}) + B(i,m,l_2,u_k^{j-1}) - B(i,m,l_2,u_k^j) - \Upsilon\Gamma \tag{38}$$

(The inequality is strict if there is loss of packets in the interval $[u_k^j, u_k^{j-1})$. Let $\mathcal{B}_1^k = \{j : 1 \leq j \leq J(k), j \text{ odd } \}$ and $\mathcal{B}_2^k = \{j : 1 \leq j \leq J(k), j \text{ even } \}$ From (34) and (36),

$$B(i,m,l_2,t) < W, \forall t \in \cup_{j\in\mathcal{B}_1^k}[u_k^j,u_k^{j-1}) \text{ and } B(i,m,l_2,t) > 0, \forall t \in \cup_{j\in\mathcal{B}_2^k}[u_k^j,u_k^{j-1}) \tag{39}$$

Also $|\mathcal{B}_1^k| \leq |\mathcal{B}_2^k| + 1$. Let $\mathcal{B}_2^k \neq \phi$. From (35), (37) and assumption (6), for $j \in\mathcal{B}_2^k$

$$P(i,m,l_1,u_k^j,u_k^{j-1}) \geq \begin{cases} P(i,m,l_2,u_k^j,u_k^{j-1}) + W - 2 - \Upsilon\Gamma & \text{if } j < J(k) \\ P(i,m,l_2,u_k^j,u_k^{j-1}) - 1 - 2\Upsilon\Gamma & \text{for } j = J(k). \end{cases} \tag{40}$$

Note that $j = J(k)$, $j \in \mathcal{B}_2^k$ implies that $|\mathcal{B}_2^k| = |\mathcal{B}_1^k|$.

$$\begin{aligned} P(i,m,l_1,s_k,t_k) &= \sum_{j=1}^{J(k)} P(i,m,l_1,u_k^j,u_k^{j-1}) \\ &= \sum_{j\in\mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \\ &\geq \sum_{j\in\mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\ &\quad + \left(|\mathcal{B}_2^k| - 1\right)(W-2) - 1 - (|\mathcal{B}_2^k|+1)\Upsilon\Gamma \\ &\quad + I_{\mathcal{B}_1^k > \mathcal{B}_2^k}(W - 1 + \Upsilon\Gamma) \text{ if } \mathcal{B} \neq \phi \end{aligned} \tag{41}$$

If $\mathcal{B}_2^k = \phi$, then clearly

$$P(i,m,l_1,s_k,t_k) = \sum_{j\in\mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_1,u_k^j,u_k^{j-1}) \tag{42}$$

From (41) and (42)

$$\begin{aligned} P(i,m,l_1,s_k,t_k) &\geq \sum_{j\in\mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j\in\mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\ &\quad + I_{|\mathcal{B}_2^k|\geq 1}\Big(\left(|\mathcal{B}_2^k| - 1\right)(W-2) - 1 - (|\mathcal{B}_2^k|+1)\Upsilon\Gamma \\ &\quad + I_{\mathcal{B}_1^k > \mathcal{B}_2^k}(W - 1 + \Upsilon\Gamma)\Big) \end{aligned} \tag{43}$$

If $B(i,m,l_2,t_k) \geq W$, then the intervals are formed somewhat differently. Now, $u_k^0 = t_k$ and u_k^1 is the earliest time t in $[s_k,t_k]$ s.t. $B(i,m,l_2,t) > 0$ for all $t \in [u_k^1, t_k)$. Now, the other points are defined similarly, e.g., u_k^2 is the earliest time in the interval $[s_k, u_k^1)$ s.t. $B(i,m,l_2,t) < W$, for all t in $[u_k^2, u_k^1)$ and u_k^3 is the earliest time in the interval $[s_k, u_k^1)$ s.t. $B(i,m,l_2,t) > 0$, for all t in $[u_k^3, u_k^2)$, etc.

$$s_k = u_k^{J(k)} \leq \qquad \ldots \qquad \leq u_k^0 = t_k \tag{44}$$

$$B(i,m,l_2,t) < W \quad \forall t \in [u_k^j, u_k^{j-1}), \quad j \text{ even} \tag{45}$$

$$B(i,m,l_2,u_k^j) = W - 1 \qquad j \text{ even, } 0 < j < J(k) \tag{46}$$

$$B(i,m,l_2,u_k^0) \geq W \tag{47}$$

$$B(i,m,l_2,t) > 0 \quad \forall t \in [u_k^j, u_k^{j-1}), \quad j \text{ odd} \tag{48}$$

$$B(i,m,l_2,u_k^j) = 1 \qquad j \text{ odd, } 1 \leq j < J(k) \tag{49}$$

Let $\mathcal{B}_1^k = \{j : 1 \leq j \leq J(k), j \text{ even }\}$ and $\mathcal{B}_2^k = \{j : 1 \leq j \leq J(k), j \text{ odd }\}$ From (45) and (48),

$$B(i,m,l_2,t) < W, \forall t \in \cup_{j \in \mathcal{B}_1^k}[u_k^j, u_k^{j-1}) \text{ and } B(i,m,l_2,t) > 0, \forall t \in \cup_{j \in \mathcal{B}_2^k}[u_k^j, u_k^{j-1}) \tag{50}$$

Also $|\mathcal{B}_1^k| \leq |\mathcal{B}_2^k|$. From (46), (47), (49) (38) and assumption (6), for $j \in \mathcal{B}_2^k$,

$$P(i,m,l_1,u_k^j,u_k^{j-1}) \geq \begin{cases} P(i,m,l_2,u_k^j,u_k^{j-1}) + W - 2 - \Upsilon\Gamma & 1 \leq j < J(k), \\ P(i,m,l_2,u_k^j,u_k^{j-1}) - 1 - 2\Upsilon\Gamma & j = J(k). \end{cases} \tag{51}$$

Note that $\mathcal{B}_2^k \neq \phi$. Proceeding as in the proof of (43),

$$\begin{aligned} P(i,m,l_1,s_k,t_k) \quad \geq \quad & \sum_{j \in \mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j \in \mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\ & + I_{|\mathcal{B}_2^k| \geq 1}(|\mathcal{B}_2^k| - 1)(W-2) - (|\mathcal{B}_2^k| + 1)\Upsilon\Gamma - 1 \end{aligned} \tag{52}$$

Note that $|\mathcal{B}_1^k| \leq |\mathcal{B}_2^k|$ in the last case considered, i.e., when $B(i,m,l_2,t_k) \geq W$. So combining (43) and (52) we have

$$\begin{aligned} P(i,m,l_1,s_k,t_k) \quad \geq \quad & \sum_{j \in \mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{j \in \mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\ + \quad & I_{|\mathcal{B}_2^k| \geq 1} \Big((|\mathcal{B}_2^k| - 2)(W-2) - 1 - (|\mathcal{B}_2^k| + 1)\Upsilon\Gamma \\ & + I_{|\mathcal{B}_1^k| > |\mathcal{B}_2^k|}(W - 1 + \Upsilon\Gamma) \Big) \end{aligned} \tag{53}$$

$$\begin{aligned} \min_{l \in \kappa(i,l_1)} B(i,m,l,t) \quad & \leq \quad B(i,m,l_2,t) \text{ (since } l_2 \in \kappa(i,l_1)) \\ & < \quad W, \forall t \in \cup_{k=1}^K \cup_{j \in \mathcal{B}_1^k} [u_k^j, u_k^{j-1}) \end{aligned} \tag{54}$$

The last inequality follows from (39) and (50).

$$\begin{aligned} & \sum_{k=1}^K P(i,m,l_1,s_k,t_k) \\ \geq \quad & \sum_{k=1}^K \sum_{j \in \mathcal{B}_1^k} P(i,m,l_1,u_k^j,u_k^{j-1}) + \sum_{k=1}^K \sum_{j \in \mathcal{B}_2^k} P(i,m,l_2,u_k^j,u_k^{j-1}) \\ & + (W-2) \sum_{k=1}^K \Big(I_{|\mathcal{B}_2^k| \geq 1}(|\mathcal{B}_2^k| - 1) \Big) - K(1 + \Upsilon\Gamma) - \sum_{k=1}^K |\mathcal{B}_2^k| \Upsilon\Gamma \end{aligned}$$

$$
\begin{aligned}
& +(W-1)\sum_{k=1}^{K} I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1} \text{ (from (53))} \\
\geq\ & r\sum_{k=1}^{K}\sum_{j\in\mathcal{B}_1^k}(u_k^j-u_k^{j-1})+\sum_{k=1}^{K}\sum_{j\in\mathcal{B}_2^k}(u_k^j-u_k^{j-1})-e_1-e_2 \\
& -\sum_{k=1}^{K}|\mathcal{B}_1^k|f_1-\sum_{k=1}^{K}|\mathcal{B}_2^k|(f_2+\Upsilon\Gamma)+(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right) \\
& -K(1+\Upsilon\Gamma)+(W-1)\sum_{k=1}^{K} I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1} \\
& \text{(from (39), (50), (54) and assumptions (2) and (3))} \\
\geq\ & r\left(\sum_{k=1}^{K}\left(\sum_{j\in\mathcal{B}_1^k}(u_k^j-u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}(u_k^j-u_k^{j-1})\right)\right)-e_1-e_2 \\
& -\sum_{k=1}^{K}\left(|\mathcal{B}_2^k|+I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1}+I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|=0}\right)f_1-\sum_{k=1}^{K}|\mathcal{B}_2^k|(f_2+\Upsilon\Gamma) \\
& +(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right)-K(1+\Upsilon\Gamma) \\
& +(W-1)\sum_{k=1}^{K} I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1} \text{ (since } |\mathcal{B}_1^k|\leq|\mathcal{B}_2^k|+1\ \forall\ k) \\
=\ & r\left(\sum_{k=1}^{K}\left(\sum_{j\in\mathcal{B}_1^k}(u_k^j-u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}(u_k^j-u_k^{j-1})\right)\right)-e_1-e_2 \\
& -\sum_{k=1}^{K}\left(|\mathcal{B}_2^k|+I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|=0}\right)f_1-\sum_{k=1}^{K}|\mathcal{B}_2^k|(f_2+\Upsilon\Gamma) \\
& +(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right)-K(1+\Upsilon\Gamma) \\
& +(W-1-f_1)\sum_{k=1}^{K} I_{|\mathcal{B}_1^k|>|\mathcal{B}_2^k|\geq 1} \\
\geq\ & r\left(\sum_{k=1}^{K}\left(\sum_{j\in\mathcal{B}_1^k}(u_k^j-u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}(u_k^j-u_k^{j-1})\right)\right)-e_1-e_2 \\
& \sum_{k=1}^{K}\left(|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}+I_{|\mathcal{B}_2^k|=0}\right)f_1-\sum_{k=1}^{K}|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}(f_2+\Upsilon\Gamma) \\
& +(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right)-K(1+\Upsilon\Gamma) \\
& \text{(from assumption (4))} \\
\geq\ & r\left(\sum_{k=1}^{K}\left(\sum_{j\in\mathcal{B}_1^k}(u_k^j-u_k^{j-1})+\sum_{j\in\mathcal{B}_2^k}(u_k^j-u_k^{j-1})\right)\right)-e_1-e_2 \\
& -\sum_{k=1}^{K}|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}f_1-\sum_{k=1}^{K}f_1 I_{|\mathcal{B}_2^k|=0}-\sum_{k=1}^{K}|\mathcal{B}_2^k|I_{|\mathcal{B}_2^k|\geq 1}(f_2+\Upsilon\Gamma)
\end{aligned}
$$

$$
\begin{aligned}
& +(W-2)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-2)\right)-K(1+\Upsilon\Gamma) \\
= \quad & r\sum_{k=1}^{K}(t_k-s_k)+(W-2-f_1-f_2-\Upsilon\Gamma)\sum_{k=1}^{K}\left(I_{|\mathcal{B}_2^k|\geq 1}(|\mathcal{B}_2^k|-1)\right) \\
& -\sum_{k=1}^{K}\left(\left((f_2+\Upsilon\Gamma)I_{|\mathcal{B}_2^k|\geq 1}+(f_2+\Upsilon\Gamma)I_{|\mathcal{B}_2^k|=0}\right)+\left(f_1 I_{|\mathcal{B}_2^k|\geq 1}\right)\right) \\
& -K(1+\Upsilon\Gamma) \\
\geq \quad & r\sum_{k=1}^{K}(t_k-s_k)-K(f_1+f_2+1+2\Upsilon\Gamma) \text{ (assumption (4))}
\end{aligned}
$$

Thus (7) follows. $\square$

Now, we generalize (7) of Lemma 1 in Corollary 1. We will use Lemma 1 and Corollary 1 in proving Lemma 2, which is a key result in proving Theorem 4. We will use Lemma 1 in proving Corollary 1.

Corollary 1: Consider layer m of session i. Let l_1 be a link on the path of session i. Let K be a positive integer and let $s_1, t_1, s_2, t_2, \ldots, s_K, t_K$ be a nondecreasing sequence of times. Let r, e_1, e_2, f_1, f_2 be some nonnegative real numbers.

For any $J(k)$ intervals $[u_k^1, v_k^1), \ldots, [u_k^{J(k)}, v_k^{J(k)})$ in $[s_k, t_k]$, s.t.
$\min_{l\in\kappa(i,l_1)} B(i,m,l,t) < W \ \forall\ t \in \cup_{k=1}^{K}\cup_{j=1}^{J(k)}[u_k^j, v_k^j)$,

$$\sum_{k=1}^{K}\sum_{j=1}^{J(k)} P(i,m,l_1,u_k^j,v_k^j) \geq r\sum_{k=1}^{K}\sum_{j=1}^{J(k)}(v_k^j-u_k^j)-e_1-\sum_{k=1}^{K}\sum_{j=1}^{J(k)} f_1 \tag{55}$$

Let $\phi \subset \mathcal{A} \subseteq \kappa(i,l_1)$ s.t. if $l \in \mathcal{A}$, then for any $J(k)'$ intervals $[u_k^1, v_k^1), \ldots,$
$[u_k^{J(k)'}, v_k^{J(k)'})$ in $[s_k, t_k]$, s.t. $B(i,m,l,t) > 0 \ \forall\ t \in \cup_{k=1}^{K}\cup_{j=1}^{J(k)'}[u_k^j, v_k^j)$,

$$\sum_{k=1}^{K}\sum_{j=1}^{J(k)'} P(i,m,l,u_k^j,v_k^j) \geq r\sum_{k=1}^{K}\sum_{j=1}^{J(k)'}(v_k^j-u_k^j)-e_2-\sum_{k=1}^{K}\sum_{j=1}^{J(k)'} f_2 \tag{56}$$

$$\text{Let } \min_{l\in\mathcal{A}} B(i,m,l,s_k) \leq W+\Upsilon\Gamma \ \forall\ k, k=1,\ldots,K \tag{57}$$

$$W > 2 \text{ and } f_1+f_2+2+\Upsilon\Gamma \leq W < G < \infty \tag{58}$$

Then the following holds.

$$\sum_{k=1}^{K} P(i,m,l_1,s_k,t_k) \geq r\sum_{k=1}^{K}(t_k-s_k)-|\mathcal{A}|(e_1+e_2)-K(f_1+f_2+1+2\Upsilon\Gamma) \tag{59}$$

Proof of Corollary 1: By (57), for any k, there exists a $l \in \mathcal{A}$, s.t. $B(i,m,l,s_k) \leq W+\Upsilon\Gamma$. Consider the least numbered link l in $\mathcal{A}$, with $B(i,j,l,s_k) \leq W+\Upsilon\Gamma$ as $l(k)$. Note that the values of $l(k)$ partition these intervals into equivalence classes. Let there be K_1 intervals in the equivalence class induced by $l(1)$. We will apply Lemma 1 for these intervals with $l_2 = l(1)$. Clearly (2) holds from (55) and (3) holds from (56). Also

(4) holds from (58). Finally $B(i, m, l(1), s_k) \leq W + \Upsilon\Gamma$, for all intervals k in this equivalence class. Thus (6) holds. From (7) of Lemma 1

$$\sum_{k:l(k)=l(1)} P(i, m, l_1, s_k, t_k) \geq r \sum_{k:l(k)=l(1)} (t_k - s_k) - (e_1 + e_2) - K_1(f_1 + f_2 + 1 + 2\Upsilon\Gamma)$$

Now consider another equivalence class. Let it be induced by $l(p)$. Again Lemma 1 applies for l_1, $l(p)$. Let there be K_p intervals in this equivalence class. Thus

$$\sum_{k:l(k)=l(p)} P(i, m, l_1, s_k, t_k) \geq r \sum_{k:l(k)=l(p)} (t_k - s_k) - (e_1 + e_2) - K_p(f_1 + f_2 + 1 + 2\Upsilon\Gamma)$$

There can be at most $|\mathcal{A}|$ distinct equivalence classes. Since the equivalence classes are disjoint, summing the result for the equivalence classes, we obtain (59). $\square$

Now we prove Lemma 2. We introduce a few more additional notations. The number of times session i is polled (offered service) in link l in time interval $[s, t)$ is $C(i, l, s, t)$. The number of times layer j of session i is polled (offered service) in link l in time interval $[s, t)$ is $C(i, j, l, s, t)$. Also, by convention, logical buffers at the session sources are always nonempty, i.e., $B(i, j, l, t) > 0$, for all sessions i, layers j and time t if l is the access link of session i. This is simply a matter of convention, and does not mean that the source always tries to send packets. Number of times a source i tries to send layer j packets over access link l in interval $[s, t)$ is given by $C(i, j, l, s, t)$.

We will use lemma 2 in proving Theorem 4. We use lemma 1 and Corollary 1 in the proof. Lemma 2 is similar to Lemma 1 of [8]. The important differences are that the assumptions in lemma 2 are more restrictive than those in Lemma 1 of [8]. For example, Lemma 2 does not assume that the polling rate of a session satisfies (60) in all links on the path of a session, but requires that the throughput be upper bounded in the links that do not satisfy (60). This relaxation of assumption is crucial to accommodate the multirate case, in which the bandwidth offered to a multicast session may be different along different paths. However, the relaxation complicates the proof significantly. This lemma shows that if the polling rate of a session is lower bounded along a source destination path, then the throughput along the same path is also lower bounded, provided that the throughput along the paths where the polling rate is not high are upper bounded. The proof shows that the flow control in a source destination path where the polling rate is high is not affected by the polling rate in other links, as the logical buffer queue lengths are high at the junctions of these different paths. In fact, the second part of the lemma actually argues that the logical buffer queue lengths are high at these junctions.

Lemma 2: Consider receiver j and layer m transmitted by the source of receiver j. Let $T \geq 0$. Let there exist nonnegative real numbers e_4, e_5, f_4, f_5 and positive real numbers r, δ, s.t. $r \geq \delta$, $\max_{l:\chi(j) \in n(l)} C_l \geq \delta$ and the following conditions hold.

