Modeling the Branching Characteristics and Efficiency Gains in Global Multicast Trees

Robert C. Chalmers, Kevin C. Almeroth

Abstract— **In this paper, we investigate two issues. First, what level of efficiency gain does multicast offer over unicast? Second, how does the shape of multicast trees impact multicast efficiency? We address the first issue by developing a metric to measure multicast efficiency for a number of real and synthetic datasets. We find that group sizes as small as 20 to 40 receivers offer a 60-70% reduction in the number of links traversed compared to separately delivered unicast streams. Addressing the second issue, we have found that almost all multicast trees have similar characteristics in terms of key parameters such as depth, degree frequency and average degree. A final contribution of our work is that we have taken multicast group membership data and multicast path data and compiled datasets which can be used to generate large, realistic multicast trees.**

Keywords—**Multicast, measurement, modeling, tree characteristics.**

I. Introduction

MULTICAST provides a conceptual advantage for designers developing applications with support for a large, distributed set of clients. Logical addressing offers a mechanism to loosely assemble groups with dynamic membership, freeing the application from the task of transmitting to each individual receiver. More importantly, however, multicast provides an operational advantage for content and network providers by reducing the overall resource demands of the application. Only a single multicast packet is sent to the group, no matter how many receivers may have joined. Multicast reduces the overall bandwidth demand of content transmission since packet duplication only occurs when paths to multiple receivers diverge. The problem, however, is that the actual benefit derived by any application is seemingly dependent upon the shape of the multicast distribution tree, and a tree's shape changes over time with the arrival and departure of receivers.

To date, multicast researchers have been hard pressed to accurately quantify how multicast performs against unicast in terms of bandwidth efficiency. Furthermore, little quantitative work has been presented which offers insight into the characteristics of multicast trees as they exist in the Internet today. In this paper, we identify several key properties which we believe sufficiently describe the shape of inter-domain multicast trees. Moreover, we present a characterization of multicast efficiency which accurately models multicast bandwidth gains as a skewed distribution related to the number of receivers in the multicast group.

We first define a metric with which to compare the bandwidth utilization of multicast and unicast in terms of links traversed, i.e. a count of duplicate packets as seen across all links of the distribution tree. Then, we provide an estimate for our metric that accurately characterizes multicast efficiency over a range of tree topologies and group dynamics. The estimate is validated by calculating the metric for a set of real groups collected from the Internet. We then loosen *temporal* and *spatial* constraints to investigate the characterization's dependence on receiver duration, inter-arrival time and receiver distribution. Consistently, the characterization shows that even for a small number of receivers, multicast out-performs unicast in terms of bandwidth utilization. Group sizes as small as 20 to 40 receivers offer a 60-70% reduction in the number of links traversed compared to separately delivered unicast streams.

The authors are with the Department of Computer Science, University of California, Santa Barbara, CA 93106-5110 (email: robertc@cs.ucsb.edu and almeroth@cs.ucsb.edu).

We investigate further by asking the question: does the consistency seen in a high-level measure, such as bandwidth efficiency, indicate a similar consistency in the underlying tree structure? For researchers developing new multicast protocols, it is critical that trees used in evaluations are as realistic as possible. Although, the efficiency characterization could potentially verify that a *generated* tree is similar to *real* multicast trees, of additional benefit is a complete set of parameters that lead to generated trees that are known to be realistic. We focus primarily on how and where branching occurs in the tree and conclude that multicast trees do share several common properties such as maximum depth, degree frequency and average degree.

The paper is organized as follows. In the next section, we further pursue the motivation for the paper. We provide background and the definition of the efficiency characterization in Section III. Then, we detail a number of experiments that validate the characterization using samples of real and generated multicast groups in Section IV. We discuss the results of our experiments and look closer at the shape of multicast trees in Section V. In Section VI, we outline future work, and we present our conclusions in Section VII.

II. Motivation

The two goals of this work, to quantify multicast efficiency and to accurately model tree topologies, are motivated by the needs of two distinct groups. Content and network providers are interested in multicast efficiency as a means to justify multicast deployment while developers and researchers require realistic tree topologies to better validate protocol designs. The pursuit of both goals depends upon the same thing: a clear understanding of the shape of multicast distribution trees.

For researchers developing congestion control algorithms and reliable multicast protocols, realistic topologies are crucial since the validity of the conclusions depend on the realism of the environment used in the evaluations. To this end, identifying a set of key properties and a range of suitable values would allow us to extend current graph generators to produce realistic sample trees for testing. Individual topologies, however, shed little light on the decisions made to control flow and loss. Eventually, a better understanding of the general shape of multicast trees may be of equal if not greater value during the design of these protocols [1].

This leads to the question: how does shape relate to efficiency? The shape of the distribution tree dictates how efficient a multicast transmission can be. If the majority of paths from the source are not shared, multicast is not much more efficient than unicast. In fact, multicast may actually be less efficient overall due to the increase in routing state and overhead required to implement it. In general, the efficiency of a particular tree is determined by a few key properties:

- **height** - As the tree grows taller and new receivers are added near the bottom, the number of links shared amongst receivers increases and thus multicast efficiency improves.
- **breadth** - Branching that occurs early in the tree decreases multicast efficiency since it represents diverging paths and packet duplication, while branching near the bottom of the tree implies a long series of heavily shared links.
- **number of receivers** - As the number of receivers increases, the likelihood that new receivers will share a part of the existing tree also increases.

Where researchers may be interested in efficiency as a means to improve communication, Internet Service Providers (ISP) are interested primarily in improving the utilization of limited resources. ISPs are faced with a decision which depends on a so-called "sweetspot" where the bandwidth cost of using unicast out-weighs the increased overhead of deploying multicast [2]. In other words, multicast's efficiency must overcome the additional state requirements of routers, as well as, the possibly more important costs of installing and managing multicast. To advance multicast deployment, then, it is critical that we quantify multicast efficiency. We must be able to quantify the benefit that multicast provides over unicast so that it may be effectively considered against its own cost. In particular, we are interested in measuring the bandwidth savings afforded by using multicast. In doing so, we can offer a quantitative measure of multicast's *sweetspot*, and consequently provide a means for network providers to evaluate the feasibility and affordability of deploying multicast.

III. MEASURING MULTICAST EFFICIENCY

The goal of this section is to characterize the bandwidth efficiency of multicast as it currently exists in the Internet. To this end, we define a simple metric to measure the reduced packet duplication achieved by multicast. Drawing from prior work [3], [4], we then present an estimate for this metric based solely on the number of group receivers.

