Efficient broadband multi-media data distribution over the Internet using Multicast

Harish Ramamurthy*, Abhay Karandikar and Rajit Gadh***

*Wireless Internet for the Mobile Enterprise Consortium, wireless@winmec.ucla.edu
University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, CA – 90095
**Department of Electrical Engineering, IIT Bombay, Powai, Mumbai, India.

Abstract

Receiver heterogeneity severely constrains the performance of multicast transmission schemes. In typical multicast transmission schemes, either the network resources are inefficiently utilized or the slowest receiver dictates the performance of the session. Intuitively, in the case where one bottleneck receiver is degrading the performance of other receivers, it would seem that isolating the bottleneck receiver and carrying out separate communication with it will improve the performance of the multicast transmission. This notion of isolation can be extended to the case where the receivers are partitioned into disjoint groups and the sender carries out separate conversation with each sub-group formed. In our earlier work [1], we had studied this problem in the context of reliable multicast data transmission. We had formalized the notion of separation and had proposed a grouping scheme that partitions receivers based on their bottleneck bandwidth and RTT. However, in the formulation of [1], a-priori knowledge of the multicast tree was assumed and hence topology was not considered while forming groups. Simulation results have shown that topology indeed plays an important role in the performance of grouping. In this paper, we extend the grouping formulation of [1] to include the effects of topology in the context of unreliable multicast. Specifically, we study the problem of joint grouping and routing for unreliable multicast. We formulate the problem of partitioning receivers using a rate-based approach and attempt to design an effective rate strategy. We present some heuristics to solve the joint routing-grouping problem and carry out simulation studies of the proposed heuristics. The preliminary simulation results of the scheme show that there is definite improvement in the performance of multicast sessions. The paper also discusses some of the important issues concerning the grouping schemes and proposes a simple framework to address these issues.

1. Introduction

In the future, video and similar high-bandwidth content like music, graphics etc. will be delivered increasingly via the Internet than other media like cable, satellites etc. Content delivery will be based on IP networks in which packets of data are used to deliver content to a set-top box or some similar interface sitting within the customer's home. Content delivery today is based primarily on two techniques: 1) Broadcast, and 2) Unicast. Both of these techniques [2] are neither efficient nor scalable, in terms of network resource utilization and number of customers respectively. Multicast is an efficient multipoint communication paradigm where data is sent simultaneously to multiple receivers.

Content distribution using multicast over Internet is plagued by the two problems of best-effort Internet, namely, efficiently adapting to network dynamics and addressing the problem of receiver heterogeneity. Network scenarios are dynamic in nature and hence data transmission techniques should self-adapt depending on the network scenario to utilize the network efficiently. Although multicast extension to the Internet (popularly called IP multicast) is reasonably scalable in terms of control overheads, it provides only a basic underlying framework for multi-point communications and thus higher-level adaptation schemes need to be built over it.

Adaptive data transmissions over the Internet are also constrained by the heterogeneous nature of receivers both in terms of numbers and their capacities. First, the size of a multicast session may vary from tens of receivers for videoconferencing to thousands of receivers for Web TV. Secondly, receivers in a network are heterogeneous in terms of their processing capabilities and capacities. Thus it is extremely difficult to come up with a one-size-fits-all adaptive transmission scheme and, thus, several solutions addressing various issues and challenges of adaptive data transmission have been proposed in recent years.

1.1 Typical Application Scenarios

Content delivery requirements differ depending on the underlying application. For example, distributing software and financial information requires reliable data delivery. In contrast, video distribution applications can tolerates some data loss. Thus, content delivery systems can be classified as reliable and unreliable. In this paper, we shall lay our focus on unreliable content delivery and we shall extend our knowledge of content delivery for reliable transmissions to unreliable transmissions.

Furthermore, user demands and preferences for media content are typically heterogeneous in nature. To illustrate more clearly let us consider the streaming of Winter Olympics. Some users may prefer to watch ice hockey while some will prefer to watch figure skating. Broadcasting select content though may be scalable to large number of receivers, but isn't network efficient and may not cater effectively the heterogeneous demands of user. Moreover, streaming personally desired content to each user through unicast addresses the problem of heterogeneity but is not network efficient. It is at such places the selective forwarding/streaming capability of multicast could be exploited. Another typical scenario is the streaming of new movie releases. Here again choices widely differ and hence selective forwarding will help improve the network utilization, while catering the heterogeneous demand of users effectively.

