Performance Analysis of Two Loss Recovery Policies for Hierarchical Reliable Multicast

Zuo Wen Wan, Michel Kadoch, Ahmed Elhakeem

***Abstract*—Hierarchical reliable multicast provides scalability as well as performance enhancement. Designated receivers (DR) are assigned for retransmission handling of each domain. Effect of DR locations on multicast performance is analyzed in this paper. The paper addresses the problem of finding the optimal placements of DRs that minimizes the bandwidth consumption. Such bandwidth consumption depends on the loss recovery scheme employed. In this respect, we compare two multicast loss recovery policies. We couple the policy choice to the problem of optimal determination of the number and places of DRs. The effects of the control traffic, independent and correlated packet losses as well as heterogeneous losses are all considered in this analysis.[1]**

Keywords — Reliable multicast, designated receivers, hierarchy, optimization, loss recovery.

I. INTRODUCTION

Much attention has been paid to reliable multicast in recent years [1][2][3][14]. Scalable loss recovery is a key factor in reliable multicast design. Two major loss recovery schemes, i.e, end-to-end schemes [1][2][8][14] and router-assisted schemes [6][7][10], have been proposed to provide scalability. Whatever loss recovery is used, it is always required to isolate the losses from these receivers that have received the packet. Hierarchical loss control is a popular approach to limit the scope of recovery of data and control messages. In these hierarchical schemes, a multicast group is partitioned into subgroups [2][4]. Within each subgroup, a special receiver or router is selected to be responsible for the retransmission of the packet and to collect the feedback from the receivers in the subgroup. Only such special receivers report the state of the subgroup to the sender. Designated receivers in reliable multicast transport protocol (RMTP) [2] or link repliers in [6] have such function of repair. The scope of loss recovery is limited to a small subgroup. Thus feedback implosion and bandwidth consumption can be substantially reduced.

The performance of reliable multicast strongly depends on the loss characteristics of tree, the topology, and the loss recovery mechanism [4][15]. The computation of the performance parameters such as bandwidth consumption and the number of transmissions is intensive for a general tree [3][4]. However, [4] only gives analysis results and optimal DR locations of one of the loss recovery policies in this paper in uniform loss and linear topologies. How to efficiently evaluate the performance of reliable multicast is significant for the design of loss recovery. Further, loss recovery should adapt the topology and loss characteristics of a tree in order to greatly reduce the cost of networks and the sender. The optimal DR locations can minimize the bandwidth consumption of networks that is important to the design of reliable multicast protocols. In this paper and different from recent work [4], we present a more general approach available to different types of general topologies for determining the optimal DR locations. We also address the two aspects of reliable multicast: an efficient evaluation of performance parameters and the optimal DR locations based on different loss scenarios, topologies and loss recovery schemes.

Motivated by the possible impact of DR locations on the performance of loss recovery mechanisms, we develop a new analysis and simulation of two new loss recovery policies. These are the main objectives of the paper. However, in the process, we provide approximations that will shade light on the effects of control traffic, independent, correlated homogeneous and heterogeneous packet losses that were not addressed in recent work.

The rest of this paper is organized as follows. In section II, we discuss work in this area. In section III, we present our model for hierarchical reliable multicast. In section IV, we study the performance of reliable multicast and give an evaluation on the number of transmissions and bandwidth consumption. In section V, we use actual bandwidth consumption to optimize the placements of DRs for the binary tree. In section VI, we present the results of the optimal placements of DRs for the binary tree. Section VII concludes the paper.

II. RELATED WORK AND CONTRIBUTIONS

Two issues are considered in this paper: performance analysis and the optimal DR locations, both of which are analyzed for two loss recovery schemes.

There is a lot of work on performance analysis of reliable multicast [3][4][11][12][15], most of which are based on the recursive equations in [11]. The cost of the sender multicast has been analyzed well by the number of transmissions E[M], to be defined shortly. Due to the intensive computation of E[M], approximations or simulations are often used [4]. We take a different approach to calculate these parameters after analyzing the transmission process of multicast packets. We discuss the effects of dependent and independent loss on

This work was supported by a CRD grant from NSERC and Bell Canada. The completion of this research was made possible thanks to the support of Bell Canada through its Bell University Research Program.
Zuo Wen Wan and Ahmed Elhakeem are with department of Electrical and Computer Engineering, Concordia University, Montreal, Canada, email: zw_wan@ece.concordia.ca. and ahmed@ece.concordia.ca.
Michel Kadoch is with department of Electrical Engineering, Ecole de Technologie Superieure, Montreal, Canada, email: kadoch@ele.etsmtl.ca.

networks and obtain the probability distribution of loss scenario over network links. Some basic characteristics are found. Further, we use them to derive E[M] and bandwidth consumption of different loss recovery schemes. By this approach, we can make some reasonable assumption and obtain an analytical evaluation of performance parameters for a multicast network with heterogeneous link loss probability.

There exist quite a few loss recovery schemes for reliable multicast. End-to-end and router assisted schemes have been employed in hierarchical loss recovery [2][6][10] and compared in [16]. Due to the lack of information about the underlying network, end-to-end schemes often multicast or unicast repair packets to all members of a subgroup, e.g. reliable multicast transport protocol (RMTP) [2] uses designated receivers (DR) to remulticast repair packets for each subgroup. Router-assisted schemes further reduce the scope of recovery, e.g., Pragmatic General Multicast (PGM) [10] conducts retransmissions to receivers that requested them by the aid of routers to remember where NAKs (negative acknowledgement) came from. If desired, suitable receivers are used to act as Designated Local Retransmitters (DLRs). Naturally, the placements of these DRs or DLRs have a noticeable influence on the bandwidth consumption of a network. Their optimal placements in hierarchy, where repair packet is always multicast to all members of each subgroup, have been obtained for a few special cases. These include uniform packet loss and linear topologies in [4] and for the general topologies and uniform packet loss in [15], where bandwidth consumption is minimized to obtain the optimal locations. Furthermore, the effects of topology on optimal locations have been analyzed in [15]. Research on the location problems can also be found for server-based reliable multicast [13] which presents only simulation results not related to any topology and for independent unicast transmissions from the sender to receivers [5]. Although the approaches proposed in [13] yield better multicast performance than the random placement method of repair servers, they do not aim to find such optimal placements in order to minimize bandwidth consumption. Their placements are determined based on the number of the shortest paths passing through the node, the number of links connected to the node and the average number of hops from the node to all other nodes. In our work, we address the problem of finding effects of loss recovery and loss characteristics on optimal DR locations. The optimal DR locations are suggested so as to minimize the bandwidth consumption for different loss characteristics of a tree and different topologies. An analytical comparison of two typical loss recovery schemes, namely, whole subgroup and selective retransmissions, are investigated. Moreover, in all analysis and simulation herein, we consider the effects of control traffic, independent, correlated, and heterogeneous losses.

III. Model

The performance analysis herein is based on a single-source multicast tree. Here we use a DR to retransmit the repair packet to a subgroup for end-to-end and router assisted loss recovery schemes. Each DR is located at the root of its subgroup. Upon receiving NAKs downstream, DR or the sender retransmits the missing packet.

A. *The functions of DRs*

- Store-and-forward. For a new data packet from the sender, DR will always store and forward to its members. Duplicate packets that the DR has already received are not forwarded to its members.
- The retransmission of data. Only if one or more members require the retransmission of a missing data packet, will the DR retransmit it. Similarly, the sender retransmits packets to the requesting members in the sender subgroup not covered by any DR.
- The feedback of the domain. If the DR does not receive a data packet, it will ask for retransmission from the sender. The DR is an ordinary member in the domain of the sender.

B. *Loss Recovery Policies*

Two loss recovery policies are discussed in this work.

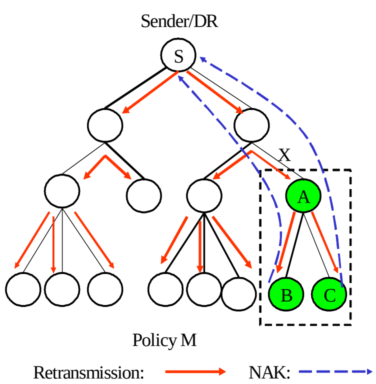

Fig. 1 The loss recovery of policy M. The loss occurs at link A.

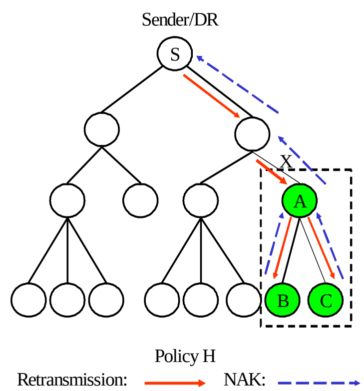

Fig. 2 The loss recovery of policy H. The loss occurs at link A.

