The Impact of Multicast Layering on Network Fairness *

Dan Rubenstein Jim Kurose Don Towsley
Department of Computer Science
University of Massachusetts at Amherst
http://www-net.cs.umass.edu/{~drubenst, ~kurose, ~towsley}

Abstract

Many definitions of fairness for multicast networks assume that sessions are single-rate, requiring that each multicast session transmits data to all of its receivers at the same rate. These definitions do not account for multi-rate approaches, such as layering, that permit receiving rates within a session to be chosen independently. We identify four desirable fairness properties for multicast networks, derived from properties that hold within the max-min fair allocations of unicast networks. We extend the definition of multicast max-min fairness to networks that contain multi-rate sessions, and show that all four fairness properties hold in a multi-rate max-min fair allocation, but need not hold in a single-rate max-min fair allocation. We then show that multi-rate max-min fair rate allocations can be achieved via intra-session coordinated joins and leaves of multicast groups. However, in the absence of coordination, the resulting max-min fair rate allocation uses link bandwidth inefficiently, and does not exhibit some of the desirable fairness properties. We evaluate this inefficiency for several layered multi-rate congestion control schemes, and find that, in a protocol where the sender coordinates joins, this inefficiency has minimal impact on desirable fairness properties. Our results indicate that sender-coordinated layered protocols show promise for achieving desirable fairness properties for allocations in large-scale multicast networks.

1 Introduction

The current Internet has few internal mechanisms to regulate the rates at which sessions should transmit data. How to achieve *fairness* within such a network, in effect allowing sessions to share bandwidth in a manner that satisfies some set of network utilization criteria, remains a challenging research problem. The problem is further complicated in networks that support both unicast and multicast delivery services. Current definitions of multicast fairness [3, 6, 13, 18, 20] typically assume that sessions are *single-rate*, requiring all receivers within a multicast session to receive data at a uniform rate. However, *layered multicast* permits *multi-rate* transmission: different receivers within a session can receive data at different rates. This is accomplished by layering data among several multicast groups and allowing each receiver to independently determine the subset of layers (i.e., multicast groups) it joins. Protocols have used a layered approach to support multicast applications ranging from live multimedia [1, 8, 9, 11] to reliable data transfer [4, 14, 19]. These protocols have the appealing property that the transmission rate to each receiver is constrained only by the bandwidth availability on the receiver's own data-path from the data source, and is not limited by other receivers' rate limitations in the same session. The fairness literature does suggest intuitions about how layering might increase the set of desirable fairness properties that hold for a particular fair allocation of receiver rates. What is lacking, however, is a formal study that examines the impact that layering has on fair allocations within a large-scale multicast network.

The goal of this paper is to contribute to the formal understanding of how layering impacts fairness in multicast networks. In particular, we focus on how layering affects properties of multicast max-min fairness in an environment in which each session has a single sender. We have chosen to use max-min fairness as our fairness measure since its formal definition is a well-accepted criterion for fairness, allowing us to proceed directly to an examination of the properties of a fair allocation. We believe that with other definitions of fairness, layered approaches will yield similar fairness advantages, and expect this work to stimulate interest in examining the impact of layering for these other definitions.

Our examination begins with a theoretical and idealized model of multi-rate sessions that does not account for the current practical limitations of layered approaches (e.g., limited number of multicast groups, pre-determined rates of layers) to achieve multi-rate max-min fairness. Hence, the results based on this model demonstrate the *potential* fairness benefits that can be gained through the use of multi-rate sessions. We show that in such networks, allowing multicast sessions to be multi-rate instead of single-rate "improves" max-min fairness within a network. We do this by identifying four desirable fairness properties of a max-min fair allocation of unicast sessions. One simple example of a property is that receiver rates should be equal for two receivers whose data transmission paths from their respective senders traverse an identical set of links. We examine mul-

*This material was supported in part by the National Science Foundation under Grant No. NCR-9508274, NCR-9527163, CDA-9502639, and by DARPA under Grant No. N66001-97-C-8513. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

ticast max-min fair allocations under the definition given by Tzeng and Siu [18], that requires that all sessions are single-rate, and find that several of these fairness properties do not necessarily hold within the max-min fair allocation (the two receiver rate example presented above is one such property). We extend the multicast max-min fair definition to permit multi-rate sessions, and formally prove that, when all sessions in a network are multi-rate, all of our identified fairness properties hold for the max-min fair allocation. We also consider networks in which not all sessions are multi-rate (e.g., a session may have an application-specific requirement that requires it to be single-rate), and examine the effect on fairness properties of the max-min fair allocation as single-rate sessions are "replaced" by identical multi-rate sessions (i.e., same session members, same topology). We demonstrate, using our identified set of fairness properties and a mathematical ordering relation of allocations that indicates an allocation's "level" of max-min fairness, that increasing the set of "replaced" sessions results in an increase in the "level" of max-min fairness and that more fairness properties hold for max-min fair allocations.

Next, we examine the impact of some current practical limitations of layering on the fairness properties of multi-rate sessions. We show that if each receiver's fair rate is restricted to what can be obtained by joining some fixed set of layers, a max-min fair allocation need not even exist. However, we do demonstrate that receivers can achieve an *average* rate that matches their fair rate by using precisely timed joins and leaves. These joins and leaves must be tightly coordinated among receivers in the same session (i.e., correlating their sets of received packets) in order to prevent excess bandwidth utilization on a shared link. To quantify bandwidth usage, we introduce the notion of *redundancy*, a ratio of bandwidth used in practice by a session on a shared link to the theoretical lower bound needed on that link to deliver fair rates to downstream receivers. While several works have indirectly identified the negative implications of redundancy, a redundancy measure has never been formally defined, and its effect on fair allocations within a network has never been studied directly. We show that increased redundancy leads to a decrease in the "level" of max-min fairness, to a decrease in the number of fairness properties that hold for the max-min fair allocation, and, usually, to a decrease in receivers' fair rates. We examine how the ideas in [8, 11, 19], that coordinate joins of receivers within a session, significantly reduce the negative effects of redundancy. The examination is performed via analytical modeling and simulation of max-min fair congestion control protocols in which receivers join and leave layers based on congestion observations. Within the model, we present three protocols that differ in the degree to which the layer joins are coordinated among session receivers. We find that although redundancy is still not optimal, coordinated joins reduce redundancy most significantly when the correlation in loss among receivers is high, and that a protocol with sender coordination keeps redundancy at low enough levels to allow layered multicast to achieve non-bandwidth-wasteful fairness within a multi-rate multicast network.

This work makes two fundamental contributions to network protocol design. First, it formally demonstrates the theoretical benefits in terms of fairness of using multi-rate (i.e., layered) sessions, and that these benefits also exist in networks that support a mix of multi-rate and single-rate sessions. Second, we formally identify and define the redundancy of layered protocols, and demonstrate the drawbacks (in terms of fairness and efficiency of using available bandwidth) of having high redundancy. This suggests that future research geared toward improving layered protocol performance for multicast (e.g., new layering or routing protocols) should aim to keep the level of redundancy low.

The paper proceeds as follows. Section 2 presents theoretical results for multicast max-min fairness with multi-rate sessions. Section 3 introduces the notion of redundancy, and Section 4 examines the effects of join coordination in several simple congestion control protocols. Related and future work are presented in Section 5, and we conclude in Section 6.

Table 1: Variables used in the network model

N	a network, $(G, \{S_1, \cdots, S_m\}, \tau, \sigma)$
G	A network graph with n links.
$l_j, 1 \leq j \leq n$	The jth link of N
$c_j, 1 \leq i \leq n$	The capacity of link l_j
$S_i, 1 \leq i \leq m$	The ith session in N
σ	a mapping onto each session S_i that indicates the session's type ($\mathcal{M}$ = multi-rate or $\mathcal{S}$ = single-rate)
$r_{i,k}$	The kth receiver in session S_i
X_i	the single sender for session S_i
τ	A topology mapping that maps session members onto network nodes.
α_i	The maximum desired rate for session S_i, $0 < \alpha_i \leq \infty$
$R_{i,j}$	The set of receivers in S_i whose data-path traverses l_j
R_j	The set of receivers over all sessions whose data-path traverses l_j
$a_{i,k}$	The data rate for transmission to receiver $r_{i,k}$
r_i	The receiver in a unicast session S_i
a_i	The data rate in a unicast or single-rate session
$u_{i,j}$	The link rate for session S_i on link l_j
u_j	The link rate for link l_j (i.e., $\sum_i u_{i,j}$)
Defined in Section 3:	
ρ	The aggregate rate of the "single-layer"
v_i	A more general session link rate function

2 Multi-rate Multicast Max-Min Fairness

In this section, we present the formal network model used to examine the max-min fairness of multicast sessions, and identify a set of desirable fairness properties of max-min fair allocations for unicast networks. We then show that in this network model, the max-min fair allocation will always achieve all of these desirable properties only if the sessions are multi-rate. For the reader's convenience, a list of all the variables used is provided in Table 1.

A session, S_i, is a tuple $(X_i, \{r_{i,1}, \cdots, r_{i,k_i}\})$ of session members: X_i is the session sender that transmits data within a network; each $r_{i,k}$ is a receiver that receives data from X_i. Each session contains exactly one sender and at least one receiver. We write $r_{i,s} \in S_i$ to indicate that receiver $r_{i,s}$ is a member of session S_i.[1] We consider two types of sessions:

- If S_i is a *single-rate* session, then data must be transmitted to all receivers in S_i at the same rate.