1.2 Current Trends

We are increasing seeing the use of the Internet for content delivery, even though most content delivery today isn't done on demand.

i. Disney has recently unveiled a new, low-cost broadband system that company executives say will become a common way for distributing movies, television shows and music around the world. Their content distribution system uses a technique called "data-casting," which assumes that the receivers are intelligent. [4]

ii. Alcatel has recently joined NDS Group Inc. to deliver secure broadband entertainment solutions. Like Disney, their system uses data-casting for content distribution. [5]

iii. Texas Instruments has announced extended support for broadband solutions. According to TI: "The broadband systems in development today are increasingly complex; the evolution to residential gateways and integrated access devices (IADs) requires integrating cable or DSL access, home and office networking, and voice over IP. TI's broadband solutions are built to help customers simplify their design and development processes and reduce the cost, complexity, and time it takes to get their products into the market." [6]

iv. Dotcast Inc., a privately-held broadband communications technology company, has developed digital broadcast technologies to leverage the existing commercial television broadcasting infrastructure for the economical distribution of content like movies, graphics, and music. [7]

1.3 Organization of the paper

In this paper, we present a survey of content distribution schemes over the Internet. We will primarily concentrate on network transport techniques, with an emphasis on using multicast for data transmission. We begin by giving a brief overview of currently used multicast schemes in Section 2 and then present our approach. We also present the implementation issues of the scheme in Section 4 and finally present some of the preliminary simulation results in Section 5. Section 6 concludes the paper and Section 7 gives future directions of work.

2. Multicast - Background

2.1 Overview

Multicast [2] is a networking transmission paradigm where multiple receivers simultaneously receive the same transmission. Unlike broadcast, here the data is sent to a specific group of receivers. We consider content distribution techniques which are maintained end-to-end i.e. we don't assume the presence of active nodes or intelligent nodes in the network. With the current Internet infrastructure, end-to-end techniques are appropriate, as they will help speedy deployment of content delivery services. The two primary approaches that have been used for multicast transmissions (end-to-end) are based on single-rate and multi-rate schemes. Based on the implementation, multi-rate schemes are further divided into simulcast and layered schemes. We now briefly discuss the different schemes used for multicast transport.

Single-rate schemes

In single-rate schemes (e.g. pgmcc [8], MTCP [9]), the source (or server) sends data to all the receivers (or users) at the same rate, which is amenable to all receivers. The transmission rate is determined by the slowest receiver of the group in order to prevent network congestion. Unfortunately this scheme has limited scalability as a single slow receiver can lead to low throughput and link utilization for the rest of the receivers. Single rate schemes are typically employed for reliable multicast as they can guarantee reliable data transmission. They are also used in video-conferencing applications, where limited number of receivers are involved but are inter-connected using high bandwidth links.

Multi-rate multicast

In multi-rate schemes (e.g. RLM [10], FLIDDL [11]), data is transmitted at different rates over multiple layers (using, for example, a layered source coding) or streams (simulcast). Depending upon their bandwidth capacities, the receivers can now subscribe to one or more layers. This strategy is well suited for typical networks, where the receiver bandwidths are correlated to each other as they use standard access interfaces—for example, a 56-Kbps phone line, a T1/E1 line (1.5Mbps/2Mbps) or a 10-Mbps switched Ethernet.

<u>Simulcast</u>

In simulcast, receiver bandwidth dependencies are exploited by using data distribution with limited number of streams, with each stream-rate being the same as the correlated cluster bandwidth. All the streams carry the same data, but they are all independent and carry data at different rates. The most common implementation of simulcast has the source transmitting three streams carrying low, medium, and high-quality versions of the content/data.