Policy M – whole subgroup retransmissions

Receivers missing packets send NAK upstream to the sender or DR to request retransmissions. The sender and/or DRs multicast the lost packet to all members in their domain once they receive the retransmission requests (NAK) from receivers below [2][4]. Fig. 1 shows the loss recovery procedure of policy M where nodes B and C have not received a certain packet. This is a typical end-to-end scheme easily employed in reliable multicast. In this policy and as a worst case treatment, we assume all users suffering loss will transmit a NAK to sender or DR as will be reflected within the subsequent analysis. We do not assume here the presence of intelligent routers that may cut the NAK traffic and reduce NAK implosion [2][8].

Policy H – selective retransmissions

When one receiver encounters a loss, it sends NAK to a sender or DR depending on which is closer. The sender and/or DR then retransmit the packet only to these receivers that did not receive data in previous transmissions. This protocol needs the assistance of routers to recover the packet loss. Some simple functions are added to intermediate routers to remember which link downstream lost the packet and needs retransmissions. Routers merge all the NAK requests downstream and send only one NAK to the parent router to avoid NAK implosion [10]. Intermediate routers will retransmit packets only to some members of their subgroups. This policy adds a small extra capability to ordinary routers (not DR). They should be able to route repair packets only to these nodes in their subgroup that have sent a NAK upstream through this router. But they would send the repair packets to all members of this subgroup if they themselves have lost the original packet. Fig. 2 shows the loss recovery of policy H. Also, nodes may randomize the transmissions times of their NAKs so as to further help the NAK implosion avoidance. These issues are not addressed in this paper. However, in the analysis, we consider the reduction of the number of NAK transmissions on the total bandwidth consumption in this policy where the routers have more intelligence.

IV. Performance Analysis

In this section, we will explain the two metrics that will be used to optimize the location of DRs, namely, the expected number of transmissions E[M] and effective bandwidth consumption. E[M], which is the average number of transmissions of a given packet by the sender until all group members receive it correctly, is one of the most important parameters for reliable multicast. It depends on the topology of trees and loss probability [11][12][15].

E[M] gives the cost of the sender in the transmission of a packet to all receivers in the tree. The actual bandwidth consumption of networks depends not only on the number of transmissions E[M] but also on the number of links traversed by each transmission. We define E[B] as the expected bandwidth consumed per link averaged over all links in one successful multicast. We define E[C] as the total expected bandwidth of a group consumed by one successful source packet multicast. Thus,

$$E[B]=E[C]/N \tag{1}$$

where N is the total number of links for a multicast group. E[M], E[B] and E[C] reflect the bandwidth performance of multicast protocols. In the following sections, one first derives their expressions of no DR. Please refer to Table 1 and Table 2 for the notation used in this paper.

A. *E[M] --Previous Exact Solutions*

The expected number of transmissions E[M] that a packet should be multicast by the source using selective reject ARQ until all group members receive it correctly can be recursively calculated. For a multicast network, one can calculate the CDF (cumulative distribution function) of the total number of transmissions and retransmissions from the sender. Let M(k) be the total number of transmission and retransmissions of a packet until received by all receivers under node k. Then CDF for node k is $F_k(m)=P[M(k)\leq m]$, i.e, $F_k(m)$=Prob.[all nodes from k and below got the packet at most in m trials]. We denote $F_r(m)$, $F_k(m)$ and $F_s(m)$ to be the CDF of the total number of transmission for leaf receivers (r), intermediate nodes (k) and the sender (s), respectively, as shown in Fig. 4. Then one can obtain the following equations for E[M] [11][12].

$$E[M(S)]=\sum_{m=o}^{\infty} mP_S(M=m)=\sum_{m=o}^{\infty}(1-F_s(m)) \tag{2}$$

$$F_S(m)=\prod_{c\in child(S)} F_c(m) \tag{3}$$

$$F_k(m)=\sum_{u=0}^{m-1}\binom{m}{u} p_k^u (1-p_k)^{m-u} \prod_{c\in child(k)} F_c(m-u) \tag{4}$$

$$F_r(m)=1-p_r^m \tag{5}$$

where $P_S(M=m)$ is the probability that the sender has to transmit the intended packets in times until all nodes receive it. p_k or p_r is the packet loss probability on the link leading to node k or leaf receiver r.

Needless to say one should use recursion starting from the bottom of nodes and the numerical evaluation of equation (2) - (5) to calculate E[M]. However, the computation of E[M] may be very intensive for a general topology [3]. Even if the reduction technique is used, the computation of E[M] can be exponential in the number of nodes N [4]. Based on the intensive computations of E[M], it is difficult to calculate other parameters such as bandwidth and delay for a large multicast network. It is necessary then to use an efficient way to estimate E[M] and E[C]. In this section, we follow a completely analytical approach to evaluate the E[M] of a general topology without the need for recursion.

B. *E[M] -- Approximate Evaluations*

Although one may recursively calculate the expected number of transmissions using equations (2)-(5), the computation of E[M] for a general topology is intensive [3][4]. Thus, we try herein to find an approximate solution for the evaluation of the expected number of transmissions for general topology.

Hierarchical reliable multicast is based on a tree topology where the loss of an intermediate link leads to the losses of all links below that link. Even if only one intermediate link loss takes place in a multicast tree, many links need retransmissions from the sender. In one multicast transmission, many losses may take place. Therefore, it is necessary to know how many losses take place and how many links need to be involved in retransmissions because only corresponding nodes care whether they receive the correct packet in the next transmission.

Table 1. NOTATIONS FOR RELIABLE MULTICAST

ID	Meaning
p_k	Packet loss probability over link k.
d_k	The number of links from the sender to node k, also the set of these links.
n_k	The total number of links under node k, also the set of these links.
D_k	The total number of links involved in retransmissions if link k loses a packet. $D_k = d_k + n_k$. Also the set of these links.
Ω	Link loss scenario, i.e. the set of independent link losses. For example, {1,2} means the losses from links 1 and 2, {0} means no loss.
D(Ω)	The number of links involved in retransmissions for multiple loss scenarios Ω, also the set of these links.
n(Ω)	The number of links under the links of the multiple loss scenarios Ω.
M, m	The total number of packet transmission times per source packet (due to losses).
C	The total bandwidth consumed in a multicast group (the successful delivery of one source packet to all nodes including all retransmission).
P(M=m)	The probability that one packet multicast is successful in the m^{th} transmission (all nodes get the same packet).
P(m|Ω)	The probability that multicast is successful in the m^{th} transmission after the loss scenarios Ω take place.
$Q_N(\Omega)$	The probability that multiple loss scenarios Ω take place in one transmission over N links.
Λ(Ω)	Bandwidth consumed by NAKs for multiple loss scenarios Ω.
Λ_k	Bandwidth consumed by NAKs if link k loses a packet.
β	The ratio of a NAK packet to a data packet.

Multicast packet losses are not independent in a general tree. In Fig.3, the loss of node A will result in the losses of all links below node A. These losses are dependent (correlated) because they involve the nodes of subtree rooted at node A. However, if one node is not under the subtree of another node, the losses such as node A and B are independent (uncorrelated). They do not affect each other. One may consider the effect of one link loss on other nodes of its subtree. Suppose that each node has a sequence number and the sender has the sequence number 0, shown in Fig.3. For a given packet transmission, if node k (e.g., node A) does not receive the packet, n_k nodes will also lose the packet resulting in a correlated loss where n_k is the total number of links below node k. The sender retransmits the repair packet to these nodes of $D_k = d_k + n_k$ where d_k is the number of links from the sender to node k. Therefore, we only concentrate if these nodes receive the packet correctly in the next retransmission. D_k reflects the effect of loss on link k on dependent losses of the network tree.

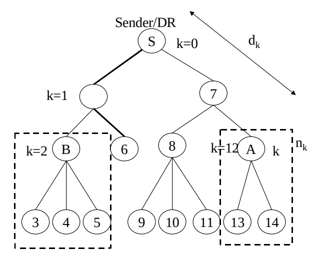

Fig.3 The links involved in retransmissions. $D_k = d_k + n_k$.