For any link $l \in L_j$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, s_2, t_2, \ldots, s_K, t_K$ of times, $s_1 \geq T$:

$$\sum_{k=1}^{K} C(\chi(j), m, l, s_k, t_k) \geq r \sum_{k=1}^{K} (t_k - s_k) - e_4 - K f_4 \tag{60}$$

For any link l on session $\chi(j)$ path, one of the following conditions hold

1. For all links on at least one source destination path including l (60) holds.

2. For any positive integer K, and for any nondecreasing sequence $s_1, t_1, s_2,$ $t_2, \ldots, s_K, t_K$ of times, $s_1 \geq T$,

$$\sum_{k=1}^{K} P(\chi(j), m, l, s_k, t_k) \leq (r - \delta) \sum_{k=1}^{K} (t_k - s_k) + e_5 + K f_5 \tag{61}$$

The window W is sufficiently large. Let $H_{\chi(j)}$ be the maximum source to destination path length for a receiver.

$$W > 2 \text{ and } W \geq H_{\chi(j)} f_4 + 2H_{\chi(j)} + 3(H_{\chi(j)} - 1)\Upsilon\Gamma \tag{62}$$

The buffer size is sufficiently large.

$$\text{If all session } \chi(j) \text{ links satisfy (60), } G > W \tag{63}$$

$$\begin{aligned}
\text{else } G \quad > \quad & \left(1 + \frac{\max_{l:\chi(j) \in n(l)} C_l}{\delta}\right) W + 2 + \frac{\max_{l:\chi(j) \in n(l)} C_l}{\delta} \Big((H_{\chi(j)} - 1) e_4 \\
& + \beta_{\chi(j)}(H_{\chi(j)} - 1, e_4) + (H_{\chi(j)} + 1)(f_4 + 2) + e_5 + f_5 \\
& + 3H_{\chi(j)} \Upsilon\Gamma \Big)
\end{aligned} \tag{64}$$

$$\begin{aligned}
\text{where } \beta_{\chi(j)}(1, e_4) &= e_4 \\
\beta_{\chi(j)}(x, e_4) &= e_4 \nu_{\chi(j)} \frac{2\nu_{\chi(j)}^{x-1} - \nu_{\chi(j)}^{x-2} - 1}{\nu_{\chi(j)} - 1} \text{ if } x > 1
\end{aligned}$$

Let $d_i(l)$ be the number of hops between the origin of link l and the farthest destination of session i, on session i path traversing link l. Let $T_{\chi(j)}(x, T)$ be the following sequence of times

$$\begin{aligned}
T_{\chi(j)}(1, T) &= T \\
T_{\chi(j)}(x+1, T) &= T_{\chi(j)}(x, T) + \frac{1}{\delta} \Big(W + 1 + (H_{\chi(j)} - x) f_4 + (H_{\chi(j)} - x) 2^{H_{\chi(j)} - x} f_4 \\
& \quad + x f_4 + x + f_5 + (H_{\chi(j)} - x) e_4 + \beta_{\chi(j)}(x, e_4) + e_5 \\
& \quad + \max_{\substack{l''': \chi(j) \in n(l''') \\ d_{\chi(j)}(l''') \leq x}} \Big(B(\chi(j), m, l''', 0) + C_{l'''} T_{\chi(j)}(x, T) \Big) \Big) \\
& \quad \text{(if a session } \chi(j) \text{ path satisfying (60) and another} \\
& \quad \text{not satsfying (60) merges at the sink of a link } l \\
& \quad \text{with } d_{\chi(j)}(l) = x + 1) \\
T_{\chi(j)}(x+1, T) &= T_{\chi(j)}(x, T) \text{ otherwise.}
\end{aligned}$$

Then for all $l \in L_j$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, s_2, t_2, \ldots, s_K, t_K$ of times, $s_1 \geq T_{\chi(j)}(H_{\chi(j)}, T)$,

$$\begin{aligned} \sum_{k=1}^{K} P(\chi(j), m, l, s_k, t_k) &\geq r \sum_{k=1}^{K} (t_k - s_k) - e - Kf \\ \text{where } e &= (H_{\chi(j)} - 1)e_4 + e_4 \beta_{\chi(j)} \left(H_{\chi(j)}, e_4\right) \\ f &= H_{\chi(j)} f_4 + 2H_{\chi(j)} - 1 + \left(3H_{\chi(j)} - 4\right) \Upsilon\Gamma \end{aligned} \tag{65}$$

Also $B(\chi(j), m, l, t) > W\Upsilon\Gamma$, for $t \geq T_{\chi(j)}(H_{\chi(j)}, T)$, for all session $\chi(j)$ links l which do not lie on a session $\chi(j)$ path satisfying (60), and have a session $\chi(j)$ sibling that lies on a session $\chi(j)$ path satisfying (60).

Proof of Lemma 2: Let $h_{\chi(j)}(l)$ be the number of hops between the origin of link l and the source of session $\chi(j)$.

$$\text{From (63) and (64), } G > W \text{ always holds.} \tag{66}$$

We first show that the following holds for every link l which lies on a session $\chi(j)$ source destination path satisfying (60).

For any positive integer K, and for any nondecreasing sequence $s_1, t_1, s_2, t_2, \ldots, s_K$, t_K of times, $s_1 \geq T$, s.t. $\min_{l' \in \kappa(\chi(j), l)} B(\chi(j), m, l', t) < W$, for all $t \in \cup_{k=1}^{K} [s_k, t_k)$,

$$\begin{aligned} \sum_{k=1}^{K} P(\chi(j), m, l, s_k, t_k) \geq\ & r \sum_{k=1}^{K} (t_k - s_k) - (h_{\chi(j)}(l) + 1)e_4 \\ & -K \left(\left(h_{\chi(j)}(l) + 1\right) (f_4 + \Upsilon\Gamma + 1) \right. \\ & \left. + \left(h_{\chi(j)}(l)(1 + \Upsilon\Gamma)\right) \right) \end{aligned} \tag{67}$$

For any positive integer K, and for any nondecreasing sequence $s_1, t_1, s_2, t_2, \ldots, s_K$, t_K of times, $s_1 \geq T_{\chi(j)} \left(d_{\chi(j)}(l), T\right)$, s.t. $B(\chi(j), m, l, t) > 0$, for all $t \in \cup_{k=1}^{K} [s_k, t_k)$,

$$\begin{aligned} \sum_{k=1}^{K} P(\chi(j), m, l, s_k, t_k) \geq\ & r \sum_{k=1}^{K} (t_k - s_k) - \beta_{\chi(j)} \left(d_{\chi(j)}(l), e_4\right) \\ & -K \left(d_{\chi(j)}(l)(f_4 + 1)\right. \\ & \left. + +(d_{\chi(j)}(l) - 1)(3\Upsilon\Gamma + 1)\right) \end{aligned} \tag{68}$$

We first prove (67). We prove it by forward induction on $h_{\chi(j)}(l)$. Let $h_{\chi(j)}(l) = 0$ and l lie on a session $\chi(j)$ path satisfying (60). Consider any positive integer K and any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T$. Now, $B(\chi(j), m, l, t) > 0$, for all t. Thus $\forall\ k = 1, \ldots, K$, $P(\chi(j), m, l, s_k, t_k) \geq C(\chi(j), m, l, s_k + \Upsilon, t_k) - 1$, if $\min_{l' \in \kappa(\chi(j), l)} B(\chi(j), m, l', t) < W$, for all $t \in [s_k, t_k)$. This is because the scheduler at the origin of link l knows that $\min_{l' \in \kappa(\chi(j), l)} B(\chi(j), m, l', t) < W$, for all $t \in [s_k + \Upsilon, t_k)$. The term $P(\chi(j), m, l, s_k, t_k)$ may still be less than $C(\chi(j), m, l, s_k + \Upsilon, t_k)$ by 1 because the former counts the number of packets which finish transmission, and the latter the number of samplings. Thus (67) follows from (60), and from the fact that $b_{i,j} \leq \Gamma$ without loss of generality, for all sessions i and layers j.

Let (67) hold for all links l that lie on a session $\chi(j)$ path satisfying (60) and have $h_{\chi(j)}(l) \leq x$. We would show that (67) holds for all links l that lie on a session $\chi(j)$ path satisfying (60) and have $h_{\chi(j)}(l) = x+1$. Consider one such link l lying on a session $\chi(j)$ path satisfying (60) and with $h_{\chi(j)}(l) = x+1$. Let l_1 be the link on session $\chi(j)$ path, terminating at the origin of l. (Note that there can be only one such link). We will apply Lemma 1 with l_1, $l_2 = l$, $e_1 = (x+1)e_4$, $f_1 = (x+1)(f_4+\Upsilon\Gamma+1)+x(\Upsilon\Gamma+1)$, $e_2 = e_4$, $f_2 = f_4+1+\Upsilon\Gamma$. Consider any positive integer K and any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T$.

$$\text{Let } \min_{l' \in \kappa(\chi(j),l)} B(\chi(j), m, l', t) < W, \forall t \in \cup_{k=1}^{K}[s_k, t_k) \tag{69}$$

Clearly $h_{\chi(j)}(l_1) = x$. Also l_1 lies on a session $\chi(j)$ path satisfying (60).

Thus (67) holds for l_1 from induction hypothesis. So (2) holds. For any positive integer $J(k)$ and any sequence of nondecreasing times, $u_k^{J(k)}, v_k^{J(k)}, \ldots, v_k^1, u_k^1$, s.t. $B(\chi(j), m, l, t) > 0$, for all $t \in \cup_{k=1}^{K} \cup_{q=1}^{J(k)} [u_k^q, v_k^q)$, $\sum_{k=1}^{K}\sum_{q=1}^{J(k)} P(\chi(j), m, l, u_k^q, v_k^q)$
$\geq \sum_{k=1}^{K}\sum_{q=1}^{J(k)} \left(C(\chi(j), q, l, u_k^q, v_k^q) - 1 - \Upsilon\Gamma\right)$ (from (69)). Now condition (3) follows from (60). Condition (4) follows from (62) and (66). Thus from (5) of Lemma 1, (67) holds for l.

Now we prove (68). We will show this by induction on the value of $d_{\chi(j)}(l)$. Let $d_{\chi(j)}(l) = 1$ and l lie on a session $\chi(j)$ path satisfying (60). Consider any positive integer K and any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T$. If $B(\chi(j), m, l, t) > 0$, $\forall\ t \in \cup_{k=1}^{K}[s_k, t_k)$, then $\sum_{k=1}^{K} P(\chi(j), m, l, s_k, t_k) \geq \sum_{k=1}^{K} \left(C(\chi(j), m, l, s_k, t_k) - 1\right)$, and hence (68) holds from (60), with
$T_{\chi(j)}(1, T) = T$.

Let (68) hold for any link l that lies on a session $\chi(j)$ path satisfying (60) and has $d_{\chi(j)}(l) \leq x$. We would show that (68) holds for all links l that lie on a session $\chi(j)$ path satisfying (60) and have $d_{\chi(j)}(l) = x+1$. Consider one such link l lying on a session $\chi(j)$ path satisfying (60) and with $d_{\chi(j)}(l) = x+1$. Let $\mathcal{A}_1$ consist of links in $\kappa(\chi(j), l)$ and lying on session $\chi(j)$ paths satisfying (60). Since l lies on a session $\chi(j)$ path satisfying (60), at least one link in $\kappa(\chi(j), l)$ lies on a session $\chi(j)$ path satisfying (60). Thus $\mathcal{A}_1 \neq \phi$. We will show that $\min_{l' \in \mathcal{A}_1} B(\chi(j), m, l', t) \leq W + \Upsilon\Gamma$, for all $t \geq T_{\chi(j)}(x+1, T)$. If $\mathcal{A}_1 = \kappa(\chi(j), l)$, then this result holds for all t. Now, let that not be so, i.e., $\mathcal{A}_1 \subset \kappa(\chi(j), l)$. Note that (64) holds in this case. We will apply Lemma 1. Let $l' \in \mathcal{A}_1$. Let $l_1 = l$, $l_2 = l'$, $e_1 = \left(h_{\chi(j)}(l)+1\right)e_4$, $e_2 = \beta_{\chi(j)}\left(d_{\chi(j)}(l'), e_4\right)$, $f_1 = (h_{\chi(j)}(l)+1)(f_4+\Upsilon\Gamma+1)+h_{\chi(j)}(l)(\Upsilon\Gamma+1)$, and $f_2 = d_{\chi(j)}(l')(f_4+1)+(d_{\chi(j)}(l')-1)(3\Upsilon\Gamma+1)$. Consider any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{\chi(j)}(x, T)$. Now, condition (2) holds from (67), since $T_{\chi(j)}(x, T) \geq T$, $\forall x \geq 1$. Condition (3) holds from induction hypothesis (68), because $d_{\chi(j)}(l') \leq x$, $l' \in \mathcal{A}_1$ and $T_{\chi(j)}(x, T)$ is nondecreasing in x, $x \leq H_{\chi(j)}$. Condition (4) holds from (62) and (66). From (5) of

Lemma 1,

$$\sum_{k=1}^{K} P(\chi(j), m, l', s_k, t_k) \geq r\sum_{k=1}^{K}(t_k - s_k) - (e_1 + e_2) - K(f_1 + f_2 + \Upsilon\Gamma + 1) \tag{70}$$

Consider any $t \geq T_{\chi(j)}(x,T)$. Result (70) holds for any positive integer K, and $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{\chi(j)}(x,T)$. So consider $K = 1$, $s \in [T_{\chi(j)}(x,T), t]$ and $t_1 = t$. It follows that

$$\begin{aligned} P\left(\chi(j), l', s, t\right) &\geq r(t-s) - (e_1+e_2) - (f_1 + f_2 + \Upsilon\Gamma + 1) && (71)\\ P\left(\chi(j), l, s - \Upsilon_l, t - \Upsilon_l\right) &\geq P\left(\chi(j), l', s, t\right) - B\left(\chi(j), l', s\right) + B(\chi(j), m, l', t) \\ &\geq P\left(\chi(j), l', s, t\right) - B\left(\chi(j), l', s\right) \\ &\geq r(t-s) - (e_1+e_2) - (f_1 + f_2 + \Upsilon\Gamma + 1) \\ &\quad -B(\chi(j), m, l', s) && (72)\\ &\quad \text{for any } s \in [T_{\chi(j)}(x,T), t] \text{ (from (71))} \end{aligned}$$

Now consider any $l'' \in \kappa(\chi(j), l) \backslash \mathcal{A}_1$. Thus l'' satisfies (61). Let $B(\chi(j), m, l'', t_1)$ $= G$, for some $t_1 \in [T_{\chi(j)}(x,T), t)$. We would show that $B(\chi(j), m, l'', t) > W$. Let that not be so, i.e., $B(\chi(j), m, l'', t) \leq W$. Let t_2 be the earliest time s.t. for all $t' \in [t_2, t)$, $B(\chi(j), m, l'', t') < G$. We know there always exists one such $t_2 \in (t_1, t)$ because $B(\chi(j), m, l'', t_1) = G > W + 1$ (by (64)), and $B(\chi(j), m, l'', t) \leq W$. Clearly $t_2 > t_1 \geq T_{\chi(j)}(x,T)$, and $t_2 < t$. Thus $t_2 \in [T_{\chi(j)}(x,T), t)$ and $B(\chi(j), m, l'', t_2) = G - 1$. From (72) and the fact that $B(\chi(j), m, l', u) \leq G$, for all times u,

$$\begin{aligned} P\left(\chi(j), l, t_2 - \Upsilon_l, t - \Upsilon_l\right) &\geq r(t - t_2) - (e_1+e_2) - (f_1 + f_2 + \Upsilon\Gamma + 1) \\ &\quad -G && (73) \end{aligned}$$

Since service is upper bounded by link capacity,

$$C_{l''}(t - t_2) \geq P\left(\chi(j), l'', t_2, t\right) - 1 \tag{74}$$

$$\begin{aligned} B(\chi(j), m, l'', t) &= B(\chi(j), m, l'', t_2) + P\left(\chi(j), l, t_2, t\right) - P\left(\chi(j), l'', t_2, t\right) \\ &\quad \text{(since } B(\chi(j), m, l'', t') < G \forall t' \in [t_2, t)) \\ &\geq B(\chi(j), m, l'', t_2) - P\left(\chi(j), l'', t_2, t\right) \end{aligned}$$

$$\begin{aligned} \text{Thus } P\left(\chi(j), l'', t_2, t\right) &\geq B(\chi(j), m, l'', t_2) - B(\chi(j), m, l'', t) \\ &\geq G - 1 - W && (75)\\ &\quad \text{(since } B(\chi(j), m, l'', t_2) = G - 1 \text{ and} \\ &\quad B(\chi(j), m, l'', t) \leq W) \\ \text{Thus } t - t_2 &\geq \frac{G - 2 - W}{C_{l''}} \text{(from (74) and (75))} && (76) \end{aligned}$$

$$\begin{aligned}
B(\chi(j), m, l'', t) &= B(\chi(j), m, l'', t_2) + P(\chi(j), l, t_2 - \Upsilon_l, t - \Upsilon_l) \\
&\quad - P(\chi(j), l'', t_2, t) \\
&\quad \text{(since } B(\chi(j), m, l'', t') < G \ \forall t' \in [t_2, t)) \\
&\geq G - 1 + \delta(t - t_2) - (e_1 + e_2) - (f_1 + f_2 + \Upsilon\Gamma + 1) \\
&\quad - G - e_5 - f_5 \\
&\quad \text{(from (73) and (61))} \\
&= \delta(t - t_2) - (e_1 + e_2) - (f_1 + f_2 + \Upsilon\Gamma + 2) - e_5 - f_5 \\
&\geq \frac{\delta(G - 2 - W)}{C_{l''}} - (e_1 + e_2) - (f_1 + f_2 + \Upsilon\Gamma + 2) \\
&\quad - e_5 - f_5 \\
&> W + \Upsilon\Gamma \text{ (from (64) and the following)} \qquad (77) \\
e_1 &\leq (H_{\chi(j)} - x) e_4 \qquad (78) \\
f_1 &\leq (H_{\chi(j)} - x)(f_4 + \Upsilon\Gamma + 1)(H_{\chi(j)} - x)(\Upsilon\Gamma + 1) \qquad (79) \\
e_2 &= \beta_{\chi(j)}(x, e_4) \qquad (80) \\
f_2 &\leq x(f_4 + 1) + (x - 1)(3\Upsilon\Gamma + 1) \qquad (81)
\end{aligned}$$

Thus this contradicts the assumption that $B(\chi(j), m, l'', t) \leq W$. Now, let $B(\chi(j), m, l'', t_1) < G \ \forall \ t_1 \in [T_{\chi(j)}(x, T), t)$. Additionally, let

$$\begin{aligned}
t &\geq T_{\chi(j)}(x, T) + \frac{1}{\delta}\Big(W + 1 + (f_1 + f_2 + f_5 + 2\Upsilon\Gamma + 1) + (e_1 + e_2 + e_5) \\
&+ B\left(\chi(j), l', T_{\chi(j)}(x, T)\right)\Big) \qquad (82)
\end{aligned}$$