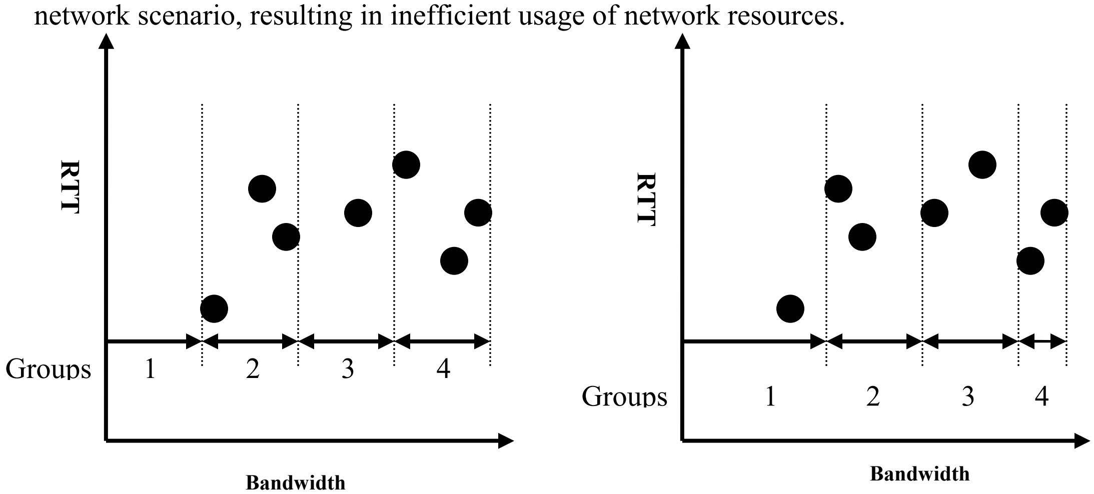
Layered schemes

Inefficient use of network resources due to data replication remains a major drawback of simulcast. In the layered approach, the content (e.g. video) is compressed into non-overlapped streams or layers. There is a base layer, which contains the data representing the most important features of the video. Additional layers (typically numbered 3-5), called enhancement layers, contain data that progressively refines the reconstructed video quality. One of the important layered schemes is the Receiver-driven Layered Multicast (RLM) [10]. In RLM, the sender transmits each video layer over a separate multicast group. It predetermines the number of layers as well as their rates. Receivers periodically join higher layer's group to explore the available bandwidth. If packet loss exceeds some threshold after the join experiment— that is, when congestion occurs—the receivers either leave the group or stay at the new subscription level.

Multi-rate schemes are used in many commercial video-streaming systems. For example, Launch, the online music subsidiary of Yahoo Inc. uses multi-rate schemes for multimedia distribution. Further, RealNetworks' RealSystem G2 supports simulcast under the name of SureStream [21], which generates a fixed number of streams at prescribed rates, and a receiver can dynamically choose a stream commensurate with its bandwidth.

Multi-rate multicast transmission is an extremely promising approach for content distribution over the Internet. Though multi-rate schemes are able to solve the problem of receiver heterogeneity effectively, the current schemes still do not use the network efficiently. Network scenarios are dynamic in nature and simple classification of receivers like low/med/high is far too restrictive. An ideal multi-rate multicast scheme is one that can respond to the changing network scenario and receivers quickly and effectively use the network resources.

The concept of multi-rates schemes can be further generalized by using the concept of partitioning receivers based on certain criteria. In our previous work [1], we had investigated such partitioning of receivers of a multicast session based on their bandwidth & Round Trip Time (RTT) in the context of reliable multicast. Receivers are partitioned into disjoint sub-groups [1] and separate communication is carried out with each sub-group formed. To illustrate the idea more clearly, let us consider a general network that has heterogeneous receivers. Receivers in a network can be characterized by their bandwidth and RTT and hence can be represented as points in a 2-dimensional space of RTT and bandwidth (see Figure 1). If the regular multi-rate schemes are used, sessions will be held with pre-defined rates (shown by dotted lines), irrespective of network scenario, resulting in inefficient usage of network resources.



Pre-defined multi-rates schemes	The Grouping scheme

Figure 1: Receivers as points in a 2D plane of RTT, Bandwidth

However in the grouping case the rates will be determined by considering the network scenario, thereby making receivers more responsive. From Figure 1, one can infer that the rates have been chosen (or the vertical lines marked) such that none of the receivers constrain the session's performance. Thus, all the receivers are serviced to their capacities. Furthermore, advancements in video codec technologies may allow sub-groups to be serviced at any specified rate.

As mentioned earlier receiver bandwidths are typically correlated as they are connected to standard interfaces such as T1/E1-line, ISDN etc. or as they share a bottleneck link. Moreover network resources (post-grouping) will be shared by lesser number of receivers; hence, such partitioning will help improve the performance of receivers. This holds true for both reliable and unreliable multicast and hence grouping will help improve the performance of a multicast session.

The proposed approach performs better than the traditional multi-rate techniques as (i) It efficiently utilizes the network resources (primarily bandwidth), (ii) It is scalable when the numbers of users becomes large, and (iii) It addresses the heterogeneous demand for data, i.e., changing demand by user and time. These are the typical requirements of multimedia distribution applications.

The grouping problem can be considered as a problem in which the aim is to divide the receivers into *n* groups in such a way that each group is as *similar* (non-constraining to each others performance) as possible. Note that on the one extreme, we can have as many groups as the number of receivers, i.e. each group has exactly one receiver. This would mean that the sender carries a unicast conversation with each receiver. But then the whole advantage of using multicast communication may be lost. Moreover, due to link sharing among receivers of the same multicast session, the performance of receivers may even worsen. On the other extreme, we can have only one multicast group. As pointed out earlier, the performance in this case is dictated by the slowest receiver or the network resources are not efficiently used.

2.2 Proposed Approach

An important component of multicast transmission is the formation of the multicast tree. A multicast tree is a distribution tree setup by the sender to transmit data to the multicast session. Traditionally, a multicast tree is formed by minimization of certain cost metric on the paths from the sender to the receivers. For example in DVMRP [12], the cost metric used is the number of hops. A multicast tree formed in conjunction with grouping will improve the performance of multicast session as the distribution tree will be set utilizing the common paths to receivers in the network. We term the problem of finding multicast tree in conjunction with grouping as the joint routing-grouping problem.