In addition to dependent (correlated) loss, we also analyze effects of independent loss. In a multicast transmission, some independent losses may take place over different links due, for example, to other session or unicast traffic. We use Ω to represent such independent loss scenario, e.g. Ω={2,8} means the packet losses from link 2 and 8 in Fig.3, Ω={0} means no loss. Here, we denote by $Q_N(\Omega)$ the probability that multiple loss scenario Ω takes place in a multicast transmission over N links of a certain part of multicast tree. $Q_N(\Omega)$ has the following probability distribution.

$$Q_N(\Omega=\{0\}) = \prod_{k\in N}(1-p_k) \tag{6}$$

$$Q_N(\Omega=\{k_1\}) = p_{k_1}\prod_{k\neq k_1,\, k\notin n_{k_1}}^{k\in N}(1-p_k) \tag{7}$$

$$Q_N(\Omega=\{k_1,k_2,\cdots,k_j\}) = \prod_{l=1}^{j} p_{k_l}\prod_{k\notin\Omega}^{k\notin n(\Omega)}(1-p_k), \tag{8}$$
$$k_1\in N, k_2\notin D_{k_1},\cdots,k_j\notin D(\Omega=\{k_1,k_2,\cdots,k_{j-1}\})$$

where p_k is the packet loss probability over link k, n_{k_1} is the total number of links below node k_1, n(Ω) is the number of links missing the packet under the multiple loss scenarios Ω, i.e. $n(\Omega=\{k_1,k_2,\cdots,k_j\}) = n_{k_1}+n_{k_2}+\cdots+n_{k_j}$. N is the total number of links in the network or a part of network under the error control of a DR. D_k was defined before and the exclusion of $k_2 \notin D_{k_1}$, for example, is to include only independent, not dependent loss. D(Ω) is the number of links involved in retransmissions for multiple loss scenarios Ω and represents the effect of Ω on dependent losses. For example, D(Ω={2,12}) represents the links of d_2, n_2, d_{12}, and n_{12}. Please refer to Table 1 for these notations. The following can be obtained from (6).

$$\sum_{\Omega\subset N} Q_N(\Omega)=1 \tag{9}$$

$$Q_N(\Omega=\{k_1,k_2,\cdots,k_j\})=\eta_{k_1}p_{k_1}Q_N(\Omega=\{k_2,\cdots,k_j\}\mid k_1) \tag{10}$$

where $\sum_{\Omega\subset N}$ stands for the summation of all possible multiple loss scenarios for N links, i.e..

$$\sum_{\Omega\subset N}=\sum_{k\in N}+\frac{1}{2!}\sum_{k_1\in N}\sum_{k_2\notin D_{k_1}}+\frac{1}{3!}\sum_{k_1\in N}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D(\{k_1,k_2\})}+\cdots \tag{11}$$

$$\eta_k=\prod_{l\in d_k,l\neq k}(1-p_l) \tag{12}$$

$Q_N(\Omega=\{k_2,\cdots,k_j\}\mid k_1)$ is the probability distribution of independent multiple loss scenarios Ω not considering the links of D_{k_1} after packet loss takes place at link k_1, i.e.

$$Q_N(\Omega=\{k_2,k_3,\cdots,k_j\}\mid k_1)=\prod_{l=2}^{j}p_{k_l}\prod_{k\notin\Omega,.k\notin d_{k_1}}^{k\notin n(\Omega)}(1-p_k), \tag{13}$$
$$k_1\in N,k_2\notin D_{k_1},\cdots,k_j\notin D(\{k_1,k_2,\cdots,k_{j-1}\})$$

Theorem 1.

Any function f(Ω) of multiple independent loss scenario Ω, which could be defined as the number of links n(Ω) as in (27), or the number of independent losses, or the number of NAKs resulting from occurrence of independent errors over links k_1, k_2,… etc as will follow, can be decomposed into the functional sum of different single independent losses, i.e.

$$f(\Omega=\{k_1,k_2,\cdots,k_j\})=f(\{k_1\})+f(\{k_2\})+\cdots+f(\{k_j\}) \tag{14}$$

In the case of no loss, i.e. Ω={0}, their values are 0, i.e.

$$f(\Omega=\{0\})=0 \tag{15}$$

Further, one can obtain its average value over different loss scenarios, i.e.

$$\overline{f_N}=\sum_{\Omega\subset N}f(\Omega)\cdot Q_N(\Omega)=\sum_k p_k\eta_k f(\{k\}) \tag{16}$$

where the summation $\sum_{\Omega\subset N}$ means all possible multiple loss scenarios whose probabilities are in (6) and (8).

Proof:

Equation (14) is evident by the independent packet losses on link k_1, k_2, … . For example, if f(Ω) is the number of NAKs generated from losses situated on link k_1, this is completely independent and adds up to the number of NAKs generated from loss on link k_2, and so on, to provide the function of sum of these events.

$$\overline{f_N}=\sum_{\Omega\subset N}f(\Omega)\cdot Q_N(\Omega)=f(\{0\})Q_N(\{0\})+\sum_{k\in N}f(\{k\})\cdot Q_N(\{k\})$$
$$+\frac{1}{2}\sum_{k_1}\sum_{k_2\notin D_{k_1}}[f(\{k_1\})+f(\{k_2\})]Q_N(\Omega=\{k_1,k_2\})+$$
$$\frac{1}{3!}\sum_{k_1}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D(\{k_1,k_2\})}[f(\{k_1\})+f(\{k_2\})+f(k_3)]Q_N(\{k_1,k_2,k_3\})+\cdots$$

$$\overline{f_N}==f(\{0\})Q_N(\{0\})+\sum_k f(k)p_k\eta_k\cdot Q_N(\{0\}\mid k)$$
$$+\sum_{k_1}f(\{k_1\})p_{k_1}\eta_{k_1}\sum_{k_2\notin D_{k_1}}Q_N(\{k_2\}\mid k_1)$$
$$+\sum_{k_1}f(\{k_1\})p_{k_1}\eta_{k_1}\frac{1}{2!}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D(\{k_1,k_2\})}Q_N(\Omega=\{k_2,k_3\}\mid k_1)+\cdots$$
$$=f(\{0\})Q_N(\{0\})+\sum_{k_1}f(\{k_1\})p_{k_1}\eta_{k_1}[Q_N(\{0\}\mid k_1)$$
$$+\sum_{k_2\notin D_{k_1}}Q_N(\{k_2\}\mid k_1)+\frac{1}{2!}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D(\{k_1,k_2\})}Q_N(\Omega=\{k_2,k_3\}\mid k_1)+\cdots]$$
$$=f(\{0\})Q_N(\{0\})+\sum_k p_k\eta_k f(\{k\}) \tag{17}$$

For f(Ω={0})=0, we have $\overline{f_N}=\sum_k p_k\eta_k f(\{k\})$.

The probability density function P(M=m) representing the number of the final successful transmission trials M can be obtained recursively, i.e.

$$P(M=1)=Q_N(\Omega=\{0\})=\prod_{k\in N}(1-p_k) \tag{18}$$

$$P(M=m)=\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)P(M=m-1\mid\Omega),\qquad m=2,3,..., \tag{19}$$

where $\sum_{\Omega}$ stands for all possible link loss scenarios. The conditional probability P(M=m-1|Ω) is the probability that m-1 retransmissions are required to recover multiple loss scenario Ω. It can be recursively calculated from (18) and (19). However, only D(Ω) not N nodes are involved in retransmissions if loss scenario Ω takes place where D(Ω) is the subset of links involved in retransmissions for the loss scenario Ω.

Thus one can easily obtain the expected number of total transmissions, which can be recursively calculated.

$$E[M]=\sum_{m=1}^{\infty}mP(M=m)=P(M=1)+\sum_{m=2}^{\infty}m\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)P(M=m-1\mid\Omega)$$
$$=P(M=1)+\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)[1+\sum_{m=2}^{\infty}(m-1)P(M=m-1\mid\Omega)]$$
$$=1+\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)E[M\mid\Omega] \tag{20}$$

where E[M|Ω] is the expected number of transmissions to correct the multiple loss scenario Ω. Only these nodes that did not receive the packet are involved in retransmissions. So E[M|Ω] can also be similarly obtained.

$$E[M\mid\Omega]=1+\sum_{\Omega_1\neq\{0\}}^{D(\Omega)}Q_{D(\Omega)}(\Omega_1)E[M\mid\Omega_1] \tag{21}$$

where Ω_1 is the set of links incurring losses after second transmissions.

Substituting (21) into (20), one may obtain the following approximation for E[M].

$$E[M]=2-Q_N(\{0\})+\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)\sum_{\Omega_1}^{D(\Omega)}Q_{D(\Omega)}(\Omega_1)E[M\mid\Omega_1] \tag{22}$$

Repeating the same procedure, i.e.