So there are no session $\chi(j)$ packet losses from session $\chi(j)$ buffer for l'' in interval $[T_{\chi(j)}(x, T), t)$. It follows that

$$\begin{aligned}
&B(\chi(j), m, l'', t) \\
=\ & B\left(\chi(j), l'', T_{\chi(j)}(x, T)\right) + P\left(\chi(j), l, T_{\chi(j)}(x, T) - \Upsilon_l, t - \Upsilon_l\right) \\
& - P\left(\chi(j), l'', T_{\chi(j)}(x, T), t\right) \\
\geq\ & P\left(\chi(j), l, T_{\chi(j)}(x, T) - \Upsilon_l, t - \Upsilon_l\right) - P\left(\chi(j), l'', T_{\chi(j)}(x, T), t\right) \\
\geq\ & \delta\left(t - T_{\chi(j)}(x, T)\right) - (e_1 + e_2 + e_5) - (f_1 + f_2 + f_5 + \Upsilon\Gamma + 1) \\
& - B\left(\chi(j), l', T_{\chi(j)}(x, T)\right) \text{ (from (61) and (72)} \\
& \qquad\qquad \text{with } s = T_{\chi(j)}(x, T)) \\
>\ & W + \Upsilon\Gamma \text{ (from (82))} \qquad (83)
\end{aligned}$$

From (77), given any $t \geq T_{\chi(j)}(x,T)$, if there exists t_1 s.t. $B(\chi(j),m,l'',t_1) = G$, for some $t_1 \in [T_{\chi(j)}(x,T),t)$, then $B(\chi(j),m,l'',t) > W + \Upsilon\Gamma$. If there does not exist such a t_1 then $B(\chi(j),m,l'',t) > W + \Upsilon\Gamma$, by (83), if t satisfies (82).

$$\text{It follows that } B(\chi(j),m,l'',t) > W + \Upsilon\Gamma, \forall\, t \geq T_{\chi(j)}(x+1,T), \tag{84}$$

where using (82), (78) to (81) and the fact that

$$B\left(\chi(j),l',T_{\chi(j)}(x,T)\right) \leq \max_{\substack{l''':\chi(j)\in n(l''')\\ d_{\chi(j)}(l''')\leq x}} \left(B(\chi(j),m,l''',0) + C_{l'''}T_{\chi(j)}(x,T)\right)$$

$$\begin{aligned}
T_{\chi(j)}(x+1,T) \quad &= \quad T_{\chi(j)}(x,T) + \frac{1}{\delta}\Big(W + 1 + (H_{\chi(j)} - x)f_4 \\
&\quad +(H_{\chi(j)} - x)2^{H_{\chi(j)}-x} f_4 + x f_4 + x + f_5 + 2\Upsilon\Gamma \\
&\quad +(H_{\chi(j)} - x)e_4 + \beta_{\chi(j)}(x,e_4) + e_5 \\
&\quad + \max_{\substack{l''':\chi(j)\in n(l''')\\ d_{\chi(j)}(l''')\leq x}} \left(B(\chi(j),m,l''',0) + C_{l'''}T_{\chi(j)}(x,T)\right)\Big) \\
&\quad \text{(if a session } \chi(j) \text{ path satisfying (60) and another} \\
&\quad \text{not satsfying (60) merges at the sink} \\
&\quad \text{of a link } l_3 \text{ with } d_{\chi(j)}(l_3) = x+1) \\
T_{\chi(j)}(x+1,T) \quad &= \quad T_{\chi(j)}(x,T) \text{ otherwise.}
\end{aligned}$$

This holds for all $l'' \in \kappa(\chi(j),l)\backslash\mathcal{A}_1$.

$$\text{Thus } \min_{l'\in\mathcal{A}_1} B(\chi(j),m,l',t) \leq W + \Upsilon\Gamma, \forall\, t \geq T_{\chi(j)}(x+1,T), \tag{85}$$

Now we will apply Corollary 1, with $l_1 = l$, $e_1 = e_4$, $f_1 = f_4 + \Upsilon\Gamma + 1$, $e_2 = \beta_{\chi(j)}(x,e_4)$, $f_2 = x(f_4+1) + (x-1)(3\Upsilon\Gamma+1)$, and $\mathcal{A}$=$\mathcal{A}_1$. Since l lies on a session $\chi(j)$ path satisfying (60), $\mathcal{A}_1 \neq \phi$. Consider any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{\chi(j)}(x+1,T)$, s.t. $B(\chi(j),m,l,t) > 0$, $\forall\, t \in \cup_{k=1}^{K}[s_k,t_k)$. Clearly $d_{\chi(j)}(l') \leq x$ for any link $l' \in\mathcal{A}_1$. Also l' lies on a session $\chi(j)$ path satisfying (60), if $l' \in\mathcal{A}_1$. Thus any $l' \in\mathcal{A}_1$ satisfies condition (56) by induction hypothesis (68), since $T_{\chi(j)}(x+1,T) \geq T_{\chi(j)}(x,T)$. Since we assume $B(\chi(j),m,l,t) > 0$, for all $t \in \cup_{k=1}^{K}[s_k,t_k)$, for any integer $J(k)$ and any nondecreasing sequence of times, $u_k^{J(k)}, v_k^{J(k)}, \ldots, v_k^1, u_k^1$ in $[s_k,t_k)$ s.t. $\min_{l'\in\kappa(i,l)} B(\chi(j),m,l',t) < W$ for all $t \in \cup_{k=1}^{K}\cup_{q=1}^{J(k)}[u_k^q, v_k)$

$$\sum_{k=1}^{K}\sum_{q=1}^{J(k)} P(\chi(j),m,l,u_k^q,v_k^q) \geq \sum_{k=1}^{K}\sum_{q=1}^{J(k)} \left(C(\chi(j),m,l,u_k^q,v_k^q) - \Upsilon\Gamma - 1\right)$$

Thus condition (55) holds from (60) and (66). Condition (57) holds from (85). Condition (58) holds from (62). Thus from (59) of Corollary 1, (68) holds for l. Thus (68) holds for all links l on session $\chi(j)$ path with $d_{\chi(j)}(l) = x+1$ and lying on a session $\chi(j)$ path that satisfies (60).

Now consider the path of receiver j. Consider any two successive links l_1 and l_2 on the path. We will use Lemma 1, with $e_1 = \left(h_{\chi(j)}(l_1)+1\right)e_4$, $f_1 = (h_{\chi(j)}(l_1)+1)(f_4+\Gamma\Upsilon+1)+h_{\chi(j)}(l_1)(\Gamma\Upsilon+1)$, $e_2 = \beta_{\chi(j)}\left(d_{\chi(j)}(l_2), e_4\right)$, $f_2 = d_{\chi(j)}(l_2)(f_4+1)+(d_{\chi(j)}(l_2)-1)(3\Gamma\Upsilon+1)$. Consider any sequence of nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{\chi(j)}(H_{\chi(j)}-1, T)$.† Note that conditions (2), (3) hold for l_1 and l_2 respectively from (67) and (68) respectively. This is because virtual session j path satisfies (60). Condition (4) holds from (62) and (66). Thus from (5),

$$
\begin{aligned}
&\sum_{k=1}^{K} P(\chi(j), m, l_2, s_k, t_k) \\
\geq\ & r\sum_{k=1}^{K}(t_k-s_k) - \left(\left(h_{\chi(j)}(l_1)+1\right)e_4 + \beta_{\chi(j)}\left(d_{\chi(j)}(l_2), e_4\right)\right) \\
&-K\left(\left(h_{\chi(j)}(l_1)+d_{\chi(j)}(l_2)+1\right)f_4 + 2\left(h_{\chi(j)}(l_1)+d_{\chi(j)}(l_2)\right)\right. \\
&\left.+1+\left(2h_{\chi(j)}(l_1)+3d_{\chi(j)}(l_2)-1\right)\Upsilon\Gamma\right) \\
\geq\ & r\sum_{k=1}^{K}(t_k-s_k) - \left((H_{\chi(j)}-1)e_4+\beta_{\chi(j)}(H_{\chi(j)}, e_4)\right) \\
&-K\left(H_{\chi(j)}f_4+2H_{\chi(j)}-1+(3H_{\chi(j)}-4)\Upsilon\Gamma\right)
\end{aligned}
\tag{86}
$$

From (86), (65) holds for all l except that originating from the source of session $\chi(j)$. For this link, (65) follows from (68). This is because $B(\chi(j), m, l, t) > 0$, for all t if l originates from the source.

Now we prove the second part of the lemma. Note that (84) shows that if (68) holds for all session $\chi(j)$ links l lying on a session $\chi(j)$ path satisfying (60), and $d_{\chi(j)}(l) \leq x$, then $B(\chi(j), m, l', t) > W + \Upsilon\Gamma$, for all $t \geq T_{\chi(j)}(x+1, T)$, and all links l' which satisfy the following properties:

Link l' does not lie on a session $\chi(j)$ path satisfying (60),

Link l' has a session $\chi(j)$ sibling that lies on a session $\chi(j)$ path satisfying (60), and

$d_{\chi(j)}(l') \leq x$.

We have already shown that (68) holds for all links on session $\chi(j)$ paths satisfying (60). Also $T_{\chi(j)}(x, T)$ is a nondecreasing sequence of times. So the second part of the lemma holds. □

Now we prove Lemma 3. We will use this lemma in proving theorem 4. This lemma states that if the throughput in a link upstream of link l is upper bounded, and the throughputs in all links between link l and this upstream link are lower bounded, then throughput in link l is also upper bounded.

Lemma 3: Consider a link l on session i path which satisfies the following properties w.r.t. layer j of session i for any time instants $t \geq s \geq T_0$, for some given T_0

1. There exists a link l_1 between session i source and link l (l may be l_1 as well) such that

$$P(i, j, l_1, s, t) \leq r(t-s)+e_1 \tag{87}$$

†Note that $d_{\chi(j)}(l_2) \leq H_{\chi(j)}-1$. Thus we use $T_{\chi(j)}(H_{\chi(j)}-1, T)$.

2. For all links l_2 on session i path between l_1 and l inclusive of link l,

$$P(i,j,l_1,s,t) \geq r(t-s) - e_2 \tag{88}$$

Let Q be the set of links between l_1 and l including l but excluding l_1. It follows that for any time instants $t \geq s \geq T_0$,

$$P(i,j,l,s,t) \leq r(t-s) + 2^{|Q|}e_1 + (2^{|Q|}-1)e_2 + |Q|\left(\max_{l\in Q} B(i,j,l,T_0) + 2\Upsilon\Gamma\right)$$

Proof of Lemma 3: Trace session i path from l_1 to l. Let l' be the link next to l_1 on this path.

$$\begin{aligned}
B(i,j,l',s) &\leq B(i,j,l',T_0) + P(i,j,l_1,T_0-\Upsilon_{l_1},s-\Upsilon_{l_1}) - P(i,j,l',T_0,s) \\
&\leq B(i,j,l',T_0) + P(i,j,l_1,T_0,s) - P(i,j,l',T_0,s) + \Upsilon\Gamma \\
&\leq B(i,j,l',T_0) + r(s-T_0) + e_1 - r(s-T_0) + e_2 + \Upsilon\Gamma \\
&\qquad \text{(from (87) and (88))} \\
&\leq B(i,j,l',T_0) + e_1 + e_2 + \Upsilon\Gamma
\end{aligned} \tag{89}$$

$$\begin{aligned}
P(i,j,l',s,t) &\leq P(i,j,l_1,s-\Upsilon_{l_1},t-\Upsilon_{l_1}) + B(i,j,l',s) - B(i,j,l',t) \\
&\leq P(i,j,l_1,s,t) + B(i,j,l',s) - B(i,j,l',t) + \Upsilon\Gamma \\
&\leq P(i,j,l_1,s,t) + B(i,j,l',s) + \Upsilon\Gamma \\
&\leq r(t-s) + B(i,j,l',T_0) + 2e_1 + e_2 + 2\Upsilon\Gamma
\end{aligned}$$

The last inequality follows from (87) and (89). Proceeding along subsequent links in this way,

$$\begin{aligned}
P(i,j,l,s,t) &\leq r(t-s) + 2^{|Q|}e_1 + (2^{|Q|}-1)e_2 \\
&\quad + |Q|\left(\max_{l\in Q} B(i,j,l,T_0) + 2\Upsilon\Gamma\right)
\end{aligned} \tag{90}$$

□

Now we prove Lemma 4. We will use this lemma in proving Theorem 4.

Lemma 4: Consider layer m of session i. Let l_1 and l_2 be two links on the path of session i with l_2 originating from the sink of l_1. Let u be a given time. Let e_7, e_8 be nonnegative real numbers and for any time $v_2 \geq v_1 \geq u$,

$$P(i,m,l_1,v_1,v_2) \geq r(v_2-v_1) - e_7 \tag{91}$$

$$P(i,m,l_2,v_1,v_2) \leq r(v_2-v_1) + e_8 \tag{92}$$

$$\text{Also } B(i,m,l_2,u) = G \tag{93}$$

$$G > W + e_7 + e_8 + 2\Upsilon\Gamma + 1 \tag{94}$$

It follows that $B(i, m, l_2, t) > W + \Upsilon\Gamma$ for all $t \geq u$.

Proof of Lemma 4: Let $B(i, m, l_2, t) \leq W + \Upsilon\Gamma$, for some $t \geq u$. Let v be the earliest time such that $B(i, l_2, t') < G$, for all $t' \in [v, t)$. We know that one such time v exists in (u, t) because $B(i, m, l_2, t) \leq W + \Upsilon\Gamma$, and $B(i, m, l_2, u) = G > W + \Upsilon\Gamma + 1$. Clearly $v > u$. Also, $B(i, m, l_2, v) = G - 1$, and there is no packet loss in session i buffer for l_2 in interval $[v, t)$. It follows that

$$\begin{aligned}
B(i, m, l_2, t) &= B(i, m, l_2, v) + P(i, m, l_1, v - \Upsilon_{l_1}, t - \Upsilon_{l_1}) - P(i, m, l_2, v, t) \\
&\geq B(i, m, l_2, v) + P(i, m, l_1, v, t) - P(i, m, l_2, v, t) - \Upsilon\Gamma \\
&\geq G - 1 + r(t - v) - e_7 - r(t - v) - e_8 - \Upsilon\Gamma \\
&\qquad \text{(from (91) and (92))} \\
&= G - 1 - e_7 - e_8 - \Upsilon\Gamma \\
&> W + \Upsilon\Gamma \text{ (from (94))}
\end{aligned}$$

This contradicts the assumption that $B(i, m, l_2, t) \leq W + \Upsilon\Gamma$. □

B. Proof of Theorems 1, 2 and 3

Now we will prove Theorem 4. The theorems in Section IV will follow from this theorem. We introduce a few more notations first. Let the virtual sessions be arranged in an increasing order of their maxmin fair rates. The first virtual session in this order has the least maxmin fair rate amongst all virtual sessions, the second one has the second minimum rate, etc. Let the index $\vartheta(s)$ of a virtual session s be its rank in this order, i.e., $\vartheta(s) = 1$ if virtual session s has the lowest maxmin fair rate. If two virtual sessions have equal maxmin fair rates and belong to equal priority sessions, they have equal indices. The rates in this order are $\hat{\lambda}_1, \hat{\lambda}_2, \ldots$, i.e., $\hat{\lambda}_1$ is the rate of all virtual sessions s with $\vartheta(s) = 1$ (Note that multiple virtual sessions may have the same index). The link index $\Delta(i, l)$ of a session i on a link l on its path is the maximum index of virtual sessions of session i traversing l. Let F be the total number of indices. Also, $\alpha(i, p)$ is the layer number such that $\sum_{m=1}^{\alpha(i,p)} b_{i,m} \leq \hat{\lambda}_p < \sum_{m=1}^{\alpha(i,p)+1} b_{i,m}$, if $b_{i,1} \leq \hat{\lambda}_p$. If $b_{i,1} > \hat{\lambda}_p$, then $\alpha(i, p) = 0$. By convention, $\alpha(i, 0) = 0$. We have made a tacit assumption in defining the $\alpha(i, j)$s, i.e., we assume that there exists a layer j such that $\hat{\lambda}_p < \sum_{m=1}^{j} b_{i,m}$, for all p. However, this may not hold always, particularly if a source transmits only a few layers. However, the necessary generalization is straightforward, and we omit it for brevity.

The proof uses Lemma 2, 3 and 4. The nature of this theorem is similar to Theorem 1 of [8]. However, the scheduling in [8] deals only with sessions. We need to deal with different layers as well, and the scheduling policy distinguishes between the different layers to provide them different qualities of service. This two level scheduling presents a significant twist to the proof. Also, the congestion related packet losses complicate the proof. For example, the scheduling policy should not serve a session at a high rate in a link if all the downstream links are bottlenecked at a low bandwidth. This is easily shown in [8] as there is no packet loss in the intermediate nodes, and the number of packets served in different links on the path of a session differ by a factor of the window size, W. This does not follow in our case on account of packet loss in the intermediate nodes. The key to overcoming this difficulty is the observation that the throughput in a link should be bounded by the maximum of the bandwidths offered by the downstream bottlenecks. Since different paths originate from the same node for multirate multicast sessions, bottleneck bandwidths of different paths may be different. The upper bound should depend on the maximum of the bandwidths offered by these bottlenecks. We showed that if the physical buffer size is sufficiently larger than the window size, then there is no packet loss in the intermediate nodes leading to the bottleneck link offering the largest bandwidth. This holds even if the physical buffer is finite. Using this result, it is possible to show that the bandwidth used by a session in a link is no greater than the maximum bandwidth offered by its downstream bottleneck links. The physical buffer sizes may equal the window sizes in [8]. But, the physical buffer sizes must be sufficiently larger than the window sizes in our case.

