Inter-Receiver Fairness: A Novel Performance Measure for Multicast ABR Sessions *

Tianji Jiang **Mostafa H. Ammar** **Ellen W. Zegura**

Networking and Telecommunications Group
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280
{tjiang,ammar,ewz}@cc.gatech.edu

Abstract

In a multicast ABR service, a connection is typically restricted to the rate allowed on the bottleneck link in the distribution tree from the source to the set of receivers. Because of this, receivers in the connection can experience *inter-receiver unfairness*, when the preferred operating rates of the receivers are different. In this paper we explore the issue of improving the inter-receiver fairness in a multicast ABR connection by allowing the connection to operate at a rate higher than what is allowed by the multicast tree's bottleneck link. Since this can result in cell loss to some receivers, we operate with the knowledge of each receiver's application-specific loss tolerance. The multicast connection rate is not allowed to increase beyond the point where the cell loss on a path to a receiver exceeds this receiver's loss tolerance. Based on these ideas we develop an inter-receiver fairness measure and a technique for determining the rate that maximizes this measure. We show possible switch algorithms that can be used to convey the parameters needed to compute the function to the connection's source. In addition we develop a global network measure that helps us assess the effect of increasing inter-receiver fairness on the total network delivered throughput. We also briefly explore improving inter-receiver fairness through the use of multiple virtual circuits to carry traffic for a single multicast session. A set of examples demonstrate the use of the inter-receiver fairness concept in various network scenarios.

1 Introduction

An available bit rate (ABR) service allows applications to access a time-varying network bandwidth capacity. Realization of such a service requires a control loop to allow information about available capacity to reach sources; further, sources must adjust their transmission rate based on feedback information from the network. As an enhanced best-effort service, the advantages of ABR include efficient use of network resources, low loss rate, minimal *a priori* specification of application behavior, and a paradigm somewhat similar to the congestion control mechanisms in TCP over best-effort IP. Given these advantages, it seems likely that an ABR service will be important in both ATM and IP-based networks [1]. In this paper, we will discuss ABR in the context of ATM's connection-oriented environment [2].

*This work is supported in part by research grants from NSF, Intel and Hitachi Telecom (USA).

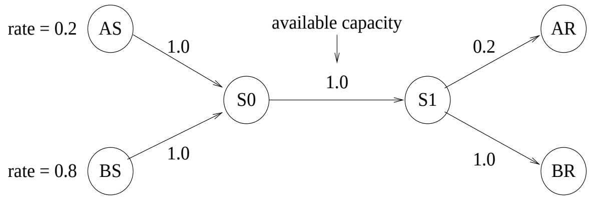

Figure 1: Unicast max-min allocation

An important feature of an ABR service is the "fair" allocation of bandwidth to the set of connections using the ABR service. Max-min allocation [3] is commonly accepted as a fair allocation for ABR. Informally, max-min allocation operates as follows: the available capacity on a link is divided equally among all ABR connections traversing the link, producing a fair share per connection. The allowed rate for the connection is determined by the minimum rate allowed on any link in the path from source to receiver. If a connection is unable to use its fair share on a particular link — e.g., due to a bottleneck link elsewhere in the path — then the excess capacity is divided evenly amongst the remaining connections sharing the link. Thus, each connection is allocated a rate based on the bottleneck link in the path from source to receiver.

An example of max-min allocation is given in Figure 1. Two unicast connections are present in the network, one labeled A (from source AS to receiver AR) and one labeled B (from source BS to receiver BR). The available capacity for each link is indicated, assuming normalization to the range $[0.0, 1.0]$. The allocated rate to connection A will be 0.2, since that is the fair share on the bottleneck link in the path. The allocated rate to connection B will be 0.8; the capacity on the link from switch S0 to switch S1 that is unused by connection A is reallocated to connection B.

When we consider extending ABR service to support point-to-multipoint applications, we must generalize the concept of max-min allocation. The most direct extension of the fair share concept is to restrict the source to the rate allowed on the bottleneck link in the distribution tree from the source to the set of receivers. For example, Figure 2 depicts

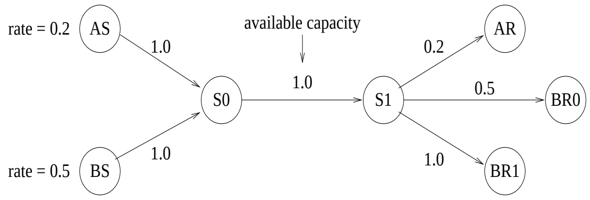

Figure 2: Multicast max-min allocation

a unicast connection A and a multicast connection B, with two receivers BR0 and BR1. The allocated rate to connection B is 0.5 since the bottleneck link in the tree (from S1 to receiver BR0) can support only this rate. Notice, however, that the path from the source to receiver BR1 could support a rate of 0.8. Indeed, if BR1 were the only receiver in the connection, then the allocated rate would be 0.8. Thus, receiver BR0 prefers an operating rate of 0.5, while receiver BR1 prefers an operating rate of 0.8. The example illustrates a key point regarding point-to-multipoint ABR allocation: receivers in a point-to-multipoint connection can experience *inter-receiver unfairness*, when the preferred operating rates of the receivers are different.

Consider operating the connection B at the rate of 0.8, with losses on the link from S1 to receiver BR0 to bring the rate on that link down to 0.5. Note that (with losses) this rate does not cause connection B to exceed its max-min fair share on any link in the network. Depending upon the characteristics of the application, this may be a preferable operating point. Clearly, to consider this operating point, the application must have some tolerance for loss to some receivers. Loss-tolerance might arise because the application is inherently robust against some losses, albeit with a degradation in quality. Real-time voice or video are examples of this type of application. Alternatively, the receivers may have some other mechanism for recovering from losses, for example using forward error correction [4] or retransmissions separate from the usual multicast transmission tree [5]. For this operating point to be preferred, the benefit to receiver BR1 (due to the increase in reception rate) must outweigh the cost to receiver BR0 (due to the losses).

Note that we assume that losses occur because of some form of per-virtual circuit policing function implemented in the switches. This policing function discards cells that exceed the allocated rate before they are placed in any common output buffer. This will insure that a connection that exceeds the rate allocated to it on a link will not adversely affect other connections sharing this link.