First, we review related work which provides the necessary background to develop the efficiency estimate. The remainder of the section provides the groundwork for our analysis. Then, in the next section, we use our proposed metric to evaluate real multicast tree data and develop synthetic case studies which have sufficient quantitative similarities to indicate that the characterization is valid.

A. Related Work

To build a foundation for the rest of the paper and to provide a litmus for evaluating the results of our analysis, we first present an overview of previous work relevant along three lines:

- the topology and shape of the Internet(unicast) and multicast trees;
- measuring multicast characteristics, e.g. bandwidth utilization, loss, delay and traffic concentration; and
- comparing multicast to unicast in terms of cost and efficiency.

A.1 Topology and Shape

Faloutsos et al. propose a number of power-law relationships describing several key network properties [5]. Power-laws are of the form $y \propto x^a$, where a is a constant and x and y are the variables of interest. These relationships describe heavily skewed distributions. Medina et al. suggest that the existence of these skewed distributions is due primarily to four key properties of real networks [6]:

- preferential connectivity,
- incremental growth,
- geographical distribution of nodes, and
- locality of edge connections.

In short, networks grow by adding new nodes incrementally. When a new node attaches, it is more likely to connect to a nearby node that already maintains a large number of links to other nodes.

These properties of general networks are particularly applicable to the construction of multicast trees; trees are constructed incrementally as new receivers join the multicast group and new branches are grafted into the nearest portions of the existing tree. This implies that many properties of multicast trees may also be characterized by skewed distributions and power-laws. One particularly applicable formula defined for networks relates outdegree and frequency, $f_d \propto d^{\mathcal{O}}$ ($\mathcal{O}$ is negative) [5]. The function maintains that the majority of the nodes in the Internet have few out-going links and that only a small number of nodes have high degrees. Applying this power-law relationship to multicast, one may infer that trees built over such an underlying network infrastructure would also consist of many nodes with small degrees, and thus produce trees that are more likely to be tall than wide. Pansiot and Grad confirm this with an experiment using a tree comprised of nearly 4,000 nodes and 5,000 edges [7]. They conclude that the majority of nodes, more than 70%, were either receivers or intermediate (relay) routers with a single out-going link.

A.2 Multicast Characteristics

Previously, a number of studies [8], [9], [10], [11], [12] have investigated multicast in terms of bandwidth utilization, loss, delay and traffic concentration. The authors also address the impact of routing protocol overhead such as dense-mode flooding, and illustrate the differences between construction methods for shortest-path and core-based trees. In this paper, we choose not to readdress these issues. Rather, we focus on how multicast performs in relation to unicast in terms of packet duplication.

A.3 Multicast versus Unicast

Chuang and Sirbu [3] provide the first definitive comparison of multicast and unicast while defining a measure for pricing multicast. They focus on the ratio between the total number of multicast links and the average unicast path length. The authors determined that the relationship can be concisely expressed as a power-law in terms of the number of receivers. They produced the formula

$$\frac{L_m}{\overline{L_u}} = N^k \quad (1)$$

where L_m is the total number of multicast links in the distribution tree, $\overline{L_u}$ is the average path length between any two nodes, N is the number of receivers in the tree, and k is an economies-of-scale factor ranging between 0 and 1. Using samples of the old ARPANET, the early MBone and a number of generated topologies, they concluded that the value of k was consistently near 0.8. This provides a simple yet unintuitive relationship. How could multicast and unicast cost be solely dependent on the number of receivers? Does it not matter where in the tree the receivers are located? Is not the shape of the tree just as important as how many receivers it serves? Either the shape of most multicast trees are in some way constrained or a wide range of shapes exhibit similar costs. In either case, the economies-of-scale factor, k, of 0.8 expresses the likelihood that a new receiver will share a portion of its path with the existing tree.

This cost relationship (1) was later confirmed with a more rigorous mathematical treatment [4]. Although the authors produced a cost function that was logarithmic rather than a power-law, it was found to behave very similarly over a wide range of generated topologies. Furthermore, they concluded that this relationship could be applicable to investigating multicast efficiency in addition to cost. Both studies [3], [4], however, used mostly generated networks and uniformly distributed receivership. Neither investigated how well this relationship holds in the current Internet for real groups.

B. Defining a Metric

In order to effectively evaluate the behavior of multicast efficiency in real networks, we must first define a metric to use as a basis for comparison. In a previous work [13], we introduced a metric using multicast and unicast hop counts. The metric is defined as

$$\delta = 1 - \frac{L_m}{L_u} \quad (2)$$

where L_m is the total number of multicast links in the distribution tree and L_u is the sum of all unicast hops. δ represents the percentage gain in multicast efficiency over unicast. As δ approaches zero, multicast and unicast are nearly equal and little to no savings in bandwidth is achieved. As δ approaches one, all receivers share a single multicast path resulting in the maximum possible bandwidth efficiency.

C. Defining the Estimate

What is interesting, however, is that we can use the relationship defined in Equation 1 to develop an estimate for the metric (2) in terms of the number of receivers, N, and an associated *efficiency factor*, ε.

$$\delta = 1 - N^{\varepsilon}, \text{ where } \varepsilon = k - 1 \quad (3)$$

Deriving the estimate in this way makes the assumption[1] that $\overline{L_u} = \frac{L_u}{N}$. For an economies-of-scale factor, k, of 0.8, the efficiency factor, ε, equals -0.2. As will be seen in the following analysis, however, ε tends to range between -0.34 and -0.30 for real and synthesized receiver distributions. Figure 1 shows how the estimate behaves as the number of receivers increases. The shape of the curve implies that multicast outperforms unicast even with a few receivers; between 20 and 40 receivers, multicast provides 60-70% increased efficiency over unicast. The estimate predicts an approximate 80% savings for 150 users, reaching 90% for large groups with 1,000 receivers or more.

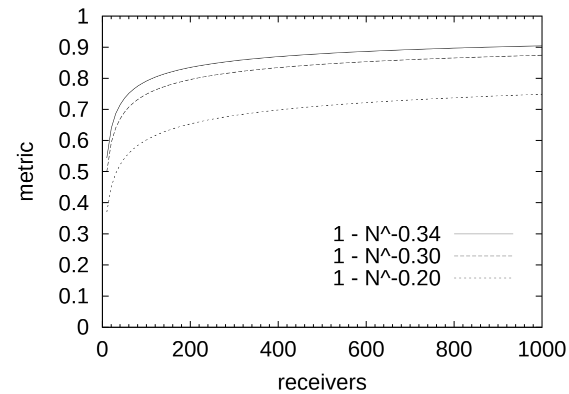

Fig. 1. Efficiency estimate shown over 1,000 receivers for a range of efficiency factors.

