On the Performance of Error-resilient End-point-based Multicast Streaming

György Dán, Ilias Chatzidrossos, Viktória Fodor and Gunnar Karlsson

School of Electrical Engineering

KTH, Royal Institute of Technology

Stockholm, Sweden

Abstract— In this paper we propose an analytical model of a resilient end-node multicast streaming architecture based on multiple minimum-depth-trees that employs path diversity and forward error correction for improved resilience to node churns and packet losses. We study the performance of the architecture in the presence of packet losses and dynamic node behavior. We show that for a given redundancy the probability that an arbitrary node possesses a packet is high as long as the loss probability in the network is below a certain threshold. After reaching the threshold the packet possession probability suddenly drops; the rate decrease gets faster as the number of nodes in the overlay grows. The value of the threshold depends on the ratio of redundancy and on the number of the distribution trees. We study the overlay structure in the presence of node dynamics and conclude that stability can be achieved only if the root node serves a large number of nodes simultaneously.

I. INTRODUCTION

The delivery of streaming media over end-point overlays has received a lot of attention recently ([1], [2] and references therein). In an end-point-based multicast distribution system end-points are organized or organize themselves into an application layer overlay and distribute the data among themselves. Such systems are easy to deploy and they reduce the load of the content provider, since the distribution cost in terms of bandwidth and processing power is shared by the nodes of the overlay. Since the success of such schemes depends on the behavior of the participating nodes, several issues have to be dealt with, such as the effects of group dynamics, stability of the system or the incentives for nodes to collaborate. Furthermore, since nodes receive data from their peer nodes only, the performance of such a scheme in an error prone environment is unclear due to possible error propagation.

The first proposed architectures focused primarily on low overhead due to control traffic and on the efficiency of the data distribution [3], [4]. Resilience to node failures and error prone transmission paths appeared as important criteria later. Robustness to node churns, i.e., node departures that disturb the data flow, was considered in SRMS by distributing packets to randomly chosen neighbors outside of the distribution tree [5]. Though this scheme provides some resilience to losses, it is known that repeating information is less efficient than using error correcting codes. SplitStream [6] and CoopNet [1] introduce multiple distribution trees and employ priority

---

encoding transmission (PET) [7] based on forward error correction (FEC) [8] to decrease the effects of node failures and to recover from packet losses. The feasibility of these overlays was studied via simulations based on measured traces of user behavior in [2]. The authors concluded that application layer multicast architectures have enough resources, are stable in spite of group dynamics, and hence can support large scale streaming content distribution. The authors in [9] proposed the use of time shifting and video patching to provide robustness to node departures. Robustness to errors using selective ARQ and error resilient video coding was considered in [10]. There are some implemented peer-to-peer multicast systems in the Internet [3], [11], but these systems suffer from large startup delays in the order of minutes, and poor stability.

Albeit an extensive literature on end-point-based multicast streaming, previous work on the behavior of these systems was mainly based on simulations. In [12] the authors presented a mathematical model for a minimum-width CoopNet like overlay employing multiple distribution trees and FEC for resilience. The model was extended to include correlated losses in [13] and applied to model dynamic node behavior. But the feasibility of the overlay considered in these works for large-scale deployment is questionable as the maximum distance of the nodes from the root node increases as $O(N)$, where N is the number of nodes in the overlay. In a large overlay, delay and delay jitter become an issue.

In this paper we present a model for a minimum-depth CoopNet like overlay combined with FEC. The maximum depth of the overlay considered in this paper grows as $O(log(N))$ and hence delay and delay jitter are less of a problem. It is recognized however that the overlay can be disconnected due to node departures and requires complex maintenance. We evaluate the performance of data distribution for a large number of nodes in the case of independent losses and investigate the effects of node dynamics on the stability of the overlay.

The rest of the paper is organized as follows. In Section II we give a brief description of the considered overlay for multicast. In Section III we evaluate the performance of the data distribution. In Section IV we analyze the stability of the overlay. In Section V we conclude our work.

II. System description

We consider an application overlay as the one described in [1], [2], [6] consisting of a root node and N peer nodes. Peer nodes are organized in t distribution trees, either by a distributed protocol like in [6] or by a central entity like in [1]. The nodes are members of all t trees, and in each tree they have a different parent node from which they receive data. Each node can have up to t children to which it forwards data in one of the t trees, called the fertile tree of the node. In all other trees the node is sterile, that is, it does not have any children. We say that each node has t cogs in its fertile tree and has no cogs in its sterile trees. Child nodes can be connected to the available (not yet taken) cogs of a node. If we denote the number of layers in the trees by L, then in a well maintained tree each node is $1 \leq i < L$ hops away from the root node in its fertile tree, and $L-1 \leq i \leq L$ hops away in its sterile trees. We denote the maximum number of children of the root node in each tree by m, and we call it the multiplicity of the root node. Hence, the number of cogs of the root node is mt. We assume that nodes do not contribute more bandwidth towards their children as they use to download from their parents. (See Fig. 1).
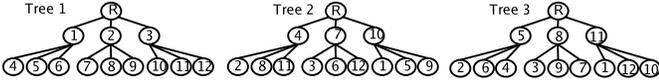
Fig. 1. Multicast tree structure for $t=3$, $m=3$ and $N=12$.