We formulate the joint routing-grouping problem using a rate-based approach and attempt to design an effective rate strategy based on the grouping framework. We focus on solving the problem in the context of unreliable multicast. An effective rate control strategy has to achieve the twin objectives of efficient usage and fair distribution of network resources. Efficient usage of network resources entails controlling congestion in the network and concurrently ensuring that the load offered to the network by different sources remain within the limits that the network can carry. The second objective entails *fair* distribution of network resources among competing streams of traffic. Ensuring fairness is simplified by the use of utility functions [13]. The utility function relates the allocated bandwidth to the "value" of the bandwidth to the receiver. The "value" could be some measure of user satisfaction; e.g. perceived video quality. Thus finding an efficient rate control strategy entails maximizing the receiver utility functions subject to network resource limits.

Grouping receivers may clearly improve the performance of the receivers, as the network resources (post-grouping) will be shared efficiently by network sources. Note that the primary motive of the grouping scheme should be to maximize the performance of receivers. We argue that if the performance maximization is kept as the primary motive, the receivers will automatically fall into an optimal grouping arrangement.

2.3 Related Work

Though several authors have studied the problem of fair allocation of resources in a unicast scenario [13, 14], the case of fair allocation in multicast sessions [15] has received less attention. It may be noted that the problem of fair allocation in multicast scenario is significantly different and more complex than its unicast version. For instance, the problem in multicast scenario is non separable, non-differentiable, unlike the unicast case. The reader is encouraged to refer [15] for more details.

In [15] the authors have formulated the optimal grouping problem by maximizing aggregate receiver utilities. The solution approach is based on the dual methods. They have formulated the optimal grouping problem using dual methods and have proposed a distributed, scalable implementation of the same.

Further, in [16], the authors have analyzed the problem of optimal grouping in the context of layered (unreliable) multicast. They formulate the optimal grouping problem as a list-partitioning problem. After proving the equivalence of optimal paging problem in wireless networks and list partitioning problem, they use standard techniques developed for solving the optimal paging problem to find solutions for the optimal grouping problem.

Though the above two schemes effectively address the problem of partitioning receivers for unreliable multicast scenarios, they are very computationally intensive over the network. We now present our formulation of the optimal grouping scheme in the context of unreliable multicast.

2.4 Problem Formulation

Consider a network consisting of a set $\boldsymbol{L}$ of unidirectional links, where a link $\boldsymbol{l} \in \boldsymbol{L}$ has capacity $\boldsymbol{\mu_l}$. The problem is to divide the receivers into sub-groups in such a way that the aggregate utility of receivers is maximized. The number of sub-groups (say $\boldsymbol{n}$) to be formed is unknown. Since the sender will maintain separate conversation with each sub-group, these sub-groups will share the network.

Each sub-group ($\boldsymbol{g_i}$, where $\boldsymbol{i=1, 2, 3 \dots n}$) is associated with a unique set of receivers and a set of links that the sub-group uses. Thus, each sub-group $\boldsymbol{g_i}$ can be specified by the set $\boldsymbol{\{R_i, L_i\}}$, where $\boldsymbol{R_i}$ is the list of receivers and $\boldsymbol{L_i}$ is the list of links. Let $\boldsymbol{R = \{R_1 \bigcup R_2 \dots R_n\}}$ denote the entire list of receivers and let $\boldsymbol{S_l}$ denote the set of sub-groups using link $\boldsymbol{l}$. We assume that the multicast or distribution tree is known in advance. We also assume fixed path routing for receivers. Each receiver $\boldsymbol{r \in R}$ is associated with a utility function $\boldsymbol{U_r(x_r)}$, where $\boldsymbol{x_r}$ is the rate at which the receiver $\boldsymbol{r}$ receives data.

We are interested in maximizing the aggregate utilities of receivers subject to the link constraints and grouping constraints. We first define the following terms:

n : Number of sub-groups

G: $\{g_i; i=1, 2 \ldots n\}$, set of sub-groups

r_{g_i} : Data reception rate of sub-group g_i

$B(g)$: Bottleneck of group g

The optimal grouping problem can be posed as an optimization problem as follows:

$$\max \sum_{\forall R_i} \sum_{r \in R_i} U_r(r_{g_i}) \tag{1}$$

subject to the link constraints:

$$\sum_{i \in S_l} r_{g_i} \leq \mu_l \tag{2}$$

Assuming same utility function ***U(.)*** for all receivers (refer [14]) the problem can be reposed as:

$$\max \sum_{i=1}^{i=n} U(r_{g_i}) \|g_i\| \tag{3}$$

subject to the link constraints:

$$\sum_{i \in S_l} r_{g_i} \leq \mu_l \tag{4}$$

here $\|g_i\|$ denotes the cardinality of the $\boldsymbol{i^{th}}$ sub-group.