The estimate provides us with a characterization of multicast efficiency as a function of the number of members in the multicast group. This characterization seemingly ignores the influence of receiver dynamics, such as duration and distribution. Having derived the estimate from previous work [3], [4], it is known to behave well with generated trees and random distributions. Over the next two sections, we attempt to validate this characterization for real multicast usage by measuring the efficiency metric for sample multicast groups. We then further extend our confidence in the estimate by evaluating synthesized group distributions.

IV. Measurement Methodology

In this section we describe the datasets collected from the Internet that are used as a basis for our analysis. Then, we apply elements of these datasets and data collected through other research efforts to generate synthetic datasets. Finally, we detail how all datasets are processed to generate our results.

A. Real Datasets

Using our metric for multicast efficiency, we must now determine how the estimate reacts to real group dynamics. We consider the efficiency characterization valid if the measured values of L_m (multicast links), L_u (unicast links) and N (number of receivers) for real multicast groups fit the power-law in Equation 3 for some values of the efficiency factor, ε. Ultimately, to be useful, ε should be similar across the entire sample range.

We have collected topologies from four multicast sessions. Using the MHealth tool [14], multicast tree information was recorded during the 43rd meeting of the IETF and the NASA shuttle launch in February of 1999. Both events consisted of

[1] This of course is not completely accurate since [3] considered $\overline{L_u}$ as the expected distance between any two nodes in the network, and we have taken it to be the average path from source to receiver. However, any difference will effect the value of ε not the form of the equation.

TABLE I

REAL AND SYNTHETIC DATASETS USED IN EVALUATION.

Name	*Description*	*Trace Period*	*Total Recv.*	*Traced Recv.*
IETF43-A	43rd IETF Audio	Dec. 7-11, 1998	257	109
IETF43-V	43rd IETF Video	Dec. 7-11, 1998	305	131
NASA-A	NASA Shuttle Launch Audio	Feb. 14-21, 1999	144	43
NASA-V	NASA Shuttle Launch Video	Feb. 14-21, 1999	209	58
SYNTH-1	Synthetic Receiver Distribution	Jan. 6-10, 2000	11,000	1,286
SYNTH-2	Synthetic Receiver Distribution	Jan. 6-10, 2000	22,000	1,950

separate audio and video channels generated by a single source. The logs of these four sessions form the source of our initial datasets: IETF43-A, IETF43-V, NASA-A and NASA-V (see Table I).

There are some problems with the datasets worth mentioning. MHealth joins a particular multicast group and collects Real-Time Control Protocol (RTCP) [15] packets to identify the sources and receivers of the group and to subsequently track each member's activity. The tool then uses the *mtrace* (multicast traceroute) [16] utility to trace the path from each receiver back to the source. MHealth traces each receiver sequentially and may actually miss receivers who are part of the group for only a short duration. Thus, only a portion of the total number of receivers in the group are successfully traced, 43% for IETF and 29% for NASA. Furthermore, not all receivers are necessarily known since some decoding tools do not implement RTCP. Also, some RTCP packets may be filtered by firewalls. It is possible that the failed or missing traces are obscuring irregularities in the tree structure or that entire subnets hidden behind firewalls have extremely different properties than what is observed in the rest of the tree. However, we feel that the results of the experiments presented in the next section are definitive enough to consider these possibilities as remote.

B. Generated Datasets

In addition to validating the multicast characterization against real multicast groups, we wanted to investigate how efficiency was affected by generalizing group size, group member composition and tree paths. We attempted to trace multicast paths using a local source for a series of 11,000 IP addresses that were known to have participated in multicast groups over a sixteen month period (Nov.'99-June'99)(SYNTH-1 in Table I). The IP addresses were collected by the *mlisten* [10] tool which joins multicast groups advertised over the MBone's session directory tool, *sdr*, and collects RTCP packets from the group members. Of the 11,000 addresses, 1,286 were traceable. We then extended that dataset in SYNTH-2 by tracing another 664 IPs from a second set of 11,000 addresses collected from June'97 through June'98. While the traces were completed recently and represent the latest multicast infrastructure topology, the group members are a relatively random sample taken from the old MBone.

C. Data Processing

For each of the datasets, the RTCP and *mtrace* data are used to re-construct multicast trees as accurately as possible. However, since the multicast distribution tree changes as receivers join and leave the group, there is no single tree that represents the entire session. Our solution was to develop a new tool called *mwalk* that builds an "activity" graph of all possible trees over time. Ultimately, these trees are traversed, or "walked", in order to calculate the efficiency metric and other properties such as the maximum height, degree frequencies, and degree distributions at depth.

The RTCP data from either the MHealth logs or *mlisten* database is used to build an activity table for each receiver which lists the intervals when the receiver is known to have been a group member. Explicit RTCP *BYE* packets are considered as the end of an active period. Otherwise, a configurable time-out value limits active periods in cases where no explicit leave is received from a group member. From the mtraces, paths are built from each receiver back to the source, adding intermediate routers to the *mwalk* graph. The timestamp of each original trace determines which of a number of possible paths are active at any given time during the life of the session. The result is a graph in which each edge is associated with an activity table. For each node, only one parent link is active at any single moment, i.e. a snapshot of the multicast tree is available.

The entire session time is then partitioned into 10,000 sample periods.[2] The graph is walked for each period, using the link and receiver activity tables to determine the *active* tree. For each walk, we calculate the number of multicast hops (L_m), unicast hops (L_u) and the number of *active* receivers (N). Since the actual unicast paths between the source and each receiver at the time of the transmission are unknown, we have made a simplifying assumption that the length of the unicast and multicast paths are equivalent and have calculated L_u accordingly. *In real topologies, this of course is not necessarily the case.* We address this issue in Section V-A.2.

V. RESULTS AND ANALYSIS

In this section we analyze the results of our experiments in two parts: (1) multicast efficiency in relation to unicast and (2) common aspects of the shape of multicast trees.

[2] The resulting granularity is between 30 seconds and one minute. Shorter sample periods resulted in equivalent results, thus 10,000 periods were used to keep the computation time manageable.

A. Efficiency: Multicast versus Unicast

A.1 Results

Using the values collected from the *mwalk* trees, we analyze how temporal and spatial constraints affect the characterization introduced earlier, $\delta = 1 - N^{\varepsilon}$ (3). We specifically look at four key variables:

1. receiver duration,
2. receiver arrival time,
3. receiver distribution, and
4. number of receivers.