We define some recursive constants which we will use in stating the Theorem.

$$\begin{aligned}
\varepsilon_1(p,i,j) &= \max\left(\xi_{i,j}, (E-1)\gamma(p-1) + 2 + \sum_{m=1}^{j-1} \varepsilon_3(p,i,m)\right) \text{ if } j \le \alpha(i,p)+1 \\
\varepsilon_1(p,i,j) &= 0 \text{ if } j > \alpha(i,p)+1 \\
\varepsilon_2(p,i,j) &= H_i \varepsilon_1(p,i,j) + 2H_i - 1 + (3H_i - 4)\Gamma\Upsilon, \text{ if } j \le \alpha(i,p)+1 \\
\varepsilon_2(p,i,j) &= 0 \text{ if } j > \alpha(i,p)+1 \\
\varepsilon_3(p,i,j) &= \varepsilon_2(p,i,j)
\end{aligned}$$

$$\begin{aligned}
\varsigma_1(p,i,j) &= \sum_{m=1}^{j-1} \varsigma_3(p,i,m) + (E-1)\kappa(p-1) \text{ if } j \le \alpha(i,p)+1 \\
\varsigma_1(p,i,j) &= 0 \text{ if } j > \alpha(i,p)+1 \\
\varsigma_2(p,i,j) &= \left(H_i - 1 + \nu_i \frac{2\nu_i^{H_i-1} - \nu_i^{H_i-2} - 1}{\nu_i - 1}\right) \varsigma_1(p,i,j) \\
\varsigma_3(p,i,j) &= \varsigma_2(p,i,j) + (2^{H_i} - 1)\left(\varsigma_2(p,i,j) + \varepsilon_2(p,i,j)\right) \\
&\quad + (H_i - 1)\left(\max_{l: i \in n(l)} B(i,j,l,T_{p,j}) + 2\Upsilon\Gamma\right) \text{ if } j \le \alpha(i,p+1) \\
\varsigma_3(p,i,j) &= \varsigma_2(p,i,j) + 2^{H_i-1}\left(\left(\varrho(p) + \tau(p)\right) E_i + 1\right) \\
&\quad + (H_i - 1)\left(\max_{l: i \in n(l)} B(i,j,l,T_{p,j}) + W + 2\Upsilon\Gamma\right) \text{ if } j > \alpha(i,p+1)
\end{aligned}$$

$$\begin{aligned}
\varepsilon_1(p) &= \max_{i: \Delta(i, l_a(i)) \ge p} \varepsilon_1\left(p, i, \alpha(i,p)+1\right) \\
\varsigma_1(p) &= \max_{i: \Delta(i, l_a(i)) \ge p} \varsigma_1\left(p, i, \alpha(i,p)+1\right)
\end{aligned}$$

Note that $\varepsilon_1(p)$ and $\varsigma_1(p)$ are nondecreasing in p (even though the respective maximizations are taken over smaller sets with increase in p). This is because $\varepsilon_1(p_1, i_1, j_1) \geq \varepsilon_1(p_2, i_2, j_2)$ and $\varsigma_1(p_1, i_1, j_1) \geq \varsigma_1(p_2, i_2, j_2)$ if $p_1 \geq p_2$, irrespective of the values of i_1, j_1, i_2, j_2.

Let $\Lambda(p) = \max_i \alpha(i, p)$.

$$
\begin{aligned}
T_{1,j} &= 0, \forall j, \\
T_{p,0} &= T_{p-1,\Lambda(p)+1} \qquad p > 1 \\
T_{p,j} &= \max_i T_i\left(H_i, T_{p,j-1}\right) \qquad j \leq 1
\end{aligned}
$$

where the function T_i is defined in Lemma 2 and

$$
\begin{aligned}
\varrho(p) &= \max_i \sum_{m=1}^{\alpha(i,p)+1} \varsigma_2(p, i, m) \qquad p > 0 \\
\tau(p) &= \max_i \sum_{m=1}^{\alpha(i,p)+1} \varepsilon_2(p, i, m) \qquad p > 0 \\
\kappa(0) &= 0 \\
\kappa(p) &= \max_i \sum_{m=1}^{\alpha(i,p)+1} \varsigma_3(p, i, m), \qquad p > 0 \\
\gamma(p) &= \max_i \sum_{m=1}^{\alpha(i,p)+1} \varepsilon_3(p, i, m) \qquad p > 0 \\
\gamma(0) &= 0
\end{aligned}
$$

$$
W_p = H\left(2 + \varepsilon_1(p)\right) + 3(H-1)\Upsilon\Gamma \tag{95}
$$

$$
\begin{aligned}
G_p = {} & \left(H + \frac{\Gamma}{\min_{r_p \neq r_q, p,q \in M} |r_p - r_q|}\right) W_p + 3 \\
& + \frac{\Gamma}{\min_{r_p \neq r_q, p,q \in M} |r_p - r_q|}\left((H-1)\varsigma_1(p) + \beta\left(H-1, \varsigma_1(p)\right)\right. \\
& \left. + (H+1)\left(\varepsilon_1(p) + 2\right) + \kappa(p-1) + \gamma(p-1) + 3H\Upsilon\Gamma\right) + \\
& E\left(\rho(p) + \tau(p)\right)
\end{aligned} \tag{96}
$$

$$
\begin{aligned}
\text{where, } \beta(1, e) &= e \\
\beta(x, e) &= e\nu\frac{2\nu^{x-1} - \nu^{x-2} - 1}{\nu - 1} \text{ if } x > 1
\end{aligned}
$$

Theorem 4: Let the traffic be well-behaved, i.e., the number of layer j packets a session i wishes to send in any interval $[s, t)$ over its access link differs from the amount $b_{i,j}(t-s)$ by at most $\xi_{i,j}$. Thus,

$$
\left|C\left(i, j, l_a(i), s, t\right) - b_{i,j}(t-s)\right| \leq \xi_{i,j}, \tag{97}
$$

For all $p = 1, \ldots, F$, if $W \geq W_p$ and $G \geq G_p$, then for all s, t, $t \geq s \geq T_{p+1,0}$, for all sessions i and links l such that $\Delta(i, l) \geq p$,

$$
P(i, l, s, t) \geq \hat{\lambda}_p(t-s) - \varrho(p) - \tau(p) \tag{98}
$$

$$P(i,j,l,s,t) \geq \min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j-1} b_{i,m}\right), b_{i,j}\right)(t-s) - \varsigma_2(p,i,j)$$
$$-\varepsilon_2(p,i,j) \text{ for each layer } j \quad (99)$$

and if $\Delta(i,l) = p$, then

$$P(i,l,s,t) \leq \hat{\lambda}_p(t-s) + \kappa(p) + \gamma(p) \quad (100)$$
$$P(i,j,l,s,t) \leq \min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j-1} b_{i,m}\right)^+, b_{i,j}\right)(t-s) + \varsigma_3(p,i,j)$$
$$+\varepsilon_3(p,i,j) \text{ for each layer } j \quad (101)$$

Proof of Theorem 4: We will prove the following for $p = 1, \ldots, F$. Let $W \geq W_p$ and $G \geq G_p$.

For each link l and for each layer j of session i traversing l and $\Delta(i,l) \geq p$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,j-1}, \infty)$,

$$\sum_{k=1}^{K} C(i,j,l,s_k,t_k) \geq \min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j-1} b_{i,m}\right), b_{i,j}\right)\sum_{k=1}^{K}(t_k - s_k)$$
$$-\varsigma_1(p,i,j) - K\varepsilon_1(p,i,j) \quad (102)$$

For each link l and for each layer j session i traversing l and $\Delta(i,l) \geq p$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,j}, \infty)$,

$$\sum_{k=1}^{K} P(i,j,l,s_k,t_k) \geq \min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j-1} b_{i,m}\right), b_{i,j}\right)\sum_{k=1}^{K}(t_k - s_k)$$
$$-\varsigma_2(p,i,j) - K\varepsilon_2(p,i,j) \quad (103)$$

If such a link l has a session i sibling l', s.t. $\Delta(i,l') < p$ and if $\sum_{m=1}^{\alpha(i,p)} b_{i,m} < \hat{\lambda}_p$, then $B(i, \alpha(i,p)+1, l', t) > W + \Upsilon\Gamma$, for all $t \geq T_{p,\alpha(i,p)+1}$, as well.

(104)

For each link l and each session i traversing l s.t. $\Delta(i,l) \geq p$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,\alpha(i,p)+1}, \infty)$,

$$\sum_{k=1}^{K} P(i,l,s_k,t_k) \geq \hat{\lambda}_p \sum_{k=1}^{K}(t_k - s_k) - \varrho(p) - K\tau(p) \quad (105)$$

For each link l and for each layer j of session i traversing l and $\Delta(i,l) = p$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,j}, \infty)$,

$$\sum_{k=1}^{K} P(i,j,l,s_k,t_k) \leq \min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j-1} b_{i,m}\right)^+, b_{i,j}\right)\sum_{k=1}^{K}(t_k - s_k)$$
$$+\varsigma_3(p,i,j) + K\varepsilon_3(p,i,j) \quad (106)$$

For each link l and each session i traversing l s.t. $\Delta(i,l) = p$, for any positive integer K, and for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,\alpha(i,p)+1}, \infty)$,

$$\sum_{k=1}^{K} P(i,l,s_k,t_k) \quad \leq \quad \hat{\lambda}_p \sum_{k=1}^{K} (t_k - s_k) + \kappa(p) + K\gamma(p) \tag{107}$$

We prove these by induction on p. We first prove (102), (103), (104), (105), (106) and (107) for $p = 1$. Let $W \geq W_1$ and $G \geq G_1$.

Note that (104) holds by vacuity for $p = 1$, since $\Delta(i,l) \geq 1$, for all links l on session i path. We will prove the others as follows. First we will prove (102), (103) and (106) for $p = 1$, for layers $1, \ldots, \alpha(i,1)$, for each session i. Next we will prove(102) and (103) for $p = 1$, and layer $\alpha(i,1)+1$, for each session i. Using these, we will prove (105) for $p = 1$ and each session i. Next, we will prove (106) for $p = 1$, and layer $\alpha(i,1)+1$, for each session i, using (105) for $p = 1$. Next, we will prove (102), (103) and (106) for $p = 1$, for layers $j > \alpha(i,1)$, for each session i. Finally, we will prove (107) for $p = 1$, using (106) for $p = 1$.

We will prove by induction on the layer number, that (102), (103), (106) hold for layers $j \leq \alpha(i,1)$ for each session i and $p = 1$. First consider a link l and a session i traversing l. Clearly $\Delta(i,l) \geq 1$. If $\alpha(i,1) = 0$, then the statement that (102), (103), (106) hold for layers $j \leq \alpha(i,1)$ hold by vacuity. So, assume that $\alpha(i,1) \geq 1$ and first consider layer 1 of session i. Under the above assumption, $b_{i,1} \leq \hat{\lambda}_1$. Let $T_{1,0} = 0$. Let l originate from the source of session i. From assumption (97), for all nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [0, \infty)$,

$$\sum_{k=1}^{K} C(i,1,l,s_k,t_k) \quad \geq \quad b_{i,1} \sum_{k=1}^{K} (t_k - s_k) - K\xi_{i,1} \tag{108}$$

Hence (102) holds in this case.

Now let l not originate from the source of session i. Let Z be the set of sessions that traverse l. Consider any nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [0, \infty)$.

$$\begin{aligned}
C(i,1,l,s_k,t_k) \quad &= \quad C(i,l,s_k,t_k) \text{ (number of chances offered to layer 1} \\
&\qquad\qquad \text{is equal to the number of chances} \\
&\qquad\qquad \text{offered to the session)} \\
&\geq \quad \max_{w \in Z} C(w,l,s_k,t_k) - 1 \text{ (property of round robin} \\
&\qquad\qquad\qquad \text{scheduling among the sessions} \\
&\qquad\qquad\qquad \text{traversing the same link)} \\
&\geq \quad \frac{1}{|Z|} \sum_{w \in Z} C(w,l,s_k,t_k) - 1 \\
&\geq \quad \frac{C_l(t_k - s_k) - 1}{|Z|} - 1 \\
\sum_{k=1}^{K} C(i,1,l,s_k,t_k) \quad &\geq \quad \frac{C_l}{|Z|} \sum_{k=1}^{K} (t_k - s_k) - K(1 + |Z|^{-1})
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{\sum_{w\in Z}\lambda_{wl}}{|Z|}\sum_{k=1}^{K}(t_k-s_k)-2K \text{ (since } |Z|\geq 1)\\
&\geq \min_{w\in Z}\lambda_{wl}\sum_{k=1}^{K}(t_k-s_k)-2K\\
&\geq \hat{\lambda}_1\sum_{k=1}^{K}(t_k-s_k)-2K
\end{aligned} \tag{109}
$$

The last inequality follows since $\hat{\lambda}_1$ is the least possible maxmin fair session link rate of a session. From (108) and (109), (102) holds for $p=1$, layer 1, with $\varsigma_1(1,i,1)=0$, $\varepsilon_1(1,i,1)=\max(\xi_{i,1},2)$.

We will use Lemma 2 to prove (103) for any link l and layer 1 of any session i traversing l, with $\Delta(i,l)\geq 1$, and $\alpha(i,1)\geq 1$. Here $r=b_{i,1}$, $e_4=\varsigma_1(1,i,1), f_4=\varepsilon_1(1,i,1)$, $T=T_{1,0}$. Clearly r,e_4,f_4 are nonnegative. Consider any link l and any session i traversing l. Clearly, $\Delta(i,l')\geq 1$ for all links l' on session i path. Now (60) holds from (102) for $p=1$, layer 1 for any session i path. Since $\varepsilon_1(x,y)$ is a nondecreasing function, (62) holds from (95) and (63) holds from (96), since $\varsigma_1(F,\tau),\varepsilon_1(F,\tau)\ \varepsilon_1(F,\tau),\varepsilon_3(F,\tau)$ are nonnegative and $H>1$. Since all session i paths satisfy (60), δ,e_5,f_5 are irrelevant. Thus from Lemma 2, (103) holds for all links for layer 1 of session i on session i path with $T_{1,1}=\max_i T_i(H_i,T_{1,0})=T_{1,0}=0$, $\varsigma_2(1,i,1)=\left(H_i-1+\nu_i\frac{2\nu^{H_i-1}-\nu^{H_i-2}-1}{\nu_i-1}\right)\varsigma_1(1,i,1)$, and $\varepsilon_2(1,i,1)=H_i\varepsilon_1(1,i,1)+2H_i-1+(3H_i-4)\Upsilon\Gamma$.

Now we prove (106) for $p=1$, layer 1, $\alpha(i,1)\geq 1$. Consider a link l and a session $i\in n(l)$, s.t. $\Delta(i,l)=1$. Consider any sequence of nondecreasing times $s_1,t_1,\ldots,s_K,t_K\in[T_{1,1},\infty)$.

$$
\begin{aligned}
\sum_{k=1}^{K}P(i,1,l,s_k,t_k) &= P(i,1,l,s_1,t_K)-\sum_{k=1}^{K-1}P(i,1,l,t_k,s_{k+1})\\
&\leq P(i,1,l,s_1,t_K)-\min\left(\hat{\lambda}_1,b_{i,1}\right)\sum_{k=1}^{K-1}(s_{k+1}-t_k)\\
&\quad+\varsigma_2(1,i,1)+(K-1)\varepsilon_2(1,i,1)
\end{aligned} \tag{110}
$$

The last step follows from (103) for $p=1$, layer 1. Consider link l^* originating from the source of session i.

$$
\begin{aligned}
P(i,1,l^*,s_1,t_K) &\leq C(i,1,l^*,s_1,t_K)\\
&\leq \min\left(\hat{\lambda}_1,b_{i,1}\right)(t_K-s_1)+\xi_{i,1}
\end{aligned} \tag{111}
$$

The last inequality follows from (97) and the fact that $b_{i,1}\leq\hat{\lambda}_1$,

We will apply Lemma 3 now, with $j=1$, $l_1=l^*$, $e_1=\xi_{i,1}$, $e_2=\varsigma_2(1,i,1)+\varepsilon_2(1,i,1)$, $T_0=0$, and $r=b_{i,1}$. Condition (87) holds from (111) and condition (88) holds from (103) for $p=1$, layer 1 since $\Delta(i,l')\geq 1$ for all links l' on the path of session i, and by assumption $\alpha(i,1)\geq 1$. From Lemma 3, and since $|Q|\leq H-1$,

$$
\begin{aligned}
P(i,1,l,s_1,t_K) &\leq \min\left(\hat{\lambda}_1,b_{i,1}\right)(t_K-s_1)+2^{H_i-1}\xi_{i,1}+(2^{H_i-1}-1)(\varsigma_2(1,i,1)\\
&\quad+\varepsilon_2(1,i,1))+(H_i-1)\left(\max_{l:i\in n(l)}B(i,1,l,0)+2\Upsilon\Gamma\right)
\end{aligned}
$$

$$\leq \min\left(\hat{\lambda}_1, b_{i,1}\right)(t_K - s_1) + 2^{H_i}\left(\varsigma_2(1,i,1) + \varepsilon_2(1,i,1)\right)$$
$$+(H_i - 1)\left(\max_{l:i\in n(l)} B(i,1,l,0) + 2\Upsilon\Gamma\right) \quad (112)$$

The last inequality follows since $\varepsilon_2(1,i,1) \geq \xi_{i,1}$. From (110) and (112)

$$\begin{aligned}
\sum_{k=1}^{K} P(i,1,l,s_k,t_k) &\leq \min(\hat{\lambda}_1, b_{i,1})\sum_{k=1}^{K}(t_k - s_k) + \varsigma_3(1,i,1) + K\varepsilon_3(1,i,1) \\
\text{with } \varsigma_3(1,i,1) &= \varsigma_2(1,i,1) + (2^{H_i} - 1)\left(\varsigma_2(1,i,1) + \varepsilon_2(1,i,1)\right) \\
&\quad +(H_i - 1)\max_{l:i\in n(l)}\left(B(i,1,l,0) + 2\Upsilon\Gamma\right) \qquad (113) \\
\text{and } \varepsilon_3(1,i,1) &= \varepsilon_2(1,i,1) \qquad (114)
\end{aligned}$$

Thus (102), (103), and (106) hold for $p = 1$, layer 1 of all sessions i if $\alpha(i,1) \geq 1$. Let (102), (103) and (106) hold for $p = 1$, layers $1, \ldots, j \leq \alpha(i,1)$ of all sessions i. We will show that (102), (103) and (106) hold for $p = 1$, layer $j+1$, of all sessions i, if $j + 1 \leq \alpha(i,1)$.

Again, consider a link l and a session i traversing l. Clearly $\Delta(i,l) \geq 1$. Let $\alpha(i,1) \geq j+1$. Consider layer $j+1$. If l originates from the source of session i, from assumption (97), for all nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [T_{1,j}, \infty)$,

$$\sum_{k=1}^{K} C(i,j+1,l,s_k,t_k) \geq b_{i,j+1}\sum_{k=1}^{K}(t_k - s_k) - K\xi_{i,j+1} \quad (115)$$

Hence (102) holds in this case.

Now let l not originate from the source of session i. Observe that in any interval $[s,t)$,

$$C(i,j+1,l,s,t) = C(i,l,s,t) - \sum_{m=1}^{j} P(i,m,l,s,t) \quad (116)$$

Consider nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [T_{1,j}, \infty)$, Proceeding as in the proof of (109),

$$\sum_{k=1}^{K} C(i,l,s_k,t_k) \geq \hat{\lambda}_1\sum_{k=1}^{K}(t_k - s_k) - 2K \quad (117)$$

$$\begin{aligned}
\sum_{k=1}^{K} C(i,j+1,l,s_k,t_k) &= \sum_{k=1}^{K} C(i,l,s_k,t_k) - \sum_{m=1}^{j}\sum_{k=1}^{K} P(i,m,l,s_k,t_k) \\
&\qquad \text{(from (116))} \\
&\geq (\hat{\lambda}_1 - \sum_{m=1}^{j} b_{i,m})\sum_{k=1}^{K}(t_k - s_k) - \sum_{m=1}^{j}\varsigma_3(1,i,m) \\
&\quad -K\left(2 + \sum_{m=1}^{j}\varepsilon_3(1,i,m)\right) \text{ (from (117) and (106)} \\
&\qquad \text{for } p = 1 \text{ and } m \leq j) \qquad (118)
\end{aligned}$$

Thus (102) holds in this case as well.