Solving the above optimization involves determining the optimal grouping arrangement, i.e. set of sub-groups g_i , and the optimal rates r_{g_i} at which these sub-groups will be serviced. Determining solutions to the above optimization problem subject to the constraints is similar to the Steiner tree problem [17], which is known to be a NP-hard combinatorial problem. Heuristics are hence needed to solve the joint routing-grouping problem. We have developed an iterative technique to solve the same and we shall present it in the next section.

3. Implementation Issues

Implementation issues are a major component of any network architecture/algorithm. The more complex the implementation gets, the lower is the probability of it being deployed in the real network. We now look at some of the implementation issues of the proposed grouping scheme, specifically in the context of multicast transmissions.

3.1 Optimal Grouping Arrangement

The optimal grouping arrangement can be determined by finding the solution to the grouping problem. However as mentioned earlier, finding a solution to the grouping problem is an *np* hard combinatorial problem. Thus, we have developed an iterative technique to solve the problem albeit sub-optimal. The basic scheme is as follows: We start by placing all receivers in separate groups and run the below algorithms on them. The final result is the optimal grouping arrangement. The proposed scheme consists of two phases.

Phase I – Removal of common bottlenecks

We determine the bottleneck link of each group. We identify groups having same bottleneck link and merge them together. This will help improve the performance of receivers in the merged group. We re-iterate the phase until no two groups have the same bottleneck link. The iterative scheme is shown in Figure 2 below. We start by keeping each receiver in separate group. Let ***B(g)*** denote the bottleneck link of group ***g***.

Phase II – Elimination of *Similar* Groups

The removal of common bottleneck phase ensures that each group has a unique bottleneck link in the network. Although two (or multiple) groups though may not have a common bottleneck link, they still might be indirectly constraining the performance of each other. This is particularly important when non-equal sharing of network resources is made. Thus there exists a need to merge and test. Furthermore, two or multiple groups may be serviced at *similar* rates. Though, merging these groups may not improve the performance of the receivers and may even marginally worsen the performance. But considering the overhead of maintaining separate groups at sender, the groups are merged together. The iterative scheme is shown in Figure 3 below:

```
Global: The network, ie., B(g) ∀g
Input: Set of sub-groups (g_i; i : 1 ≤ i ≤ n)
Output: The optimal grouping arrangement

foundOptimal=true
for all i : 1 ≤ i ≤ n do
   for all j : i + 1 ≤ j ≤ n do
      if ||g_i||!=null && ||g_j||!=null then
         if B(g_i) == B(g_j) then
            foundOptimal=false
            merge(g_i, g_j) → g_j
         end if
      end if
   end for
end for
if (foundOptimal) then
   return optimise(G)
else
   Apply Algorithm 1 with present grouping arrangement G
end if

Here ||g_i|| denotes the cardinality of the sub-group g_i.
The merge(g_i, g_j) function returns the merged sub-group of g_i and g_j.
```

Figure 2: Algorithm for removing the common bottlenecks

Function: optimise(.)

Inputs: Grouping arrangement G, δ: Tolerance

Output: Grouping arrangement after minimizing the number of groups

```
DONE=true
for all i : 1 ≤ i ≤ n do
    for all j : i+1 ≤ j ≤ n do
        if ||g_i||!=null && ||g_j||!=null then
            if |g_i − g_j| ≤ δ then
                DONE=false
                merge(g_i, g_j) → g_j
            end if
        end if
    end for
end for
if (DONE) then
    return G
else
    return(optimise(G))
end if
```

Figure 3: Algorithm for eliminating similar groups

3.2 Estimation of network parameters

Network parameters like bandwidth, RTT are used for varied purposes by data transmission schemes. For example, round trip time (RTT) is generally used to set the re-transmission timeout in **pgmcc** [8]. Network parameters also form an important metric to classify receivers. This concept is exploited by grouping schemes. Our grouping scheme for reliable multicast, for example, uses bottleneck bandwidth and RTT for grouping receivers. All grouping schemes, depending on the grouping criteria used, may require estimation of certain (or some combination of) network parameters. For example, grouping used for error control purposes [18], which uses packet loss rate as grouping criteria, requires an estimation of the packet loss rate. In general, estimates of network parameters may be required for numerous functions. Although several authors have proposed several techniques for the accurate estimation of such network parameters in unicast scenarios, these solutions generate far too many control message exchanges to be scalable in a multicast scenario.

We present a hierarchical technique for efficient and scalable estimation of network parameters on a single path. The scheme can be easily extended to the case where the network parameters of multiple-paths need to be determined. By utilizing router support, our scheme avoids re-estimation of network parameters at common paths. We explain the scheme for the estimation of bottleneck bandwidth and RTT. Though we present the hierarchical scheme for bottleneck bandwidth and RTT, the scheme may be extended, with some minor modifications, to estimate other network parameters such as delay and throughput too.