Starting from the most restrictive environment, we incrementally loosen the constraints placed on the tree instances. At each stage, we test whether the efficiency metric still conforms to the characterization, determining whether a dependence exists upon the variable of interest.

Figure 2 illustrates the path of our analysis. We begin (step a) with our real datasets (Table I). When evaluating the efficiency metric, we initially take into account the actual receiver distribution as well as the join time and duration of each member. Then, we stretch the duration (step b) until the receiver eventually remains joined from his/her initial arrival through the end of the session. Next, the receiver activity is randomized (step c) using a uniform distribution, thus ignoring the time domain completely. Finally, we stretch the spatial domain (step d) by randomizing the receiver distribution with synthesized datasets. Throughout the path so far, the tree topologies have been based on the actual multicast infrastructure of the Internet. The final step (step e) takes randomly generated topologies into account. Since the development of the characterization (see Section III-B) was originally based on randomly generated trees this step has already been validated in previous works.

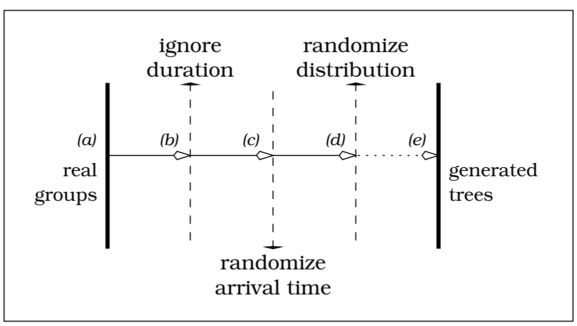

Fig. 2. An illustration of the path followed by the analysis in this section: transitions from real data to synthetic data.

Our initial objective, then, is to validate the efficiency estimate with as restrictive a case as possible (step a). With this in mind, we collect L_m, L_u and N using the actual inter-arrival pattern of the receivers and a small 1-minute time-out value to minimize any effects of artificially extending receiver durations. Using the IETF43 and NASA datasets, we plot the average metric at each receiver count. Figure 3 shows results for the IETF43-A dataset. Though not shown, the other datasets exhibit similar behavior. As can be seen from the graph, the data roughly fits the characterization with an efficiency factor, ε, of -0.34 (99% confidence intervals are shown with error bars).

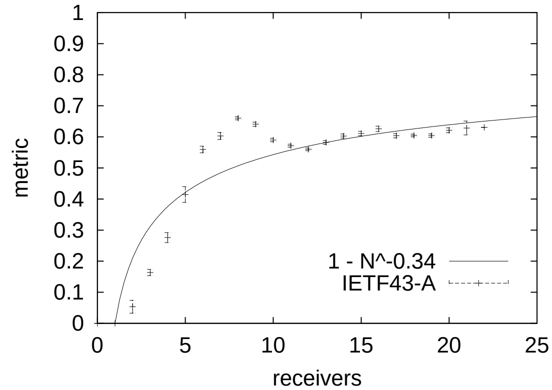

Fig. 3. (step a) Average metric calculated using a 1 minute time-out in the IETF43-A dataset (99% confidence intervals).

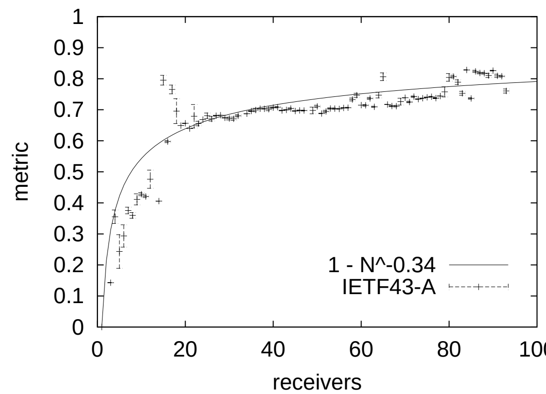

Fig. 4. (step b) Average metric calculated ignoring duration in the IETF43-A dataset (99% confidence intervals).

The resulting graph is not smooth, however, due to the rather low number of simultaneous receivers present in any one sample period (in this case around 20). This leads to an interesting observation about Figure 3. The higher-than-estimated efficiency gains noticeable between five and ten receivers is likely due to receiver clustering. Several receivers may join the group from the same or a relatively nearby leaf router. Consequently, they share a large portion of the same links and artificially inflate the metric. However, as the number of receivers increases to greater than fifteen, we can see that the efficiency metric begins to follow the characterization quite closely.

The next step (step b from Figure 2) in our analysis is to lengthen the time-out value to determine whether receiver duration is a contributing factor to the behavior of the metric. This also allows us to incrementally increase the number of active receivers since longer durations mean that more receivers are likely to be joined in any given period.

Lengthening receiver duration to 5 and then 20 minutes results in graphs similar to Figure 3 which are not included due to space constraints. We ignore the duration completely in Figure 4 by extending the time-out beyond the length of the actual session, i.e. once receivers join they remain members for the duration of the session. Each receiver still joins the group according to the original dataset, thus the inter-arrival time is not

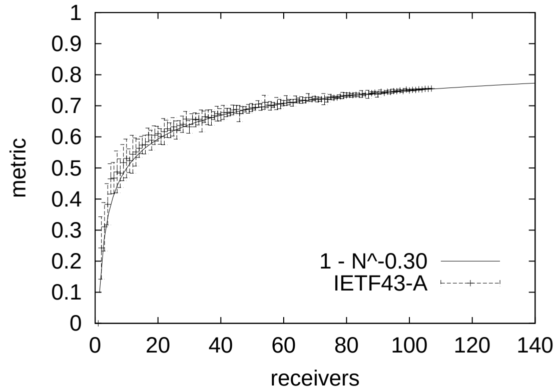

Fig. 5. (step c) Average metric calculated using random receiver activity and primary paths in the IETF43-A dataset (99% confidence intervals).

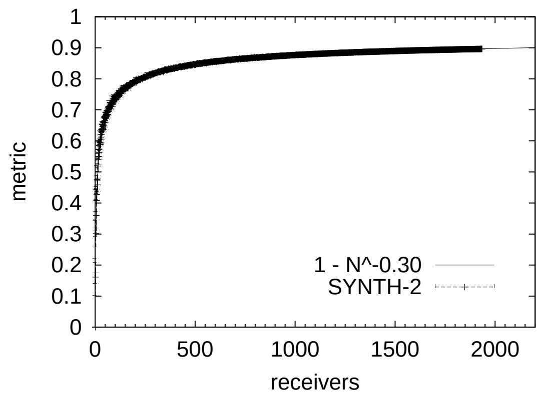

Fig. 6. (step d) Average metric calculated using the random receiver distribution of SYNTH-2 (99% confidence intervals).

adjusted. In all three cases, the metric stays fairly consistent with the efficiency characterization. The coefficient of determination[3] for each experiment is quite high, ranging between 0.94 and 0.97. From this, we conclude that the multicast characterization is not affected by the the total amount of time receivers stay in the group.