[1] We assume that each receiver is a member of a single session. A receiver that is a member of two sessions can simply be viewed as two distinct receivers.

- If S_i is a *multi-rate* session, then receivers within S_i can receive data at multiple (arbitrary) rates.

A *network graph*, G, consists of a set of nodes connected together by n links in some arbitrary fashion. The links are labeled $l_1, \cdots, l_n$. Each link l_i has a *capacity*, c_i, that limits the aggregate rate of flow it can transmit in either direction between the two nodes it connects.[2] We define a network, $N = (G, \{S_1, \cdots, S_m\}, \tau, \sigma)$ to be a tuple containing a network graph, G, a set of sessions, $\{S_1, \cdots, S_m\}$, a mapping, τ, that maps each member of each session to a node in the network graph, and a second mapping, σ, that maps each session S_i to its type. We write $\sigma(S_i) = \mathcal{S}$ to indicate that session S_i is single-rate, and $\sigma(S_i) = \mathcal{M}$ to indicate that session S_i is multi-rate.

The mapping, τ, of a session onto the network graph has one restriction: no two members of a single session are mapped to the same node. However, there is no restriction that forbids two members of different sessions to be mapped to the same node. The network employs a routing algorithm, such that for each receiver $r_{i,k} \in S_i$, there is a sequence of links $(l_{j_1}, \cdots, l_{j_s})$ that carries data from X_i to $r_{i,k}$. We refer to this set of links in this sequence as the receiver's *data-path*. The data-path for a session is defined to be the set of all links that carry data to any receiver within the session.

For a network N, we define $R_{i,j}$ to be the set of receivers in session S_i whose data-path includes link l_j, and define R_j to be the set of all receivers whose data-path includes link l_j, i.e., $R_j = \cup_i R_{i,j}$. We account for the fact that session S_i might choose a maximum rate, α_i, at which it will transmit data (α_i can be infinite). An *allocation* is an assignment of receiver rates within a network. Once an allocation has been determined, we use $a_{i,k}$ to represent the rate at which data is transmitted to receiver $r_{i,k}$ (that equals the rate at which the data is received by $r_{i,k}$, barring loss). We let $u_{i,j}$ represent an absolute measure of bandwidth (e.g., in bytes/sec) used by session S_i on link l_j to transmit data to its receivers, and u_j the amount of bandwidth used by all sessions across link j, $u_j = \sum_{i=1}^{m} u_{i,j}$. We refer to $u_{i,j}$ as the *session link rate* of l_j for session S_i, and u_j simply as the *link rate* of l_j. Since bandwidth for each flow is non-negative, we have $0 \leq u_{i,j} \leq u_j$. We say a link is *fully utilized* if the total bandwidth used by all sessions across the link matches its capacity, i.e., l_j is fully utilized iff $u_j = c_j$.

We require that $u_{i,j} \geq a_{i,k}$ whenever $r_{i,k} \in R_j$, i.e., any bandwidth received by a receiver must traverse its data-path. In this section, we make an additional assumption that $u_{i,j} = \max\{a_{i,k} : r_{i,k} \in R_{i,j}\}$, which is the minimum value for $u_{i,j}$ that satisfies the above requirement. The reader that is familiar with layered approaches should see that if there is no restriction on the number of layers that a session can use, such a session link rate is easily achieved using a layered approach. In later sections, we examine the implications if $u_{i,j}$ is larger than this value. The assumption also allows us to model a unicast session as either a multi-rate session with a single receiver, or as a single-rate session with a single receiver. Thus, any results given in this section for networks containing a mix of single-rate and multi-rate sessions also holds for networks that contain a mix of single-rate, multi-rate, and unicast sessions.

An allocation is *feasible* if each receiver $r_{i,k}$ is assigned a rate $0 \leq a_{i,k} \leq \alpha_i$, and all receivers can receive at these rates without overutilizing any link's capacity in the network, i.e., $\forall i, k, 0 \leq a_{i,k} \leq \alpha_i$, and $\forall j, u_j \leq c_j$ (Hence, in this section, we require $u_j = \sum_i u_{i,j} = \sum_i \max_{\{a_{i,k}: r_{i,k} \in R_{i,j}\}} a_{i,k} \leq c_j$). The additional requirement imposed on each single-rate session S_i that all of its receivers' rates must be equal means that for any pair of receivers, $r_{i,k}, r_{i,k'} \in S_i$, when $\sigma(S_i) = \mathcal{S}$, then $a_{i,k} = a_{i,k'}$. When S_i is a single-rate session, or a session of either type containing a single receiver (i.e., a unicast session), we can write the single rate at which all receivers within the session receive data simply as a_i. We stress that in this section, receiver rates in multi-rate sessions are not constrained by practical limitations of layering. In effect, one can assume that a multi-rate session has at its disposal an unlimited supply of multicast groups, and can configure the rates on the layers to the exact needs and desires of its receivers.

Note that the feasibility of a particular allocation of receiver rates is a function of the link capacities of the network graph, G, the mapping τ, and also of the mapping σ. The dependence of an allocation's feasibility on σ is important: we will be examining how varying σ (i.e., varying sessions' types between single-rate and multi-rate) affects which allocation within a network is max-min fair.

Definition 1 (Max-min Fairness) *An allocation of receiver rates is said to be* **max-min fair** *if it is feasible, and for any alternative feasible allocation of rates (where for each receiver $r_{i,k}$ we define $\bar{a}_{i,k}$ as an alternative feasible rate) where $\bar{a}_{i,k} > a_{i,k}$, there is some other receiver $r_{i',k'} \neq r_{i,k}$ such that $a_{i,k} \geq a_{i',k'} > \bar{a}_{i',k'}$.*

In other words, if any receiver $r_{i,k}$'s rate is increased beyond its max-min fair rate to obtain some other feasible allocation, then there is some other receiver whose max-min fair rate is no larger than that of $r_{i,k}$, and whose adjusted rate (to account for the increase in $r_{i,k}$'s rate) must be decreased.

When all sessions within N are single-rate (i.e., $1 \leq i \leq m, \sigma(S_i) = \mathcal{S}$), we say that N is a single-rate network, and the max-min fair allocation is called the single-rate max-min fair allocation. A similar naming convention holds when all sessions are multi-rate. The definition of max-min fairness in [18] holds only for single-rate networks,[3] and involves a comparison of session rates rather than of receiver rates as in our definition. It is easy to show that the max-min fair allocation in a single-rate network is identical under both definitions. In a network that contains multi-rate sessions, their definition is not well defined.

Just as there is always one and only one unicast max-min fair allocation [2] and one and only one single-rate max-min fair allocation [18], there is one and only one multi-rate max-min fair allocation. In fact, for any choice of σ, the network has one and only one max-min fair allocation. We show the existence of a max-min fair allocation for a network with an arbitrary σ by constructing an algorithm that achieves a max-min fair allocation for that network. The algorithm that yields the max-min fair allocation is given in Appendix A. The proof of its max-min fairness appears in [15] (Lemma 5, Appendix B). The uniqueness of the max-min fair allocation is given by Corollary 5 in Appendix C of [15].

2.1 Fairness Properties

Let us first examine some desirable properties of a unicast max-min fair allocation, i.e., a max-min fair allocation in

[2] Assigning capacity per direction is a simple extension: simply extend a bidirectional link into two unidirectional links.

[3] [18] also permits a multicast session to consist of distinct unicast connections. We model this inherently via separate unicast sessions. Such a session differs significantly from a multi-rate session achieved through layering.

a network where all sessions are unicast. It is well known that the following properties hold for a unicast max-min fair allocation [2].

Unicast Fairness Property 1 (Unicast Max-min Fairness) *For each session S_i, $1 \leq i \leq m$, either $a_i = \alpha_i$, or else there is at least one fully utilized link, l_j, where for all $1 \leq i' \leq m, 0 < u_{i',j} \leq u_{i,j}$ (or, equivalently for the unicast case, $a_{i'} \leq a_i$ whenever $r_{i'} \in R_j$).*

Unicast Fairness Property 2 (Unicast Same Path Receiver Fairness) *If two unicast sessions, S_i and $S_{i'}$, within a unicast network have identical data-paths, then either $a_i = \alpha_i < a_{i'}$, or $a_{i'} = \alpha_{i'} < a_i$, or $a_i = a_{i'}$.*

Let us consider what makes these fairness properties desirable. To do this, we consider two *perspectives* of fairness of an allocation. From a receiver perspective, an allocation should be fair to receiver rates: a receiver's rate should be as large as possible without "stealing" bandwidth from receivers with lower rates. This is guaranteed by Unicast Property 1: there is no unused available bandwidth since some link on the receiver's data-path is fully utilized. Also, there is a fully utilized link over which the receiver receives at as high a rate as any other receiver whose data-path crosses the link. Increasing this receiver's rate further would result in "stealing" bandwidth from these other receivers sharing the link. From a session perspective, a link's capacity should be used "fairly" by sessions. In other words, a session's allocation on a link should be as large as possible without "stealing" bandwidth from other sessions that utilize the link.