$E[M\mid\Omega_1]=1+\sum_{\Omega_2\neq\{0\}}^{D(\Omega_1)}Q_{D(\Omega_1)}(\Omega_2)E[M\mid\Omega_2]$ and we obtain

$$E[M]=2-Q_N(\{0\})+\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)[1-Q_{D(\Omega)}(\Omega_1=\{0\})] \quad (23)$$
$$+\sum_{\Omega\neq\{0\}}^{N}Q_N(\Omega)\sum_{\Omega_1\neq\{0\}}^{D(\Omega)}Q_{D(\Omega)}(\Omega_1)\sum_{\Omega_2\neq\{0\}}^{D(\Omega_1)}Q_{D(\Omega_1)}(\Omega_2)E[M\mid\Omega_2]$$

Taking $Q(\Omega_1=\{0\})=\prod_{k=1}^{D(\Omega)}(1-p_k)\approx 1-\sum_{k=1}^{D(\Omega)}p_k$ for low p_k and ignoring the 4th term on the right hand side of (23) relative to the first few terms, one obtains the following approximation.

$$E[M]\approx 2-Q_N(\{0\})+\sum_{l=1}^{N}p_l\eta_l\sum_{k}^{D_l}p_k \quad (24)$$

where $Q_N(\Omega=\{0\})=\prod_{k\in N}(1-p_k)$ depends on link loss probability and is independent of topology. $\sum_{l=1}^{N}p_l\eta_l\sum_{k}^{D_l}p_k$ is a term dependent on the topology. Assume $q=\sum_{l=1}^{N}p_l$. For a small value of q, one obtains the following approximation.

$$E[M]\approx 1+q-\frac{1}{2}q^2+\frac{1}{6}q^3+\sum_{l=1}^{N}p_l\eta_l\sum_{k}^{D_l}p_k \quad (25)$$

C. The Evaluation of Bandwidth Consumption

Although E[M] gives the expected number of transmissions, the number of links affected by each transmission is different. Thus, it is necessary to evaluate the bandwidth consumed on all links until final success. By final success, we mean to include all link transmissions used until all nodes receive the packet correctly. The first original transmission always affects the whole multicast nodes regardless of which loss recovery mechanism is used. If node k does not receive the packet, n_k links will lose the packet as shown in Fig.3. Then the actual bandwidth consumption is N-n(Ω) for multiple loss scenarios Ω where n(Ω) is the number of dependent losses for multiple loss scenarios Ω. Thus, we obtain the expected bandwidth consumption over N links as follows,

$$E[C]=\sum_{\Omega\subset N}Q_N(\Omega)\{N-n(\Omega)+\beta\Lambda(\Omega)+E[C\mid\Omega]\} \quad (26)$$

where $Q_N(\Omega)$ is the probability of having multiple loss scenarios Ω in a transmission trial as defined in Table 1, and E[C|Ω] is the extra bandwidth consumed in all retransmission trials to correct the loss scenarios Ω that take place in the first transmission. Λ(Ω) is the bandwidth consumed by NAKs for the loss scenario Ω. β is the ratio of NAK packet size to regular data packet size.

Due to $n(\Omega=\{k_1,k_2,\cdots,k_j\})=n_{k_1}+n_{k_2}+\cdots+n_{k_j}$ and $n(\Omega=\{0\})=0$, one has the following from (16).

$$\sum_{\Omega\subset N}Q_N(\Omega)n(\Omega)=\sum_{k\in N}p_k\eta_k n_k=\sum_{k\in N}p_k\eta_k\sum_{l\in n_k}1 \quad (27)$$

Using (12) and (45) of the appendix A where we take $\varphi_k=p_k\eta_k$ and $\xi_k=1$, one obtains

$$\sum_{\Omega\subset N}Q_N(\Omega)n(\Omega)=\sum_{k\in N}\sum_{l\in d_k,l\neq k}p_l\eta_l=\sum_{k\in N}(1-\eta_k) \quad (28)$$

Substituting (28) into (26), one has

$$E[C]=\sum_{k\in N}\eta_k+\sum_{\Omega\subset N}Q_N(\Omega)\{\beta\Lambda(\Omega)+E[C\mid\Omega]\} \quad (29)$$

(29) suggests that the bandwidth consumption of original data packets in the first transmission is $\sum_{k\in N}\eta_k$. In order to evaluate the expected conditional bandwidth consumption of retransmissions, i.e. E[C|Ω], we encounter two cases depending on the recovery methodology, i.e., policy M and H to follow.

1) Policy M

Due to the lack of router assistance, each receiver losing a packet unicasts a NAK packet to the sender/DR, which costs a bandwidth d_k for receiver k. All receivers within the loss scenario Ω send NAK upstream, i.e. $\Lambda(\Omega=\{k\})=\sum_{l\in R_k}d_l$ where R_k is the set of receivers within the loss scenario Ω={k}. Thus, we have the following result using (50) of the appendix with $\varphi_k=p_k\eta_k$ and $\xi_k=d_k$,

$$\sum_{\Omega\subset N}Q_N(\Omega)\Lambda(\Omega)=\sum_{k\in N}p_k\eta_k\sum_{l\in R_k}d_l=\sum_{k\in R}d_k\sum_{l\in d_k}p_l\eta_l \quad (30)$$
$$=\sum_{k\in R}d_k[1-\eta_k(1-p_k)]=\sum_{k\in R}d_k(1-\eta_k+p_k\eta_k)$$

where $\sum_{l\in R_k}d_k$ is the sum of affected links between the sender or DR and the location of the k^{th} loss since the corresponding NAKs sent upstream transverse these links. The derivation of (30) follows steps similar to these leading to (28).

In this loss recovery scheme, the packet is always multicast to the whole node population by the sender or DR, i.e., each transmission or retransmission affects all members of the same subgroup. The actual bandwidth consumption of each retransmission is $\sum_{k\in N}\eta_k$ (does not change in policy M) as can be seen by repeated application of (29) to each retransmission trial. Thus, noting we have E[M|Ω] transmission trials and each has a corresponding transmission cost of $\sum_{k\in N}\eta_k$ plus cost of E[M|Ω]-1 NAK messages, one obtains the following equation:

$$E[C\mid\Omega]=E[M\mid\Omega]\cdot\sum_{k\in N}\eta_k+\beta(E[M\mid\Omega]-1)\sum_{k\in R}d_k(1-\eta_k+p_k\eta_k) \quad (31)$$

where E[M|Ω] is the expected number of retransmissions after loss patterns Ω take place, the first term is bandwidth consumption of data packets in retransmissions and the 2nd term accounts for the cost of NAKs in retransmission trials.

By substituting (31) into (29) and using (20), one may get E[C].

$$E[C]=E[M]\cdot\sum_{k\in N}\eta_k+\beta(E[M]-1)\sum_{k\in R}d_k(1-\eta_k+p_k\eta_k) \quad (32)$$

2) Policy H

Policy H uses routers to recover losses. Receivers losing a packet send a NAK to the upstream router, and routers merge all received NAKs and send only one NAK upstream. Thus the bandwidth cost of NAKs is $\Lambda(\Omega) = D(\Omega) \le \sum_k D_k$ for the loss scenario Ω. The upper bound of NAK cost is obtained similar to (30),

$$\sum_{\Omega \subset N} Q_N(\Omega)\Lambda(\Omega) = \sum_{k \in N} p_k \eta_k D_k \quad (33)$$

For each retransmission, the sender and/or DRs will retransmit the packet to only these receivers that did not receive data in previous transmissions, e.g. if link k loses a packet, only D_k links are involved in retransmissions. If a loss scenario Ω with multiple independent losses takes place, one can sum up the bandwidth of single independent loss, i.e. $E[C \mid \Omega] = \sum_k E[C \mid k]$.

We can evaluate E[C|k] in the following equation for the season of data and NAKs transmissions.

$$E[C \mid k] \le E[M \mid k] \sum_{l \in D_k} \eta_l + \beta(E[M \mid k] - 1) D_k \quad (34)$$

The upper bound stems from the fact that only the first retransmission will lead to $loss \in D_k$ and further retransmissions are expected to handle less loss. Considering (16) and (34) with f({k})=E[C|k], one obtains the bandwidth consumption of retransmissions for policy H as follows:

$$\begin{aligned} &\sum_{\Omega} Q_N(\Omega) E[C \mid \Omega] = \sum_k p_k \eta_k E[C \mid k] \\ &\le \sum_k p_k \eta_k \{E[M \mid k] \sum_{l \in D_k} \eta_l + \beta(E[M \mid k] - 1) D_k\} \end{aligned} \quad (35)$$

One may estimate the bandwidth consumed in protocol H by substituting (33) and (35) into (29).

$$\begin{aligned} E[C] &= \sum_{k \in N} \eta_k + \beta \sum_{k \in N} p_k \eta_k D_k \\ &+ \sum_k p_k \eta_k \{E[M \mid k] \sum_{D_k} \eta_i + \beta(E[M \mid k] - 1) D_k\} \\ &= \sum_{k \in N} \eta_k + \sum_k p_k \eta_k E[M \mid k] \cdot \sum_{l \in D_k} (\eta_l + \beta) \end{aligned} \quad (36)$$

V. EFFECT OF DRS ON BANDWIDTH

Hierarchical loss recovery is often used to provide the scalability in which case some special receivers or routers such as DR have the FEC/ARQ capacity and retransmit the repair packet to reduce the load of the sender. Each DR is the representative of a multicast group and sends NAK upstream to the sender on behalf of receivers downstream, e.g. DR_1 is responsible for the retransmissions of N_1 links in Fig. 4 and reports the state to the sender. Here we denote by N_i the number of links covered by the DR_i. N_0 is the number of links not covered by any DRs. Fig. 4 is an example of a multicast group having 2 DRs.