Next, we dismiss the time element entirely (step c from Figure 2). The results are shown in Figure 5. By ignoring the receiver arrival times, we explore whether the pattern in which the receivers originally joined the group has an effect on multicast efficiency. To determine receiver activity, we select ten sets of n randomly chosen receivers for each possible value of n, $1 <= n <= N$, where N is the total number of distinct receivers throughout the session. As can be seen by comparing Figures 4 and 5 (transition from step b to step c), ignoring actual receiver arrival time has no appreciable effect on the character of the efficiency metric. One discernible result, however, is that the magnitude of ε drops from 0.34 to 0.30 and the coefficient of determination rises to 0.99. Randomization in the time domain reduces outliers and seems to underestimate the potential benefit of multicast efficiency. This is possibly due to a reduction in receiver clustering as a result of temporal dependencies between joining members. Therefore, as we randomly distribute active receivers, we inadvertently reduce the amount of temporal clustering and in turn underestimate the actual multicast benefit.

Working with real data offers confidence in the efficiency estimate. The estimate correctly characterizes the data from real multicast groups and it does not appear to depend on group dynamics such as duration and inter-arrival time. However, to verify that the characterization is truly consistent with what might be found for most groups in the Internet, we felt it was necessary to further investigate how receiver distribution might affect multicast efficiency. In truth, one might expect receiver distribution to be one of the major factors governing how well multicast performs in respect to unicast since where the receivers are placed in the tree dictates how much sharing can possibly occur. Changing the placement of receivers in the network changes where branching occurs in the multicast tree.

The next step (step d in Figure 2) in our analysis, then, is to consider purely synthetic data, ignoring both temporal and spatial domains. The previous work [3], [4] cited earlier was based on random distributions for networks other than the modern Internet. Should their conclusions necessarily hold for randomly distributed receivers throughout the Internet? The SYNTH-1 and SYNTH-2 datasets consist of trees constructed from randomly chosen receivers and their corresponding paths traced to a local source through the current multicast infrastructure (see Section IV-B for details on the procedure used to select receivers). For each dataset, we select ten sets of n randomly chosen receivers for each possible value of n, $1 <= n <= N$, where N is the size of the dataset. This approach allows for much larger group trees to be constructed, nearly 2,000 receivers in the case of SYNTH-2. More importantly, the multicast topology is accurately represented. As can be seen in Figure 6, the characterization holds extremely well with a coefficient of determination of 0.99. As in the previous discussion, we see that randomization does reduce the efficiency factor slightly.

A.2 Discussion

What exactly have we accomplished? Most importantly, we have identified a characterization of multicast efficiency that is independent of the group distribution and behavior. The efficiency estimate, $\delta = 1 - N^{\varepsilon}$ (3), has been shown to hold rather well for real group dynamics, real group distributions and random distributions for the modern multicast infrastructure. Furthermore, we have found that the efficiency factor, ε, tends to stay within a fairly tight range between -0.30 and -0.34 for real networks.

This characterization is beneficial for the current set of ISPs who are considering whether to implement multicast. Using the metric estimate to calculate possible bandwidth savings, a network provider can compare the savings afforded by multicast against the cost of its implementation. In this vain, we have focused only on the bandwidth efficiency of data transmission. Other problems that face network providers considering multicast deployment, such as the increase in router state and other overhead specific to routing protocols, domain independence, address allocation and billing, have been covered in detail in other work [8], [9], [3], [2].

Another result of validating the efficiency metric has been

[3] The coefficient of determination is the square of the correlation coefficient, in the range 0 to 1. High values reflect a good linear fit on a log-log scale.

to indirectly support the conclusions of previous work [3], [4] with real data. However, the economies-of-scale factor, k, of Equation 1 seems to be nearer 0.7 than 0.8 for our experiments. The trouble with comparing these two numbers directly is that Chuang and Sirbu [3] computed L_m, $\overline{L_u}$ and N somewhat differently. They were interested in the cost of providing multicast to the leaf router, and, consequently, counted any router with at least one member to be a single receiver no matter how many actual receivers it served. For this paper, we wanted to look at the efficiency of the entire multicast tree. Ignoring a large number of receivers at the last-hop seriously underestimates multicast's efficiency. This difference in approach explains much of the difference in scale factors.

A point of divergence that still needs to be addressed is the assumption made in Section IV-C. We simplified the calculation of L_u by assuming that the unicast and multicast paths were of equivalent length. To test this assumption, we performed a series of unicast traceroutes [17] from the local source to each of the receivers in the SYNTH-1 dataset. Of the 1286 receivers, 1198 unicast paths were successfully recorded. Comparing these paths to the multicast paths used in the prior analysis (see Figure 7) we find that the assumption holds. The majority of multicast paths were in the range of 0.7 to 1.2 times the length of the corresponding unicast path with the worst case being twice the length. An interesting note about this graph is that most multicast paths were actually shorter than their unicast counterpart. This may be an indication that tunnels are still in use in the multicast infrastructure.

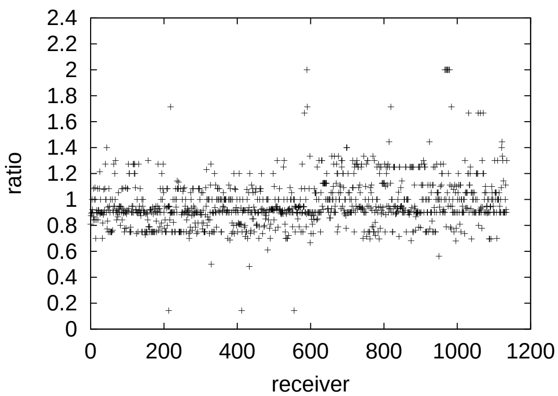

Fig. 7. Ratio of unicast to multicast path lengths for 1198 receivers in the SYNTH-1 dataset.