For a unicast network, the receiver and session perspectives are identical because a session's data-path is identical to its receiver's data-path, and the share of bandwidth used on each link by the session equals the receiving rate of its receiver. This is not always true in a multicast network: a receiver's data-path is only part of the session's data-path, and, in a multi-rate session, when two receivers within the session receive at different rates, there is some pair of links that have differing session link rates for that session. Hence, an allocation might be "fair" from the session perspective without being "fair" from the receiver perspective, or vice versa. One possibility is to only consider fairness properties from a single perspective (e.g., [18] considers only the session perspective). However, in this section we will assume that it is more desirable to satisfy fairness properties from both perspectives. We extend the properties of a unicast max-min fair allocation, as described in Unicast Properties 1 and 2, to multicast networks from both a session and receiver perspective.

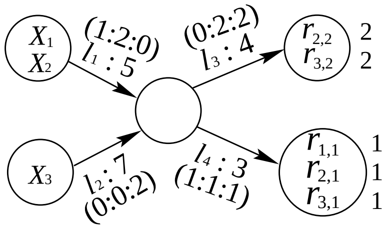


Figure 1: A sample network

Before presenting the desirable fairness properties for multicast networks, we introduce an example network that we will use to illustrate these different properties. Figure 1 presents a simple network with three sessions; sender X_1 in session S_1 sends to a single receiver, $r_{1,1}$. In session S_2, sender X_2 sends to two receivers $r_{2,1}$ and $r_{2,2}$. In session S_3, sender X_3 sends to two receivers, $r_{3,1}$ and $r_{3,2}$. The receiving rate of a receiver, $a_{i,k}$, is indicated to the immediate right of the receiver. Each link l_j has its capacity indicated next to the link labeling, separated by a colon (e.g., $l_1 : 5$ means that $c_1 = 5$). Adjacent to the link labeling for each l_j are the session link rates, appearing in the form, $(u_{1,j} : u_{2,j} : u_{3,j})$.

Fairness Property 1 (Fully-Utilized-Receiver-Fairness) *A receiver's rate $a_{i,k}$ is fully-utilized-receiver-fair if either $a_{i,k} = \alpha_i$, or there is at least one fully utilized link, l_j, where $r_{i,k} \in R_{i,j}$ and $a_{i',k'} \leq a_{i,k}$ for all receivers $a_{i',k'} \in R_j$. A session's allocation is defined to be fully-utilized-receiver-fair if the rate for each receiver in the session is fully-utilized-receiver-fair. An allocation of rates throughout the network is fully-utilized-receiver-fair if each session is fully-utilized-receiver-fair.*

Fully-utilized-receiver-fairness is the multicast extension of Unicast Property 1's prevention of "stealing bandwidth" from other receivers. For instance, in Figure 1, link l_3 is fully utilized and lies on receiver $r_{2,2}$'s data-path. Because $r_{2,2}$ receives at a rate that is no less than any other receiver whose data-path traverses l_3, its rate is fully-utilized-receiver-fair. Because all other receivers' rates in S_2 are fully-utilized-receiver-fair, session S_2's allocation of rates is fully-utilized-receiver-fair. Because S_1's and S_3's allocations are also fully-utilized-receiver-fair, the allocation (of rates for the entire network) is fully-utilized-receiver-fair.

Fairness Property 2 (Same-Path-Receiver-Fairness) *A pair of receivers $r_{i,k}$ and $r_{i',k'}$ are same-path-receiver-fair if their data-paths traverse the same set of links ($r_{i,k} \in R_j \iff r_{i',k'} \in R_j$), and either one receiver's rate is constrained by its session's maximum desired rate (i.e., either $a_{i,k} = \alpha_i < a_{i',k'}$ or $a_{i',k'} = \alpha_{i'} < a_{i,k}$), or else $a_{i,k} = a_{i',k'}$.*

Same-path-receiver-fairness states that if two receivers' data-paths traverse identical links, then the receivers should receive at identical rates (unless a receiver's rate reaches its application's maximum desired rate, α_i or $\alpha_{i'}$). In Figure 1, receivers $r_{1,1}$ and $r_{2,1}$ are a pair of receivers whose rates are same-path-receiver-fair. The reader should note that same-path-receiver-fairness is also a property of TCP-fairness [10]. If S_1 is a unicast TCP session, then, in order for $r_{2,1}$'s rate to be TCP-fair, same-path-receiver-fairness must hold for these two receivers.

Fairness Property 3 (Per-Receiver-Link-Fairness) *A session S_i's allocation is per-receiver-link-fair if for each receiver $r_{i,k} \in S_i$, either 1) $a_{i,k} = \alpha_i$, or 2) there is a link l_j that is fully utilized ($\exists j, r_{i,k} \in R_j, u_j = c_j$), and for other sessions $S_{i'}, u_{i',j} \leq u_{i,j}$. An allocation of rates throughout the network is per-receiver-link-fair if each session's allocation is per-receiver-link-fair.*

Fairness Property 4 (Per-Session-Link-Fairness) *An allocation is per-session-link-fair for a session S_i if $a_{i,k} = \alpha_i$ for each receiver in S_i or there exists a fully utilized link l_j in S_i's data-path where for other sessions $S_{i'}, u_{i',j} \leq u_{i,j}$. A allocation of rates throughout the network is per-session-link-fair if each session's allocation is per-session-link-fair.*

Per-receiver-link-fairness requires that session S_i gets a "fair share" of link rate along every path from sender X_i to its receivers. Per-session-link-fairness is a weaker version of this: a session must get a "fair share" of link rate on at least one link in its data-path (i.e., along the data-path of at least one receiver). In Figure 1, session S_2 is per-session-link-fair: on the data-path to receiver $r_{2,2}$, link l_3 is fully utilized and session S_2's link rate on l_3 is no less than the link rates of other sessions on l_3. It is also per-receiver-link-fair, because similar conditions hold on the data-path of its other receiver, $r_{2,1}$. Sessions S_1 and S_3 are also both per-receiver-link-fair and per-session-link-fair, making the network allocation both per-receiver-link-fair and per-session-link-fair.

2.2 Multi-rate Session Impact on Fairness Properties

It is fairly easy to see that in a unicast network, Fairness Property 2 and Unicast Property 2 are identical, and the remaining multicast fairness properties are identical to Unicast Property 1. We now proceed to establish properties of max-min fair allocations in terms of the types of sessions (multi-rate or single-rate) within the network. Due to lack of space, we omit proofs of any theorems, lemmas, and corollaries. All proofs appear in [15].

Theorem 1 *A multi-rate max-min fair allocation satisfies the Fairness Properties 1, 2, 3, and 4. In other words, the multi-rate max-min fair allocation is fully-utilized-receiver-fair, same-path-receiver-fair, per-receiver-link-fair, and per-session-link-fair.*

Theorem 1 tells us that if all sessions are multi-rate, then the max-min fair allocation satisfies all of our desired fairness properties. We now introduce a mathematical ordering among allocations that allows us to comparatively examine the "max-min fairness" of an allocation within a network:

Definition 2 *We say a vector $(x_1, x_2, \cdots, x_k)$ is* ordered *if for all $i, 1 \le i < k, x_i \le x_{i+1}$. Let $X = (x_1, x_2, \cdots x_k)$ and $Y = (y_1, y_2, \cdots y_k)$ be ordered vectors. We write $X \le_m Y$ (and say X is min-unfavorable to Y) if no i exists such that $x_i > y_i$, or for any i where $x_i > y_i$, there is some $j < i$ where $x_j < y_j$. We write $X <_m Y$ to indicate $(X \le_m Y) \wedge (X \ne Y)$.*

Note that under the above definition, $\le_m$ is reflexive ($X \le_m X$), non-symmetric ($X = Y \iff X \le_m Y \wedge Y \le_m X$), and transitive ($W \le_m X \wedge X \le_m Y \Rightarrow W \le_m Y$). Furthermore, for any pair, X and Y, of ordered vectors of identical length, either $X \le_m Y$ holds, or $Y \le_m X$ holds, or both. Min-unfavorability is similar to alphabetizing two text strings of the same length. Let x_i represent the ith character of the first string, and y_i represent the ith character of the second string. Then $X \le_m Y$ if and only if $X = Y$ or an alphabetization places X before Y. A more general version of this ordering has been applied specifically within unicast networks [5]. Let us now see how this ordering relation relates to multicast max-min fairness:

Lemma 1 *Let $A = (a_1, \cdots, a_s)$ be the ordered vector of receiver rates in a max-min fair allocation in a network $N = (G, \{S_1, \cdots, S_m\}, \tau, \sigma)$, and let $B = (b_1, \cdots, b_s)$ be the ordered vector of receiver rates for some other feasible allocation in N. Then $B <_m A$.*

Note that the network N in Lemma 1 can have any arbitrary session type mapping, σ (i.e., some sessions can be multi-rate, while others are single-rate). However, σ must be fixed when applying the lemma. Lemma 1 along with the definition of min-unfavorability can be combined to show that the max-min fair allocation maximizes the minimum rates in a network: since all allocations are min-unfavorable to the max-min fair allocation, there exists a threshold rate x' such that for any rate $z < x'$, the number of receivers that receive at or below z is minimal (smaller or equal) within the max-min fair allocation. Furthermore, the number of receivers that receive at or below x' is minimized (strictly smaller) within the max-min fair allocation. This result can be stated more formally as a general property of min-unfavorability:

Lemma 2 *$X <_m Y \iff \exists x'$ such that $\forall z < x'$, $|\{x_i \in X : x_i \le z\}| \ge |\{y_i \in Y : y_i \le z\}|$ and $|\{x_i \in X : x_i \le x'\}| > |\{y_i \in Y : y_i \le x'\}|$.*

Because the min-unfavorable relation is transitive, it gives a strict ordering among the feasible allocations for a network, where the max-min fair allocation is the maximum under the ordering. Thus, one can quantitatively compare the max-min fairness of two allocations A and B, where $A \le_m B$ means that B is "more max-min fair" than A, and the minimum receiver rates are larger in B than in A.[4]

2.3 Fairness limitations of single-rate sessions

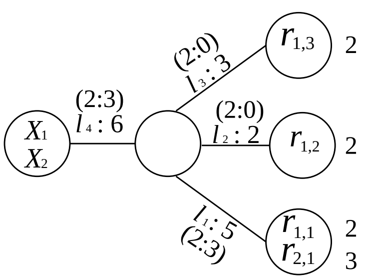


Figure 2: An example where a single-rate session would fail all but one of the fairness properties.