In this paper, we develop an inter-receiver fairness measure in the context described above. We then develop and analyze a scheme for achieving a source operating rate that maximizes the inter-receiver fairness. In Section 2 we formally define inter-receiver fairness and state the associated optimization problem. Section 3 gives an overview of the scheme to achieve the desired operating rate in an ABR environment. In Sections 4 and 5 we describe the details of the scheme for small and large groups, respectively. In Section 6 we use simulation to investigate the behavior of the inter-receiver fairness measure in various contexts. We comment briefly in Section 7 on an alternative approach that maximizes inter-receiver fairness by allowing the receivers to be partitioned into different connections. This approach has a more complex implementation but offers the potential for better performance. We conclude in Section 8.

2 Inter-Receiver Fairness

2.1 Single Receiver Fairness

Ideally the receivers in a multicast group are isolated from the limitations of the other receivers. In the available bit rate context, a receiver has an *isolated rate* that is the rate that would be achieved if this were the only receiver in the multicast group. To be more specific, for receiver i, the isolated rate r_i is the rate that the ABR source would transmit under the following conditions:

- The path from the source to destination i in the multicast tree is used.
- The source is transmitting only to this destination. Moreover, the other network traffic conditions are the same as the environment in which the multicast connection is to operate.

Note that the isolated rate is equal to the fair share rate on the bottleneck link for the path used from the source to the destination.

If the actual rate of the connection exceeds the isolated rate, then the receiver experiences performance degradation in the form of losses. If the actual rate is less than the isolated rate, then the receiver experiences performance degradation in the form of decreased throughput. In either case, there is a form of unfairness to the receiver.

We define a *single receiver fairness function* that maps from the actual operating rate to a fairness value. For receiver i, the fairness function is denoted $F_i(r)$, where r is the actual operating rate of the connection. The function $F_i(r)$ must satisfy the following constraints:

- $F_i(r) \in [0.0, 1.0]$. That is, the fairness is normalized to a quantity between 0.0 and 1.0.
- $F_i(r_i) = 1.0$, where r_i is the isolated rate of receiver i.
- $F_i(r) < 1.0$ if $r \neq r_i$. In other words, the fairness function has a single maximum at $r = r_i$.
- $F_i(r)$ is non-decreasing in the range $[0, r_i)$. That is, as the throughput increases, the fairness remains constant or increases.
- $F_i(r)$ is non-increasing in the range $(r_i, \infty]$. That is, as losses increase, the fairness remains constant or decreases.

It should be clear that many different functions are possible for $F_i(r)$. The appropriate function will generally depend on application and receiver characteristics.

Our mechanisms to optimize inter-receiver fairness apply to any fairness function meeting the above criteria. For our simulation, we use a particular fairness function of the form:

$$F_i(r) = \begin{cases} \frac{r}{r_i} & \text{if } r \leq r_i \\ \frac{r_i}{r} & \text{if } r > r_i \end{cases} \tag{1}$$

where r_i is the isolated rate of receiver i and r is the multicast group operating rate. Equation (1) can be written as:

$$F_i(r) = \frac{\min(r_i, r)}{\max(r_i, r)} \tag{2}$$

The function in Equation (1) attempts to capture a measure of "utility" for the receiver. When $r \leq r_i$, the receiver

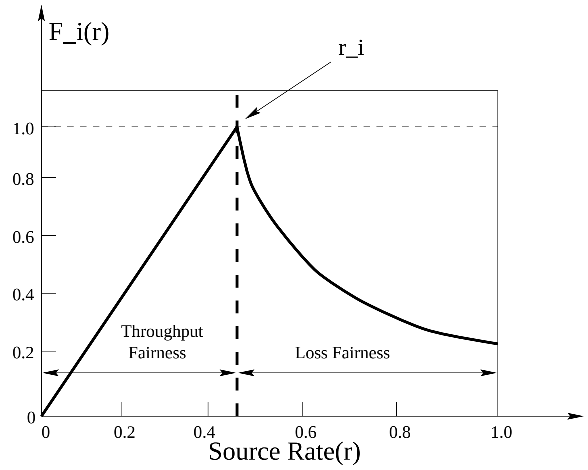

Figure 3: Example Receiver Fairness Function

utility is directly related to the ratio of its achieved throughput to its isolated rate. When $r > r_i$, the receiver utility decreases as it loses data in direct proportion to the ratio of received data.

Figure 3 shows the shape of this individual fairness function. The x axis is the value of actual operating rate and the y axis is the fairness value for the individual destination. The point r_i on the x axis is the optimal point for destination i, where i experiences no unfairness. For $r < r_i$, receiver i experiences unfairness because the rate is lower than the receiver can handle. For $r > r_i$, receiver i experiences unfairness due to losses.

Beyond the fairness function, receiver i is also allowed to specify an additional performance constraint, the *maximum acceptable loss tolerance*, ℓ_i, indicating the maximum fraction of transmitted data that can acceptably be lost. The loss tolerance restricts the operating rate r to be at most $x_i = r_i/(1-\ell_i)$. We call the range $[0, x_i]$ receiver i's *operational range* [1]. In general we allow a receiver's operational range to vary over time and to be receiver-specified. For a group of n receivers, a *group operational range* $[0, X]$ is defined where $X = \min_{1 \le i \le n} x_i$.

Note that for the special case when $\ell_i = 0$ for all i, we have $x_i = r_i$. This is exactly the original multicast ABR context [2] that requires the source to transmit at the minimum of the isolated rates of all receivers.

2.2 Inter-Receiver Fairness Function

We define the *inter-receiver fairness* achieved by the multicast group to be the weighted sum of the individual fairness values. That is,

$$IRF(r) = \sum_{i=1}^{n} \alpha_i F_i(r), \qquad (3)$$

$$subject\ to \quad \sum_{i=1}^{n} \alpha_i = 1 \quad and \quad 0 \le \alpha_i \le 1.0, \ \ i = 1, \ldots, n$$

[1] We assume that the minimum cell rate for the ABR connection is zero, thus the operational range has lower bound of 0.0.

[2] Since ATM Forum has not yet standardized how multicast ABR will work, we will refer to the multicast ABR mechanisms proposed in [6, 7, 8] as the original multicast ABR.

where n is the number of receivers, $F_i()$ is the fairness function for receiver i, and r is the transmission rate used by the source. Note that this definition allows receivers to contribute unequally to the total inter-receiver fairness, via unequal α_i's. These weights can be statically assigned or may vary over time.