A final consideration, then, is the effect that changes in the multicast infrastructure, such as the migration from tunnels to native support, might have on the efficiency characteristic. Is the characterization simply a snapshot of the properties of the current topology or is it more fundamental? Can we expect the characterization to hold as multicast deployment continues? We note that our datasets have been gathered over a two-year period that has seen many changes in how multicast is implemented in the Internet. Furthermore, the characterization estimate was derived from previous work [3] that included the early MBone architecture(1996) as one of their sample topologies. As more backbone routers become multicast capable, the presence of tunnels has declined. As a result it has become more likely that unicast and multicast routing paths are similar. However, with the increased prevalence of inter-domain routing through MBGP, policy routing is beginning to become a major factor in resultant multicast topologies. The fact that the efficiency characteristic and even more remarkably the efficiency factor have not changed through the datasets is an indication that future changes in the multicast infrastructure should not cause significant change in how multicast efficiency compares to unicast

B. *Multicast Tree Shape*

In addition to understanding multicast efficiency, an important issue for many multicast researchers is to establish a method for generating trees that resemble real multicast distribution trees. Current research efforts in multicast routing protocols, congestion control and reliability would benefit greatly since the validity of their conclusions depend on the realism of the environment used in their evaluations. For this purpose, a characterization such as Equation 3 is very powerful in verifying that generated trees conform to the properties of real trees.

Much work [18], [19], [20], [5], [6] has been performed to determine the factors important in constructing realistic Internet topologies, but little effort [7], [21] has been made to classify what real multicast trees look like. We feel that it is important to look more closely at the trees from our datasets to investigate whether the shape of the tree is as universal as some of the higher-level relationships, e.g cost and efficiency. In other words, does the consistency seen in the efficiency metric imply that all real trees are similarly shaped, or that different shapes produce relatively similar efficiencies? If one were to use the efficiency metric to validate a generated tree, would it be possible to say that not only does this tree act (in terms of efficiency) like a real tree, but it also looks like a real tree? Our hypothesis is that efficiency alone does not capture the essence of a multicast tree; there are other factors.

As previously discussed in Section II, the important factors to consider are the height and breadth of the multicast tree. The key to relating this back to efficiency, however, is to know where in the height of the tree does the branching occur. It is not enough to say that the tree is tall, or that the tree is wide. Rather, one needs to understand when the tree widens since branching indicates duplication. Branching nearer the receiver improves efficiency; branching close to the source reduces efficiency.

In the following discussion of branching, we separate a node's *total* out-degree into two distinct components: (a) *internal* degree and (b) *leaf* degree (see Table II). The internal degree counts the number of out-going links from a router that connect to other routers. The leaf degree refers to the number of incident links directly connected to receivers. It is important to note that receivers do not just occur in the deepest part of the tree. By separating the total out-degree, we can investigate how branching is distributed throughout the body of the tree independent of the branching that occurs due to receiver clustering.

Figure 8 depicts the average of the total out-degree for nodes occurring at different levels of the distribution tree. Each of the real and random group distributions produces a similar graph. The majority of branching appears to occur near the top of the tree, but focusing only on the total out-degree can be misleading. Figure 9 shows the same graph but for the internal degree

TABLE II

DEFINITION OF TERMS INTRODUCED TO SEPARATE A NODE'S OUT-DEGREE INTO TWO DISTINCT COMPONENTS.

total out-degree	total number of out-going links incident to a node (internal + leaf degrees)
internal degree	number of out-going links from a router to other routers in the distribution tree
leaf degree	number of out-going links from a router to directly connected receivers

TABLE III

PROPERTIES OF PRIMARY TREES FROM EACH DATASET.

		Graph	*Max*	*Total Degree*			*Internal Degree*			*Leaf Degree*		
Dataset	*Nodes*	*Edges*	*Depth*	*Mean*	*Std.Dev*	*Max*	*Mean*	*Std.Dev*	*Max*	*Mean*	*Std.Dev*	*Max*
NASA-AV	333	397	23	1.34	0.80	5	1.34	0.82	5	1.28	0.70	5
IETF43-A	477	555	24	1.38	0.92	7	1.36	0.91	6	1.30	0.66	4
IETF43-V	533	614	24	1.40	0.98	7	1.40	0.95	7	1.30	0.77	5
SYNTH-1	2094	2113	22	2.64	5.40	108	1.57	1.18	12	4.17	8.21	108
SYNTH-2	2936	2986	22	3.07	6.79	133	1.57	1.27	13	5.21	10.28	133

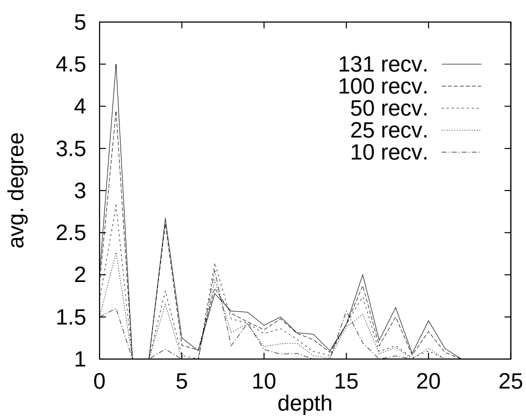

Fig. 8. Average **total** out-degree distribution for nodes at differing depths across a range of receiver set sizes of IETF43-V.

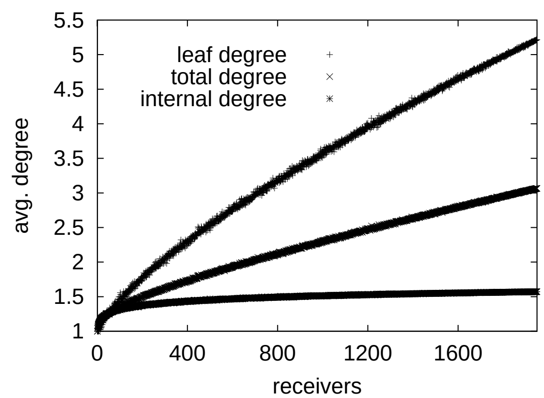

Fig. 10. Distribution of total, internal and leaf out-degree as the receiver set size grows from the SYNTH-2 dataset.

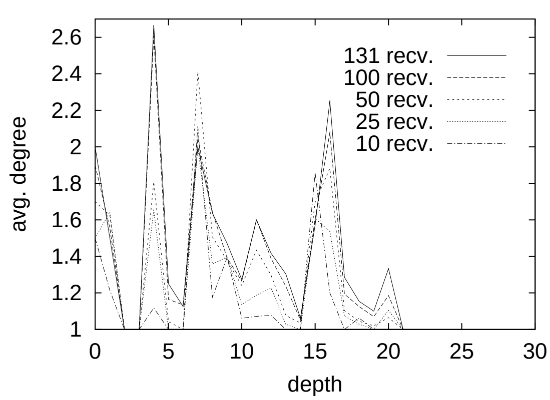

Fig. 9. Average **internal** out-degree distribution for nodes at differing depths across a range of receiver set sizes of IETF43-V.