Theorem 1 states that a multi-rate max-min fair allocation satisfies our four desirable fairness properties. Let us now see where a single-rate max-min fair allocation fails to do so. The fact that a single-rate max-min allocation is per-session-link-fair is a direct consequence of the results in [18]. However, the single-rate max-min fair allocation can fail to satisfy the other fairness properties. Consider the simple example in Figure 2, whose labeling is performed in an identical manner to that of Figure 1. Here, we have a network with two sessions, S_1 and S_2, whose respective senders, X_1 and X_2, are located at the same point in the network. We assume that the maximum desired rates are large, $\alpha_1 = \alpha_2 = 100$, such that they do not bound receiving rates in this network. Session S_1 is a single-rate session containing three receivers $r_{1,1}, r_{1,2}, r_{1,3}$, session S_2 is a unicast session whose receiver $r_{2,1}$ is located at the same point in the network as receiver $r_{1,1}$. In the max-min fair allocation, receivers in session S_1 receive at a rate of 2 (since this fully utilizes link l_2 and all receivers must receive at the same

[4] If one prefers to think in terms of utility rather use an ordering relation, it is fairly easy to construct a utility function, U, for allocations within a network, such that for any two allocations $A \ne B$, it is the case that $U(A) < U(B) \iff A <_m B$. For such a utility function, the max-min fair allocation is Pareto-optimal [16].

rate in a single-rate session), the receiver in S_2 receives at a rate of 3. Receivers $r_{1,1}$ and $r_{2,1}$ fail to achieve same-path-receiver-fairness, since they have the same data-paths, but differing receiving rates. Receiver $r_{1,3}$'s rate does not satisfy fully-utilized-receiver-fairness, because there is no fully utilized link along its data-path on which its rate is the largest compared to other receivers whose data-paths cross the same link. It follows that fully-utilized-receiver-fairness does not hold for session S_1, nor does it hold for the network. Last, per-receiver-link-fairness fails to hold for session S_1 (hence for the network as well) on the data-path to receiver $r_{1,3}$, since no link on this data-path is fully utilized. Per-receiver-link-fairness also fails to hold on the data-path to receiver $r_{1,1}$. This is because link l_1 is the only fully utilized link on $r_{1,1}$'s data-path, and the link rate of session S_1 on l_1 is smaller than that of session S_2. This example demonstrates that three out of the four desirable properties can fail to hold for single-rate max-min fair allocations.

2.4 Combining Multi-rate and Single-rate Sessions

We have examined the extent to which our four desirable properties hold for networks in which all sessions are the same type. Let us now consider these properties in the context of a network that contains a combination of multi-rate and single-rate sessions. Single-rate sessions are likely to always exist due to application constraints, such as a requirement that all receivers must complete receipt of data at approximately the same time.

Theorem 2 *Consider a network $N = \{G, \{S_1, \cdots, S_m\}, \tau, \sigma\}$ in which session types can differ, i.e., there can exist a pair of sessions, $S_i, S_{i'} \in N$ such that $\sigma(S_i) \neq \sigma(S_{i'})$. Then, the following are properties of the max-min fair allocation of N:*

(a) Fully-utilized-receiver-fairness holds for each receiver $r_{i,k} \in S_i$ where $\sigma(S_i) = \mathcal{M}$.

(b) per-receiver-link-fairness holds for each session S_i where $\sigma(S_i) = \mathcal{M}$.

(c) Per-session-link-fairness holds for all sessions S_i.

(d) Same-path-receiver-fairness holds between any two receivers $r_{i,k}$ and $r_{i',k'}$ where $\sigma(S_i) = \sigma(S_{i'}) = \mathcal{M}$.

(e) If $\sigma(S_i) = \mathcal{M}$ and $\sigma(S_{i'}) = \mathcal{S}$, and $r_{i,k} \in S_i$ and $r_{i',k'} \in S_{i'}$ have identical data-paths, then either $a_{i,k} = \alpha_i$ or $a_{i,k} \geq a_{i',k'}$.

Theorem 2 states that, even with single-rate sessions within the network, all four desirable fairness properties continue to hold within the max-min fair allocation for session link rates of multi-rate sessions, and for receiver rates of receivers belonging to multi-rate sessions. Hence, multi-rate sessions maintain their desirable fairness properties even when there are single-rate sessions within the network.

Let us also examine another way in which multi-rate multicast makes the max-min fair allocation for a network "more max-min fair". Recall that if an allocation A is min-unfavorable to an allocation B, then B is "more max-min fair" than A. Let us now consider how the max-min fair allocations compare for any two networks, N and $\bar{N}$ that differ only in their sessions' types.

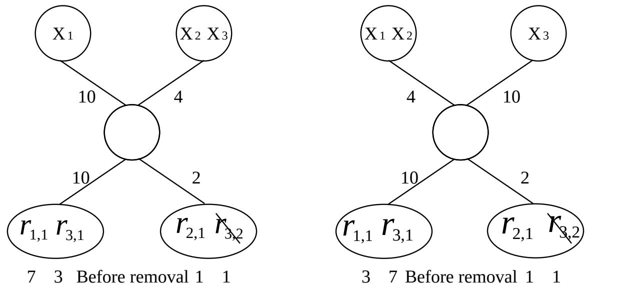
Lemma 3 *Let $N = (G, \{S_1, \cdots, S_m\}, \tau, \sigma)$ and $\bar{N} = (G, \{S_1, \cdots, S_m\}, \tau, \bar{\sigma})$ be networks where the set of multi-rate sessions in $\bar{N}$ is a subset of the set of multi-rate sessions in N, (i.e., $\forall i, \bar{\sigma}(S_i) = \mathcal{M} \Rightarrow \sigma(S_i) = \mathcal{M}$). If A is the ordered vector of max-min fair receiver rates in N, and $\bar{A}$ is the ordered vector of max-min fair receiver rates in $\bar{N}$, then $\bar{A} \leq_m A$.*

Lemma 3 tells us that as we "replace" single-rate sessions with identical multi-rate sessions (i.e., the only difference between the single-rate session and its replacement is the session type), then the max-min fair allocation is "more max-min fair". Hence, the "most max-min fair" allocation is the one in which all sessions are multi-rate:

Corollary 1 *Let $N = (G, \{S_1, \cdots, S_m\}, \tau, \sigma)$ be a multi-rate network ($\forall i, \sigma(S_i) = \mathcal{M}$), and let $\bar{N} = (G, \{S_1, \cdots, S_m\}, \tau, \bar{\sigma})$ be identical to N, except that $\bar{\sigma}(S_i) = \mathcal{S}$ for some sessions. Let A be the ordered vector of receiver rates for a multi-rate max-min fair allocation within N, and let B be the ordered vector of receiver rates in $\bar{N}$. Then $B \leq_m A$.*

2.5 Impact of Session Type on Receiver Rates

Last, let us consider how varying session types affects receiving rates on a session-by-session basis. We can prove that if all sessions' types are fixed except for session S_i, then if S_i is multi-rate, all of its receivers will receive at rates that are no less than what they would receive at if S_i is single-rate (see Lemma 9 in [15]). Unfortunately, this result does not extend to the case when several sessions can switch types. In fact, it is rather difficult to say what happens to receiver rates due to changes in the session type or the network topology. For example, one might conjecture that removing a receiver $r_{i,k}$ from a session would only increase other receivers' fair rates. Our intuition was that this would be the case since the removal frees up bandwidth that can then be used by other receivers whose data-path crosses $r_{i,k}$'s data-path. However, the max-min fair allocation of bandwidth after the receiver is removed can cause receiver rates (both in session S_i and in other sessions) to vary in either direction.



(a) intra-session decrease (b) intra-session increase

Figure 3: The change in max-min fair rates due to a removal of a receiver from a session.