We consider first the case of 1 DR. For the multiple loss scenarios Ω in the first transmission, these losses occur within, i.e. Ω_1 or outside, i.e. Ω_0, the DR coverage. We denote by $Q_{N_1}(\Omega_1)$ ($Q_{N_0}(\Omega_0)$) the probability that multiple loss scenarios Ω_1 (Ω_0) take place within (outside) DR coverage. Ω_1 and Ω_0 are two independent events occurring in different loss recovery domains. Thus, the probability $Q_N(\Omega_0,\Omega_1)$ that the loss scenarios Ω_0 and Ω_1 take place over N links is given by the following:

$$Q_N(\Omega_0, \Omega_1) = Q_{N_0}(\Omega_0) Q_{N_1}(\Omega_1) \quad (37)$$

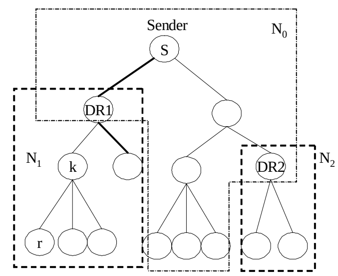

Fig. 4 The partition of a multicast group

Table 2. NOTATIONS FOR RELIABLE MULTICAST

ID	Meaning
N	The total number of links within a multicast group, also the set of these links.
N_i	The number of links covered by DR_i, also the set of these links.
C	The total bandwidth consumed in the multicast group N (the successful delivery of one source packet to all nodes including all retransmission.)
B	Average bandwidth consumption per link for the multicast group. B=C/N.
$E[M_i]$	The expected number of transmissions for N_i links to receive a correct packet.
$E[C_i]$	The total expected bandwidth consumed in N_i links over all transmissions.

The total bandwidth consumed in N links can be obtained recursively

$$\begin{aligned} E[C] = &\sum_{\Omega_0 \subset N_0} \sum_{\Omega_1 \subset N_1} Q_N(\Omega_0, \Omega_1) \{N_0 - n(\Omega_0) + N_1 - n(\Omega_1) \\ &\beta\Lambda(\Omega_0) + \beta\Lambda(\Omega_1) + E[C \mid (\Omega_0, \Omega_1)]\} \end{aligned} \quad (38)$$

where N_0 and N_1 are the number of links outside and within DR coverage, respectively. $\Lambda(\Omega_0)$ (or $\Lambda(\Omega_1)$) is the bandwidth consumed by NAKs for the loss scenario Ω_0 (or Ω_1). β is the ratio of NAK packet size to regular data packet size. $E[C|(\Omega_0,\Omega_1)]$ is the extra bandwidth consumption when multiple loss scenarios Ω_0 and Ω_1 take place. These losses are recovered by either the DR or the sender, depending on their locations. DR recovers only the losses under this DR while the sender recovers the losses outside all DRs. The sender and the DR work independently in the next retransmissions. Thus, the total bandwidth $E[C|(\Omega_0,\Omega_1)]$ can be divided into the summation of bandwidth $E[C| \Omega_0]$ and $E[C|\Omega_1]$ for two subgroups:

$$E[C \mid \Omega_0, \Omega_1] = E[C \mid \Omega_0] + E[C \mid \Omega_1] \quad (39)$$

where $E[C|\Omega_1]$ ($E[C|\Omega_0]$) is the total conditional bandwidth consumption to recover the loss scenario Ω_1 (Ω_0).

One obtains the following results for the total bandwidth consumption by substituting equation (39) into (38) and using (29):

$$\begin{aligned} E[C] &= \sum_{\Omega_0 \subset N_0} Q_{N_0}(\Omega_0) \sum_{\Omega_1 \subset N_1} Q_{N_1}(\Omega_1)\{N_0 - n(\Omega_0) + \beta\Lambda(\Omega_0) + E[C \mid \Omega_0]\} \\ &+ \sum_{\Omega_0 \subset N_0} Q_{N_0}(\Omega_0) \sum_{\Omega_1 \subset N_1} Q_{N_1}(\Omega_1)\{N_1 - n(\Omega_1) + \beta\Lambda(\Omega_1) + E[C \mid \Omega_1]\} \\ &= E[C_0] + E[C_1] \end{aligned} \quad (40)$$

where $E[C_k]$ (k=0,1) is the expected bandwidth consumed in N_k links and can be calculated by (32) or (36) depending on the loss recovery policy. The above equation shows that the total bandwidth consumed in the whole group is equal to the summation of bandwidth consumed in every subgroup.

Based on the discussion above, the total bandwidth is equal to the summation of bandwidth consumed in N_0 and N_1 links, i.e., $E[C_0]$ and $E[C_1]$. In the same way, we can extend this result to the case of r DRs. The total bandwidth consumption is the summation of bandwidth consumed by all subgroups.

$$E[C] = \sum_{i=0}^{r} E[C_i] \quad (41)$$

where $E[C_i]$ is the expected bandwidth consumed in N_i links handled by the i^{th} DR. The optimal DR locations make the network consume the minimal bandwidth consumption E[C].

As previously discussed, the evaluation of bandwidth $E[C_i]$ depends on the policy of loss recovery. Similarly, the following discusses two policies of repair packet retransmissions.

A. Policy M

The sender or DRs multicast repair packets to the subgroup after losses take place. We assume that each subgroup is independent of each other. The losses within the DR_1 have influence only on N_1 links. The losses outside the DR_1 do not have influence on the bandwidth consumption of N_1 links but they affect N_0 links. Therefore, one may obtain the following bandwidth consumption of N_0 (or N_1) links from (32).

$$E[C_i] = E[M_i] \cdot \sum_{k \in N_i} \eta_k + \beta(E[M_i] - 1) \sum_{k \in R_i} d_k (1 - \eta_k) \quad (42)$$

where $E[M_i]$ is the expected number of transmissions for N_i links and can be easily evaluated from (24) (i=0,1,…,r).

B. Policy H

The sender or DRs retransmit the repair packet to only those receivers that did not receive the packet. We can estimate the bandwidth consumption of N_i links (i=0, 1, …, r) using (36).

$$E[C_i] = \sum_{k \in N_i} \eta_k + \sum_{k \in N_i} p_k \eta_k E[M \mid k] \cdot \sum_{l \in D_k} (\eta_l + \beta) \quad (43)$$

VI. RESULTS AND DISCUSSIONS

A. The number of transmissions

We use an example of a general topology in Fig. 6 to compare the exact equations (2)-(5) and the approximate equation (24) with given link heterogeneous loss probability. First, we compare one special case where all router links have the same loss probability p_router while all receiver links have the same loss probability p_receiver. The approximate results with different p_router and p_receiver in Fig. 7 show that they match well to the exact results.

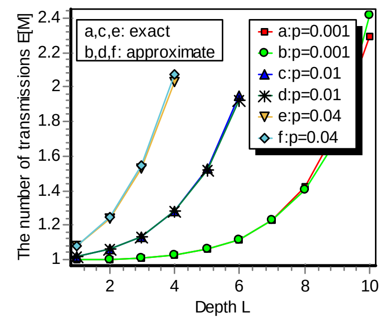

Fig. 5 A comparison of E[M] for binary topologies.

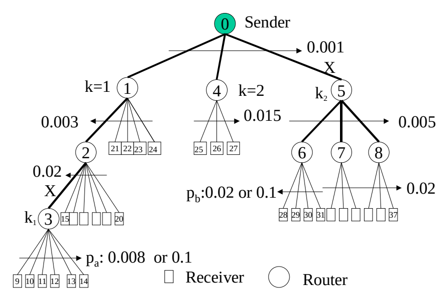

Fig. 6 One example of a general multicast network. Table 4 gives the values of p_a and p_b for different scenarios HL0, HL1 and HL2.