The root uses block based FEC, e.g., Reed-Solomon codes [14], so that nodes can recover from packet losses due to network congestion and node departures. To every k packets of information c packets of redundant information are added resulting in a block length of $n = k + c$. If a source would like to increase the ratio of redundancy while maintaining its bitrate unchanged, then it has to decrease its source rate. We denote this FEC scheme by FEC(n,k). Using this FEC scheme one can implement UXP, PET, or the MDC scheme considered in [1]. Lost packets can be reconstructed as long as no more than c packets are lost out of n packets. The root splits the data stream into t stripes, with every t^{th} packet belonging to the same stripe, and it sends every t^{th} packet to its children in a given tree. If $n \leq t$ then at most one packet of a block is distributed over the same distribution tree. Peer nodes relay the packets upon reception to their respective child nodes in the tree corresponding where they are fertile, and once they received at least k packets of a block of n packets they recover the remaining c packets. If a packet, which should have been received in the tree where the node is fertile, is recovered, then it is sent to the respective children. A packet received from the parent node after it has been decoded based on other packets in the block is discarded.

III. Data distribution

In the following section we evaluate the performance of the overlay in the absence of group dynamics. In this scenario the only source of failure is the loss of packets between peer-nodes.

A. Mathematical model

In this section our goal is to calculate the probability $\pi(i)$ that a node, which is in layer i in the tree where it is fertile, receives or can reconstruct an arbitrary packet. The mathematical model we present describes the behavior of the overlay in the presence of independent packet losses. We denote the probability that a packet is lost between two adjacent nodes by p. We assume that the probability that a node is in possession of a packet is independent of that a node in the same layer is in possession of a packet. We also assume that nodes can wait for redundant copies to reconstruct a packet for an arbitrary amount of time. For the model we consider a tree with the maximum number of nodes in the last layer. We will comment on the possible effects of our assumptions later.

To calculate the probability $\pi(i)$ we have to calculate the probability $\pi_f(i)$ that a node, which is in layer i in its fertile tree, receives or can reconstruct an *arbitrary packet in its fertile tree*. Since the root node possesses every packet, we have that $\pi_f(0) = 1$. The probability that a node in layer i receives a packet in a tree is $\pi_a(i) = (1-p)\pi_f(i-1)$. Every node is fertile in one tree and is sterile in the other $t-1$ trees. A node can possess a packet in its fertile tree either if it receives the packet or if it can reconstruct it using the packets received in its sterile trees. Reconstruction can take place if the number of received packets is at least $n-c$ out of the remaining $n-1$, hence we can write for $0 \leq i < L-1$

$$
\begin{aligned}
\pi_f(i+1) = {} & \pi_a(i+1) + \Big\{(1-\pi_a(i+1)) \\
& \sum_{j=n-c}^{n-1} \binom{n-1}{j} \pi_a(L)^j (1-\pi_a(L))^{n-1-j}\Big\}.
\end{aligned} \tag{1}
$$

Based on the probabilities $\pi_f(i)$ we can express the probability $\pi(i)$ ($0 \leq i < L-1$) as the probability of possessing an arbitrary packet in a block of n packets, i.e., the mean number of packets possessed after FEC reconstruction in a block of n packets. If a node receives at least k packets in a block of n packets then it can use FEC to reconstruct the lost packets, and hence possesses all n packets. Otherwise, FEC cannot be used to reconstruct the lost packets. Hence for $\pi(i)$ we get the equation

$$
\begin{aligned}
\pi(i+1) = {} & \frac{1}{n}\Big\{\pi_a(i+1) \sum_{j=1}^{n} \tau(j) \binom{n-1}{j-1} \\
& \pi_a(L)^{j-1}(1-\pi_a(L))^{n-1-(j-1)}\Big\} + \\
& \frac{1}{n}\Big\{(1-\pi_a(i+1)) \sum_{j=0}^{n-1} \tau(j) \binom{n-1}{j} \\
& \pi_a(L)^j (1-\pi_a(L))^{n-1-j}\Big\},
\end{aligned} \tag{2}
$$

where $\tau(j)$ indicates the number of packets after FEC reconstruction if j packets have been received; it is given as

$$
\tau(j) = \begin{cases} j & 0 \leq j < k \\ n & k \leq j \leq n. \end{cases}
$$

To calculate the probabilities $\pi_f(i)$ we use an iterative method. First, we set $\pi_f(L-1)^{(0)} = 1$ and calculate the probabilities $\pi_f(i)^{(1)}$, $1 \le i \le L$. Then, in iteration r, we calculate $\pi_f(i)^{(r)}$, $1 \le i \le L$ using $\pi_f(L-1)^{(r-1)}$. The iteration stops when $\pi_f(L-1)^{(r-1)} - \pi_f(L-1)^{(r)} < \varepsilon$, where $\varepsilon > 0$.