To see this, consider the examples in Figure 3. Both networks contain three multi-rate sessions, S_1, S_2, and S_3. S_1 and S_2 each contain a single receiver, S_3 contains two receivers, the second ($r_{3,2}$) is subsequently removed. The max-min fair rates for receivers are indicated before and after this removal. Note that in Figure 3(a), $r_{3,1}$'s max-min fair rate decreases and $r_{1,1}$'s rate increases as a result of the removal. In Figure 3(b), $r_{3,1}$'s rate increases and $r_{1,1}$'s rate decreases. This demonstrates that removing receivers from sessions can have a non-obvious impact on the max-min fair rates of the remaining receivers in the network. Additional results appear in Appendix D of [15].

2.6 Section Summary

We now summarize the main results of this section. We have shown that if multicast sessions are multi-rate, then the max-min fair allocation is "more max-min fair" than if the sessions are restricted to being single-rate. This is demonstrated by showing that there are four desirable fairness properties that hold in the multi-rate max-min fair allocation that do not necessarily hold in a single-rate max-min fair allocation. We also examined networks in which some of the sessions are single-rate, while the remaining are multi-rate. By examining fairness properties on a per-session basis, we find that all of the fairness properties hold in general only in multi-rate sessions. Last, we use the min-unfavorable relation to comparatively examine which of any two allocations for a network is "more max-min fair". We find that "replacing" single-rate sessions by multi-rate sessions makes the max-min fair allocation more "max-min fair", which means that when all sessions are multi-rate, the max-min fair allocation is the "most max-min fair".

3 Achieving Multi-Rate Max-Min Fairness with Layering

In the previous section, we motivated the use of multi-rate sessions by showing that in theory they yield more desirable max-min fair allocations. One way to then obtain these rates in practice is to have the sender configure layers so that each receiver can obtain its fair rate by joining some subset of layers. However, the number of layers can be as large as the number of receivers in the session, making such an approach infeasible for large multicast sessions. Furthermore, the number of layers and the rate per layer is often beyond the control of the session itself, due to application-specific requirements, a limitation in the availability of multicast groups, or because it is too difficult for the sender to obtain the feedback needed to appropriately configure the rates of each of the layers. In this section, we examine how receivers can obtain their long term average max-min fair rates by repeated joins and leaves from multicast groups on which data is sent at a restricted set of rates. We will see that such a mechanism will force us to reconsider our previous assumption of how receiver rates impact link rates in the network.

Let us first discuss the implementation of a layered multicast approach. Data to be transferred is split into M *layers* by the sender, where layers are transmitted on separate multicast groups, each at some rate. The layers are ordered $L_1, \cdots, L_M$, such that all receivers desiring transmission join the group containing layer L_1, and any receiver that joins the group containing layer L_j must also join or already be joined to layer L_i for all $1 \leq i < j$ (henceforth, this is implied when we say that the receiver joins the layer or joins *up to* the layer). A receiver joined up to layer L_i receives data from the sender at an aggregate rate equal to the sum of the rates of layers L_1 through L_i. Joining layers increases the aggregate rate, while leaving layers decreases the aggregate rate.[5]

[5] We make the assumption that there is some utility in receiving at a faster rate, e.g., audio and video transmissions increase in clarity, reliable data transmissions take less time.

Let us examine why receivers must join and leave layers to obtain max-min fair rates. An obvious alternative is to require receivers to choose rates that can be obtained by joining up to a given layer and remaining at that rate for the duration of the session. This makes a finite set of rates available to the receiver. However, if these layers cannot be configured to the needs of receivers for reasons described above, the max-min fair allocation might not even exist! As an example, consider a simple network that consists of a single link with capacity c, and let there be two layered multicast sessions, S_1, and S_2 that traverse this link. Each session contains a single receiver, respectively denoted r_1 and r_2. The sender for session S_1 provides three layers, and sends at a rate of $c/3$ per layer. The sender for session S_2 provides two layers, and sends at rate $c/2$ per layer. The set of feasible allocations is $\{(0,0),(0,c/2),(0,c),(c/3,0),(c/3,c/2),(2c/3,0),(c,0)\}$, where (a_1,a_2) implies receiver r_i receives at a rate of a_i. None of these allocations are max-min fair. For instance, $(a_1,a_2) = (c/3,c/2)$ is not max-min fair since $(\bar{a}_1,\bar{a}_2) = (2c/3,0)$ is feasible, and $a_1 < \bar{a}_1$, but $a_2 > a_1$, hence there is no j where $\bar{a}_j < a_j \leq a_1$[6] (contradicting the defined requirement for max-min fairness). The reader can easily verify that none of the other feasible allocations is max-min fair.

Although it is not possible to achieve a max-min fair rate allocation when receivers are restricted to joining some arbitrarily chosen fixed set of layers for the entire length of a session, it is possible to achieve long-term average max-min fair rates through joins and leaves. The idea of using long term average rates also appears in current definitions of TCP-fairness [3, 10, 13, 19]. We define the *quantum*, Δt, to be the minimum amount of time over which a receiver's average rate is computed. We say that a rate of r is obtained through a link during the ith quantum if $r\Delta t$ bytes pass through the link between times $i\Delta t$ and $(i+1)\Delta t$. We say that a link l_j can support a capacity of c_j if it is able to forward $c_j\Delta t$ bytes within each time quantum.

Let us now consider an idealized network where a receiver can use joins and leaves to obtain its fair rate. The network is ideal in that we assume that network propagation delays and leave latencies are negligible compared to Δt and to packet inter-arrival times for each session. In this model, a packet traverses a link l_j only if it is received by some receiver $r_{i,k} \in R_j$. We also assume that all packets are of equal size, and for any receiver $r_{i,k}$, let $a_{i,k} \leq \alpha_i$ be its fair packet rate (in packets/sec) within the network. Consider a single layer (multicast group), where the transmission rate on the layer, ρ, satisfies $\rho \geq \max\{a_{i,k} : r_{i,k} \in S_i\}$. Receiver $r_{i,k}$ joins the single layer so that it receives the first $a_{i,k}\Delta t$ packets within the quantum,[7] then leaves the group. This is clearly possible, since $a_{i,k} \leq \alpha_i \leq \rho$, and $\rho\Delta t$ packets are transmitted on the layer during the quantum.

In this scenario, for any link l_j and session S_i where $|R_{i,j}| > 0$, there is some receiver $r_{i,k'}$ that receives $a_{i,k'} = \max\{a_{i,k} | r_{i,k} \in R_{i,j}\}$ packets per time quantum. Hence, this is the minimum number of packets that traverse link l_j for session S_i per quantum. Transmitting exactly this number of packets requires that all other receivers $r_{i,k} \in R_{i,j}$ receive a subset of the packets that are received by $r_{i,k'}$ per quantum. When this is not the case, $u_{i,j} > a_{i,k'}$.

[6] Or less formally, r_1's increase in rate does not result in a decrease in any receiver's rate whose original rate was less than r_1's.

[7] If $a_{i,k}\Delta t$ is not an integer, then it can elect to receive $\lfloor a_{i,k}\Delta t \rfloor$ packets in each quantum, and periodically receive $\lceil a_{i,k}\Delta t \rceil$ to come arbitrarily close to $a_{i,k}\Delta t$.

Definition 3 *We define the* **redundancy** *of a link l_j for a session S_i to be $u_{i,j}/\max\{a_{i,k}|r_{i,k} \in R_{i,j}\}$, where $u_{i,j}$ is the long-term average link rate l_j by session S_i, and $a_{i,k}$ is the long-term average rate for receiver $r_{i,k}$. We say a session's bandwidth utilization of a link is* **efficient** *for session S_i if the link's redundancy for that session is one, and define a session S_i's* **efficient link rate** *to equal $\max\{a_{i,k}|r_{i,k} \in R_{i,j}\}$.*

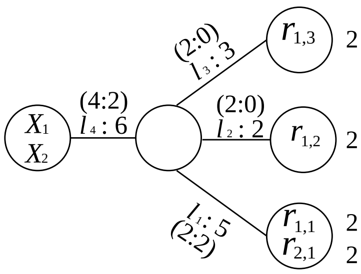


Figure 4: An example where a network fails to achieve session-perspective fairness properties due to redundancy.

Note that our assumption in Section 2 that $u_{i,j} = \max\{a_{i,k} : r_{i,k} \in R_{i,j}\}$ amounts to an assumption that multi-rate sessions are efficient (i.e., on all links in the network, a multi-rate session's link rate equals its efficient link rate). When there are multi-rate sessions that are not efficient, a multi-rate max-min fair allocation might not satisfy per-session-link-fairness (and hence might not satisfy per-receiver-link-fairness). To show this, we consider the network shown in Figure 4, whose labeling is similar to that of Figures 1 and 2. We again assume that the maximum desired rates are large so as not to bound receiving rates, e.g., let $\alpha_1 = \alpha_2 = 100$. Here, session S_1 is multi-rate with a redundancy of 2 over the shared link, l_4. Since the maximum receiving rate for receivers in S_1 (all of whose data-paths traverse l_4) is 2, $u_{1,4} = 4$. Since this is the only link that is fully utilized, and $u_{1,4} > u_{2,4}$, per-session-link-fairness fails to hold for session S_2. It follows that per-receiver-link-fairness fails to hold for session S_2 as well.

It is trivial to show that the fairness properties that do not compare session link rates, (specifically same-path-receiver-fairness and fully-utilized-receiver-fairness), continue to hold even when sessions are not efficient.