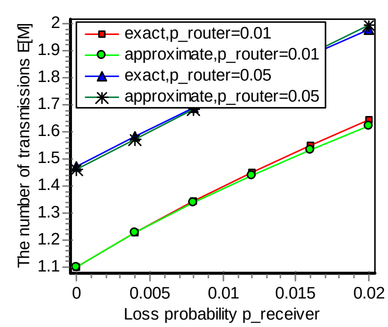

Fig. 7 Comparison of E[M] between exact and approximate solutions

We then compare exact and approximate E[M] where the various loss probability are given in Fig. 6. However, in Fig. 8, we increment all heterogeneous loss probabilities in Fig. 6 by the same variable amount delta p. In addition, two link losses p_a and p_b are varied as per Table 4 leading to the HL0, HL1, HL2 cases. The exact and approximate results in Fig. 8 match closely.

In Fig. 9, We compare E[M] values of 3 different topologies for a uniformly distributed random link loss probability in the range {0, 0.01}. These are linear, star and full binary trees. We take N and all loss probability to be the same for all these topologies under all random loss scenarios. Fig. 9 shows that even if these topologies have the same value of N, E[M] differs greatly depending on the type of topology.

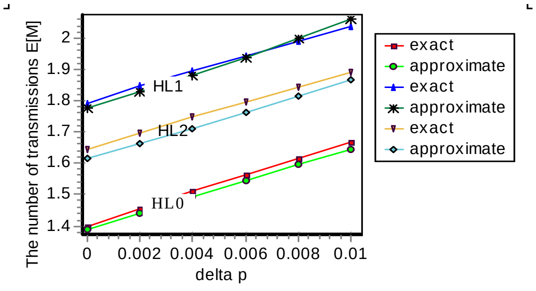

Fig. 8 Comparison of exact and approximate E[M] for a multicast network with heterogeneous loss probability where the loss probability of each link is the summation of value of the link in Fig. 6 and delta p.

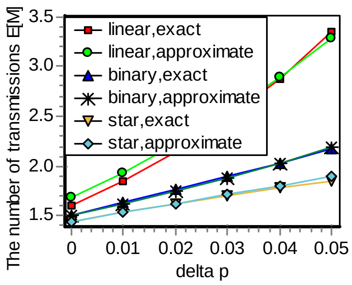

Fig. 9 Comparison of E[M] for different topologies. N=14.

B. *Homogeneous link loss probability*

In a large multicast tree, if the background traffic (unicast traffic and other multicast sessions) is randomized over all links, and if link channel error effects are ignored, then same link loss probability will prevail over all links. There will be minimal or no loss changes from link to link since the traffic over all links is almost the same. Multicast retransmissions from one sender and no DR will lead to the same uniform traffic above. Even in the case of using DR for error correction, these DRs will handle the retransmissions in their subgroup while the sender handles the rest of retransmissions, leading to the same uniform traffic above, i.e. total original and retransmissions traffic is uniformly distributed over all links. However, if the background traffic, and /or link errors are not uniform over all links, then one has to assume different loss probability for each link. In this section, we will compute the optimal DR locations for the network with homogeneous (same) link loss probability.

As an example of optimizing the location of DRs, we use a binary tree to illustrate the performance of uniform loss probability. Suppose that the binary tree has depth L and homogeneous link loss probability p. Thus, the total number of links is $N=2^{L+1}-2$. For hierarchical reliable multicast, the sender and/or DRs need to transmit the packet several times to have a successful multicast transmission.

Based on the estimate of bandwidth consumption, the optimal placements of DRs are obtained further. Fig. 10 - Fig.15 give the optimal placements of DRs for two policies. Fig. 10 and Fig.11 show that the best location of 1 DR is in level 1 for the two policies. When 2 DRs are used, the optimal locations are: one DR in level 1, another DR in level 2 for policy M, as shown in Fig.12. For policy H, the best locations of 2 DRs are in level 1. When 3 DRs are used, all 3 DRs are placed in level 2 for policy M to obtain the minimal bandwidth consumption, as shown in Fig.14. The best locations of 3 DRs for policy H are 1,2,2. The best locations of two policies are very close. These results are exactly the same for the two methods of computing E[M], i.e. the exact and approximate solutions.

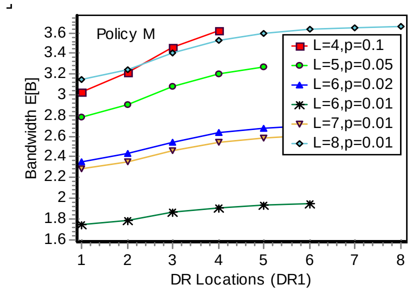

Fig. 10 The bandwidth E[B] of 1 DR for policy M.

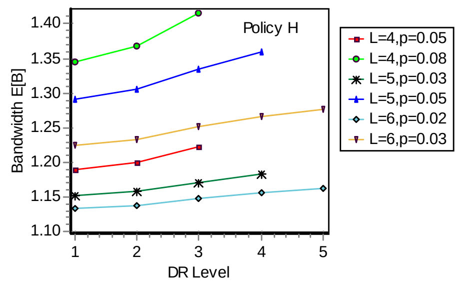

Fig.11 The bandwidth E[B] of 1 DR for policy H

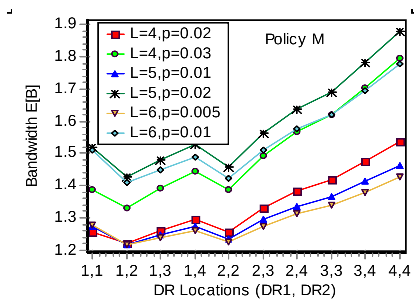

Fig.12 The bandwidth E[B] of 2 DRs for policy M. x axis is assumed levels of the 2 DRs.

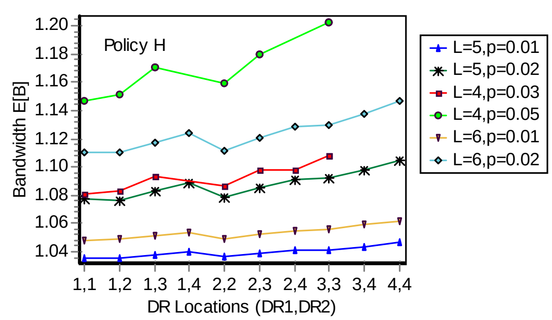

Fig.13 The bandwidth E[B] of 2 DRs for policy H

Some results with different tree depth and different loss probability are also given in Fig.12 - Fig.15. It is obvious that one can greatly reduce the bandwidth consumption by placing DRs in appropriate levels. If these DRs are not placed in appropriate locations, bandwidth consumption is large. From these curves, one can easily see that the bandwidth consumption increases with the increasing depth L and loss probability p. That is because the number of retransmissions will become larger and larger as N and p increase. However, the optimal locations do not change with depth L and/or loss probability p although they change with the number of DRs. As has been proved in [15], this is true only for homogeneous binary topologies.

This goes well with saying that the best DR allocation policy is the one that divides the total network into r+1 equal parts, where r is the number of DRs such that the duty of error correction is fairly shared in homogeneous topologies. This equalized distribution is clearly independent of the number of links.

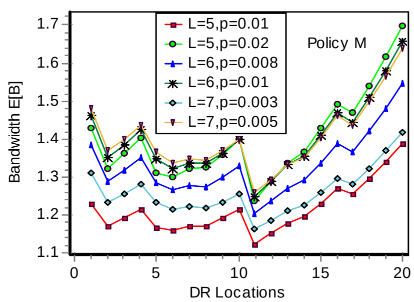

Fig.14 The bandwidth E[B] of 3 DRs for policy M. x axis is assumed levels of the 3 DRs as shown in Table 3.

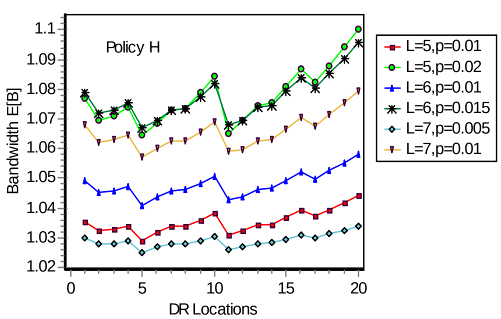

Fig.15 The bandwidth E[B] of 3 DRs for policy H. x axis is assumed levels of the 3 DRs as shown in Table 3.