The proof of (103) follows from (102) and Lemma 2 in this case, similar to the proof of (103), for layer 1. Here, (103) holds with $T_{1,j+1} = 0$, $\varsigma_2(1,i,j+1) = \left(H_i - 1 + \nu_i \frac{2\nu_i^{H_i-1} - \nu_i^{H_i-2} - 1}{\nu_i - 1}\right)\varsigma_1(1,i,j+1)$, and $\varepsilon_2(1,i,j+1) = H_i\varepsilon_1(1,i,j+1) + 2H_i - 1 + (3H_i - 4)\Upsilon\Gamma$.

The proof of (106) is similar to the previous one, in this case. Using Lemma 3 as before, and the fact that $\hat{\lambda}_1 - \sum_{i=1}^{j} b_{i,1} \geq b_{i,j+1} \geq 0$, (since $\alpha(i,1) \geq j+1$).

Now we prove (102) and (103) for $p = 1$, and layer $\alpha(i,1)+1$ of each session i. Note that $\sum_{m=1}^{\alpha(i,1)} b_{i,m} \leq \hat{\lambda}_1 < \sum_{m=1}^{\alpha(i,1)+1} b_{i,m}$. As before, consider a link l and a session i traversing l. Again, $\Delta(i,l) \geq 1$. If l originates from the source of session i, from assumption (97), for all nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [T_{1,\alpha(i,1)}, \infty)$,

$$\sum_{k=1}^{K} C(i,\alpha(i,1)+1,l,s_k,t_k) \;\geq\; b_{i,\alpha(i,1)+1}\sum_{k=1}^{K}(t_k - s_k) - K\xi_{i,\alpha(i,1)+1} \tag{119}$$

Hence (102) holds in this case.

Now let l not originate from the source of session i. Consider nondecreasing times $s_1, t_1, \ldots, s_K, t_K \in [T_{1,\alpha(i,1)}, \infty)$, Proceeding as before, and using the fact that $\sum_{m=1}^{\alpha(i,1)} b_{i,m} \leq \hat{\lambda}_1$, and that (106) hold for all layers $1, \ldots, \alpha(i,1)$ (if $\alpha(i,1) = 0$, then (106) holds for layers, $1, \ldots, \alpha(i,1)$ by vacuity), we have

$$\begin{aligned}\sum_{k=1}^{K} C(i,j+1,l,s_k,t_k) \;\geq\; & \left(\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m}\right)\sum_{k=1}^{K}(t_k - s_k) - \sum_{m=1}^{\alpha(i,1)} \varsigma_3(1,i,m) \\ & - \left(2 + \sum_{m=1}^{\alpha(i,1)} \varepsilon_3(1,i,m)\right) K \end{aligned} \tag{120}$$

From (119) and (120), (102) holds for $p = 1$, session i and layer $\alpha(i,1)+1$.

The proof of (103) follows from (102) and Lemma 2 as before. Here, (103) holds with $T_{1,\alpha(i,1)+1} = 0$, $\varsigma_2(1,i,\alpha(i,1)+1) = \left(H_i - 1 + \nu_i \frac{2\nu_i^{H_i-1} - \nu_i^{H_i-2} - 1}{\nu_i - 1}\right)\varsigma_1(1,i,\alpha(i,1)+1)$,
and $\varepsilon_2(1,i,\alpha(i,1)+1) = H_i\varepsilon_1(1,i,\alpha(i,1)+1) + 2H_i - 1 + (3H_i - 4)\Upsilon\Gamma$.

Now, we prove (105) for $p = 1$. Consider a link l and session $i \in n(l)$. Clearly, $\Delta(i,l) \geq 1$. So, (103) holds for all layers $1, \ldots, \alpha(i,1)+1$. Hence, for any $s_1, t_1, \ldots, s_K, t_K \in [T_{1,\alpha(i,1)+1}, \infty)$, and any positive integer K,

$$\begin{aligned}\sum_{k=1}^{K} P(i,l,s_k,t_k) \;\geq\; & \sum_{m=1}^{\alpha(i,1)+1}\sum_{k=1}^{K} P(i,m,l,s_k,t_k) \\ \geq\; & \sum_{m=1}^{\alpha(i,1)} b_{i,m}\sum_{k=1}^{K}(t_k - s_k) + (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})\sum_{k=1}^{K}(t_k - s_k) \\ & - \sum_{m=1}^{\alpha(i,1)+1} \varsigma_2(1,i,m) - K\sum_{m=1}^{\alpha(i,1)+1} \varepsilon_2(1,i,m) \\ =\; & \hat{\lambda}_1\sum_{k=1}^{K}(t_k - s_k) - \sum_{m=1}^{\alpha(i,1)+1} \varsigma_2(1,i,m) - K\sum_{m=1}^{\alpha(i,1)+1} \varepsilon_2(1,i,m)\end{aligned}$$

Thus (105) holds for $p = 1$.

Now we will prove (106) for $p=1$, and layer $\alpha(i,1)+1$ of session i, Consider a link l and a session i traversing l, with $\Delta(i,l)=1$. Consider layer $\alpha(i,1)+1$. Consider any $s_1, t_1, \ldots, s_K, t_K \in [T_{1,\alpha(i,1)+1}, \infty)$. Initially, assume that link l is the bottleneck link[‡] of virtual session s of session i. (So l is not the link originating from the source of session i.) Since l is the bottleneck link of s, under maxmin fair allocation, its capacity is fully consumed by sessions w with $\Delta(w,l) \le \vartheta(s) \le \Delta(i,l) = 1$. Let $X = \{w : w \in n(l), \Delta(w,l) = 1, w \ne i\}$.

$$C_l = \sum_{w \in X} \lambda_{wl} + \lambda_{il} \tag{121}$$

$$\begin{aligned}
& \text{Also } P(i, \alpha(i,1)+1, l, s_1, t_K) \\
\le\ & \left(C_l(t_K - s_1) + 1 - \sum_{w \in X} P(w, l, s_1, t_K) - \sum_{m=1}^{\alpha(i,1)} P(i, m, l, s_1, t_K) \right)^+ \\
& \text{(since the same turn can not be used} \\
& \text{to transmit packets of two different sessions} \\
& \text{or two different layers of the same session} \\
& \text{in the same link and at most } C_l(t_K - s_1) + 1 \\
& \text{turns are offered at the origin of } l \text{ during } [s_1, t_K)) \\
\le\ & \left(C_l - \sum_{w \in X} \hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m} \right)^+ (t_K - s_1) \\
& + \left(\sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) \right) + (\varrho(1) + \tau(1)) |X| + 1 \\
& \text{(from (103) for layers } 1, \ldots, \alpha(i,1) \text{ and (105) for } p = 1) \\
\le\ & \left(C_l - \sum_{w \in X} \lambda_{wl} - \sum_{m=1}^{\alpha(i,1)} b_{i,m} \right)^+ (t_K - s_1) \\
& + \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(i,m) + \varepsilon_2(i,m)) + (\varrho(1) + \tau(1)) (|n(l)| - 1) + 1 \\
& \text{(since } \lambda_{wl} = \hat{\lambda}_1 \forall w \in X, \text{ and } X \subset n(l)) \\
=\ & (\lambda_{il} - \sum_{m=1}^{\alpha(i,1)} b_{i,m})^+ (t_K - s_1) + \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) \\
& + (\varrho(1) + \tau(1)) (|n(l)| - 1) + 1 \qquad \text{(from (121))} \\
=\ & (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) \\
& + (\varrho(1) + \tau(1)) (|n(l)| - 1) + 1 \qquad (122) \\
& \text{(since } i \in n(l) \text{ and } \Delta(i,l) = 1, \hat{\lambda}_1 = \lambda_{il} \text{and } \sum_{m=1}^{\alpha(i,1)} b_{i,m} \le \hat{\lambda}_1)
\end{aligned}$$

Now let l not be the bottleneck link of any virtual session of session i traversing l. Thus all the virtual sessions

[‡]From [14] every virtual session has a bottleneck lenk under maxmin fair allocation. The virtual session has the highest rate amongst all virtual sessions of its session traversing the bottleneck link. Also, the capacity of the bottleneck link is fully utilized.

in $m(i,l)$ are bottlenecked somewhere else. Let all of them have bottleneck links downstream. Consider the sequence of links on session i path constructed as follows. The first link in the sequence originates from the sink of l and has the minimum queue length at t_K amongst all session i links originating from the sink of l. Any other link in the sequence originates from the sink of the previous link and has the minimum queue length at t_K amongst all session i links originating from the sink of the previous link. Since all virtual sessions in $m(i,l)$ have their bottleneck links downstream of l, bottleneck link of some virtual session in $m(i,l)$ will be in the sequence. That virtual session has index 1 because it traverses l and $\Delta(i,l) = 1$. Let the bottleneck link be l^*. By (122),

$$\begin{aligned} P(i,\alpha(i,1)+1,l^*,s_1,t_K) \;\; &\leq \;\; (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(i,m) + \varepsilon_2(i,m)) \\ &\quad + (\varrho(1) + \tau(1))\,(E_i - 1) + 1 \end{aligned} \tag{123}$$

for any s_1, t_K, s.t. $t_K \geq s_1 \geq 0$.

$$\text{Note that } B(i,\alpha(i,1)+1,l_q,t_K) \leq W + \Upsilon\Gamma, \text{ for all links } l_q \text{ in the sequence.} \tag{124}$$

This is because $B(i,\alpha(i,1)+1,l_q,t_K) = \min B(i,\alpha(i,1)+1,l',t_K)$, with the minimum being taken over all session i links l' originating from the source of l_q. We will use Lemma 4 to show that $B(i,\alpha(i,1)+1,l_q,u) < G$, $\forall\ u \in [s_1,t_K)$, for all links l_q in this sequence. First consider the links l^* and the link l_q just preceding l^* in the sequence. Let $B(i,\alpha(i,1)+1,l^*,u) = G$, for some $u \in [s_1,t_K)$. Let $l_1 = l_q$, $l_2 = l^*$, $j = \alpha(i,1)+1$, $r = \hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m}$, $e_7 = \varsigma_2(1,i,\alpha(i,1)+1) + \varepsilon_2(1,i,\alpha(i,1)+1)$, $e_8 = \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) + (\varrho(1) + \tau(1))\,(E_i - 1) + 1$. Session i has index 1 in all links in the sequence, because all of these links are downstream of l and $\Delta(i,l) = 1$. From (103) for $p = 1$, layer $\alpha(i,1)+1$, (91) holds. From (123), (92) holds. Clearly, (93) holds from the assumption that $B(i,\alpha(i,1)+1,l^*,u) = G$. Also (94) holds from (96). From Lemma 4, $B(i,\alpha(i,1)+1,l^*,t) > W + \Upsilon\Gamma$, for all $t \geq u$. This contradicts (124) since $t_K > u$. Thus $B(i,\alpha(i,1)+1,l^*,u) < G\ \forall\ u \in [s_1,t_K)$. Thus there is no packet loss in the session i buffer for l^* in the interval $[s_1,t_K)$. It follows that,

$$\begin{aligned} & P(i,\alpha(i,1)+1,l_q,s_1,t_K) \\ \leq\ & P(i,\alpha(i,1)+1,l_q,s_1-\Upsilon_{l_q},t_K-\Upsilon_{l_q}) + \Upsilon\Gamma \\ =\ & P(i,\alpha(i,1)+1,l^*,s_1,t_K) + B(i,\alpha(i,1)+1,l^*,t_K) \\ & -B(i,\alpha(i,1)+1,l^*,s_1) + \Upsilon\Gamma \\ \leq\ & P(i,\alpha(i,1)+1,l^*,s_1,t_K) + B(i,\alpha(i,1)+1,l^*,t_K) + \Upsilon\Gamma \\ \leq\ & P(i,\alpha(i,1)+1,l^*,s_1,t_K) + W + 2\Upsilon\Gamma \text{ (from (124))} \\ \leq\ & (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + W + 2\Upsilon\Gamma + (\varrho(1)+\tau(1))\,E_i + 1 \end{aligned} \tag{125}$$

The last inequality follows from (123) using the fact that $\sum_{m=1}^{\alpha(i,1)} \varsigma_1(1,i,m) \leq \rho(1)$, and $\sum_{m=1}^{\alpha(i,1)} \varepsilon_2(1,i,m) \leq \tau(1)$. Now using (125) in place of (123), a similar upper bound as (125) can be proved for the link preceding l_q in this sequence. Proceeding in this manner,

$$\begin{aligned} P(i,\alpha(i,1)+1,l,s_1,t_K) \ \leq\ & (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + (W + 2\Upsilon\Gamma)(H_i - 1) + \\ & (\varrho(1) + \tau(1))\,(E_i - 1 + H_i - 1) + 1 \end{aligned} \tag{126}$$

Now let at least one virtual session in $m(i,l)$ have a bottleneck link upstream. Let the link be l^* and the virtual session s. As discussed before, $\vartheta(s) = 1$. Since l^* is the bottleneck link of virtual session s, index of any other session i virtual session traversing l^* is 1. Thus $\Delta(i,l^*) = 1$.

We will use Lemma 3, with $j = \alpha(i,1)+1$, $l_1 = l^*$,
$e_1 = \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) + (\varrho(1) + \tau(1))\,(E_i - 1) + 1$,
$e_2 = \varsigma_2\,(1,i,\alpha(i,1)+1) + \varepsilon_2\,(1,i,\alpha(i,1)+1)$, $T_0 = 0$, and $r = \hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m}$. Condition (87) follows from (123), condition (88) follows from (103) since $\Delta(i,l') \geq 1$ for all links l' on the path of session i. Hence,

$$\begin{aligned} & P(i,\alpha(i,1)+1,l,s_1,t_K) \\ \leq\ & (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + 2^{H_i-1}\Big((E_i - 1)\,(\varrho(1) + \tau(1)) \\ & + \sum_{m=1}^{\alpha(i,1)} (\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)) + 1\Big) \\ & + (2^{H_i-1} - 1)\,(\varsigma_2\,(1,i,\alpha(i,1)+1) + \varepsilon_2\,(1,i,\alpha(i,1)+1)) \\ & + (H_i - 1) \max_{l: i \in n(l)} (B(i,\alpha(i,1)+1,l,0) + 2\Upsilon\Gamma) \\ \leq\ & (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K - s_1) + 2^{H_i-1}\,(E_i\,(\varrho(1) + \tau(1)) + 1) \\ & + (H_i - 1) \max_{l: i \in n(l)} (B(i,\alpha(i,1)+1,l,0) + 2\Upsilon\Gamma) \end{aligned} \tag{127}$$

The last inequality follows since $\sum_{m=1}^{\alpha(i,1)+1} \varsigma_2(1,i,m) \leq \varrho(1)$ and
$\sum_{m=1}^{\alpha(i,1)+1} \varepsilon_2(1,i,m) \leq \tau(1)$. Similar to (110),

$$\begin{aligned} & \sum_{k=1}^{K} P(i,\alpha(i,1)+1,l,s_k,t_k) \\ =\ & P(i,\alpha(i,1)+1,l,s_1,t_K) - \sum_{k=1}^{K-1} P(i,\alpha(i,1)+1,l,t_k,s_{k+1}) \\ \leq\ & P(i,\alpha(i,1)+1,l,s_1,t_K) - (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m}) \sum_{k=1}^{K-1} (s_{k+1} - t_k) \\ & + \varsigma_2\,(1,i,\alpha(i,1)+1) + (K-1)\varepsilon_2(1,i,\alpha(i,1)+1) \end{aligned} \tag{128}$$

The last step follows from (103) for $p = 1$, layer $\alpha(i,1)+1$.

From (126), (127) and (128)

$$\begin{aligned}
\sum_{k=1}^{K} P(i,\alpha(i,1)+1,l,s_k,t_k) &\leq (\hat{\lambda}_1 - \sum_{m=1}^{\alpha(i,1)} b_{i,m}) \sum_{k=1}^{K}(t_k - s_k) + \varsigma_3(1,i,\alpha(i,1)+1) \\
&\quad + K\varepsilon_3(1,i,\alpha(i,1)+1) \\
\text{with } \varsigma_3(1,i,\alpha(i,1)+1) &= \varsigma_2(1,i,\alpha(i,1)+1) + (W+2\Upsilon\Gamma)(H_i-1) \\
&\quad + 2^{H_i-1}\left(E_i\left(\varrho(1)+\tau(1)\right)+1\right) \\
&\quad + (H_i-1)\max_{l:i\in n(l)} B(i,\alpha(i,1)+1,l,0) \qquad (129) \\
\text{and } \varepsilon_3(1,i,\alpha(i,1)+1) &= \varepsilon_2(1,i,\alpha(i,1)+1) \qquad (130)
\end{aligned}$$

We have used the fact that $E_i + H_i - 2 \leq 2^{H_i-1}E_i$, since $E_i \geq 1$, in deriving (129).

Now, we prove (102), (103) and (106) for layers $j > \alpha(i,1)+1$. Observe that $\hat{\lambda}_1 < \sum_{m=1}^{\alpha(i,1)+1} b_{i,m} \leq \sum_{m=1}^{j-1} b_{i,m}$. Thus $\hat{\lambda}_1 - \sum_{m=1}^{j-1} b_{i,m} < 0$. Hence, (102) and (103) follow trivially, since $\varsigma_1(1,i,j)$, $\varepsilon_1(1,i,j)$, $\varsigma_2(1,i,j)$, $\varepsilon_2(1,i,j)$ are nonnegative by definition. Now, we prove (106) for layers $j > \alpha(i,1)+1$. Consider a link l and a session i traversing l, with $\Delta(i,l) = 1$. Consider layer $j > \alpha(i,1)+1$ and nondecreasing times, $s_1, t_1, \ldots, s_k, t_k \in [T_{1,\alpha(i,1)+1}, \infty)$. Initially, assume that link l is the bottleneck link of virtual session s of session i. As before,

$$\begin{aligned}
P(i,j,l,s_1,t_K) &\leq \Bigg(C_l(t_K - s_1) + 1 - \sum_{w\in X} P(w,l,s_1,t_K) \\
&\qquad - \sum_{m=1}^{\alpha(i,1)+1} P(i,m,l,s_1,t_K) \Bigg)^{+}
\end{aligned}$$

Using (103) for layers $1,\ldots,\alpha(i,1)+1$, (105) for $p=1$, and proceeding as in the proof of (122) we have,

$$\begin{aligned}
P(i,j,l,s_1,t_K) &\leq \sum_{m=1}^{\alpha(i,1)+1} \left(\varsigma_2(1,i,m) + \varepsilon_2(1,i,m)\right) + \left(\varrho(1)+\tau(1)\right)(E_i - 1) + 1 \\
&\leq \left(\varrho(1)+\tau(1)\right)E + 1
\end{aligned}$$

The last inequality follows since $\sum_{m=1}^{\alpha(i,1)+1} \varsigma_2(1,i,m) \leq \varrho(1)$ and $\sum_{m=1}^{\alpha(i,1)+1} \varepsilon_2(1,i,m) \leq \tau(1)$.