2.3 The Optimal Operating Point

The optimal operating point problem can be stated as follows:

> Given the fairness function $F_i()$, maximum loss tolerance ℓ_i, and weight α_i for each receiver i, find the operating rate $r*$ (i.e., the rate at which the source should transmit) that maximizes the inter-receiver fairness function in Equation (3), subject to the constraints imposed by the ℓ_i's.

In the next section we present an approach that allows the multicast source to compute and optimize the Inter-Receiver Fairness function in order to determine a transmission rate.

2.4 A Global Network Measure

Improving inter-receiver fairness for a multicast ABR connection is typically accomplished by increasing its transmission rate beyond the bottleneck fair share for that connection. This typically increases the utilization of this connection on some links in the network, which in turn can result in a reduction of the throughput experienced by other connections in the network. In order to judge the effect of improved inter-receiver fairness for multicast connections on other ABR connections, we define a global network delivered throughput (GDT) measure as

$$GDT = \sum_{\text{all connections } c} \overline{\gamma_c} * (n_c)^y \qquad (4)$$

where $\overline{\gamma_c}$ is the average throughput delivered to receivers in connection c, n_c is the number of receivers in connection c and y is a constant between 0 and 1 that can be used to weight multicast connections relative to unicast connections.

For unicast connections, $\overline{\gamma_c}$ is the actual bottleneck fair share of the connection and $n_c = 1$. For multicast connections $\overline{\gamma_c} = (\sum_{\text{receivers } j} \gamma_j)/n_c$. When $y = 0$ the multicast connection counts as just a single connection with an average rate $\overline{\gamma_c}$. When $y = 1$ the multicast connection's delivered throughput to each receiver is counted in GDT. For $0 < y < 1$ we obtain other relative weightings of multicast connections.

It is important to emphasize that the GDT measure in Equation (4) gives a network-wide measure that allows us to judge the effect of increasing a multicast connection's rate on other connections. The global objective of ABR (and our own extensions to it) is to provide inter-connection fairness, not to optimize total delivered throughput.

3 Overview of Approach

We now give a brief overview of our approach to the design of an ABR congestion control algorithm that aims to maximize inter-receiver fairness. Our approach can best be viewed as an extension to existing multicast ABR techniques [6, 7, 8]. In these techniques, Resource Management (RM) cells are used by receivers to convey back to the source an explicit rate indicating their bandwidth fair share on the bottleneck of the path leading to them. Switches in the multicast tree aggregate the information in RM cells received from downstream[3] switches or receivers and forward this aggregated information to their upstream switch. One point of difference among the various multicast ABR algorithms is how this aggregation is performed. They all, however, aggregate the information about downstream paths into a single explicit rate representing the minimum rate stated in the RM cells they receive.

Recall that our goal is to determine an explicit rate at which the sender should transmit. In addition to the usual ABR criterion of allocating a rate that represents the connection's fair share on a bottleneck link, this rate should be chosen to maximize inter-receiver fairness as described in the previous section. We propose to achieve this by conveying to the source (in the RM cells) enough information for it to determine (either exactly or approximately) the Inter-Receiver Fairness (IRF) function shown in Equation (3). For this to be possible, the current approach employed by the switches of summarizing rate information into a single (minimum) rate is not adequate. Ideally, we would like to have a switch accumulate (without loss) all the rate information from all its downstream receivers in the RM cell it forwards upstream. This would allow the source to receive all the isolated rates needed to exactly compute the IRF function. In case where there are a large number of receivers, some aggregation of information is needed at the switches because of the limited size of the RM cell and also for scalability of the source computation of the optimum rate.

The source also needs the group operational range and receiver weights to completely compute the IRF function. The group operational range can be computed by each switch as it receives the individual receiver's operational range much in the same way as current proposed multicast ABR mechanisms [6, 7, 8] determine the minimum rate of all receivers. The receiver weights can be conveyed along with the isolated rate information for each receiver.

In the next section we describe our approach as it would be applied to ABR connections with a relatively small number of receivers. We then describe how it can be extended to handle the case with large number of receivers.

4 Maximizing Inter-Receiver Fairness for Small Groups

4.1 Delivering Receiver Information to the Source

We now describe an algorithm for conveying the appropriate information to the source in the context of an ABR connection with a small number of receivers [4]. The algorithm is an extension of Algorithm 4 in [8] and can be described as follows.

1. The source periodically sends Forward RM cells (FRM) that are multicast over the entire multicast ABR connection.

2. Switches receiving an FRM compute the connection's fair share on each of their outgoing links that belong to the multicast tree. An FRM is forwarded on each outgoing link with an indication of the minimum of the computed rate for that link and the rate indicated in the received FRM at that switch (see Figure 4). FRMs that reach receivers contain the receiver's fair share of bandwidth on the bottleneck link on the path from the source to the receiver. This is what we have dubbed the receiver's isolated rate.

3. The reception of an FRM triggers the generation and transmission of a Backward RM (BRM) cell at each receiver. A BRM at the receiver contains the bottleneck fair share rate that was received in the triggering FRM. It will also contain the receiver's operational range and weight.

4. A switch processes BRM cells it receives as shown in the algorithm in Figure 4. In the algorithm, each switch will receive BRM cells from all its directly attached downstream switches or hosts on the multicast connection's tree. The contents of each BRM cell are stored at the switch.

   A switch maintains a flag for each downstream link. These flags are set for each BRM cell the switch receives on that link. Once all the flags are set, a BRM is generated and sent upstream to the parent switch, and all the flags are reset. If a BRM is received from a downstream link while its flag is set, the contents of this BRM will overwrite the contents stored from the previous BRM received on the same link.

   The BRM generated by the switch (when all the flags are set) contains a concatenation of all the rate and weight information that this switch received in BRM cells received from its downstream links as well as an updated operational range.

5. The source will receive BRM cells (from the switch to which it attaches). Each BRM cell will contain the necessary information for the source to compute the IRF function. In the traditional ABR algorithm, the source will compute a new explicit rate for the connection for each BRM cell it receives. In our case, however, we would like the source to compute an explicit rate for the connection that would maximize inter-receiver fairness. In the next section we describe a low-complexity optimization technique that can be used to determine the optimum rate at which the source should operate.

4.2 Optimizing the IRF Function at the Source

Once the IRF information is available at the source, it is necessary to determine the point at which this function is maximized to determine the appropriate transmission rate at the source.