The approach is as follows. The intermediate network elements store two states, namely the RTT and bottleneck bandwidth (BBW). Unless explicitly specified, we will use RTT and BBW to signify the RTT and BBW of the sender-receiver pair in consideration. Whenever a receiver needs to estimate RTT and BBW along a certain path, it sends a request to its parent router (the parent router is the next hop router on the path to the sender). Upon receiving the estimation packets, the router checks whether the BBW and RTT states are set or not. If the states have been set, it replies back to the receiver with RTT and BBW information. The receiver then uses a filtering algorithm (explained later) to get a good estimate of BBW and RTT.

Every intermediate router maintains the information (RTT, BBW) till sender and runs the following routine:

```
bbw(sender) = Bottleneck bandwidth till sender
rtt(sender) = RTT till sender
request(sender): Request to calculate the (RTT,BBW) till sender
present: The intermediate router
estimate_on = false, flag to check whether for any ongoing
estimation
reply(sender): Estimation packets received from the sender.
W = Set of nodes waiting for estimation of parameters

if (request(sender))
    if(bbw(sender)!=null AND rtt(sender)!=null)
        send(rtt,bbw) to receiver
    else
        if(estimate_on = true)
            W <- wait(receiver);
        else
            send(request(present,sender)) to the parent node
            estimate_on=true;
        end
    end
elseif (reply(sender))
    process_packet();
    for all W's
        send(rtt,bbw) to w
end
```

Figure 5: Hierarchical technique to estimate network parameters

If the request reaches the sender, the sender replies back with rtt=0 and bbw=infinite to the requesting node.

The proposed scheme being a hierarchical one, distributes the network load for estimating network parameters over the entire network. The problem of feedback implosion is avoided; the intermediate routers and sender are limited to feedback only from their immediate children. This ensures the scalability of the scheme in large multicast scenarios. In our proposed scheme, all calculations involved in the estimation of RTT and BBW are done at one place thereby avoiding the problem of clock synchronization. The RTT and BBW states set by the intermediate network elements and are in soft states that expire after periodic time intervals. The expiry period or state timeout can be set according to application needs, i.e. by considering how routinely the BBW and

RTT need to be dynamically updated. But there exists a tradeoff in choosing an appropriate state timeout value. If the timeout is low, more number of control messages has to be sent to update the RTT and BBW, but the estimates will be more accurate. In our proposed scheme, the child nodes use their parent node's RTT and BBW estimates in order to estimate their own RTT and BBW. These results in saving large number of control messages that would have been used for estimating the BBW and RTT of common paths.

In conventional RTT measurement algorithms, the receivers spend RTT amount of time to estimate the RTT. Schemes which estimate bandwidth using the packet pair technique also require and RTT amount of time to make their estimates. However, in our proposed scheme, the time required to estimate RTT and the bandwidth are always less than the RTT. In the best case, the time required would be equal to the delay between the receiver and its parent link and in the worst case; the time required would be RTT, when the request packet reaches goes till the sender. Additionally, the two cases are not equally probable as the estimations are carried out only once on common paths. Hence it is more likely that the receivers (in a multicast scenario) would require an amount of time less than the RTT in order to estimate the RTT. In the proposed scheme, the bottleneck bandwidth and RTT are estimated on the multicast path used by the data in a multicast session. The above advantages make the scheme efficient and scalable for multicast sessions.

In [19], the authors propose an alternative hierarchical technique for estimating receiver RTTs. Unfortunately, their technique has high resource demands because the intermediate routers have to store a number of states that is equal to the number of immediate children. In contrast, our proposed technique only stores one state per session in the intermediate routers. Using [19], estimations require the exchange of two sets of control messages between the sender and receiver whereas in our proposed scheme an exchange of one set of control messages is sufficient.

In our previous work [23] we examined grouping scheme implementation issues such as dynamic grouping, efficient feedback etc. As the network scenarios change, the grouping decisions need to be dynamic and should align appropriately to the network conditions. Thus, efficient dynamic grouping is an important issue in grouping scheme development. Readers are encouraged to read [23] for more details.

4. Results

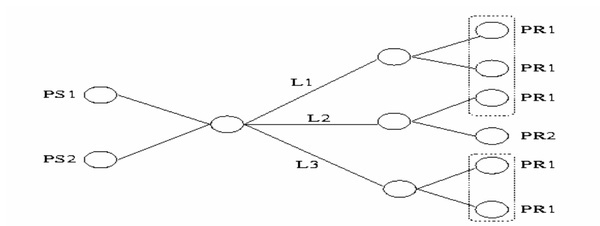
We have evaluated the performance of our scheme in various scenarios in the context of reliable multicast. Simulations for unreliable multicast are being performed currently and will be reported separately. But we believe as grouping will help utilize the network resources efficiently, it will help improve the performance of unreliable multicast too.