If l is not the bottleneck link, and all the virtual sessions of session i traversing link l are bottlenecked downstream, then constructing the sequence of links l_q as before, and subsequently proceeding similarly,

$$P(i,j,l,s_1,t_K) \leq (W+2\Upsilon\Gamma)(H_i-1) + \left(\varrho(1)+\tau(1)\right)E_i + 1$$

Proceeding similarly in the other remaining case, i.e., at least one virtual session is bottlenecked upstream, using Lemma 3, (103) for $p=1$, layer j, session i and the facts that $\varsigma_2(1,i,j) = \varepsilon_2(1,i,j) = 0$,

$$P(i,j,l,s_1,t_K) \leq 2^{H_i-1}\left(E_i\left(\varrho(1)+\tau(1)\right)+1\right) + (H_i-1)\max_{l:i\in n(l)} B(i,j,l,0)$$

It follows that for all these cases,

$$\begin{aligned} P(i,j,l,s_1,t_K) &\leq 2^{H_i-1}\left(E_i\left(\varrho(1)+\tau(1)\right)+1\right) \\ &\quad +(H_i-1)\left(\max_{l:i\in n(l)} B(i,j,l,0)+W+2\Upsilon\Gamma\right) \end{aligned}$$

$$\text{Clearly, } \sum_{k=1}^{K} P(i,\alpha(i,1)+1,l,s_k,t_k) \leq P(i,\alpha(i,1)+1,l,s_1,t_K)$$

Thus, for $j > \alpha(i,1)+1$,

$$\begin{aligned} \sum_{k=1}^{K} P(i,j,l,s_k,t_k) &\leq \varsigma_3(1,i,j)+K\varepsilon_3(1,i,j) \\ \text{with } \varsigma_3(1,i,j) &= 2^{H_i-1}\left(E_i\left(\varrho(1)+\tau(1)\right)+1\right) \\ &\quad +(H_i-1)\left(\max_{l:i\in n(l)} B(i,j,l,0)+W+2\Upsilon\Gamma\right) && (131) \\ \text{and } \varepsilon_3(1,i,j) &= 0 && (132) \end{aligned}$$

Thus (106) follows for $j > \alpha(i,1)+1$.

Thus (106) holds for $p=1$ for all layers. Now we will prove (107) for $p=1$. Consider a link l and a session i traversing l, with $\Delta(i,l)=1$. Consider any $s_1,t_1,s_2,t_2,\ldots,s_K,t_K \in [T_{1,\alpha(i,1)+1},\infty)$

$$\begin{aligned} \sum_{k=1}^{K} P(i,l,s_k,t_k) &= \sum_j \sum_{k=1}^{K} P(i,j,l,s_k,t_k) \\ &= \sum_{j\leq\alpha(i,1)} \sum_{k=1}^{K} P(i,j,l,s_k,t_k) + \sum_{k=1}^{K} P(i,\alpha(i,1)+1,l,s_k,t_k) \\ &\quad + \sum_{j>\alpha(i,1)+1} \sum_{k=1}^{K} P(i,j,l,s_k,t_k) \\ &\leq \sum_{j\leq\alpha(i,1)} b_{i,j} \sum_{k=1}^{K}(t_k-s_k) + (\hat{\lambda}_1 - \sum_{j\leq\alpha(i,1)} b_{i,j}) \sum_{k=1}^{K}(t_k-s_k) \\ &\quad + \sum_j \left(\varsigma_3(1,i,j)+K\varepsilon_3(1,i,j)\right) \\ &\quad \text{(from (106) for } p=1) \\ &= \hat{\lambda}_1 \sum_{k=1}^{K}(t_k-s_k) + \sum_j \varsigma_3(1,i,j) + K\sum_j \varepsilon_3(1,i,j) \end{aligned}$$

Thus (107) follows.

Now we assume that (102), (103), (104), (105), (106) and (107) hold for $q=1,\ldots,p$. and show that these hold for $p+1$. Let $W \geq W_p$ and $G \geq G_p$. Let $T_{p+1,0} = T_{p,\beta(p)+1}$.

We will prove by induction on the layer number, that (102), (103), (105), (106) hold for layers $j \leq \alpha(i,p+1)$ for each session i and $p+1$. Note that it is not necessary to prove (104) for $j \leq \alpha(i,p+1)$. Consider a link l

and a session i traversing l, with $\Delta(i,l) \geq p+1$. Consider nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K \in [T_{p+1,0}, \infty)$. We will mention the differences with the case for $p = 1$ (base case). As before, it is worthwhile to look at the case $\alpha(i, p+1) \geq 1$. We first consider layer 1. Let l originate from the source of session i. From assumption (97),

$$\sum_{k=1}^{K} C(i,l,s_k,t_k) \quad \geq \quad b_{i,1} \sum_{k=1}^{K} (t_k - s_k) - K\xi_{i,1} \tag{133}$$

Hence (102) holds in this case.

Now let l not originate from the source of session i. Let $Y = \{w \in n(l), \Delta(w,l) \leq p\}$ and $Z = \{w \in n(l), \Delta(w,l) \geq p+1\}$. Now, Z includes i.

$$\begin{aligned}
C(i,1,l,s_k,t_k) \quad &= \quad C(i,l,s_k,t_k)\ \forall k \text{ (number of chances offered to layer 1}\\
&\qquad\qquad \text{is equal to the number of chances}\\
&\qquad\qquad \text{offered to the session)}\\
&\geq \quad \max_{w \in Z} C(w,l,s_k,t_k) - 1\\
&\qquad \text{(property of round robin scheduling)}\\
&\geq \quad \frac{1}{|Z|} \sum_{w \in Z} C(w,l,s_k,t_k) - 1\\
&\geq \quad \frac{C_l(t_k - s_k) - 1 - \sum_{w \in Y} P(w,l,s_k,t_k)}{|Z|} - 1\\
&\qquad \text{(chances not used by sessions in } Y\\
&\qquad \text{are offered to sessions in } Z)
\end{aligned}$$

$$\begin{aligned}
&\sum_{k=1}^{K} C(i,1,l,s_k,t_k)\\
\geq\ &\frac{C_l \sum_{k=1}^{K}(t_k - s_k) - \sum_{w \in Y} \sum_{k=1}^{K} P(w,l,s_k,t_k)}{|Z|} - K(1 + |Z|^{-1})\\
\geq\ &\frac{(C_l - \sum_{w \in Y} \lambda_{wl}) \sum_{k=1}^{K}(t_k - s_k)}{|Z|} - \frac{|Y|}{|Z|}\kappa(p) - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1 + |Z|^{-1}\right)\right)\\
&\text{(from (107) by induction hypothesis and using the facts that}\\
&\Delta(w,l) \leq p \text{ for } w \in Y, \lambda_{wl} = \hat{\lambda}_{\Delta(w,l)}, \kappa(r) \geq \kappa(r-1),\ \ \gamma(r) \geq \gamma(r-1) \forall r)\\
\geq\ &\frac{\sum_{w \in Z} \lambda_{wl} \sum_{k=1}^{K}(t_k - s_k)}{|Z|} - \frac{|Y|}{|Z|}\kappa(p) - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1 + |Z|^{-1}\right)\right)\\
&\text{(since } C_l \geq \sum_{w \in Z} \lambda_{wl} + \sum_{w \in Y} \lambda_{wl})\\
\geq\ &\min_{w \in Z} \lambda_{wl} \sum_{k=1}^{K}(t_k - s_k) - \frac{|Y|}{|Z|}\kappa(p) - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1 + |Z|^{-1}\right)\right)
\end{aligned}$$

$$\geq \hat{\lambda}_{p+1}\sum_{k=1}^{K}(t_k - s_k) - \frac{|Y|}{|Z|}\kappa(p) - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1+|Z|^{-1}\right)\right) \tag{134}$$
$$\text{(since } \Delta(w,l) \geq p+1, \text{ for } w \in Z)$$

From (133) and (134), (102) holds for $p+1$ and layer 1.

Now we prove (103) for $p+1$, layer 1. Consider a link l and a session i traversing l s.t. $\Delta(i,l) \geq p+1$. Observe that if $b_{i,1} \leq \hat{\lambda}_p$, then $\min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{0} b_{i,1}\right), b_{i,1}\right) = b_{i,1}$, and hence for any nondecreasing sequence $s_1, t_1, \ldots, s_K, t_K$ of times in $[T_{p,0}, \infty)$,

$$\sum_{k=1}^{K} P(i,1,l,s_k,t_k) \geq b_{i,1}\sum_{k=1}^{K}(t_k - s_k) - \varsigma_2(p,i,1) - K\varepsilon_2(p,i,1)$$

from the induction hypothesis that (103) holds for p. Now, (103) follows in this case since $\varsigma_2(p,i,1) \leq \varsigma_2(p+1,i,1)$ and $\varepsilon_2(p,i,1) \leq \varepsilon_2(p+1,i,1)$. So let $b_{i,1} > \hat{\lambda}_p \geq 0$. Since $\Delta(i,l) \geq p+1$, there is a session i virtual session s traversing l, with $\vartheta(s) \geq p+1$. We will apply Lemma 2 to receiver s of session i, layer 1 traffic, $T = T_{p+1,0}$, $r = b_{i,1} > 0$, $\delta = b_{i,1} - \hat{\lambda}_p > 0$, $e_4 = \varsigma_1(p+1,i,1)$, $f_4 = \varepsilon_1(p+1,i,1)$, $e_5 = \varsigma_3(p,i,1)$, $f_5 = \varepsilon_3(p,i,1)$. Now, $\Delta(i,l') \geq p+1$, for all links l' on the path to receiver s, and link l lies on this path. From (102) for $p+1$, layer 1, (60) holds for all links on the path. Consider any link l'' traversed by session i. If $\Delta(i,l'') \geq p+1$, repeating the same arguement it lies on a session i source destination path satisfying (60). Let $\Delta(i,l'') \leq p$. By induction hypothesis (106), and using the fact that $\hat{\lambda}_{\Delta(i,l'')} \leq \hat{\lambda}_p < b_{i,1}$, for $1,\ldots,p$, $T_{q,j} \leq T_{p,j}$ for all $q \leq p$, and layer j, for any positive integer K, any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{p+1,0}$

$$\begin{aligned}\sum_{k=1}^{K} P(i,1,l'',s_k,t_k) &\leq \hat{\lambda}_{\Delta(i,l'')}\sum_{k=1}^{K}(t_k - s_k) + \varsigma_3\left(\Delta(i,l''),i,1\right) + K\varepsilon_3\left(\Delta(i,l''),i,1\right) \\ &\leq \hat{\lambda}_p\sum_{k=1}^{K}(t_k - s_k) + \varsigma_3(p,i,1) + K\varepsilon_3(p,i,1)\end{aligned} \tag{135}$$

The last inequality follows because ς_3 and ε_3 are nondecreasing in p and also because from $\Delta(i,l'') \leq p$, $\hat{\lambda}_{\Delta(i,l'')} \leq \hat{\lambda}_p$. Thus from (135), (61) holds for l''. Also (62) holds from (95), since $\varepsilon_1(q)$ is nondecreasing, and (64) holds from (96), since $\varsigma_1(q), \varsigma_3(q), \varepsilon_1(q), \varepsilon_3(q)$ are nondecreasing. Thus Lemma 2 applies in this case. Thus, from Lemma 2, (103) holds for $p+1$ for session i on link l, with $T_{p+1,1} = \max_{s:\vartheta(s)\geq p+1} T_{\chi(s)}(H_{\chi(s)}, T_{p+1,0})$.

Now we prove (106) for $p+1$, layer 1. This proof is similar to the case for base case ($p=1$), with layer 1 and we do not repeat it for brevity. The only point to note is that Lemma 3 holds in this case as well, because condition (88) holds from (103) for $p+1$ layer 1 as $\Delta(i,l') \geq p+1$, for any link l' on the path of any session i, if there is a link l downstream of l' and $\Delta(i,l) = p+1$.

Thus (102), (103), (104) and (106) hold for $p+1$, layer 1 of all sessions i if $\alpha(i,p+1) \geq 1$. Let (102), (103) and (106) hold for $p+1$, layers $1,\ldots,j \leq \alpha(i,1)$ of all sessions i. We will show that (102), (103) and (106) hold for $p+1$, layers $j+1$, of all sessions i, if $j+1 \leq \alpha(i,p+1)$.

We first show that (102) holds in this case. Note that $\alpha(i,p+1) \geq j+1$. Consider a link l and a session i traversing l, and a nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K \in [T_{p+1,j}, \infty)$. Let $\Delta(i,l) \geq p+1$. Consider layer $j+1$. If l originates from the source of session i, from assumption (97),

$$\sum_{k=1}^{K} C(i, j+1, l, s_k, t_k) \quad \geq \quad b_{i,j+1} \sum_{k=1}^{K} (t_k - s_k) - K\xi_{i,j+1} \tag{136}$$

Hence (102) holds in this case.

Now let l not originate from the source of session i. The proof follows that for $p = 1$ and layer $j+1 \leq \alpha(i,1)$. We will use the relation that

$$\sum_{k=1}^{K} C(i, l, s_k, t_k) \geq \hat{\lambda}_{p+1} \sum_{k=1}^{K} (t_k - s_k) - \frac{|Y|}{|Z|}\kappa(p) - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1 + |Z|^{-1}\right)\right)$$

for any $T_{p+1,j} \leq s_1 \leq t_1 \leq \ldots \leq s_K \leq t_K$, This follows from (134). Now using (106) for $p+1$, layers $1, \ldots, j$

$$\begin{aligned}
\sum_{k=1}^{K} C(i, j+1, l, s_k, t_k) \quad \geq \quad & (\hat{\lambda}_{p+1} - \sum_{m=1}^{j} b_{i,m}) \sum_{k=1}^{K} (t_k - s_k) \\
& - \left(\sum_{m=1}^{j} \varsigma_3(p+1, i, m) + \frac{|Y|}{|Z|}\kappa(p) \right) \\
& - K\left(\frac{|Y|}{|Z|}\gamma(p) + \left(1 + |Z|^{-1}\right) \sum_{m=1}^{j} \varepsilon_3(1, i, m) \right)
\end{aligned}$$

Hence (102) holds in this case as well.

Now we prove (103) for $p+1$, session i, layer $j+1$, $\alpha(i,p+1) \geq j+1$. The proof will be similar to that for $p+1$, layer 1. Consider a link l and a session i traversing l s.t. $\Delta(i,l) \geq p+1$. Observe that if $\sum_{m=1}^{j+1} b_{i,m} \leq \hat{\lambda}_p$, then $\min\left(\left(\hat{\lambda}_p - \sum_{m=1}^{j} b_{i,m}\right), b_{i,j+1}\right) = b_{i,j+1}$, and hence (103) follows from the induction hypothesis for p, as in the proof for $p+1$, layer 1. So let $\sum_{m=1}^{j+1} b_{i,m} > \hat{\lambda}_p$. We will apply Lemma 2 to layer $j+1$ traffic with $T = T_{p+1,j}$, $r = b_{i,j+1}$, $\delta = \min\left(\sum_{m=1}^{j+1} b_{i,m} - \hat{\lambda}_p, b_{i,j+1}\right) > 0$, $e_4 = \varsigma_1(p+1, i, j+1)$, $f_4 = \varepsilon_1(p+1, i, j+1)$, $e_5 = \varsigma_3(p, i, 1)$, $f_5 = \varepsilon_3(p, i, 1)$. Consider any link l'' traversed by session i. Let $\Delta(i, l'') \leq p$. Since $\hat{\lambda}_{\Delta(i,l'')} \leq \hat{\lambda}_p < \sum_{m=1}^{j+1} b_{i,m}$, for $1, \ldots, p$, $\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{j} b_{i,m} < b_{i,j+1}$. From induction hypothesis (106) for $\Delta(i, l'')$, and the fact that $\varsigma_3\left(\Delta(i,l''), i, j+1\right) \leq \varsigma_3\left(p, i, j+1\right)$, and $\varepsilon_3\left(\Delta(i,l''), i, j+1\right) \leq \varepsilon_3\left(p, i, j+1\right)$, (since $\Delta(i,l'') \leq p$), for any positive integer K, any nondecreasing sequence of times $s_1, t_1, \ldots, s_K, t_K$, $s_1 \geq T_{p+1,j}$

$$\sum_{k=1}^{K} P(i, j, l'', s_k, t_k) \leq \left(\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{j} b_{i,m}\right)^{+} \sum_{k=1}^{K} (t_k - s_k) + \varsigma_3\left(p, i, 1\right) + K\varepsilon_3\left(p, i, 1\right) \tag{137}$$

Thus from (137), (61) holds for $l''(r - \delta = \left(\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{j} b_{i,m}\right)^{+})$. The arguements for the remaining conditions of Lemma 2 are similar to those for layer 1, $p+1$. Thus Lemma 2 applies in this case. Thus, from Lemma 2, (103) holds for $p+1$ for session i on link l, with

$T_{p+1,j+1} = \max_{s:\vartheta(s) \geq p+1} T_{\chi(s)}\left(H_{\chi(s)}, T_{p,j}\right)$.

The proof for (106) is similar to that for $p = 1$, layer 1. We would use the fact that $\hat{\lambda}_{p+1} - \sum_{i=1}^{j} b_{i,1} \geq b_{i,j+1} \geq 0$, since $j+1 \leq \alpha(i,p+1)$.

It follows that (102), (103), (104) and (106) hold for $p+1$, layer j of any session i if $j \leq \alpha(i,p+1)$.

Now we will prove (102), (103) and (104) for $p+1$, layer $\alpha(i,p+1)+1$, for any session i. The proof for (102) is similar to that for $p+1$, layer $j \leq \alpha(i,p+1)$. We use the fact that (106) holds for all layers $1,\dots,\alpha(i,p+1)$.