Note that the IRF function is a linear combination of the individual receiver fairness functions, each of which has a maximum (and a discontinuity) at the receiver's isolated rate. The IRF function will, therefore, also be discontinuous at the points r_i for $i = 1, \ldots, n$.

The following proposition makes the optimization task at the source very simple:

Proposition 1:

The IRF function as shown in Equation (3) achieves a maximum at some r_i, $i = 1, \ldots, n$.

[3] The *downstream* direction from a switch is towards the receivers and the *upstream* direction is towards the source.

[4] The number of receivers this exact algorithm can handle is limited by the size of an RM cell. RM cells typically have around 36 unused bytes [2]. In this space we need to encode each receiver's isolated rate and its weight as well as an accumulated operational range. If one assumes that encoding this information can be done in 3 bytes per receiver, then we can include information about as many as 12 receivers.

```
/********** Comments for Switch Algorithm ***********/
/*FwdER field in forward RM(FRM) cell records the    */
/* minimal fair share so far between source and current  */
/* switch when FRM is traversing forward.             */
/*IRF-list is a list of tuples (r_i, α_i), where r_i is the  */
/* isolated rate and α_i is the weight of ith receiver.   */
/*X is the maximum acceptable rate.                  */
/***************************************************/

  Upon the receipt of a forward RM(FwdER) cell:
       Calculate fair-share;
                 FwdER=min(FwdER, fair-share);
       Multicast RM cells containing FwdER to all
                 downstream tree branches;

  Upon the receipt of a BRM(IRF-list, X) cell;
       Set the corresponding flag for that branch;
       Store the RM cell's contents;
       If (all flags have been set) then
               IRF-list= concatenate all IRF-lists
                       stored;
               X = min of received X values;
               Forward RM(IRF-list, X) cell
                       to upstream node;
               Reset all flags;
       else
               Discard this RM cell;
```

Figure 4: Switch Algorithm

Proof: See Appendix.

The procedure for determining the optimum transmission rate is, therefore, as follows:

1. Determine the group operational range $[0, X]$ from the received BRM cell.
2. Evaluate the IRF function at 0, X and at all r_i, $i = 1, \ldots, n$ where $0 < r_i < X$ [5].
3. The optimum rate is the one that maximizes the IRF function from among the ones considered.

5 Maximizing Inter-Receiver Fairness for Large Groups

5.1 Aggregating feedback information

If a multicast group has a large number of receivers, it is not possible to accumulate the information necessary for the computation of the IRF function in a single RM cell. One possibility would be to accumulate the information in multiple BRM cells and have the source wait until it receives all the information it needs before it determines a new rate at which it should transmit. This approach may cause excessive delay in adjusting to network conditions since the source has to process many BRM cells for each rate adjustment. This can lead to problems in the convergence of the ABR algorithm. Also the technique will not scale very well as the number of receivers gets very large.

[5] We do not expect that the complexity of the optimization at the source is a problem. Suppose the source rate is 5 Mbps, which means roughly 10000 cells/sec. Assume the source sends out an RM cell every 30 cells. Then, the time between receiving consecutive BRM cells at the source (which is the time the source has to compute the optimal rate) is 30/10000=3 ms.

We propose an approach in which we aggregate information about many receivers with similar isolated rates. In this approach the range of possible rates from zero to the maximum link rate on the network is subdivided into intervals. Ultimately the information fed back to the source is the number of receivers whose rates fall within each interval and the average isolated rates of these receivers.

More formally, if we normalize the maximum link bandwidth in the network to 1.0, the bandwidth range 0 to 1.0 is divided into k intervals: $(0, s_1], (s_1, s_2], \ldots, (s_{k-1}, 1.0]$. We allow the intervals to be of unequal lengths.

For each interval, $j = 1, \ldots, k$, the source receives the following:

- $count_j$: The number of receivers in the multicast group whose isolated rate is within interval j (such receivers are said to *belong to interval* j),
- ave_j: The average isolated rate of receivers who belong to interval j,
- awt_j: The average weight for receivers that belong to interval j, and
- X: The upper bound on the rate for all receivers.

The IRF function is approximated at the source using this information as

$$approxIRF(r) = \sum_{j=1}^{k} count_j * awt_j * \frac{\min(ave_j, r)}{\max(ave_j, r)} \qquad (5)$$

In essence the approximation in Equation (5) assumes that each receiver that belongs to an interval is equivalent to a receiver whose isolated rate is the average of all receivers belonging to the interval. We address the accuracy of the approximation in Section 6.

To determine the transmission rate that would maximize inter-receiver fairness, we appeal to the following proposition:

Proposition 2:

The approximate IRF function as shown in Equation (5) achieves a maximum at some ave_j, $j = 1, \ldots, k$.

Proof: Same as Proposition 1 (see Appendix).

5.2 Switch Algorithm

The switch algorithm described for small groups needs to be modified to achieve the aggregation described above. The modification is in the way the isolated rates and weights are accumulated by the switch in the BRM cells. For small groups the switch simply concatenated this information. Here, whenever the switch is ready to generate a BRM cell to forward upstream, it calculates for each interval j, $count_j$, the number of (directly or indirectly connected) destinations whose isolated rates fall in interval j; ave_j, the average of the these rates and awt_j, the average weight of the receivers in this interval. This information is easy to calculate at a switch based on similar information included in incoming BRM cells. Based on similar assumptions as in the previous section, we estimate that a single BRM cell can include information on approximately 10-12 intervals.