We now present the simulation results for reliable multicast. We have tested the performance of our scheme by simulating it in the presence of the **`pgmcc`** [8] protocol. **`pgmcc`** is a single-rate representative-based multicast congestion control scheme that sets its window according to a group representative called the *acker*. The acker is chosen as the receiver with the worst throughput among group members. A window based congestion control scheme similar to TCP is employed between the acker and the sender. Since our scheme requires the source to carry out separate conversations with each group, for the purpose of simulations, the source maintains separate **`pgmcc`** window sizes for each group. Although we have evaluated the performance of our scheme with **`pgmcc`**, the scheme is also expected to perform well with other single-rate schemes like MTCP [9].

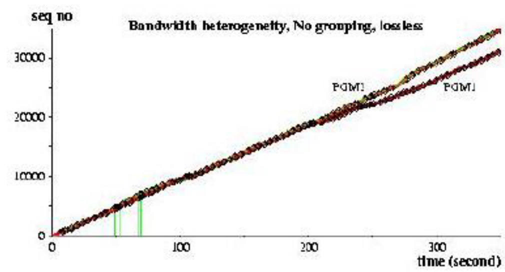
The simulations have been performed using the **ns v2.1b5** network simulator [20]. We use the **pgmcc** patch files available at [21]. We now present the simulation results of our scheme in one of the scenarios as shown in Figure 4.1. Figure 4.2, 4.3 and 4.4 represent the performance of receivers in the No grouping, Static Grouping and Dynamic Grouping case. Here we have two senders designated by PS1, PS2 and set of associated receivers PR1 and PR2. A multicast session is characterized by the set of a sender and group of receivers (PGM1 and PGM2 here). Figure 4.2, 4.3 and 4.4 are the plot of sequence number of packets received over time.

The link parameters are as shown in the table below:

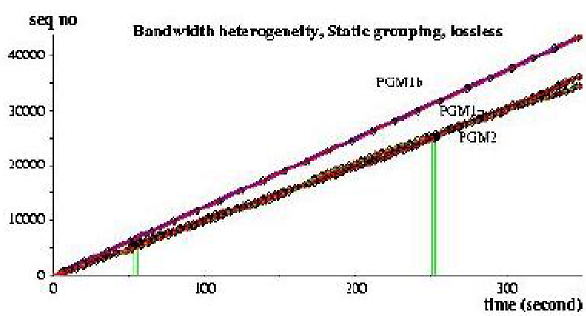
Link	L1	L2	L3	Others
Bandwidth (Mbps)	1.4	Variable	2.4	1.0
Delay (ms)	50	50	50	10



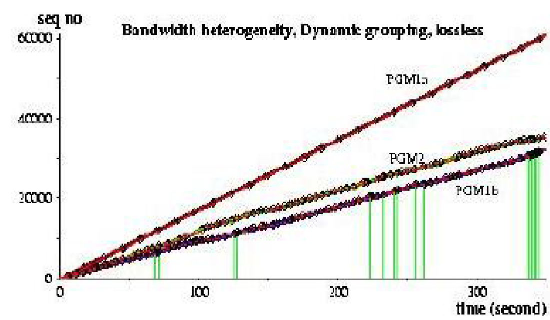
(1) – Network Scenario



(2) – No grouping



(3) – Static grouping



(4) – Dynamic grouping

Figure 4: Simulation topology and results

After 300 seconds:

	No Grouping		Static Grouping			Dynamic Grouping		
Sessions	PGM1	PGM2	PGM1		PGM2	PGM1		PGM2
			PGM1a	PGM1b		PGM1a	PGM1b	
Total packets	26395	31457	31546	36302	31457	22638	52307	32466
Aggregate	27238.67		33116.5			39110.5		
Enhancement %	-		21.58			43.48		

The simulation results show that our scheme definitely improves in the overall performance of receivers of the multicast session. Although the amount of enhancement differs between scenarios, the worst grouping scheme performs almost as well as scenarios that have no grouping. To get more accurate results one can simulate the grouping scheme for various scenarios and check their performance. The readers are refereed to [23] for more detailed simulation studies and discussion. We are using grouping in a generic sense and hence the performance of our scheme will be better than those of the existing solutions of today.

5. Summary

High bandwidth content such as video, graphics and music is increasingly distributed using IP networks. Traditional techniques like broadcast or unicast are neither efficient nor scalable for large number of users. Multicasting content is an extremely promising approach for content distribution over broadband as it is efficient as well as scalable.

Receivers/Users in a network are heterogeneous in terms of their content demands, bandwidth, delay etc. Detailed studies for evaluating the performance of multicast session in the presence of heterogeneous receivers have been carried out and it has been found that grouping substantially improves the overall performance of receivers. Novel algorithms to accomplish optimal grouping have also been developed.

Efficient estimation of network parameters in a multicast scenario, determining the multicast tree, dynamic grouping and other such implementation issues are important components of the grouping scheme. The proposed hierarchical method is scalable and efficient, in terms of number of receivers and network utilization respectively.