Now we prove (103) for $p+1$, session i, layer $(\alpha(i,p+1)+1$. Consider a link l and a session i traversing l s.t. $\Delta(i,l) \geq p+1$. If $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} = \hat{\lambda}_{p+1}$, then (103) holds trivially. So let $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < \hat{\lambda}_{p+1}$. Since $\Delta(i,l) \geq p+1$, there is a session i virtual session s traversing l, with $\vartheta(s) \geq p+1$. We will apply Lemma 2 to virtual session s of session i, layer $\alpha(i,p+1)+1$ traffic with $T = T_{p+1,\alpha(i,p+1)}$, $r = \hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m} > 0$, $\delta = \hat{\lambda}_{p+1} - \max(\hat{\lambda}_p, \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}) > 0$, $e_4 = \varsigma_1\left(p+1, i, \alpha(i,p+1)+1\right)$, $f_4 = \varepsilon_1\left(p+1, i, \alpha(i,p+1)+1\right)$, $e_5 = \varsigma_3\left(p, i, \alpha(i,p+1)+1\right)$, $f_5 = \varepsilon_3\left(p, i, \alpha(i,p+1)+1\right)$. Thus, $\Delta(i,l') \geq p+1$, for all links l' on the path to receiver s and link l lies on this path. From (102) for $p+1$, layer 1, (60) holds for all links on the path. Consider any link l'' traversed by session i. If $\Delta(i,l'') \geq p+1$, repeating the same arguement it lies on a session i source destination path satisfying (60). Let $\Delta(i,l'') \leq p$. Note that $\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < \hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < b_{i,\alpha(i,p+1)+1}$. Since $b_{i,\alpha(i,p+1)+1} \geq 0$, it follows that $\left(\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}\right)^+ \leq b_{i,\alpha(i,p+1)+1}$. Thus by induction hypothesis (106) for $1,\dots,p$, and since $T_{q,j}$ is nondecreasing in q for all layers j, for any positive integer K, any nondecreasing sequence of times $s_1, t_1, \dots, s_K, t_K \in [T_{p+1,\alpha(i,p+1)}, \infty)$

$$\begin{aligned}\sum_{k=1}^{K} P(i,\alpha(i,p+1)+1,l'',s_k,t_k) \quad &\leq \quad \left(\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}\right)^+ \sum_{k=1}^{K}(t_k - s_k) \\ &+\varsigma_3\left(\Delta(i,l''), i, \alpha(i,p+1)+1\right) \\ &+K\varepsilon_3\left(\Delta(i,l''), i, \alpha(i,p+1)+1\right) \end{aligned} \tag{138}$$

Thus (61) holds for l'', holds for (138) and the fact that $r - \delta \geq$ $\left(\hat{\lambda}_{\Delta(i,l'')} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}\right)^+$. Also (62) holds from (95), since $\varepsilon_1(q)$ is nondecreasing, and (64) holds from (96), since $\varsigma_1(q), \varsigma_3(q), \varepsilon_1(q), \varepsilon_3(q)$ are nondecreasing. Thus Lemma 2 applies in this case. Thus, from Lemma 2, (103) holds for $p+1$ for session i on link l, with
$T_{p+1,\alpha(i,p+1)+1} = \max_{s:\vartheta(s)\geq p+1} T_{\chi(s)}\left(H_{\chi(s)}, T_{p+1,\alpha(i,p+1)}\right)$.

Now we prove (104) for $p+1$, session i, layer $\alpha(i,p+1)+1$. Consider a link l and a session i traversing l with $\Delta(i,l) \geq p+1$. Let $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < \hat{\lambda}_{p+1}$. We have shown that Lemma 2 applies in this case with $T = T_{p+1,\alpha(i,p+1)}$. It follows from Lemma 2 that $B(i,\alpha(i,p+1)+1,l',t) > W$, for $t \geq T_{p+1,1}$, with $T_{p+1,\alpha(i,p+1)+1} = \max_{s:\vartheta(s)\geq p+1} T_{\chi(s)}(H_{\chi(s)}, T_{p+1,\alpha(i,p+1)+1})$. Thus (104) holds in this case.

Now, we prove (105) for $p+1$. The proof is similar to the case $p = 1$, and uses the fact that (103) holds for layers $1,\dots,\alpha(i,p+1)+1$, for any link l and a session i traversing l with $\Delta(i,l) \geq p+1$.

Now we will prove (106) for $p+1$, session i, layer $\alpha(i,p+1)+1$. Consider a link l and a session i traversing l, with $\Delta(i,l)=p+1$. Consider layer $\alpha(i,p+1)+1$. Consider any $s_1,t_1,\ldots,s_K,t_K \in [T_{1,\alpha(i,p+1)+1},\infty)$. Initially, assume that link l is the bottleneck link of virtual session s of session i. (So l is not the link originating from the source of session i.) Since l is the bottleneck link of s, under maxmin fair allocation, its capacity is fully consumed by sessions w with $\Delta(w,l) \le \vartheta(s) \le \Delta(i,l) = p+1$. Let $X = \{w : w \in n(l), \Delta(w,l) \le p+1, w \ne i\}$.

$$C_l = \sum_{w\in X} \lambda_{wl} + \lambda_{il} \tag{139}$$

Also

$$\begin{aligned}
& P(i,\alpha(i,p+1)+1,l,s_1,t_K) \\
\le\ & \left(C_l(t_K-s_1)+1-\sum_{w\in X}P(w,l,s_1,t_K)-\sum_{m=1}^{\alpha(i,p+1)}P(i,m,l,s_1,t_K)\right)^+ \\
& \text{(since the same turn can not be used} \\
& \text{to transmit packets of two different sessions} \\
& \text{or two different layers of the same session} \\
& \text{in the same link and at most } C_l(t_K-s_1)+1 \\
& \text{turns are offered at the origin of } l \text{ during } [s_1,t_K)) \\
\le\ & \left(C_l-\sum_{w\in X}\hat{\lambda}_{\Delta(w,l)}-\sum_{m=1}^{\alpha(i,p+1)}b_{i,m}\right)^+(t_K-s_1) \\
& +\left(\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(1,i,m)+\varepsilon_2(1,i,m))\right)+(\varrho(p+1)+\tau(p+1))\,|X|+1 \\
& \text{(from (103) for layers } 1,\ldots,\alpha(i,p+1),\ (105) \\
& \text{for } 1,\ldots,p,\ \varrho(x),\tau(x) \text{ are nondecreasing)} \\
\le\ & \left(C_l-\sum_{w\in X}\lambda_{wl}-\sum_{m=1}^{\alpha(i,p+1)}b_{i,m}\right)^+(t_K-s_1) \\
& +\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(i,m)+\varepsilon_2(i,m))+(\varrho(p+1)+\tau(p+1))\,(|n(l)|-1)+1 \\
& \text{(since } \lambda_{wl}=\hat{\lambda}_{\Delta(w,l)}\forall w\in X, \text{ and } X\subset n(l)) \\
=\ & (\lambda_{il}-\sum_{m=1}^{\alpha(i,p+1)}b_{i,m})^+(t_K-s_1)+\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(1,i,m)+\varepsilon_2(1,i,m)) \\
& +(\varrho(p+1)+\tau(p+1))\,(|n(l)|-1)+1 \qquad \text{(from (139))} \\
=\ & (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)}b_{i,m})(t_K-s_1)+\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(1,i,m)+\varepsilon_2(1,i,m)) \\
& +(\varrho(p+1)+\tau(p+1))\,(|n(l)|-1)+1
\end{aligned} \tag{140}$$

$$\text{(since } i\in n(l) \text{ and } \Delta(i,l)=p+1, \hat{\lambda}_{p+1}=\lambda_{il} \text{ and } \sum_{m=1}^{\alpha(i,p+1)}b_{i,m}\le\hat{\lambda}_{p+1})$$

Now let l not be the bottleneck link of any virtual session of session i traversing l. Thus all the virtual sessions in $m(i,l)$ are bottlenecked somewhere else. Let all of them have bottleneck links downstream. Consider the sequence of links on session i path constructed as follows. The first link in the sequence originates from the sink of l and has the minimum queue length at t_K amongst all session i links originating from the sink of l. Any other link in the sequence originates from the sink of the previous link and has the minimum queue length at t_K amongst all session i links originating from the sink of the previous link. Let $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < \hat{\lambda}_{p+1}$. Let $\Delta(i, l_q) < p+1$, for some link l_q in the sequence. Consider the link l' closest to l such that $\Delta(i, l') < p+1$. Clearly l' will have a session i sibling l'' s.t. $\Delta(i, l'') = p+1$. From (104), $B(i, \alpha(i,p+1)+1, l', t) > W + \Upsilon\Gamma$, for all $t \geq T_{p+1,\alpha(i,p+1)+1}$. But,

$$\begin{aligned} & B(i, \alpha(i,p+1)+1, l_q, t_K) \leq W + \Upsilon\Gamma, \text{ for all links } l_q \text{ in the sequence} \\ & \text{and } t_K \geq T_{p+1,\alpha(i,p+1)+1} \end{aligned} \tag{141}$$

This is because $B(i, \alpha(i,p+1)+1, l_q, t_K) = \min B(i, \alpha(i,p+1)+1, l_r, t_K)$, with the minimum being taken over all session i links l_r originating from the source of l_q. This is a contradiction. Hence, $\Delta(i, l_q) \geq p+1$, for all links l_q in the sequence. Since $\Delta(i, l) = p+1$, and l is upstream of any link l_q in the sequence, $\Delta(i, l_q) = p+1$, for all links l_q in the sequence. Since all virtual sessions in $m(i, l)$ have their bottleneck links downstream of l, bottleneck link of some virtual session in $m(i, l)$ will be in the sequence. Let the bottleneck link be l^*. As argued before, $\Delta(i, l^*) = p+1$. By (140),

$$\begin{aligned} & P(i, \alpha(i,p+1)+1, l^*, s_1, t_K) \\ \leq\ & (\hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m})(t_K - s_1) + \sum_{m=1}^{\alpha(i,p+1)} (\varsigma_2(p+1, i, m) + \varepsilon_2(p+1, i, m)) \\ & + (\varrho(p+1) + \tau(p+1))(E_i - 1) + 1 \end{aligned} \tag{142}$$

First consider the links l^* and the link l_q just preceding l^* in the sequence. We will use Lemma 4 to show that $B(i, \alpha(i,p+1)+1, l^*, u) < G$, $\forall\ u \in [s_1, t_K)$. Let $B(i, \alpha(i,p+1)+1, l^*, u) = G$, for some $u \in [s_1, t_K)$. Let $l_1 = l_q$, $l_2 = l^*$, $j = \alpha(i,p+1)+1$, $r = \hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}$, $e_7 = \varsigma_2(p+1, i, \alpha(i,p+1)+1) + \varepsilon_2(p+1, i, \alpha(i,p+1)+1)$, $e_8 = \sum_{m=1}^{\alpha(i,p+1)} (\varsigma_2(p+1, i, m) + \varepsilon_2(p+1, i, m))$
$+ (\varrho(p+1) + \tau(p+1))(E_i - 1) + 1$. As argued before, session i has index $p+1$ in all links in the sequence. From (103) for $p+1$, (91) holds. From (142), (92) holds. Clearly, (93) holds from the assumption that $B(i, \alpha(i,p+1)+1, l^*, u) = G$. Also (94) holds from (96). From Lemma 4, $B(i, \alpha(i,p+1)+1, l^*, t) > W + \Upsilon\Gamma$, for all $t \geq u$. This contradicts (141) since $t_K > u$. Thus $B(i, \alpha(i,p+1)+1, l^*, u) < G\ \forall\ u \in [s_1, t_K)$. Thus there is no packet loss in the session i buffer for l^* in the interval $[s_1, t_K)$. It follows that,

$$\begin{aligned} & P(i, \alpha(i,p+1)+1, l_q, s_1 - \Upsilon_{l_q}, t_K - \Upsilon_{l_q}) \\ \leq\ & P(i, \alpha(i,p+1)+1, l_q, s_1, t_K) + \Upsilon\Gamma \\ =\ & P(i, \alpha(i,p+1)+1, l^*, s_1, t_K) + B(i, \alpha(i,p+1)+1, l^*, t_K) \\ & -B(i, \alpha(i,p+1)+1, l^*, s_1) + \Upsilon\Gamma \end{aligned}$$

$$
\begin{aligned}
&\leq P(i,\alpha(i,p+1)+1,l^*,s_1,t_K)+B(i,\alpha(i,p+1)+1,l^*,t_K)+\Upsilon\Gamma \\
&\leq P(i,\alpha(i,p+1)+1,l^*,s_1,t_K)+W+2\Upsilon\Gamma \text{ (from (141))} \\
&\leq (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m})(t_K-s_1)+W+2\Upsilon\Gamma+ \\
&\quad (\varrho(p+1)+\tau(p+1))\,E_i+1 \text{ (from (142))} \qquad (143)
\end{aligned}
$$

The last inequality follows from (142) using the fact that $\sum_{m=1}^{\alpha(i,p+1)} \varsigma_2(p+1,i,m) \leq \rho(p+1)$, and $\sum_{m=1}^{\alpha(i,p+1)} \varepsilon_2(p+1,i,m) \leq \tau(p+1)$. Now using (143) in place of (142), a similar upper bound as (143) can be proved for the link preceding l_q in this sequence. Proceeding in this manner,

$$
\begin{aligned}
P(i,\alpha(i,p+1)+1,l,s_1,t_K) &\leq (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m})(t_K-s_1) \\
&\quad +(W+2\Upsilon\Gamma)(H_i-1) \\
&\quad (\varrho(p+1)++\tau(p+1))\,(E_i+H_i-2)+1 \qquad (144)
\end{aligned}
$$

Now consider the case in which $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} = \hat{\lambda}_{p+1}$. (By definition, $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} \leq \hat{\lambda}_{p+1}$. We have considered the case in which $\sum_{m=1}^{\alpha(i,p+1)} b_{i,m} < \hat{\lambda}_{p+1}$, previously). If $\Delta(i,l_q)=p+1$, for all links l_q in the sequence, then (144) holds, and can be argued as before. So, let $\Delta(i,l_q)<p+1$ for some link l_q in the sequence ($\Delta(i,l_q)\leq\Delta(i,l)=p+1$, for all links l_q in the sequence, since all these links are downstream of l). Let l' be the link closest to l in the sequence, such that $\Delta(i,l')<p+1$ (Note that $l'\neq l$ as $\Delta(i,l)=p+1$). Since $\Delta(i,l')<p+1$, (106) holds for session i, layer $\alpha(i,p+1)+1$ by induction hypothesis. Also, $\hat{\lambda}_{\Delta(i,l')}<\hat{\lambda}_{p+1}=\sum_{m=1}^{\alpha(i,p+1)} b_{i,m}$. Thus $\left(\hat{\lambda}_{\Delta(i,l')}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m}\right)^+=0$. Hence, from (106), and the fact that $b_{i,\alpha(i,p+1)+1}\geq 0$,

$$
\begin{aligned}
P(i,\alpha(i,p+1)+1,l',s_1,t_K) &\leq \varsigma_2\left(\Delta(i,l'),i,\alpha(i,p+1)+1\right) \\
&\quad +\varepsilon_2\left(\Delta(i,l'),i,\alpha(i,p+1)+1\right) \qquad (145)
\end{aligned}
$$

First, consider the link l_q preceding l' in the sequence. We will use Lemma 4 to show that $B(i,\alpha(i,p+1)+1,l',u)<G$, $\forall\ u\in[s_1,t_K)$. Let $B(i,\alpha(i,p+1)+1,l',u)=G$, for some $u\in[s_1,t_K)$. Let $l_1=l_q$, $l_2=l'$, $j=\alpha(i,p+1)+1$, $r=0$, $e_7=0$, $e_8=\varsigma_2\left(\Delta(i,l'),i,\alpha(i,p+1)+1\right)+\varepsilon_2\left(\Delta(i,l'),i,\alpha(i,p+1)+1\right)$. Here, (91) holds trivially. From (145), (92) holds. Clearly, (93) holds from the assumption that $B(i,\alpha(i,p+1)+1,l^*,u)=G$. Also (94) holds from (96). From Lemma 4, $B(i,\alpha(i,p+1)+1,l',t)>W+\Upsilon\Gamma$, for all $t\geq u$. This contradicts (141) since $t_K>u$. Thus $B(i,\alpha(i,p+1)+1,l',u)<G\ \forall\ u\in[s_1,t_K)$. Thus there is no packet loss in the session i buffer for l' in the interval $[s_1,t_K)$.

$$
\begin{aligned}
&P(i,\alpha(i,p+1)+1,l_q,s_1-\Upsilon_{l_q},t_K-\Upsilon_{l_q}) \\
&\leq P(i,\alpha(i,p+1)+1,l_q,s_1,t_K)+\Upsilon\Gamma \\
&= P(i,\alpha(i,p+1)+1,l',s_1,t_K)+B(i,\alpha(i,p+1)+1,l^*,t_K) \\
&\quad -B(i,\alpha(i,p+1)+1,l^*,s_1)+\Upsilon\Gamma
\end{aligned}
$$

$$
\begin{aligned}
&\leq P(i,\alpha(i,p+1)+1,l^*,s_1,t_K)+B(i,\alpha(i,p+1)+1,l^*,t_K)+\Upsilon\Gamma \\
&\leq P(i,\alpha(i,p+1)+1,l^*,s_1,t_K)+W+2\Upsilon\Gamma \text{ (from (141))} \\
&\leq W+2\Upsilon\Gamma+(\varrho(p+1)+\tau(p+1))\,E_i+1 \text{ (from (145))} \qquad (146)
\end{aligned}
$$

The last inequality follows from (145) using the fact that $\varsigma_2(\Delta(i,p+1),i,\alpha(i,p+1)+1) \leq \rho(p+1)$, and $\varepsilon_2(\Delta(i,p+1),i,\alpha(i,p+1)+1) \leq \tau(p+1)$. Now using (146) in place of (145), a similar upper bound as (146) can be proved for the link preceding l_q in this sequence. Proceeding in this manner,

$$
\begin{aligned}
& P(i,\alpha(i,p+1)+1,l,s_1,t_K) \\
\leq\ & (W+2\Upsilon\Gamma)(H_i-1)+(\varrho(p+1)+\tau(p+1))\,(E_i-1+H_i-1)+1 \\
=\ & (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m})(t_K-s_1)+(W+2\Upsilon\Gamma)(H_i-1)+ \\
& (\varrho(p+1)+\tau(p+1))\,(E_i-1+H_i-1)+1 \qquad (147)
\end{aligned}
$$

The last inequality follows since $\hat{\lambda}_{p+1}=\sum_{m=1}^{\alpha(i,p+1)} b_{i,m}$, by assumption.

Now let at least one virtual session in $m(i,l)$ have a bottleneck link upstream. Let the link be l^* and the virtual session s. Now, $\Delta(i,l') \geq \Delta(i,l)=p+1$, for all links l' upstream of l. So, $\Delta(i,l^*) \geq p+1$. Now, $\vartheta(s) \leq \Delta(i,l)=p+1$, since $s \in m(i,l)$. Also, $\vartheta(s)=\Delta(i,l^*)$, since l^* is the bottleneck link[§] of s. Hence, $\Delta(i,l^*)=p+1$. We will use Lemma 3, with $j=\alpha(i,p+1)+1$, $l_1=l^*$, $e_1=\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(p+1,i,m)+\varepsilon_2(p+1,i,m))+(\varrho(p+1)+\tau(p+1))\,(E_i-1)+1$, $e_2=\varsigma_2(p+1,i,\alpha(i,p+1)+1)+\varepsilon_2(p+1,i,\alpha(i,p+1)+1)$, $T_0=T_{p+1,\alpha(i,p+1)+1}$ and $r=\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m}$. Condition (87) follows from (142), condition (88) follows from (103) since $\Delta(i,l') \geq p+1$ for all links l' between l and l^*. Hence,

$$
\begin{aligned}
& P(i,\alpha(i,p+1)+1,l,s_1,t_K) \\
\leq\ & (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,p+1)} b_{i,m})(t_K-s_1)+2^{H_i-1}\left((E_i-1)\,(\varrho(p+1)+\tau(p+1))\right. \\
& \left.+\sum_{m=1}^{\alpha(i,p+1)}(\varsigma_2(p+1,i,m)+\varepsilon_2(p+1,i,m))+1\right) \\
& +(2^{H_i-1}-1)\,(\varsigma_2(p+1,i,\alpha(i,p+1)+1)+\varepsilon_2(p+1,i,\alpha(i,p+1)+1)) \\
& +(H_i-1)\left(\max_{l:i\in n(l)} B(i,\alpha(i,p+1)+1,l,T_{p+1,\alpha(i,p+1)+1})+2\Upsilon\Gamma\right) \\
\leq\ & (\hat{\lambda}_{p+1}-\sum_{m=1}^{\alpha(i,1)} b_{i,m})(t_K-s_1)+2^{H_i-1}\,(E_i\,(\varrho(p+1)+\tau(p+1))+1) \\
& +(H_i-1))\left(\max_{l:i\in n(l)} B\left(i,\alpha(i,p+1)+1,l,T_{p+1,\alpha(i,p+1)+1}\right)+2\Upsilon\Gamma\right) \qquad (148)
\end{aligned}
$$

[§]Recall that every virtual session has a bottleneck link under maxmin fair allocation [14]. A virtual session has the largest rate amongst all virtual sessions of its session traversing the bottleneck link.