Table 3 The DR levels corresponding to the x axis of Fig.14 and Fig.15

X axis	1	2	3	4	5	6	7
DR levels	1,1,1	1,1,2	1,1,3	1,1,4	1,2,2	1,2,3	1,2,4
	8	9	10	11	12	13	14
	1,3,3	1,3,4	1,4,4	2,2,2	2,2,3	2,2,4	2,3,3
	15	16	17	18	19	20	
	2,3,4	2,4,4	3,3,3	3,3,4	3,4,4	4,4,4	

The optimal DR locations for the two loss recovery policies give comparable results although policy H has less bandwidth consumption than policy M. The optimal partitioning for policy M also gives a bandwidth consumption saving up to 40%. Policy H can work efficiently for binary trees without DR. However, in the case of linear-like topologies, DRs are required to further aid the loss recovery of Policy H. Fig.14 shows the importance of DR location in Policy H for a linear topology where the bandwidth saving is up to 30%.

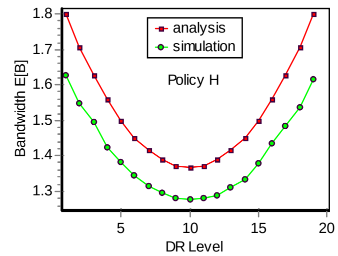

Fig.16 The bandwidth E[B] of 1DR for policy H in the linear topology.

One may also see an obvious result that the closer the DR is to the receivers, the more bandwidth is consumed. If bandwidth consumption is to be minimized, one should not place DRs too close to the receivers for binary trees. The explanation for this is that the sender will consume more bandwidth while DRs consume less leading to a total bandwidth consumption which is still large. Thus, the optimal locations are actually a trade-off of bandwidth consumption between the sender and these DRs.

C. Heterogeneous loss probability

In this section, we discuss the optimal DR locations of the multicast network with non-uniform link loss probability, i.e. each link may have different packet loss probability. In this case, we also simulate the multicast transmissions with two loss recovery schemes and compare the analysis and simulation results as shown in Fig.17 and Fig.18. In the simulation, we average the number of transmissions and bandwidth consumption over 100000 packets. This was found enough to provide us with a confidence interval of 99.9% about the true mean of each result in all simulation figures. These results show that the analysis and simulations give the same optimal DR locations for both loss recovery schemes. Simulation results prove that bandwidth consumption can be evaluated by our analysis equations.

Table 4 Packet loss probability and optimal DR location of Fig. 6 for different scenarios shown in Fig.17 and Fig.18.

Scenario		HL0	HL1	HL2
Loss probability p_a		0.008	0.1	0.008
Loss probability p_b		0.02	0.02	0.1
The Optimal RR Location	Policy M	1	2	5
	Policy H	2	3	6

We assume different packet loss probability for some links of the network in Fig. 6. One finds from Fig.17 and Fig.18 that the optimal RR location depends on the policy and loss scenario. DR should be located closer to the links of high packet loss probability for both loss recovery schemes. The RR locations depend on the distribution of packet loss probability over different links. The high loss links frequently request retransmissions. In order to reduce the effects of such retransmissions on other nodes, the DR moves to the high loss links to support efficient loss recovery, which applies to the two schemes. The two policies yield similar DR locations.

Fig.19 gives the effects of NAK overhead on the bandwidth consumption of multicast networks. NAK packets are usually much smaller than data packets and require less bandwidth consumption. Due to the existence of DRs, NAK packets are noticeably reduced for policy M and are minimized for policy H.

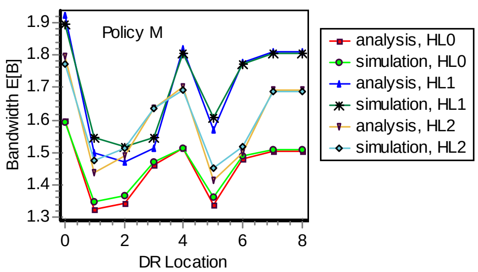

Fig.17 The bandwidth E[B] of heterogeneous packet loss probability in Fig. 6 for policy M

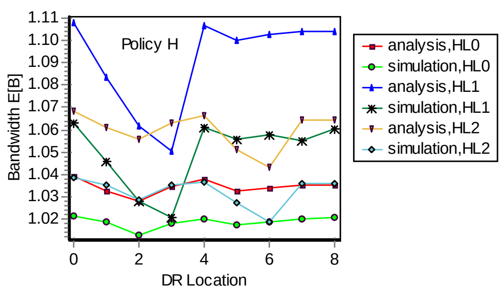

Fig.18 The bandwidth E[B] of heterogeneous packet loss probability in Fig. 6 for policy H

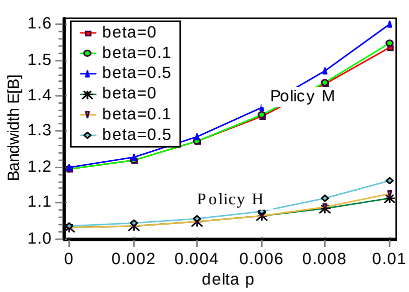

Fig.19 The effect of NAKs.

VII. Conclusions

In this paper, we have derived an analytical approximation for the number of transmissions for general multicast trees. Furthermore, we have analyzed the bandwidth consumption of two loss recovery polices, i.e., policy M and H. We have also compared the values of bandwidth consumption provided by the analysis and simulation. The bandwidth consumption of policy H was found to be much smaller than that of policy M. The bandwidth consumption of NAKs packets was found to be negligible compared to data bandwidth.

Based on the estimation of bandwidth consumed by the two policies, the optimal DR placements have been investigated in this work. The results based on the binary tree of uniform loss probability show that the best placements of DRs are closer to the sender rather than the receivers. For policy M, the best location is in level 1 when only 1 DR is used to recover loss. When 2 DRs are used, the optimal locations are: 1 DR in level 1, and another DR in level 2. When 3 DRs are used, 3 DRs are placed in level 2 to obtain the minimal consumption of bandwidth. For policy H, the best location is in level 1 for 1 DR. When 2 DRs are used, the optimal locations are: both DRs in level 1. When 3 DRs are used, 2 DRs are placed in level 2 and one DR is placed in level 1 to obtain the minimal bandwidth consumption. The optimal placements for the two policies are very close though slightly different. We have also investigated the optimal DR locations for the multicast network with heterogeneous loss probability. The optimal DR location should be located closer to the links of high loss probability for both loss recovery schemes.

VIII. Appendix A

To prove lemma (45) of this appendix, we start with a simple example where we define ξ_k as loss probability of link k and φ_k as the number of links from the sender to link k, i.e. $\xi_k=p_k$ and $\varphi_k=d_k$, (see Fig.20). We obtain the following equation:

$$\begin{aligned}
&\sum_{k=1}^{N} p_k \sum_{j\in d_k}^{j\neq k} d_j = p_2 d_1 + p_3 d_1 + p_4(d_1+d_3) + p_5(d_1+d_3) \\
&+ p_6(d_1+d_3+d_5) + p_7(d_1+d_3+d_5) + p_8(d_1+d_3) \\
&+ p_{10}d_9 + p_{11}d_9 = d_1(p_2+p_3+p_4+p_5+p_6+p_7+p_8) \\
&+ d_3(p_4+p_5+p_6+p_7+p_8) + d_5(p_6+p_7) \\
&+ d_9(p_{10}+p_{11}) = \sum_{k=1}^{N} d_k \sum_{j\in n_k} p_j
\end{aligned} \tag{44}$$

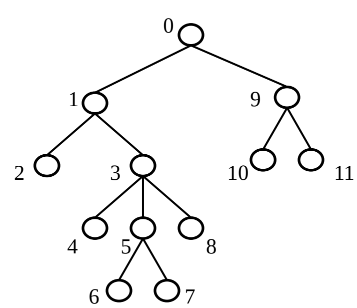

Fig.20 An example of tree topology

Lemma 1:

For an arbitrary tree, we define two functions ξ_k and φ_k

affiliated with node k. For example, ξ_k or φ_k or both could be the number of links from sender to node k. In a different utilization, they could represent the number of links under node k, or the number of NAKs, or loss probability p_k, or $p_k\eta_k$ defined before, or just 1. Many such combinations are possible scenarios for the application of (45), e.g. $\xi_k=n_k$ and $\varphi_k=p_k$, or $\xi_k= d_k$ and $\varphi_k=n_k$, or $\xi_k=1$ and φ_k could be any function of p_k such as $\varphi_k= p_k\eta_k$ as in (27), or $\xi_k=d_k$ and $\varphi_k=p_k$, or $\xi_k=d_k$ and $\varphi_k=d_k$, or $\xi_k=p_k$ and $\varphi_k=p_k$, or $\xi_k=d_k$ and $\varphi_k=1$, or $\xi_k=1$ and $\varphi_k= p_k\eta_k$, and so on, to name but a few.