To understand the impact on redundancy of coordination between receiver joins and leaves, let us examine what happens on a shared link when there is no implicit join/leave coordination. Assume each receiver $r_{i,k}$ within session S_i randomly chooses the $a_{i,k}\Delta t$ packets it should receive within the quantum, with each packet having an equally likely chance of being chosen as any other in that quantum. In this case, $E[U_{i,j}] = \rho(1 - \prod_{t=1}^{s}(1 - a_{i,k_t}/\rho))$, where $\{a_{i,k_1}, \cdots, a_{i,k_s}\}$ are the rates of receivers that are members of the set $R_{i,j}$ (derivation in Appendix B).

Figure 5 shows how the number of receivers within a session that utilize a link (i.e., $|R_{i,j}|$) impacts the redundancy of a layer in this scenario. The number of receivers is shown on the x-axis, while the session's redundancy is indicated on the y-axis. The curves represent various configurations of $\{a_{i,1}, \cdots, a_{i,s}\}$. For curves labeled **All** z, ($z = 0.1, 0.5$, or 0.9), $a_{i,k}$ is set respectively to $.1\rho$, $.5\rho$, and $.9\rho$ for all receivers. For curves labeled **1st** w **rest** z, $a_{i,1} = w\rho$, and $a_{i,s} = z\rho$ for $1 < s \leq |R_{i,j}|$. Note that in each plot, the efficient link rate remains constant as the number of receivers is varied.

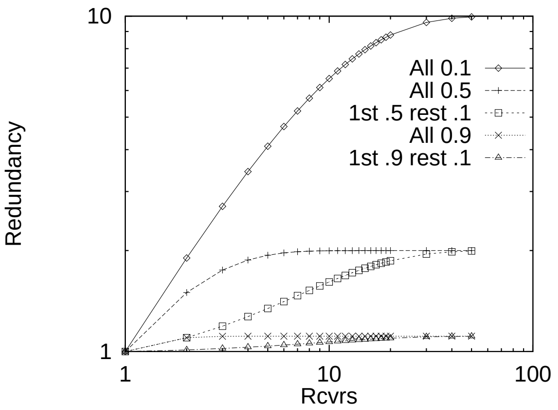


Figure 5: Redundancy of a single layer with random joins

We find that for redundancy to be high, the ratio of the efficient link rate to the transmission rate (i.e., $\max_{r_{i,k} \in R_{i,j}}\{a_{i,k}\}/\rho$) must be small. In fact, the redundancy can only be as large as the multiplicative inverse of this value (e.g., $\max\{a_{i,k}|r_{i,k} \in R_{i,j}\}/\rho = .1$ bounds redundancy from above by 10), and asymptotically reaches this value with an increase in the number of receivers that share the link. In other words, for redundancy to be high, all receivers must require only a small percentage of packets per quantum from a layer.

A second result is that for a fixed efficient link rate, redundancy increases most rapidly as a function of the number of receivers when all receivers receive at the same rate. In other words, an upper bound on how additional receivers impact redundancy is obtained by considering a network in which all receivers within a session have identical fair rates.

These results gives a preliminary indication as to what impacts the magnitude of redundancy within a network. We find that having additional layers often leads to a reduction in redundancy that is sometimes substantial, and that it never increases redundancy beyond that exhibited for the single-layer case. Details of these results can be found in Appendix E of [15].

3.1 The impact of redundancy on fair rates

Let us now examine the impact that redundancy has on fairness within a network. We now demonstrate why sessions with lower redundancy are "more max-min fair" than corresponding ones with high redundancy. We begin by relaxing our assumption made in Section 2 that $u_{i,j} = \max\{a_{i,k} : r_{i,k} \in R_{i,j}\}$. We extend our definition of a session to be a tuple $S_i = (X_i, \{r_{i,1}, \cdots, r_{i,k_i}, \}, v_i)$ that now includes a redundancy function v_i. Here, v_i maps a set (of arbitrary size) of receiver rates to a link rate. Given an allocation of receiver rates, A, session S_i's link rate for link l_j is computed as $u_{i,j} = v_i(\{a_{i,k} : r_{i,k} \in R_{i,j}\})$. In Section 2, v_i is simply the max operation. Since $u_{i,j} \geq a_{i,k}$ must hold whenever $r_{i,k} \in R_j$ (for reasons discussed in Section 2), it is necessary that $v_i(\{a_{i,k} : r_{i,k} \in R_{i,j}\}) \geq \max\{a_{i,k} : r_{i,k} \in R_{i,j}\}$.

Lemma 4 *Let $N = (G, \{S_1, \cdots, S_m\}, \tau, \sigma)$, and $\bar{N} = (G, \{\bar{S}_1, \cdots, \bar{S}_m\}, \tau, \sigma)$ be identical networks, where each session S_i in N is identical to $\bar{S}_i$ in $\bar{N}$, except for their respective redundancy functions, v_i and $\bar{v}_i$. Assume* *sessions in $\bar{N}$ exhibit higher redundancy than those in N, (i.e., for each session S_i and any set of real numbers, X, $v_i(X) \leq \bar{v}_i(X)$). Let A be the max-min fair allocation in N and $\bar{A}$ the max-min fair allocation in $\bar{N}$. Then $\bar{A} \leq_m A$.*

Lemma 4 states the following: assume that sessions are "replaced" by sessions that are identical, except that the session link rates required to support a given set of receiver rates are higher (e.g., the amount of coordination of joins and leaves between receivers within a session is reduced). It follows that the resulting max-min fair allocation is "less max-min fair" than the max-min fair allocation for the network with the sessions prior to the "replacement".

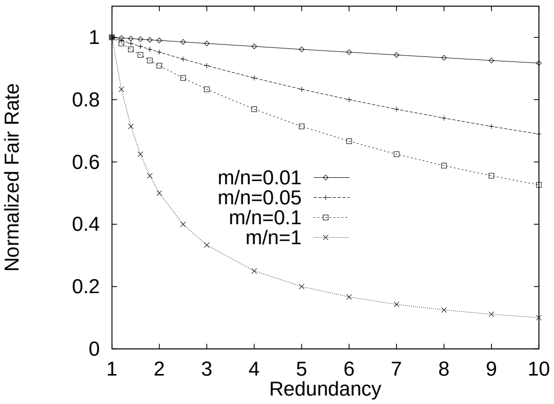


Figure 6: The impact of redundancy on fair rates.

We know that a redundancy greater than one produces max-min fair rate allocations within the network that might not exhibit the session-perspective fairness properties, per-receiver-link-fairness and per-session-link-fairness. Also, using the min-unfavorable relation, we have shown that increased redundancy might reduce the "max-min fairness" of a max-min fair allocation. Let us now quantitatively examine how redundancy may impact fair rates. Consider a set of n sessions whose receiver rates are constrained by the same link, l with capacity c. Let m of these sessions be multi-rate with a redundancy of v on link l, and the remaining $n - m$ sessions have redundancy 1. Since we assume that all receivers' rates are constrained by link l, their max-min fair rates are all equal to $\frac{c}{(n-m)+mv}$. Figure 6 shows the receivers' rates as a function of the redundancy, v. The x-axis indicates v, the various curves represent various values of the ratio of sessions, m/n, that exhibit redundancy v. The y-axis presents the fair rate normalized by c/n, the fair rate for all the receivers in the network when all sessions are efficient.

Figure 6 indicates that even modest levels of redundancy can substantially reduce the fair receiver rates for all sessions in the network. From this we can draw two conclusions: first, it is important to maintain low redundancy on network links to keep fair rates high. Second, when multi-rate sessions make up a small percentage of sessions in the network, they have less of an impact on the fair rates of sessions. Due to the current proliferation of unicast traffic within the network, we expect that less than 5% of sessions within the network will be multi-rate. This means that low levels of redundancy greater than one can be tolerated.

These results raise an interesting dilemma: should multi-rate protocols be used to achieve fairness from the receivers' perspectives, even if it means failing to achieve per-session-link-fairness (a fairness property that holds when all network sessions are single-rate and unicast)? We argue that yes, multi-rate protocols should still be used, because the "unfair" additional usage of link bandwidth due to redundancy can be justified in that the session is transmitting data to multiple receivers. A similar argument is used in [7] to allocate link bandwidth to sessions in a manner that is proportional to the number of receivers within the session.

The reduction in rate due to redundancy can occur whenever a multi-rate session tries to achieve some form of fairness using joins and leaves of layers. For example, in [19], receiver join experiments are coordinated within a network where TCP-fairness is the fairness criteria. The coordination prevents "bottleneck bandwidth allocated to [the] protocol instance [from] not being fully exploited." This lack of "exploitation" is, in effect, an artifact of redundancy.

4 Redundancy in Practical Congestion Control Protocols

In Section 3, we showed that a lack of join and leave coordination within a session increases the session's redundancy on links shared by that session's receivers. This in turn is likely to reduce their fair receiving rates. Our final contribution is to show that redundancy can easily be kept quite low in practice. We show this by measuring the redundancy of several Internet layered congestion control protocols that vary in the degree to which joins are coordinated among receivers. In these protocols, receivers react to congestion by leaving layers, and probe for available bandwidth by joining layers. We compute each protocol's redundancy using analysis and simulation of simple network models. Because of the simplicity of the models, there may be some differences between what we observe and what will actually occur in practice. However, we do not expect these differences to alter results significantly enough to change our conclusions.