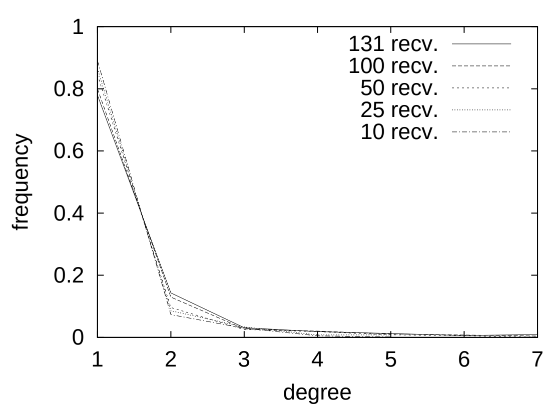

Fig. 11. Frequency of out-degrees across nodes for a number of receiver set sizes of IETF43-V.

rather than total out-degree. Comparing the two graphs shows that the initial spike in Figure 8 is due to receiver clustering near the source and is not representative of the body of the tree. From visual representations of the sample trees we see that branching does occur throughout the height of the tree. Although this observation in itself is not quantitatively compelling, it does have some qualitative value. It is valuable to note that looking at internal and leaf degrees independently may give a better view of how the body of multicast trees are shaped in relation to receiver clustering at the leaves.

Continuing to ignore leaf branching, we see in Table III that the average internal degree is consistent across datasets and quite low, ranging between 1.36 and 1.57[4]. Once we take the leaf links into account, however, the average total out-degree begins to climb quite rapidly, almost doubling in the case of SYNTH-2. Figure 10 illustrates how the average total out-degree depends on the internal and the leaf components. The leaf degree grows almost linearly with the number of receivers in the tree while the internal degree grows logarithmically. The internal degree tapers off quickly around 1.5. The total out-degree shadows the internal degree until the two degree components pass around 60 receivers, then the leaf degree dominates. The character of this relationship holds for all datasets, the primary difference being the steepness of the leaf degree distribution and thus at which point it begins to dominate the total out-degree.

The observation that the average internal degree is low begs the question: how frequently do large degrees occur within the tree? Figure 11 shows total out-degree plotted against frequency for the IETF43-V dataset. The graph is almost identical for every dataset, both real and generated, as well as across degree components. The majority of nodes in multicast trees have an out-degree of one. This supports previous work which claims that a large percentage of nodes in real multicast trees are routers with a single out-going link [7], [21]. These "relay" nodes simply pass the packet along the path to the receiver. In an experiment conducted in 1995 [7], almost 60% of non-leaf nodes were found to have an out-degree of only one. In our results, the proportion of relay nodes is closer to 80%. This increase may be due to the further development of native multicast routing and specifically to the fact that the majority of path divergence occurs in a limited number of MBGP peering points. An AS acting as a transit for a multicast tree is unlikely to allow branching within its domain.

These observations of degree frequency follow from the conclusions that the majority of routers in the Internet have a very low degree [5][5]. *The possible range of tree shapes are necessarily constrained by the underlying network connectivity.*

Existing graph generation techniques already take a number of factors into consideration to determine whether a graph is realistic [19]: hierarchy, average node degree, network diameter, bi-connected components. When generating random trees for testing multicast protocols, these factors can be extended to include those discussed here to better evaluate the realism of the shape of generated trees. In short, real multicast trees appear to exhibit:

- a high frequency of "relay" nodes through the body of each path;
- low average internal degrees that grow logarithmically with the number of receivers in the tree;
- average leaf degrees that grow almost linearly with the number of receivers in the tree; and
- a maximum height of approximately 23 nodes.

Finally, a high-level characterization such as cost (1) or efficiency (3) can be employed as a metric for validating the behavior of the tree.

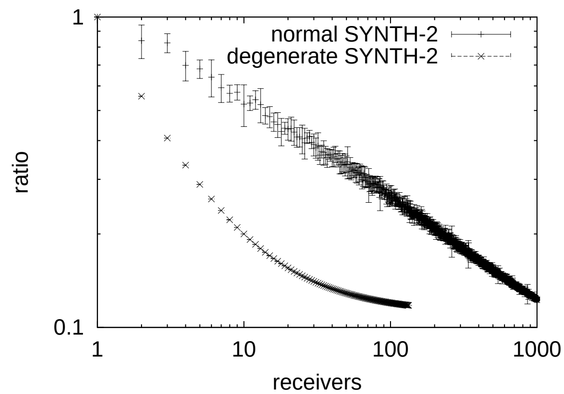

Fig. 12. Efficiency ratio $\frac{L_m}{L_u}$ in log-log scale for the normal SYNTH-2 dataset and a degenerate subset (99% confidence intervals).

VI. Future Work

With any useful investigation, there are often more new questions raised than existing questions answered. The more we come to understand multicast trees, the more we feel there is to know. As a result, there are a number of open issues left for future work.

A. Degenerate Cases

So far, we have focused on rather "normal" receiver distributions attempting to characterize how multicast behaves in the common case. We have not yet looked closely at truly degenerate cases such as extreme affinity or dis-affinity between receivers. It has been postulated [4] that affinity and dis-affinity generally do not affect the form of the cost function (1). One might assume that it would affect the economies-of-scale factor while maintaining the power-law relationship. However, looking at a case from the SYNTH-2 dataset where a collection of 133 receivers were chosen from a single subnet, we see in Figure 12 that the ratio of multicast links to unicast links is no longer linear when plotted on a log-log scale, i.e. no longer a power-law. This implies that extreme affinity does affect the general form of the relationship. Degenerate behavior like this might possibly occur in real multicast groups if extreme temporal or spatial dependencies exist between members of the group.

In future work, we plan to further investigate how degenerate cases affect the efficiency characterization. This will include understanding at what level of abnormality the relationship begins

[4]The average internal degree is fairly stable as the number of receivers increases, but the standard deviation is quite large. The actual range of degrees grows with the number of receivers. We did not feel that modeling the internal degree as a power-law was appropriate.