6. Future Directions

In this paper, we have addressed the issue of grouping receivers for multicast congestion control. Grouping is just one aspect of the complete problem. A complete solution would involve the development of a protocol that runs in conjunction with grouping. Such a protocol should include a congestion control algorithm that exploits the grouping structure. Ideally the protocol should be scalable, efficient and should be a distributed implementation or decentralized. Unfortunately, these demands may not be compatible with the grouping structure. As mentioned earlier, determining an optimal grouping arrangement requires the knowledge of each receiver's status (or characterization) at a single point in the network, thus making it infeasible to implement it in a complete distributed fashion. Thus, there exists a tradeoff in accuracy and scalability of the scheme and hence search for solutions that offer better tradeoff should be made. This is also an area of future research.

Acknowledgements

We would like to thank Mr. Rahul Verma for his contributions in developing the grouping scheme. We would also like to thank Saurabh Jain for compiling the simulation results for the reliable multicast scenario.

References

1. Harish Ramamurthy, Abhay Karandikar and Rahul Verma, "A Grouping scheme for Reliable Multicast Congestion Control", 8th IEEE International Conference on Communication Systems (ICCS 2002), Singapore.
2. S. Keshav, An engineering approach to computer networking: ATM networks, the internet, and the telephone network, Addison-Wesley, 1st edition, 1997.
3. IP multicast. RFC 2365, available at `http://www.ietf.org/rfc/rfc2365.txt`
4. MSNBC news, URL: `http://www.msnbc.com/news/973798.asp?cp1=1`
5. NDS and Alcatel sign agreement to deliver secure broadband IP entertainment solutions. URL: `http://www.home.alcatel.com/vpr/vpr.nsf/080dff217f9c5f1fc12567a900582617/29b802584741e872c1256d9d0034bec0!OpenDocument`
6. Texas Instruments – Broadband solution, `http://www.ti.com/sc/access`
7. Dotcast. URL: `http://www.dotcast.com`
8. Luigi Rizzo, "pgmcc: A TCP-friendly single rate multicast congestion control scheme", SIGCOMM 2000, August 2000.
9. Injong Rhee, Nallathambi Balaguru and George N. Rouskas, "MTCP: Scalable TCP-like congestion control for reliable multicast", IEEE Infocomm '99, March 1999.
10. S. McCanne, V. Jacobson and M. Vetterli, "Receiver-driven Layered Multicast", ACM SIGCOMM '96, October 1996.
11. J. Byers, M. Frumin, G. Horn, M. Luby, M. Mitzenmacher, A. Roetter and W. Shaver, "FLIDDL: Congestion Control for Layered Multicast", in Proceedings of NGC 2000, November 2000.
12. The Steiner Tree Problem. URL: `http://ganley.org/steiner/`
13. F. P. Kelly, A. Maulloo and D. Tan, "Rate-control for communication networks: Shadow prices, proportional fairness and stability", Journal on Operation Research, Vol. 49, pp. 237-252 1998.
14. E. Graves, R. Srikant and D. Towsley, "Decentralized computation of weighted Max-Min fair bandwidth allocation in networks with multicast flows", Proc. of TIWDC 2001, Sept. 2001.
15. Koushik Kar, Saswati Sarkar and Leandros Tassiulas, "A scalable low-overhead rate control algorithm for multi-rate multicast sessions", IEEE JSAC, Vol. 20, No. 8, October 2002.
16. Rung-Hung Gau, Zygmunt J. Haas and Bhaskar Krishnamachari, "On multicast flow control for heterogeneous receivers", IEEE ACM Trans. on Networking, Vol. 10, No. 1, February 2002.
17. M. H. Ammar and L. R. Wu, "Improving the throughput of point-to-multipoint ARQ protocols through Destination Set Splitting", IEEE Infocomm 92, May 1992.
18. A. Basu and S. J. Golestani, "Estimation of Receiver Round Trip Times in Multicast Communications", RMRG meeting, December 1998. Technical Report available at: `http://citeseer.nj.nec.com/basu98estimation.html`
19. Available at `http://www.isi.edu/nsnam/ns`
20. Luigi Rizzo, "A PGM Host implementation for FreeBSD", Available at: `http://www.iet.unipi.it/~luigi/pgm.html`
21. SureStream: `http://computing.ee.ethz.ch/sepp/rvplayer-8.0.3.423-mo/surestrm.htm`
22. J. De Lameillieure and S. Pallavicini. "A comparative study of simulcast and hierarchical coding", Technical report, RACE project R2110 Hamlet, subgroup WG2, Feb. 1996.
23. Harish Ramamurthy, "A Grouping Framework for Reliable Multicast Congestion Control", Masters Thesis (under Dual Degree Program), Department of Electrical Engineering, IIT Bombay, India, June 2003.