In each protocol, a receiver leaves the highest layer joined (unless only joined to one layer) whenever it observes a *congestion event*: an indication that some part of its data-path is being overutilized. In practice, a congestion event may be the loss of a packet by the receiver, or a bit set within a packet by the network used to indicate that the receiving rate should be lowered [12]. If no congestion events are observed by a receiver within a sequence of packet arrivals, it joins an additional layer (unless already joined to all layers). Using these protocols, a receiver repeatedly adjusts the set of layers to which it is joined for the duration of the session. The protocols differ in the degree to which joins are coordinated within a session.

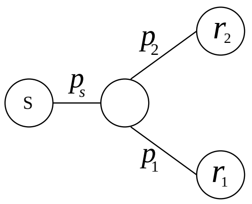
- In the *Uncoordinated* protocol, there is no inherent coordination: upon receiving a packet, a receiver randomly decides whether to join an additional layer.

- In the *Deterministic* protocol, there is also no inherent coordination; a receiver joins an additional layer after receiving a fixed number of packets without loss since its last join or leave event.

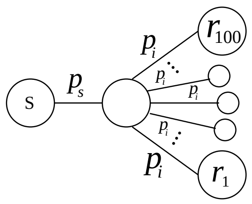
- In the *Coordinated* protocol, the sender indicates (e.g., through a field within its transmitted packet) when receivers should join an additional layer. This is done in such a way so that when the field indicates that receivers joined up to layer i should join layer $i+1$, it also indicates that receivers joined up to layer $j < i$ should join layer $j+1$.

The additional details of the protocols (layer rates, join-period) are based on the choices made in [19]. For instance, we require that the aggregate rate of layers 1 through i equals 2^{i-1}, and that the expected number of packets received by a receiver between a previous join/leave event to its join to layer $i+1$ equals $2^{2(i-1)}$.[8] Because of these protocols' similarities to the protocol in [19], we anticipate these protocols are suitable for the same set of continuous stream and reliable bulk data transfer applications described in [19]. Due to a lack of round-trip-time dependence, these protocols come closer to achieving max-min fair rates than TCP-fair rates. See Appendix F in [15] for a more precise description of these protocols and how they differ from the protocol in [19].

We model packet loss (or equivalently, congestion marking of packets) as a Bernoulli loss process. The reader can consider the loss process to be fairly accurate for a network where the number of flows across links is large, so that there is little correlation between the rate of an individual flow and the link loss rate [21]. Our model also assumes that receivers' reactions to coordinated events (shared loss, coordinated joins) take effect at the same time: two receivers that see identical loss patterns would be joined to the same set of layers. Under these conditions, it can be argued that these protocols come "close" to achieving the max-min fair rates, i.e., the expected rate does not exactly equal the max-min fair rate, but the difference is fairly small.

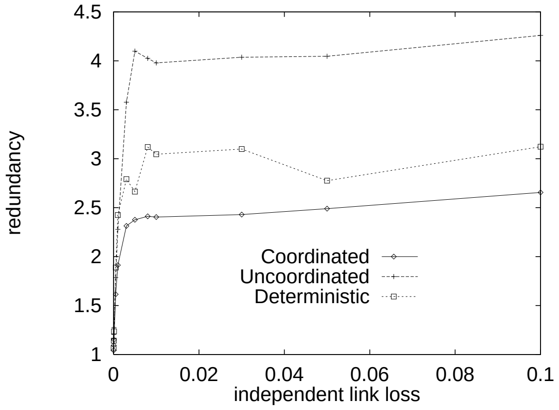


(a) Analysis Model

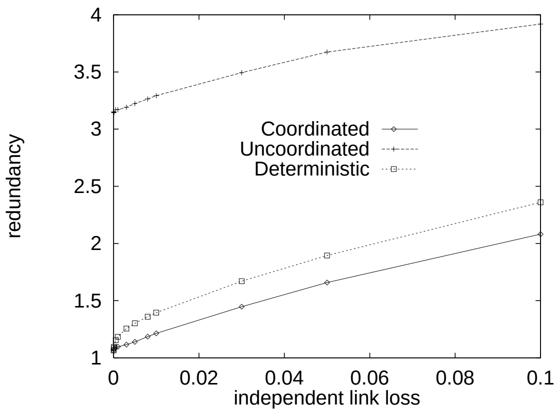


(b) Simulation Model

Figure 7: Network models for coordination experiments



(a) Low shared loss .0001



(b) High shared loss .05

Figure 8: 100 receivers, 8 layers

Our experiments use modified star networks, as shown Figure 7, to examine how shared loss (i.e., loss on the shared link abutting the sender) and independent loss (i.e., loss on the fanout links) impact redundancy. The initial set of experiments uses the topology in Figure 7(a). Using Markov models of the protocols over this network, we examine how different values of shared and independent loss impact the redundancy of a session on the shared link. The details of these models appear in Appendix F of [15] We summarize the most important finding: redundancy is highest when receivers experience the same end-to-end loss rates. This result follows intuitively from our observation in Section 3 that redundancy is highest when all receivers' receiving rates are equal.

Our Markov models are too computation-intensive to allow us to examine sessions with large sets of receivers. Instead, we turn to simulation. Figure 8 shows simulations of the protocols using 8 layers with 100 receivers in the session that have identical end-to-end loss rates, configured in the modified-star topology of Figure 7(b). In Figure 8(a), the shared loss rate is fixed to 0.0001 (i.e., very low shared loss), and the loss rate on each of the fanout links is given on the x-axis. Each curve shows the redundancy for the three protocols we consider. Each point plotted is the mean of 30 experiments where the sender transmits 100,000 packets, the variance is less than 1% with 95% confidence. Figure 8(b) plots similar results, but where the shared loss rate is .05. We see that for all protocols, redundancy remains fairly low (below 5) for reasonable loss rates. By having the sender coordinate joins as in the Coordinated protocol, redundancy remains below 2.5, even when there are 100 receivers within the session, each of whose data-path contains the shared link. We observed negligible changes in the results when we increased the number of receivers beyond 100. Since our previous results indicate that redundancy is highest when all receivers have identical end-to-end loss rates, we can conclude that sender-coordinated congestion control protocols can keep redundancy below 2.5. This is low enough so that, in networks where multi-rate protocols make up a small percentage of sessions, multi-rate protocols will yield fair allocations with sufficiently desirable fairness properties.

[8] In [19], the number of packets received equals $2^{2(i-1)}$ (i.e., it is a deterministic value).

5 Related / Future Work

The application of layering, in the context of video transmission, to maximize usage of available bandwidth and the benefits of coordination of receiver join events within a session is discussed in [11], and further explored in [8]. Clever use of parity coding techniques extend layering's applicability to reliable multicast [4, 14, 19]. Preliminary experiments and definitions of various forms of fairness for layered approaches are explored in [8, 19], as well as in [9], which discusses at a high level how using a layered approach can change the max-min fair allocation. An examination that uses fairness metrics to compare various allocation strategies for layered multicast protocols is presented in [7]. There, the authors argue that link bandwidth should be allocated to sessions in some manner that is proportional to the number of receivers in the session because doing so increases the average "receiver satisfaction". However, none of these works look consider how multi-rate approaches affect fairness properties (in comparison to single-rate approaches) throughout a large-scale network.

Much of the remaining work that deals with multicast fairness assumes that sessions are single-rate [3, 13, 18, 20], and therefore compromise fairness from the receiver perspective, due to tight binding of receiver rates within a session. There has been some work that discusses how one might choose a single-rate session's rate in order to maximize a measure of fairness on a per-receiver basis [6].

There are numerous issues that remain open with regard to using layering to achieve multi-rate max-min fairness. The effects of layering on desirable fairness properties for other definitions of fairness is one possible avenue for examination. We believe that many of our results can be directly applied to TCP-fairness by constructing a definition of max-min fairness where receiver rates are assigned weights (i.e., a receiver's rate is weighted by the inverse of round trip time). It would also be interesting and useful to extend definitions of fairness to multicast sessions with multiple senders. There are also many issues that deal with the practicality of using layering to achieve fairness. One question that comes to mind is whether priority dropping schemes for layered approaches [1] might aid in reducing redundancy by increasing coordination among receivers. Also, multicast routing technology must be improved to make layered approaches practical for congestion control and fairness purposes. For instance, join and leave latencies complicate coordination among various receivers within a session, which is likely to increase redundancy. We believe that long leave latencies will also increase redundancy (a link continues to receive at the rate prior to the leave, until the leave takes effect, while the receiver's rate reduces immediately). We expect that many such problems are solvable, perhaps with the aid of active routing technology [17]. For instance, placing the decision to add and drop layers at the active nodes, rather than at receivers, should increase the coordination of the joins and leaves of layers by downstream receivers, thereby reducing redundancy. Such an approach would make a redundancy of one feasible for a layered multi-rate session.

It is also unclear whether bandwidth can be shared fairly by sessions that measure fairness on different timescales (i.e., use different quanta), especially in networks like the Internet, where a session's fair allocation may vary due to startup and/or termination of other sessions within the network. Finally, our models contain numerous simplifications of what exists in practice; they are merely used to illustrate concepts, identify challenges, and provide a basic understanding of what can be expected in practice. Extensive development and testing is still necessary to verify that our hypotheses presented here do in fact occur in practice.