[5]Although the distribution is heavily skewed, our experiments do not indicate the existence of a power-law relationship for degree frequency in multicast trees similar to that seen in the underlying network.

to break down, and whether the change in the characterization could be useful in identifying the type of abnormality. The second point could prove very beneficial when generating trees for testing. If a tree does not conform to the efficiency characterization, it would be useful to know how the tree could be changed to ensure conformity.

B. Advanced Metrics

Another area of interest is in developing more advanced metrics for measuring multicast efficiency. In particular, we plan to look at techniques for weighting unicast streams to better capture the efficiency gains available through multicast. In the current model, each additional unicast stream that passes over a given link has an additive impact. In actuality, duplicating a stream over a link has implications beyond the extra bandwidth allocated to the stream, since that bandwidth is no longer available to other multicast and unicast streams. Applying metrics that more aggressively penalize duplicate unicast streams, such as multiplicative or logarithmic metrics [22] may give a more appropriate view of multicast's benefits. Another possible avenue is to weight the links themselves. Rather than simply using hop counts (weight of one), properties of the individual links such as capacity and delay may give a more accurate view of the impact of using multicast over unicast. Can these advanced metrics be characterized in the same way as the simple metric we have introduced here? Do changes in weighting simply affect the efficiency factor or do they radically change the form of the characterization?

VII. CONCLUSIONS

The focus of this paper has been to quantify multicast's bandwidth gain over unicast and to better understand what factors determine the realism of multicast trees. To this end, we have accomplished the following:

- developed a characterization of multicast that accurately describes its efficiency over a wide range of group dynamics and receiver distributions;
- validated previous work [3], [4] defining cost metrics for pricing multicast with real group data;
- expanded on previous work [5], [6] investigating skewed distributions in network properties to include properties of multicast trees;
- discovered that the range of possible shapes of inter-domain multicast trees is constrained by the underlying network connectivity;
- identified several properties of our sample groups that seem representative of real trees and that could prove useful for tree-generation tools; and
- collected a dataset, SYNTH-2, with nearly 2,000 receivers that has been shown to be representative of real multicast groups in both its efficiency characteristic and general shape. This collection of possible multicast trees will be made available to other researchers for generating realistic trees of varying sizes[6].

[6] Available at http://www.nmsl.cs.ucsb.edu/mwalk.

ACKNOWLEDGMENTS

The authors would like to acknowledge the suggestions of Josep M. Blanquer and Anurag Acharya which went a long way to making this work possible.

REFERENCES

[1] S. Jagannathan, K. Almeroth, and A. Acharya, "Topology sensitive congestion control for real-time multicast," in *Proceedings of the 10th International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV'00)*, Chapel Hill, North Carolina, USA, June 2000.

[2] C. Diot, B. N. Levine, B. Lyles, H. Kassem, and D. Balensiefen, "Deployment issues for the IP multicast service and architecture," *IEEE Network Magazine special issue on Multicasting*, vol. 14, no. 1, Jan. 2000.

[3] J. Chuang and M. Sirbu, "Pricing multicast communication: A cost based approach," in *Proceedings of INET'98*, Geneva, SWITZERLAND, July 1998.

[4] G. Phillips, S. Shenker, and H. Tangmunarunkit, "Scaling of multicast trees: Comments on the Chuang-Sirbu scaling law," in *Proceedings of ACM SIGCOMM'99*, Cambridge, Massachusetts, USA, Aug. 1999.

[5] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proceedings of ACM SIGCOMM'99*, Cambridge, Massachusetts, USA, Aug. 1999.

[6] A. Medina, I. Matta, and J. Byers, "On the origin of power laws in Internet topologies," *ACM Computer Communication Review*, vol. 30, no. 2, pp. 18–28, Apr. 2000.

[7] J. Pansiot and D. Grad, "On routes and multicast trees in the Internet," *ACM Computer Communication Review*, vol. 28, no. 1, pp. 41–50, Jan. 1998.

[8] L. Wei and D. Estrin, "Multicast routing in dense and sparse modes: Simulation study of tradeoffs and dynamics," in *Proceedings of IEEE Conference on Computer Communications and Networks (ICCCN'95)*, Las Vegas, Nevada, USA, Sept. 1995.

[9] K. L. Calvert, E. W. Zegura, and M. J. Donahoo, "Core selection methods for multicast routing," in *Proceedings of IEEE Conference on Computer Communications and Networks (ICCCN'95)*, Las Vegas, Nevada, USA, Sept. 1995.

[10] K. Almeroth, "A long-term analysis of growth and usage patterns in the multicast backbone (MBone)," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'00)*, Tel Aviv, ISREAL, Mar. 2000.

[11] M. Handley, "An examination of MBone performance," Tech. Rep., USC/ISI, Jan. 1997.

[12] M. Yajnik, J. Kurose, and D. Towsley, "Packet loss correlation in the MBone multicast network," in *Proceedings of IEEE Global Internet Conference (GLOBECOM'96)*, London, UK, Nov. 1996.

[13] R. Chalmers and K. Almeroth, "Developing a multicast metric," in *Proceedings of IEEE Global Internet (GLOBECOM'00)*, San Francisco, California, USA, Nov. 2000.

[14] D. Makofske and K. Almeroth, "MHealth: A real-time graphical multicast monitoring tool," in *Proceedings of Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV'99)*, Basking Ridge, New Jersey, USA, June 1999.

[15] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," Tech. Rep. draft-ietf-avt-rtp-new-*.ps, Internet Engineering Task Force (IETF), Feb. 1999.

[16] W. Fenner and S. Casner, "A 'traceroute' facility for IP multicast," Tech. Rep. draft-ietf-idmr-traceroute-ipm-*.txt, Internet Engineering Task Force (IETF), Aug. 1998.

[17] V. Jacobson, *Traceroute*, Lawrence Berkeley Laboratory (LBL), Feb. 1989, available at ftp://ee.lbl.gov/traceroute.tar.Z.

[18] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, pp. 1617–1622, Dec. 1988.

[19] E. W. Zegura, K. L. Calvert, and M. J. Donahoo, "A quantitative comparison of graph-based models for Internet topology," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 770–783, Dec. 1997.

[20] M. B. Doar, "A better model for generating test networks," in *Proceedings of IEEE Global Internet (GLOBECOM'97)*, London, UK, Nov. 1997.

[21] I. Stoica, T.S. Eugene Ng, and H. Zhang, "REUNITE: A recursive unicast approach to multicast," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'00)*, Tel Aviv, ISREAL, Mar. 2000.

[22] A. Legout, J. Nonnenmacher, and E. W. Biersack, "Bandwidth allocation policy for unicast and multicast flows," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM'99)*, New York, USA, Mar. 1999.