We illustrate the operation of this algorithm by the very simple example shown in Figure 5. Suppose the range [0.0,

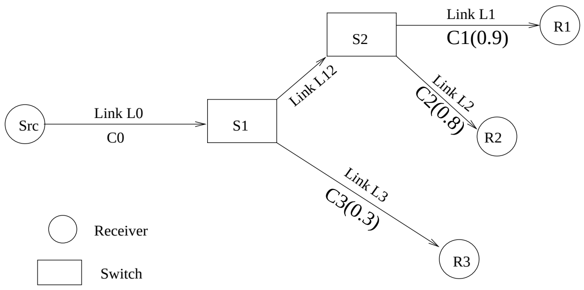

Figure 5: An approximate algorithm is required

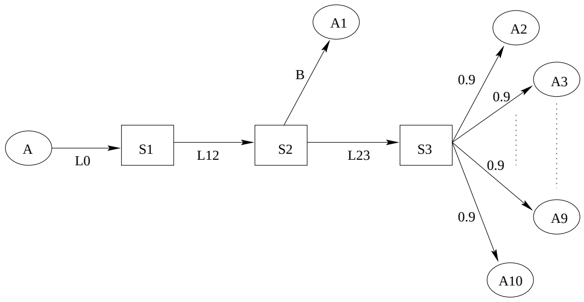

Figure 6: Small group configuration

1.0] is divided into 2 intervals, (0.0, 0.5] and (0.5, 1.0]. Then switch S2 shall generate a BRM cell which contains $count_1$=0, ave_1=0.0 and $count_2$=2, ave_2=0.85 (=(0.9+0.8)/2). After switch S1 receives BRM cells from both S2 and host R3, it can generate a BRM cell which contains $count_1$=1, ave_1=0.3 and $count_2$=2, ave_2=0.85 (=(0.9+0.8)/2). Finally, when this BRM cell arrives at the source, it can be interpreted as: there are 3 receivers in the multicast group, with two receivers having the same isolated rate (0.85) and one with isolated rate of 0.3.

6 Performance of Inter-Receiver Fairness Scheme

In this section, we examine the performance of our scheme under a variety of network topologies and group characteristics. We compare to the original multicast ABR in which the rate of a multicast group is determined by the bottleneck fair share on any link in the multicast tree. Unless otherwise stated, all configurations have the following properties:

- All connections are uni-directional. No traffic flows from the receiver(s) to the source, except for RM cells.
- Link bandwidth is normalized to be in the range $[0.0, 1.0]$. Unless otherwise stated, the link bandwidth is 1.0.

6.1 Small Group Results

We start by examining the performance for small groups. We use the topology shown in Figure 6, which has ten receivers. Receiver 1 is reached over a link with bandwidth capacity B, while Receivers 2–10 are reached over links with bandwidth capacity 0.9. All other links have bandwidth 1.0, thus $r_1 = B$ and $r_i = 0.9$ for $i = 2, \ldots, 10$. We assume that all receivers have the same loss tolerance. Initially, all receivers have the same weight, $\alpha_i = 0.1$.

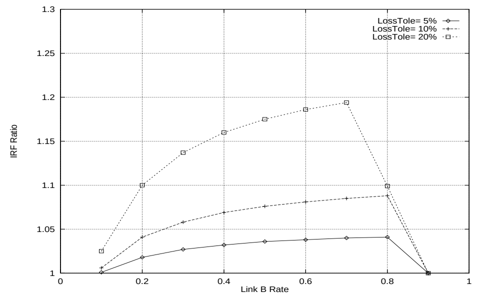

Figure 7: Small group: Ratio of IRF for equal weight

Figure 7 shows the ratio of the inter-receiver fairness under our scheme to the original multicast ABR as the bandwidth capacity B is varied. Each curve corresponds to a different loss tolerance. As expected, the improvement is greater when the loss tolerance is higher. Because of the configuration, the optimal operating rate is as high as the loss tolerance of Receiver 1 will allow. All of the curves reach a maximum when the capacity B is about 0.7 or 0.8. As the capacity B increases further, there is little possibility of improving on the fairness of the original multicast ABR scheme.

The IRF ratio decreases linearly with an increase in the weight α_1 associated with the slow receiver, when all other parameters are kept constant. This can be seen directly from the equation for inter-receiver fairness. Thus the fairness may be improved more significantly if the slowest receivers are less important than the faster receivers.

6.2 Large Group Results

For the large group experiments, we do not specify a particular topology for the network. Instead, we directly model the isolated rates, r_i, as coming from different distributions. In particular, we use the following three distributions:

Uniform The isolated rates are evenly distributed over the interval $[0.1, 0.9]$.

Gaussian The isolated rates come from a Gaussian distribution with mean 0.5 and variance 1.0. Values greater than 0.9 are rounded down to 0.9; values less than 0.1 are rounded up to 0.1.

Bi-modal The isolated rates come from two disjoint ranges, one towards the low end of the range and one towards the high end. Within a range, the values are uniformly distributed. The model has as a parameter the fraction of receivers in the low-end range.

For small groups, the source has complete information on the values of r_i and α_i needed to optimize the inter-receiver fairness. For large groups, this information is summarized by the switches, and the source must approximate the operating rate to optimize fairness. We begin by assessing the error introduced by the approximation.

We examine the accuracy of the approximation for a group containing 100 receivers. The error is measured by taking $(IRF(r*) - IRF(r))/IRF(r*)$, where $r*$ is the optimal operating point, and r is the operating point identified by the approximation algorithm. Figure 8 shows the error when the values of r_i are Uniform; Figure 9 shows the error when the values of r_i are Gaussian. The x-axis is the number of intervals used in the approximation.

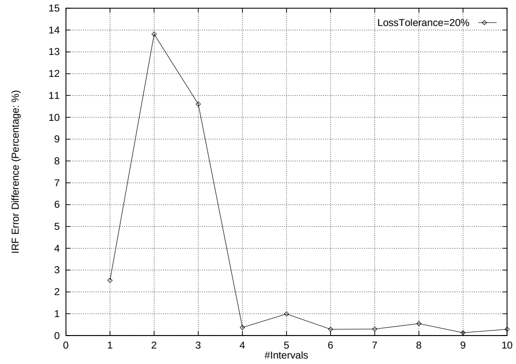

Figure 8: Error for Uniform distribution

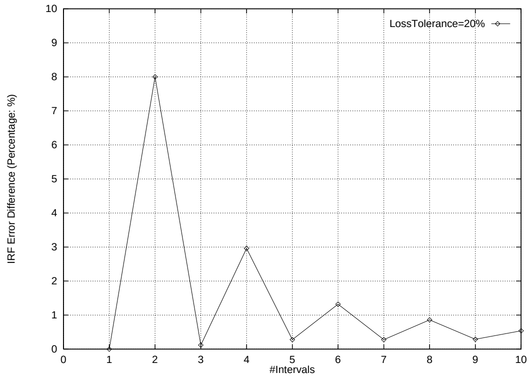

Figure 9: Error for Gaussian distribution

The figures have four salient features. First, as expected, the general tendency is for the error to decrease as the number of intervals increases. Second, the error does not strictly decrease as the number of intervals increases. This is a result of the aggregation of isolated rates into intervals; each interval tends to produce a local maximum in the IRF, with the largest value used to select an operating rate. For small numbers of intervals, the error can increase if the optimal operating point tends towards the boundary of an interval. Third, the error tends to be small, below 3% with four intervals, and below 1% as the number of intervals approaches 10. Finally, the error is larger under the Gaussian distribution than the uniform distribution. This is likely caused by the fact that the intervals are all of equal size, thus they "match" the uniform distribution well. The Gaussian distribution has most of its mass in a few intervals, increasing the error caused by summarizing the results into a single mean value for the interval.