6 Conclusion

We have explored how multi-rate multicast, achievable using layered multicast approaches, can impact fairness within a network. In particular, we showed that in theory, multi-rate sessions can achieve several desirable fairness properties that cannot be achieved in general networks using single-rate sessions. In a practical environment, we demonstrate how receivers can join and leave layers so that their rates are max-min fair over a long term average. Unfortunately, this join-leave process has several practical difficulties. One difficulty that we address is redundancy: an excessive use of bandwidth by a session over a link shared by multiple receivers in the session. High redundancy not only leads to failure of several fairness properties from a session perspective (i.e., fairness of session link rates), but is also likely to reduce most receivers' fair rates. However, our subsequent analysis shows that based on the portion of network sessions that are expected to be multi-rate, practical solutions can keep the amount of redundancy low enough such that layering can be used to improve fairness within multicast networks.

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions. We are also grateful to several UMass colleagues: Timur Friedman and Jitendra Padhye provided thorough comments on previous drafts. They, as well as Supratik Bhattacharya, Sneha Kasera, Sue Moon, Sambit Sahu and Maya Yajnik, provided key insights.

Appendix

A Max-min Fair Construction Algorithm

The following algorithm constructs a max-min fair allocation for a network, N. In plain English, the algorithm iterates over a set of receivers, each step increasing those receivers' rates uniformly as much as possible without overutilizing any links in the network. A receiver is removed from this set once some link on its data-path reaches full capacity, or, if the receiver is part of a single-rate session, the data-path of some receiver in the session contains a link that has reached full capacity. We define

$$\phi_{i,j}(T) = \left\{ \begin{array}{ll} 1 & |R_{i,j} \cap T| > 0 \\ 0 & \text{otherwise} \end{array} \right\}$$

1. $T_0 = \{r_{i,k}\}; \forall r_{i,k}, a^0_{i,k} = 0; \forall i,j, u^0_{i,j} = 0, u^0_j = 0; b = 0$
2. While $|T_b| > 0$
3. $\quad t_{b+1} = \sup\{t : \forall j, u^b_j + \sum_i \phi_{i,j}(T_b)t \leq c_j \wedge \forall r_{i,k} \in T_b \Rightarrow a^b_{i,k} + t \leq \alpha_i\}$
4. $\quad \forall r_{i,k} \in T_b, a^{b+1}_{i,k} = a^b_{i,k} + t_{b+1}$. For all other $r_{i,k}, a^{b+1}_{i,k} = a^b_{i,k}$.

5. $u_{i,j}^{b+1} = \sum_{r_{i,k} \in R_j} a_{i,k}^{b+1}, u_j^{b+1} = \sum_i u_{i,j}^{b+1}$.

6. $T' = T_b - \{r_{i,k} \in T_b : a_{i,k}^{b+1} = \alpha_i \vee (\exists j, r_{i,k} \in R_{i,j} \wedge u_j^{b+1} = c_j)\}$

7. $T_{b+1} = T' - \{r_{i,k} \in T' : \sigma(S_i) = \mathcal{S} \wedge \exists r_{i,k'} \notin T'\}$

8. $b++$

9. end while

10. $\forall r_{i,k}, a_{i,k} = a_{i,k}^b, \forall i, j, u_{i,j} = u_{i,j}^b, u_j = u_j^b$

Step 3 the largest value that all receivers' rates in T_b can be incremented while maintaining feasibility of the allocation. Steps 4 and 5 apply this increase to the "current" receiver rates and link rates respectively. Step 6 removes any receivers from T_{b+1} whose rates cannot be incremented any further, or else they would be larger than the maximum session rate, or would cause overutilization of some link. Step 7 removes any receivers in single-rate sessions from T_{b+1}, given that some other receiver in that session has been removed (so that all receiver rates in this session remain identical).

B Expected Bandwidth with Random Joins

We compute the expected bandwidth for session S_i on a link l_j. For simplicity, we write $R = |R_{i,j}|$, and denote the set of receivers from session S_i whose data-path utilizes this link (i.e., $R_{i,j}$) as $\{r_1, \cdots, r_R\}$, and let a_t be the number of packets that receiver r_t must receiver per quantum.

Let ρ packets be transmitted in a time quantum, and let X_i be a random variable that equals 1 if any receiver is joined when packet i is transmitted, and 0 otherwise ($1 \le i \le \rho$). Let $Y_{i,t}$ be a random variable that equals 1 if receiver r_t joins to receive packet i, and 0 otherwise. Since we assume a receiver chooses the packets it is to receive from a uniform distribution, we have $\Pr(Y_{i,t} = 1) = a_t/\rho$.

$$\begin{aligned} E[X_i] &= 1 - \prod_{t=1}^{R} \Pr(Y_{i,t} = 0) \;=\; 1 - \prod_{t=1}^{R} (1 - a_t/\rho) \\ E[U_{i,j}] &= E[\sum_{i=1}^{\rho} X_i] \;=\; \sum_{i=1}^{\rho} E[X_i] \\ &= \rho(1 - \prod_{t=1}^{R} (1 - a_t/\rho)) \end{aligned}$$

References

[1] S. Bajaj, L. Breslau, and S. Shenker, *Uniform Versus Priority Dropping for Layered Video*, Proceedings of ACM SIGCOMM 98, Vancouver, CA, September 1998.

[2] D. Bertsekas and R. Gallagher, *Data Networks*, Englewood Cliffs, NJ, Prentice-Hall, 1992.

[3] S. Bhattacharyya, D. Towsley, and J. Kurose, *The Loss Path Multiplicity Problem for Multicast Congestion Control*, Proceedings of IEEE INFOCOM 99, New York, NY, March 1999.

[4] J. Byers, M. Luby, M. Mitzenmacher, and A. Rege, *A Digital Fountain Approach to Reliable Distribution of Bulk Data*, Proceedings of ACM SIGCOMM 98, Vancouver, CA, September 1998.

[5] Z. Cao and E. Zegura, *Utility Max-Min: An Application-Oriented Bandwidth Allocation Scheme*, Proceedings of IEEE INFOCOM 99, New York, NY, March, 1999.

[6] T. Jiang, M. Ammar, and E. Zegura, *Inter-Receiver Fairness: A Novel Performance Measure for Multicast ABR Sessions*, Proceedings of ACM SIGMETRICS 98, Madison, Wisconsin, June 1998.

[7] A. Legout, J. Nonnenmacher, and E. Biersack, *Bandwidth Allocation Policies for Unicast and Multicast Streams*, Proceedings of IEEE INFOCOM 99, New York, NY, March, 1999.

[8] X. Li, S. Paul, and M. Ammar, *Layered Video Multicast with Retransmissions (LVMR): Evaluation of Hierarchical Rate Control*, Proceedings of INFOCOM 98, March 1998, San Francisco, CA.

[9] X. Li, S. Paul, and M. Ammar, *Multi-Session Rate Control for Layered Video Multicast*, Proceedings of Symposium on Multimedia Computing and Networking, San Jose, CA, January 1999.

[10] J. Mahdavi and S. Floyd, *TCP-Friendly Unicast Rate-Based Flow Control*, Note sent to e2e mailing list, January, 1997.

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

[12] K.K. Ramakrishnan and S. Floyd, *A Proposal to add Explicit Congestion Notification (ECN) to IP.* RFC 2481, January 1999.

[13] I. Rhee, N. Balaguru, G. Rouskas, *MTCP: Scalable TCP-like Congestion Control for Reliable Multicast*, Proceedings of IEEE INFOCOM 99, New York, NY, March 1999.

[14] L. Rizzo and L. Vicisano, *RMDP: An FEC-based Reliable Multicast Protocol for Wireless Environments*, Mobile Computing and Communications Review, Volume 2, Number 2, April 1998.

[15] D. Rubenstein, J. Kurose, and D. Towsley, *The Impact of Multicast Layering on Network Fairness*, UMASS CMPSCI Technical Report 99-08. Available on-line at *http://www-net.cs.umass.edu/papers/papers.html*

[16] S. Shenker, *Making Greed Work in Networks: A Game-Theoretic Analysis of Switch Service Disciplines*, Proceedings of ACM SIGCOMM 94, London, UK, August 1994.

[17] Tennenhouse, D., Smith, J., Sincoskie, D., Wetherall, D., and Minden, G., *A Survey of Active Network Research*, IEEE Communications Magazine, January 1997.

[18] H. Tzeng and K. Siu, *On Max-Min Fair Congestion Control for Multicast ABR Service in ATM*, IEEE JSAC, Vol. 15, No. 3, April 1997.

[19] L. Vicisano, J. Crowcroft, and L Rizzo, *TCP-like Congestion Control for Layered Multicast Data Transfer*, Proceedings of IEEE INFOCOM 98, San Francisco, CA, March, 1998.

[20] H. Wang and M. Schwartz, *Achieving bounded fairness for multicast and TCP traffic in the Internet*, Proceedings of ACM SIGCOMM 98, Vancouver, CA, September 1998.

[21] M. Yajnik, S.B. Moon, J. Kurose, and D. Towsley, *Measurement and Modeling of the Temporal Dependence in Packet Loss*, Proceedings of IEEE INFOCOM 99, New York, NY, March, 1999.