We claim that

$$\sum_{k=1}^{N}\xi_k \sum_{j\in d_k}^{j\neq k}\varphi_j = \sum_{k=1}^{N}\varphi_k \sum_{j\in n_k}\xi_j \quad (45)$$

where d_k represents the links from the sender to node k, n_k represents the links under node k.

Proof:

Suppose that each node of a multicast tree has one counter whose initial value is 0. The k^{th} tick of the counter adds a value (k=1,2,…, N) $\xi_k\varphi_j$ as in the left hand of (45) to the counters of all nodes belonging to the path from the sender to the parent node of node k, i.e., the counter of node j will be incremented by $\xi_k\varphi_j$ if $j\in d_k, j\neq k$, the counters of other nodes are incremented by 0. The left hand side of (45) is sum of final count of all counters when k=N. Noting that node j belongs to many paths, i.e., the paths from the sender to these nodes under node j, after summing up all nodes (i.e, the first summation over k in the left hand of (45)), only these nodes under node j will affect the counter value of node j, i.e, the counter value of node j is $\sum_{k\in n_j}\xi_k\varphi_j$. Thus the summations $\sum_{j=1}^{N}\sum_{k\in n_j}\xi_k\varphi_j$ of all counters equal to $\sum_{k=1}^{N}\xi_k \sum_{j\in d_k}^{j\neq k}\varphi_j$.

In addition, from (45), it follows,

$$\sum_{k=1}^{N}\xi_k \sum_{j\in D_k}\varphi_j = \sum_{k=1}^{N}\xi_k \sum_{j\in d_k}^{j\neq k}\varphi_j + \sum_{k=1}^{N}\xi_k \sum_{j\in n_k}\varphi_j + \sum_{k=1}^{N}\xi_k\varphi_k$$
$$= \sum_{k=1}^{N}\varphi_k \sum_{j\in n_k}\xi_j + \sum_{k=1}^{N}\varphi_k \sum_{j\in d_k}^{j\neq k}\xi_j + \sum_{k=1}^{N}\xi_k\varphi_k = \sum_{k=1}^{N}\varphi_k \sum_{j\in D_k}\xi_j, \quad (46)$$

$$\sum_{k=1}^{N}\xi_k \sum_{j\notin D_k}\varphi_j = \sum_{k=1}^{N}\xi_k \sum_{j\in N}\varphi_j - \sum_{k=1}^{N}\xi_k \sum_{j\in D_k}\varphi_j = \sum_{k=1}^{N}\varphi_k \sum_{j\notin D_k}\xi_j, \quad (47)$$

$$\text{also} \quad \sum_{k=1}^{N} p_k\eta_k n_k = \sum_{k=1}^{N} p_k\eta_k \sum_{l\in n_k}1 = \sum_{k=1}^{N}\sum_{l\in d_k}^{l\neq k} p_l\eta_l, \quad (48)$$

all of which we have actually used within the paper.

In the same way as (45), we obtain the following equations:

$$\sum_{k}\sum_{j\notin D_k} f(j)Q(\Omega=\{k,j\}) = \sum_{k} f(k)\sum_{j\notin D_k}Q(\Omega=\{k,j\}) \quad (49)$$

$$\sum_{k\in N}\xi_k \sum_{l\in R_k}\varphi_l = \sum_{k\in R}\varphi_k \sum_{l\in d_k}\xi_l \quad (50)$$

where R_k is the set of receivers within the loss scenario $\Omega=\{k\}$, R is the set of all receivers in the multicast group.

IX. REFERENCES

[1] Sally Floyd, Van Jacobson, Steven McCanne, Ching-Gung.Liu, Lixia.Zhang, "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing", *Proceedings of ACM SIGCOM* '95, pp.342-356, Oct.1995.

[2] John Lin, Sanjoy Paul, "RMTP: A Reliable Multicast Transport Protocol", *Proc. of IEEE INFOCOM*, San Francisco, USA, pp.1414-1424, 1996.

[3] Joerg Nonnenmacher, Ernst W. Biersack, "Parity-Based Loss Recovery for Reliable Multicast Transmission", *IEEE/ACM Transactions on Networking*, pp.349-361, Vol.6, No.4, August 1998.

[4] Athina P. Markopoulou, Fouad A.Tobagi, "Hierarchical Reliable Multicast: performance analysis and placement of proxies. *Proceedings of the Second International Workshop on Networked Group Communication (NGC 2000)*, Palo Alto, USA, ACM Press, pp. 27-35, Nov. 2000.

[5] Bo Li, M.Golin, G.Italiano, et al, " On the optimal placement of web proxies in the Internet", *Proc. of IEEE INFOCOM*, pp.1282-1290, 1999.

[6] Christos Papadopoulos, Guru Parulkar and George Varghese, "An Error Control Scheme for Large-Scale Multicast Applications", *Proc. of IEEE INFOCOM*, pp.1188-1196,San Francisco, USA,1998.

[7] Adam M.Costello, Steven McCanne, "Search Party: Using Randomcast for Reliable Multicast with Local Recovery", *Proc. Of IEEE INFOCOM*, pp.1256-1264, 1999.

[8] Zhen Xiao, Kenneth P.Birman, "A Randomized Error Recovery Algorithm for Reliable Multicast", *Proc.* of *IEEE INFOCOM*, pp. 239-248, 2001.

[9] Luigi Rizzo, "Effective erasure codes for reliable computer communication protocols", *Comput. Commun. Rev.*, vol.27, No.2, pp. 24-36, Apr. 1997.

[10] Dino Farinacci, "PGM Reliable Transport Protocol Specification", *Internet draft rfc3208*, Aug. 1998.

[11] Pravin Bhagwat, Partho P.Mishra, Satish K.Tripathi, "Effect of Topology on Performance of Reliable Multicast Communication", *Proc. of IEEE INFOCOM.* 1994.

[12] Joerg Nonnenmacher, Ernst W.Biersack, "Performance Modeling of Reliable Multicast Transmission", *Proc. of IEEE INFOCOM*, 1997.

[13] Hwa-Chun Lin, Kuen-Feng Yang, "Placement of Repair Servers to Support Server-Based Reliable Multicast", *IEEE ICC*, pp.1802-1806, vol.16, 2001.

[14] Markus Yavatkar, James Griffioen, and Madhu Sudan, "A reliable dissemination protocol for interactive collaborative applications", *Proceedings of ACM Multimedia*, 1995.

[15] Zuo Wen Wan, Michel Kadoch, and Ahmed Elhakeem, "Performance Evaluation Of Tree-based Reliable Multicast", *the 12th IEEE International Conference on Computer Communications and Networks (ICCCN'03)*, Dallas Texas, USA, October 20 - 22, pp.67-73, 2003.

[16] Pavlin Radoslavov, Christos Papadopoulos, R.Govindan, and et al, "A comparison of application-level and router-assisted hierarchical schemes for reliable multicast", *Proceedings of IEEE INFOCOM*, pp.229-238, 2001.

Ahmed K. Elhakeem received his Ph.D. degree from the Southern Methodist University, Dallas, TX, in 1979. He spent the next two years working as a Visiting Professor in Egypt, after which he moved to Ottawa, Canada, in 1982. He assumed research and teaching positions at Carleton and Manitoba Universities, and later moved to Concordia University, Montreal, Canada, where he has been a Professor in the Electrical and Computer Engineering Department since 1983. He has published numerous papers in IEEE and in international journals in the areas of spread spectrum and networking. He is a well-known expert in these areas and serves as a consultant to various companies. His current research interests include interconnected wireless LANs, Error correction for IP Multicast, wideband networks, switching architectures, CDMA networks, multi protocols, software Radios, and reconfigurable networks. He is a co-author of the book *Fundamentals of Telecommunications Networks* (New-York: Wiley, 1994). He has chaired

and organized numerous technical sessions in IEEE conferences, and was the Technical Program Chairman for IEEE Montech '86 in Montreal, Canada. More recently, he was the Key Guest Editor for four issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on Code Division Multiple Access, namely CDMA I, II, III and IV, appearing in May '94, June '94, October '96, and December '96. Dr. Elhakeem was the Communications chair of IEEE Montreal and the TCCC Rep. to ICC '99. He served as an associate editor for the IEEE Communications Letters Journal (1996-1999). He is a professional engineer in Ontario and a senior member of IEEE. Dr. Elhakeem has served as a consultant to many companies, and to Communication Research Center and Industry Canada.