Next we examine the improvement in inter-receiver fairness when using our method for a large group. In Figures 10 and 11 we show the distribution of individual fairness values with the original multicast ABR and our inter-receiver fair ABR respectively. There are 200 receivers, with isolated rates drawn from a Gaussian distribution, and loss tolerance of 20%. Ten intervals are used in the approximation. Note that our scheme increases the total fairness (observed as a shift in the distribution to the right) and also tends to make the individual fairness values more equal (observed as a flattening of the distribution). The ratio of inter-receiver fairness for our scheme as compared to the original multicast ABR is 1.24 for this configuration.

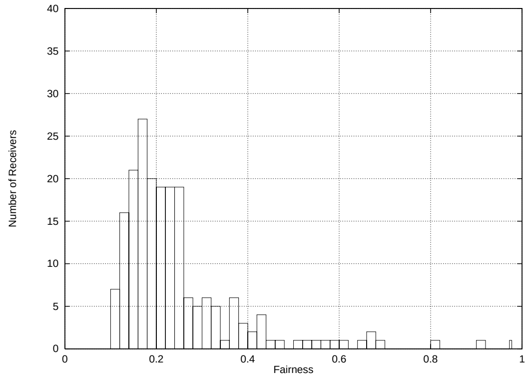

Figure 10: Original multicast ABR

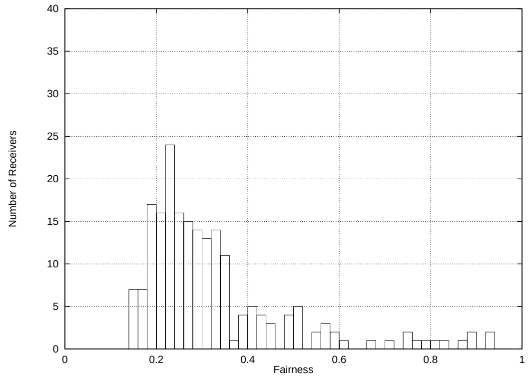

Figure 11: Inter-receiver fair multicast ABR

To get a better understanding of the behavior in a large group, we consider varying the number of receivers and the loss tolerance (still, ten intervals are used). Figure 12 shows the ratio of the inter-receiver fairness using our scheme to the inter-receiver fairness using the original multicast ABR. The number of receivers is increased along the x-axis; each curve corresponds to a different loss tolerance. The isolated rates come from a Gaussian distribution. We see that the majority of the benefit to our scheme comes about when there are approximately 70 receivers; after this point, the IRF ratio approaches a constant. As expected, the improvement is greater for a larger loss tolerance, approaching 25% for a loss tolerance of 20%.

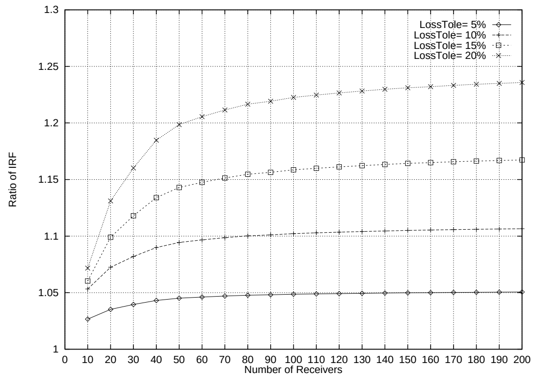

Figure 12: IRF ratio for Gaussian distribution

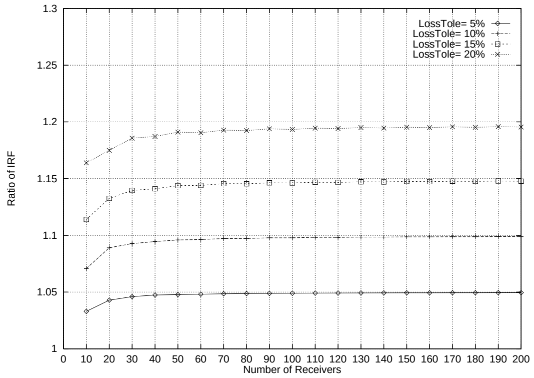

Figure 13: IRF ratio for Bi-modal distribution

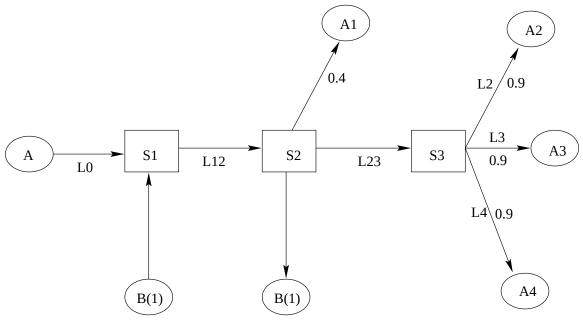

Figure 14: Topology with interacting connections

Figure 13 shows similar results for a Bi-Modal distribution with a low-end range of $[0.2, 0.3]$ and a high-end range of $[0.4, 0.5]$. 25% of the receivers are in the low-end and 75% are in the high end. The curves exhibit the same basic behavior as for the Gaussian distribution, however the majority of the benefit is evident with fewer receivers (approximately 50), and the improvement is slightly lower. If we shift the high-end range to $[0.8, 0.9]$ we see similar results, though the improvement is again slightly reduced.