The last inequality follows since $\sum_{m=1}^{\alpha(i,p+1)+1} \varsigma_2(p+1,i,m) \leq \varrho(1)$ and $\sum_{m=1}^{\alpha(i,p+1)+1} \varepsilon_2(p+1,i,m) \leq \tau(p+1)$. Similar to (110),

$$\begin{aligned}
& \sum_{k=1}^{K} P(i,\alpha(i,p+1)+1,l,s_k,t_k) \\
= \;& P(i,\alpha(i,p+1)+1,l,s_1,t_K) - \sum_{k=1}^{K-1} P(i,\alpha(i,p+1)+1,l,t_k,s_{k+1}) \\
\leq \;& P(i,\alpha(i,p+1)+1,l,s_1,t_K) - (\hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}) \sum_{k=1}^{K-1}(s_{k+1}-t_k) \\
& + \varsigma_2\left(1,i,\alpha(i,p+1)+1\right) + (K-1)\varepsilon_2(1,i,\alpha(i,p+1)+1)
\end{aligned} \tag{149}$$

The last step follows from (103) for $p+1$, layer $\alpha(i,p+1)+1$.

From (144), (147), (148) and (149)

$$\begin{aligned}
& \sum_{k=1}^{K} P(i,\alpha(i,p+1)+1,l,s_k,t_k) \\
\leq \;& (\hat{\lambda}_{p+1} - \sum_{m=1}^{\alpha(i,p+1)} b_{i,m}) \sum_{k=1}^{K}(t_k-s_k) + \varsigma_3(1,i,\alpha(i,p+1)+1) \\
& + K\varepsilon_3(1,i,\alpha(i,p+1)+1)
\end{aligned}$$

$$\begin{aligned}
& \text{with } \varsigma_3\left(p+1,i,\alpha(i,p+1)+1\right) = \varsigma_2\left(p+1,i,\alpha(i,p+1)+1\right) \\
& +(W+2\Upsilon\Gamma)(H_i-1) + 2^{H_i-1}\left(E_i\left(\varrho(p+1)+\tau(p+1)\right)+1\right) \\
& +(H_i-1)\max_{l:i\in n(l)} B\left(i,\alpha(i,p+1)+1,l,T_{p+1,\alpha(i,p+1)+1}\right)
\end{aligned} \tag{150}$$

$$\text{and } \varepsilon_3\left(1,i,\alpha(i,1)+1\right) = \varepsilon_2\left(1,i,\alpha(i,1)+1\right)$$

We have used the fact that $E_i + H_i - 2 \leq 2^{H_i-1}E_i$, since $E_i \geq 1$, in deriving (150).

Now, we prove (102) and (103) for layers $j > \alpha(i,p+1)+1$. Observe that $\hat{\lambda}_{p+1} < \sum_{m=1}^{\alpha(i,p+1)+1} b_{i,m} \leq \sum_{m=1}^{j-1} b_{i,m}$. Thus $\hat{\lambda}_{p+1} - \sum_{m=1}^{j-1} b_{i,m} < 0$. Hence, (102) and (103) follow trivially, since $\varsigma_1(p+1,i,j)$, $\varepsilon_1(p+1,i,j)$, $\varsigma_2(p+1,i,j)$, $\varepsilon_2(p+1,i,j)$ are nonnegative by definition.

Now, we prove (106) for layers $j > \alpha(i,p+1)+1$. Consider a link l and a session i traversing l, with $\Delta(i,l) = p+1$. Consider layer $j > \alpha(i,p+1)+1$ and nondecreasing times, $s_1, t_1, \ldots, s_k, t_k \in [T_{p+1,\alpha(i,p+1)+1}, \infty)$. Initially, assume that link l is the bottleneck link of virtual session s of session i. Let $X = \{w : \Delta(w,l) \leq p+1\}$. As before,

$$\begin{aligned}
P(i,j,l,s_1,t_K) \;\leq\; & \Bigg(C_l(t_K-s_1) + 1 - \sum_{w\in X} P(w,l,s_1,t_K) \\
& - \sum_{m=1}^{\alpha(i,p+1)+1} P(i,m,l,s_1,t_K)\Bigg)^+
\end{aligned}$$

Using (103) for $p+1$, layers $1, \ldots, \alpha(i, p+1)+1$, (105) for $1, \ldots, p+1$, and proceeding as in the proof of (140) we have,

$$\begin{aligned} P(i,j,l,s_1,t_K) &\leq \sum_{m=1}^{\alpha(i,p+1)+1} (\varsigma_2(p+1,i,m) + \varepsilon_2(p+1,i,m)) \\ &\quad + (\varrho(p+1) + \tau(p+1))\,(E_i - 1) + 1 \\ &\leq (\varrho(p+1) + \tau(p+1))\,E + 1 \end{aligned}$$

The last inequality follows since $\sum_{m=1}^{\alpha(i,p+1)+1} \varsigma_2(p+1,i,m) \leq \varrho(p+1)$ and $\sum_{m=1}^{\alpha(i,p+1)+1} \varepsilon_2(p+1,i,m) \leq \tau(p+1)$.

If l is not the bottleneck link, and all the virtual sessions of session i traversing link l are bottlenecked downstream, then construct the sequence of links l_q as before, and consider two separate cases. In one, $\Delta(i, l_q) = p+1$ for all links l_q in the sequence, and in the other $\Delta(i, l_q) < p+1$, for some link l_q in the sequence. Proceeding as for $p+1$, layer $\alpha(i,p)+1$, in both cases,

$$P(i,j,l,s_1,t_K) \leq (W + 2\Upsilon\Gamma)(H_i - 1) + (\varrho(p+1) + \tau(p+1))\,E_i + 1$$

Proceeding similarly in the other remaining case, i.e., at least one virtual session is bottlenecked upstream, using Lemma 3, (103) for $p = 1$, layer j, session i and the facts that $\varsigma_2(p+1,i,j) = \varepsilon_2(p+1,i,j) = 0$,

$$\begin{aligned} P(i,j,l,s_1,t_K) &\leq 2^{H_i-1}\,(E_i\,(\varrho(p+1) + \tau(p+1)) + 1) \\ &\quad + (H_i - 1)\left(\max_{l: i \in n(l)} B(i,j,l,T_{p+1,\alpha(p+1)+1}) + 2\Upsilon\Gamma\right) \end{aligned}$$

It follows that for all these cases,

$$\begin{aligned} P(i,j,l,s_1,t_K) &\leq 2^{H_i-1}\,(E_i\,(\varrho(p+1) + \tau(p+1)) + 1) \\ &\quad + (H_i - 1)\left(\max_{l: i \in n(l)} B(i,j,l,T_{p+1,\alpha(p+1)+1}) + W + 2\Upsilon\Gamma\right) \end{aligned}$$

$$\text{Clearly, } \sum_{k=1}^{K} P(i, \alpha(i,p+1)+1, l, s_k, t_k) \leq P(i, \alpha(i,p+1)+1, l, s_1, t_K)$$

Thus, for $j > \alpha(i, p+1) + 1$,

$$\sum_{k=1}^{K} P(i,j,l,s_k,t_k) \leq \varsigma_3(p+1,i,j) + K\varepsilon_3(p+1,i,j) \tag{151}$$

Thus (106) follows for $j > \alpha(i,1) + 1$.

Thus (106) holds for $p = 1$ for all layers. Now we will prove (107) for $p+1$. This proof follows from (106) for $p+1$, similar to the proof for for $p+1$.

Note that (98) follows from (105), and (100) folllows from (107). Also, (99) folllows from (103) and (101) folllows from (106). $\square$

Note that Theorems 1, 2 and 3 follow directly from Theorem 4. Let $l(j)$ be the link serving packets to receiver j (last link on the path of receiver j). Theorem 1 follows from (98) and (100) of Theorem 4 with $W^* = W_F$, $G^* = G_F$ and $\delta_j = C_{l(j)} T_{\vartheta(j)+1,0} + \kappa(\vartheta(j)) + \gamma(\vartheta(j))$. We use the facts that $\varrho(p) \leq \kappa(p)$ and $\tau(p) \leq \gamma(p)$. Theorem 2 follows from (99) and (101) of Theorem 4 with $W^* = W_F$, $G^* = G_F$ and $\iota_{j,k} = b_{i,k} T_{\vartheta(j)+1,0} + \xi_{i,k} + \varsigma_3(\vartheta(j), i, k) + \varepsilon_3(\vartheta(j), i, k)$. We have used the facts that $\varsigma_3(p, i, k) \geq \varsigma_2(p, i, k)$ and $\varepsilon_3(p, i, k) \geq \varepsilon_2(p, i, k)$. Theorem 3 follows from Theorem 4 with W_is and G_is given by (95) and (96).

VI. Conclusion

We discuss some attractive features of this scheduling policy in this concluding section.

If the maximum rates of sessions form a feasible rate allocation, then the maximum rates are maxmin fair. Thus the proposed scheduling policy attains the maximum rates, as long as these rates are feasible. In other words, resource limitation permitting, the proposed scheduling policy satisfies all users.

As we have mentioned before, the scheduling policy is adaptive as it does not require any knowledge of maximum rates of users, layer bandwidths, or the user traffic statistics. The results are independent of the hierarchical structure of signals and unequal bandwidth layers are permitted. Also, a malicious session can not increase the throughput of its receivers, by choosing layer bandwidths suitably. Maxmin fairness of receiver throughputs are guaranteed irrespective of the layer bandwidths. This is useful because in real networks these parameters are not generally known.

The proposed scheduling policy is simple to implement. No intensive computation is involved anywhere.

A link scheduler takes scheduling decisions whenever it is free to transmit a packet. It need not synchronize with schedulers for other link.

Note that this policy offers different quality of service to different layers. Layered traffic is a special case of priority traffic, with the lowest layer traffic having the highest priority, and the higher layers lower priority. It is possible to generalize this scheduling policy to attain maxmin fairness with priorities, by considering different priority sessions instead of considering different layers. This would allow service differentiation within the framework of fairness. Service differentiation is a major feature of modern communication networks. Network may offer different priorities to different sessions. For example, network may decide to distribute the available bandwidth fairly between high priority sessions. Any residual bandwidth is distributed fairly between remaining low priority sessions. In this scenario, it is useful to have a scheduling policy which can attain prioritized maxmin fairness, i.e., maxmin fairness among sessions with highest priority, and maxmin fair allocation of the residual bandwidth (if any) among sessions with low priority, etc. This has not been addressed even in the unicast context. Layered traffic is a special case of priority traffic, with lowest layer traffic having the highest priority, and the higher layers lower priority. It is possible to generalize this scheduling policy to attain maxmin fairness with priorities, by considering different priority sessions instead of considering different layers.

The scheduling policy does not assume any particular drop policy. Physical buffers can follow any drop pattern. When a packet arrives and finds the buffer full, it is not necessary that the new packet is dropped, but an old packet may also be dropped to make place for the new packet. Dropping an old packet may be a better option for real time transmission, because packets arriving after a certain delay become useless. The routers may follow a drop tail (drop the new packet) or random drop (drop a randomly chosen packet in the queue) or drop head (drop the oldest packet) policy. The allocated rates will converge to the maxmin fair rates in all these cases.

The disadvantages of this policy are as follows. Link schedulers need congestion information of the neighbors. If a link scheduler is located at the source of a link, it would mostly need local information to make the scheduling decision. Specifically, it needs to know whether the number of packets of a session at the destination node of the link is less than the credit value or not. Whenever, queue length of a session crosses the threshold of W, a message indicating the same can be sent to the source node. This message can be piggy backed in information packets as well. Thus overhead is low. As we have mentioned earlier, this hop by hop congestion feedback has certain advantages over end to end congestion control only, and is used in LAN based networks[15]. Also, we have shown analytically and experimentally that feedback delays do not alter the throughput, as long as the delays are bounded. This makes it suitable for implementation in networks where propagation delay is significant, e.g., networks with satellite links. It is worthwhile to point out that even though this is a back pressure based policy, it does not suffer from any deadlock since the back pressure on a session is exerted by the accumulation of the packets of the same session. Also, livelock can be avoided by using first come first serve scheduling among packets of the same layer of the same session. The other disadvantage is that this scheduling policy requires per flow states in the routers. The complexity increase is not as drastic as it seems, though. Arguing in the lines of Grossglausser and Bolot[6], implementing a multicast/multilayer service requires per-flow state in the routers anyway. So, the incremental cost of maintaining some more information for each flow and using this additional information in the scheduling policy is much smaller than in the case of unicast. However, if these additional flow states become an issue, then this policy can be used in the VPNs and intranets, and state aggregation may be resorted to in the backbones. In a nutshell, this scheduling policy is suitable for use in large, dynamic, high speed networks that have access to decentralized, delayed and partial information only.

This scheduling policy assumes that intermediate routers are able to make class based distinction in multicast networks. Current day multicast capable routers are not able to do so. However, this is likely to happen in near future, especially as multicast communication becomes more popular. Newly suggested internet policies like diffserv and intserv already propose class based distinction in unicast networks to guarantee requisite quality of services. It is likely that these would be extended to multicast in future.

We also assume that receiving a portion of the packets of a layer improves the reception quality as compared to receiving no packet of the layer. This assumption is realistic as packet loss causes a graceful degradation in signal quality in many coding schemes[5]. For example, in [5] signal exhibits a reasonable enhancement in quality above the lowest layer for 10% packet loss in the next higher layer. This transition is gradual up until 100% packet loss in the next higher layer. Real time transmission in internet uses thse coding schemes, because packet losses are frequent. However, if the coding scheme does not show a reception quality improvement, unless most of the packets of a layer are received, then we need to use a scheduling policy which can detect partial transmission of a layer, and subsequently drop such a layer. The residual bandwidth can be used to serve complete layers for at least a few sessions. The objective in this discrete case should be to attain a maximally fair allocation. Developing such a scheduling policy is a topic of future research.

We have presented a scheduling policy which attains maxmin fair rates in multirate multicast networks, without computing the rates beforehand. The scheduling policy concentrates the packet loss in the highest layer served. This improves quality of service in view of the characteristic of the hierarchical coding that reception of higher layer conveys useful information, only when all the lower layers have been successfully decoded. As discussed, this scheduling policy has several other attractive properties, which render it suitable for use in large, dynamic, high speed networks that have access to decentralized, delayed and partial information only. It allows the network to enforce fair allocations, without trusting the receivers to attain fairness. This scheme is computationally simple, has low message overhead and hence does not unduly overload the network. Also, we have assumed multirate transmission. Class based distinction is more difficult for multirate transmission than for unirate transmission. With a minor modification, this scheduling policy will work for unirate networks. In fact an extension of the policy in [8] will attain fairness in unirate multicast networks, as unirate multicast networks are not significantly different from unicast networks.

References

[1] T. Ballardie, P. Francis, and J. Crowcroft. Core based trees: an architecture for scalable inter-domain multicast routing, *Proceedings of ACM SIGCOMM,* Sept. 1993, pp. 85-95

[2] D. Bertsekas and R. Gallager, *Data Networks*, Englewood Cliffs, NJ:Prentice-Hall, 1987

[3] J. Bennett and H. Zhang. Hierarchical packet fair queueing algorithms *Proceedings of ACM SIGCOMM'* 96, *pp* 143 − 156, Palo Alto, CA, August ′96

[4] S. Deering and D. Cheriton. Multicast routing in datagram internetworks and extended LANs, ACM Transactions on Computer Systems, vol 8, no. 2, pp. 54-60, Aug. 1994

[5] M. Ghanbari. Two-Layer Coding of Video Signals for VBR Networks. *IEEE Journal on Selected Areas in Communications*, VOL. 7. No. 5. June 1989

[6] M. Grossglausser and J. Bolot. On service models for Multicast Transmission in Heterogeneous Environments. *Proceedings of IEEE' INFOCOM'* 2000*,* Tel-Aviv, Israel, March, 2000

[7] P. Goyal, H. Vin and H. Chen. Start-time Fair Queueing: A scheduling algorithm for integrated services. *Proceedings of the ACM-SIGCOMM* 96 *pp* 157 − 168, Palo Alto, CA, August ′96

[8] E. Hahne. "Round-Robin Scheduling for Max-Min Fairness in Data Networks," *IEEE Journal on Selected Areas in Communications,* Vol. 9, No. 7, Sept. 1991, pp $1024 - 1039$

[9] X. Li, S. Paul and M. H. Ammar. Layered Video Multicast with Retransmission (LVMR): Evaluation of Hierarchical rate control, *Proceedings of IEEE Infocom' 98*, March 1998

[10] X. Li, S. Paul and M. H. Ammar. Multi-session Rate Control for Layered Video Multicast *Technical Report* GT-CC-$98 - 21$, College of Computing, Georgia Institute of Technology, 1998

[11] S. McCanne, V. Jacobson and M. Vetterli. Receiver-Driven Layered Multicast, *Proceedings of ACM SIGCOMM '96*, Stanford, CA, September 1996

[12] D. Rubenstein, J. Kurose and D. Towsley. The Impact of Multicast Layering on Network Fairness *Proceedings of ACM SIGCOMM '*99, *Cambridge, MA, September,* 1999

[13] S. Sarkar and L. Tassiulas: Distributed Algorithms for Computation of Fair Rates in Multirate Multicast Trees *Proceedings of IEEE INFOCOM'* 2000, Tel Aviv, Israel, March 2000

[14] S. Sarkar and L. Tassiulas, "Back pressure based multicast scheduling for fair bandwidth allocation," *Technical report*, Electrical and Systems Engineering Department, University of Pennsylvania, Available at http://www.seas.upenn.edu/ ~swati

[15] F. A. Tobagi and W. K. Noureddine, "Back-Pressure Mechanisms in Switched LANs Carrying TCP and Multimedia Traffic," *submitted to IEEE GLOBECOM'* 99, *Symposium on High-Speed Networks,* December, 1999

[16] H. Y. Tzeng and K. Y. Siu. On Max-Min Fair Congestion Control for Multicast ABR Service in ATM, *IEEE Journal on Selected Areas In Communications*, Vol 15, No. 3, 1997