Finally, we examine the interaction between different connections when our scheme is used. Figure 14 shows a topology with two connections, one multicast (connection A) and one unicast (connection B). These connections share one common link. If we use the original multicast ABR, connection A will operate at 0.4, leaving bandwidth 0.6 on the shared link for connection B. The global delivered throughput is thus 1.0 when $y = 0$ and 2.2 when $y = 1$. Using our scheme, with 20% loss tolerance, connection A will operate at 0.5, using its fair share of the bandwidth on the shared link. Connection B will also operate at 0.5. The global delivered throughput under this scheme decreases slightly to 0.98 when $y = 0$ and increases to 2.4 when $y = 1$. In general, if y is chosen to be closer to 1.0, then an improvement in inter-receiver fairness is likely to be accompanied by an improvement in global delivered throughput. The other key factor in improving the global performance measure —not illustrated in this example— is the presence of under-utilized links. Our scheme will often increase the utilization of these links to improve fairness, with the additional effect of increasing the global delivered throughput. This phenomenon will be illustrated in a slightly different context in the next section.

7 Improving Inter-Receiver Fairness via Destination Set Grouping

In the development so far we have focused on how to improve inter-receiver fairness by exploiting receivers' loss tolerance. We observe that if receivers have low or zero loss tolerance, improving inter-receiver fairness is not possible. In this section we briefly explore another possible technique, *destination set grouping (DSG)*, that can be used to improve inter-receiver fairness for multicast ABR connections. The DSG technique was first proposed in [9]. It has since been used to provide flow control for multicast sessions in [10, 11]. Our objective here is to introduce this idea in the context of multicast ABR connections and demonstrate its potential to improve inter-receiver fairness. An in-depth examination of its use is relegated to future research.

With the destination set grouping technique [9], the set of destinations in a single multicast session are split into disjoint subgroups. The source then carries independent conversations with each subgroup. By judicious splitting of the original destination set we can improve the inter-receiver fairness for the multicast ABR session. This improvement will be achieved at the expense of increased bandwidth on some links and perhaps a reduced allocation to other ABR connections sharing such links.

We demonstrate the use of DSG through two examples:

Example 1: Consider the network in Figure 15 where we assume that all link bandwidths are equal to one unit. Connection sources are the nodes in the lower part of the figure and connection destinations are in the upper part. There are nine separate ABR connections between the nodes labeled C, and two separate ABR connections between the nodes labeled B. A single multicast connection delivers cells from A to both A1 and A2. We assume that all receivers have zero loss tolerance and that the minimum cell rate for all connections is also zero. Moreover, all receivers in multicast connection have the same weight, $\alpha_i = 0.5$.

We consider two approaches to handling the multicast connection from A to A1 and A2. In the first (non-DSG), a single multicast VC carries the traffic and in the second (DSG) approach two VCs are used to carry the traffic from A one for A1 and the other for A2. The resulting throughput in the non-DSG approach are shown in the first row of Table 1. We also show the IRF value for the multicast connection (A) and the GDT value with $y = 1$.

With the DSG approach we impose the following additional constraint:

> The total bandwidth consumed on any link by the VCs resulting from splitting an original multicast VC cannot exceed the maximum isolated rate of a receiver.

This constraint is necessary to insure that a split connection does not result in excessive bandwidth demands.

For the example at hand, this results in a maximum single link bandwidth consumption of 0.33 for the multicast traffic from A. If this is carried over two separate VCs (one to A1 and another to A2), the resulting connection throughputs are shown in the second row of Table 1.

Note that splitting the multicast connection has resulted in an increase in inter-receiver fairness for this multicast session. This resulted in a slight decrease in our global delivered throughput (GDT) measure.

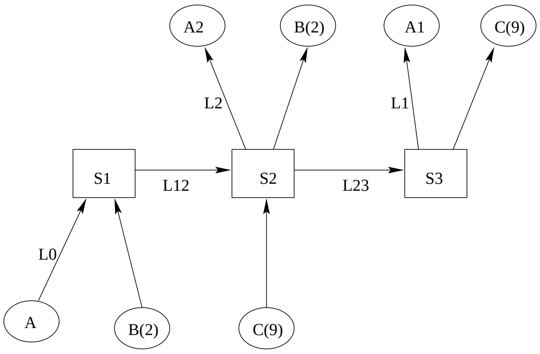

Figure 15: Example One

Table 1: For Simple Configuration

	A_1	A_2	B	C	IRF (A)	GDT
non-DSG	0.1	0.1	0.45	0.1	0.65	2.0
DSG	0.1	0.23	0.33	0.1	0.85	1.9

Example 2: With this example we consider the effect of the choice of subgrouping of a multicast session. The example is based on the network in Figure 16. The basic assumptions are carried over from Example 1 above. In this network there is 1 unicast connection between the B nodes, 3 unicast connections between the C nodes, and 8 unicast connections between the D nodes. Target link utilization is 99%. Moreover, all receivers in multicast connection have the same weight, $\alpha_i = 0.33$.

Note that for the multicast connection the isolated rates in this example are 0.89, 0.22, and 0.1 for A_1, A_2, A_3, respectively. The achieved connection rates (based on bottleneck fair share allocation and zero loss tolerance) are shown in the first row of Table 2 for the case where the multicast connection is handled by a single VC. We also show the IRF value for the multicast connection as well as the global delivered throughput in the network.

We consider next using DSG on the multicast connection and using one multicast VC to A_1 and A_2 and another for A_3. the resulting rates, IRF and GDT values are shown in the second row of Table 2. Note that using DSG has resulted in an increase in inter-receiver fairness for the multicast connection with no decrease in the GDT value. Using a separate VC for each receiver results in the rates shown in the third row of the table (we maintain the restriction on the total bandwidth usage of the split multicast connection as discussed above). Note that in this case we achieve an increase in the inter-receiver fairness as well as an increase in the GDT.

The fact that we can increase inter-receiver fairness without adverse effects on the global operation of the network derives from the fact that the link from S1 to S2 in our example was severely underutilized at 20% when the multicast connection was handled by a single VC. Using DSG increases the utilization to 99% in the case represented by the 3rd row in the Table 2.

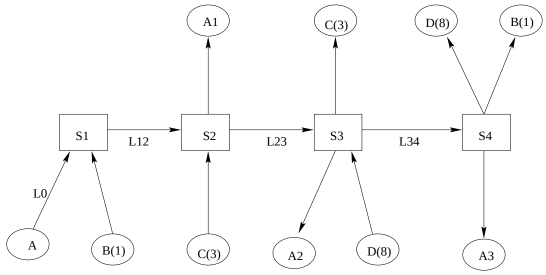

Figure 16: Example Two

8 Concluding Remarks

This work addresses the issue of the fair treatment of a multicast session's receivers when they have differing reception capabilities. Our work focuses on providing such fairness in the context of ATM's Available Bit Rate (ABR) service. We define the properties of an inter-receiver fairness function and propose one specific function with these properties. We propose ABR algorithms at the multicast sources, receivers and network switches that would allow the multicast source to compute this function. Determination of the rate to be used for the connection is made by the source through a simple optimization of the function.

Through the use of application and receiver-specific cell loss tolerance our approach yields a generalization of the original multicast ABR techniques. In particular if all receivers in a multicast session have no tolerance to cell loss then our schemes yield identical results to previous multicast ABR proposals. Increasing inter-receiver fairness for a multicast connection typically involves an increase in its transmission rate. This in turn can have an effect on other ABR connections (unicast and multicast) that are using the network. We develop a global network measure that captures the total delivered throughput in the network with parameters permitting a weighting of multicast connections by the number of receivers. Our experiments demonstrate that inter-receiver fairness can be achieved without severe degradation in this global measure. In fact, in some instances, the total network delivered traffic increases as a result of an increase of a multicast connection's inter-receiver fairness.

Finally we introduce the use of destination set grouping in the context of multicast ABR connections. We illustrate the potential for such an approach to improve inter-receiver fairness, even when receivers have no loss tolerance. Our future work will examine this technique in more detail. In particular we will examine the issue of determining optimal groupings and of developing protocols that will allow the splitting of a multicast connection's receivers in a dynamic fashion.

References

[1] C. Lefelhocz, B. Lyles, S. Shenker, and L. Zhang, "Congestion control for best-effort service: Why we need a new paradigm," *IEEE Networks*, pp. 10–19, Jan. 1996.

[2] "ATM Forum traffic management specification 4.0," 1996.

[3] J. M. Jaffe, "Bottleneck flow control," *IEEE Transactions on Communications*, vol. 29, pp. 954–962, July 1981.

Table 2: For Chain Configuration

	A_1	A_2	A_3	B	C	D	IRF (A)	GDT
non-DSG	0.1	0.1	0.1	0.1	0.27	0.1	0.52	2.0
$\{A_1, A_2\}, \{A_3\}$	0.2	0.2	0.1	0.1	0.2	0.1	0.71	2.0
$\{A_1\}, \{A_2\}, \{A_3\}$	0.59	0.2	0.1	0.1	0.2	0.1	0.85	2.39

[4] J. Nonnenmacher, E. W. Biersack, and D. Towsley, "Parity-based loss recovery for reliable multicast transmission," in *Proceedings of SIGCOMM '97*, 1997.

[5] H. Holbrook, S. Singhal, and D. Cheriton, "Log-based receiver reliable multicast for distributed interactive simulation," in *Proceedings of SIGCOMM '95*, 1995.

[6] 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, 1997.

[7] L. Roberts, "Rate based algorithm for point to multipoint ABR service," 1994. ATM Forum Contribution 94-0772.

[8] W. Ren, K. Siu, and H. Suzuki, "On the performance of congestion control algorithm for multicast ABR service in ATM," in *IEEE ATM '96 Workshop*, 1996.

[9] M. H. Ammar and L. R. Wu, "Improving the throughput of point-to-multipoint ARQ protocols through destination set splitting," in *IEEE INFOCOM'92*, pp. 262–271, IEEE Communications Society, 1992.

[10] S. Y. Cheung and M. H. Ammar, "Using destination set grouping to improve the performance of window-controlled multipoint connections," *Computer Communications*, vol. 19, pp. 723–736, 1996.

[11] S. Bhattacharyya, J. Kurose, D. Towsley, and R. Nagarajan, "Efficient rate-controlled bulk data transfer using multiple multicast groups," in *IEEE INFOCOM'98*, 1998.

Appendix

We prove that the optimal operating rate is achieved at one of the isolated receiver rates r_i. The reader will recall that this implies an efficient algorithm for determining the operating rate: the source need only evaluate the inter-receiver fairness for each potential operating point r_i, and select the best rate.

The proof requires that each individual fairness function meet the following conditions:

1. Only one point of discontinuity at $r = r_i$
2. Second derivative greater than or equal to 0 in the region $r < r_i$
3. Second derivative greater than or equal to 0 in the region $r > r_i$

These conditions are slightly more restrictive than the list given in Section 2. The conditions are met by the individual fairness function used for examples in the paper, as well as by other fairness functions.

Let n denote the number of receivers, and r_i denote the isolated rate for receiver i. Without loss of generality, assume that $r_i \le r_j$ if $i < j$ and that $\alpha_i = 1.0/n$, where n is the number of receivers in multicast group. For consistent treatment of the boundary case, define $r_{n+1} = 1.0$. We will show that in the range $[r_i, r_{i+1}]$, the second derivative of the inter-receiver fairness is greater than or equal to zero, thus the IRF function achieves a maximum at the endpoints of the range. A singular maximum is not possible in the interior of the range. This implies that the optimal operating rate is one of the r_i values.

The proof has two cases. In the first case, $r_i = r_{i+1}$, and thus we trivially find that the maximum must be at the (common) endpoints of the range. In the second case, $r_i < r_{i+1}$. In the range (r_i, r_{i+1}), we know that receivers $1, \ldots, i$ are experiencing loss, since their isolated rates are at most r_i. Receivers $i+1, \ldots, n$ are experiencing throughput degradation, since their isolated rates are at least r_{i+1}. The inter-receiver fairness is thus:

$$IRF(r) = \sum_{j=1}^{i} F_j(r) + \sum_{j=i+1}^{n} F_j(r),$$

where the first term is for receivers whose $r_j < r$ and the second term is for receivers whose $r_j > r$. The second derivative of $IRF(r)$ over the range $[r_i, r_{i+1}]$ is clearly greater than or equal to zero, since the second derivative of each term is greater than or equal to zero (see conditions 2 and 3 above).

Intuitively, the set of receivers is partitioned into those whose individual fairness is decreasing over the interval (r_i, r_{i+1}), namely receivers $1, \ldots, i$, and those whose individual fairness is increasing, namely receivers $i+1, \ldots, n$. Since the individual fairness functions obey conditions 2 and 3 above, the decrease in fairness is happening more slowly than the increase in fairness. It is straightforward to see that as a result, the inter-receiver fairness function between the points r_i and r_{i+1} can consist only of a (possibly empty) decreasing portion, followed by a (possibly empty) increasing portion. Thus, one of the endpoints must be a maximum.