Based on the $\pi(i)$ we can calculate the probability π of packet possession for an arbitrary node in the overlay by weighting the $\pi(i)$ with the portion of nodes that are in layer i of their fertile tree

$$\pi = \sum_{i=1}^{L-1} \frac{t^{i-1}}{(t^{L-1}-1)/t-1} \pi(i). \tag{3}$$

1) Lower bound: In the following we give an asymptotic lower bound on the values of the above probabilities to better understand their evolution. Let us consider an overlay where the number of layers can be arbitrarily high. It is clear that $\pi_f(i)$ is a non-increasing function of i and $\pi_f(i) \ge 0$. Hence $\lim_{i \to \infty} \pi_f(i) = \pi_f(\infty)$ exists, and instead of eq. (1) we get the following nonlinear recurrence equation

$$\pi_f(i+1) = \pi_a(i+1) + \{(1 - \pi_a(i+1)) \tag{4}$$
$$\sum_{j=n-c}^{t-1} \binom{n}{j} \pi_a(i+1)^j (1 - \pi_a(i+1))^{n-1-j} \}.$$

This equation is the same as eq. (4) in [12], and thus the analysis shown there can be applied to describe the evolution of $\pi_f(i)$. For brevity, we only state the main results regarding $\pi_f(i)$, for a detailed explanation see [12], [13]. For every (n,k) there is a loss probability p_{max} below which the packet possession probability $\pi_f(\infty) > 0$ and above which $\pi_f(\infty) = 0$. Furthermore, for any $0 < \delta < 1$ there is (n,k) such that $\pi_f(\infty) \ge \delta$. Consequently, in the overlay considered in this paper, we have $\pi_f(i) \ge \pi_f(\infty) > 0$ and $\pi(i) \ge \pi_f(\infty) > 0$ for $p < p_{max}$. For loss probabilities $p > p_{max}$ such a positive lower bound cannot be given, and the packet possession probability approaches 0 as the number of layers increases. We refer to the system as stable if $p < p_{max}$ and call it unstable otherwise. To get a lower bound on $\pi(i)$ we substitute $\pi_f(\infty)$ in eq. (2) instead of $\pi_f(i)$ and $\pi_f(L-1)$. The lower bound is 0 for $p > p_{max}$ and is positive otherwise.

2) Discussion: In the following we discuss the validity of certain assumptions made in the model. The model does not take into account the correlations between packet losses in the Internet. Losses occurring in bursts on the output links of the nodes influence the performance of the overlay if several packets of the same block are distributed over the same tree, that is if $n > t$. Bursty losses in the backbone influence the performance if packets of different distribution trees traverse the same bottleneck. The effects of correlated losses on the input links of the nodes has been considered in [13], and the analysis showed that correlated losses slightly decrease the performance of the overlay.

In the analysis we assume that the number of nodes in the last layer of the tree is maximal. If the number of nodes in the last layer of the tree is not maximal then some nodes are in layer $L-1$ in their sterile trees, and the overlay's performance
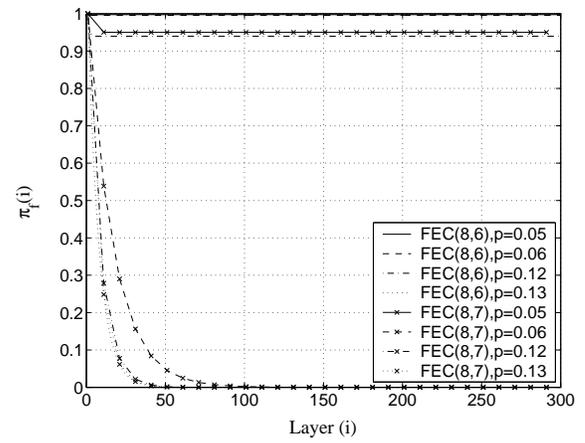
Fig. 2. $\pi_f(i)$ vs. i for $t = 8$, $m = 8$, $n = 8$. $\pi_f(\infty) > 0$ for $p < p_{max}$, $\pi_f(i)$ decreases quickly in i otherwise.

is slightly better. The results of the asymptotic analysis still hold however.

The model does not take into account node departures, an important source of disturbances for the considered overlay. Following the arguments presented in [13] node departures can be incorporated in the model as an increase of the loss probability by $p_\omega = N_d/N \times \theta$, where N_d is the mean number of departing nodes per time unit and θ is the time nodes need to recover from the departure of a parent node. The simulation results presented in [13] support this hypothesis.

B. Performance evaluation

In this section we first show results obtained with the mathematical model and then verify the results via simulations. In all scenarios we set $t = n$ and consider the streaming of a 112.8 kbps stream to nodes with link capacity 128 kbps. The packet size is set to 1410 bytes. Each node has a playout buffer of 70 packets, which corresponds to 7 seconds of playout buffer delay. Each node has an output buffer of 40 packets to absorb the bursts of outgoing packets in its fertile tree. The simulation time is 4000 seconds and the presented results are the averages over 10 simulation runs.

Figure 2 shows $\pi_f(i)$ as a function of i for $t = 8$, $m = 8$, $n = 8$, $L = 300$, and different ratios of redundancy and packet loss probability p. The figure shows that $\pi_f(i) > 0$ for any i as long as the packet loss probability is below p_{max}, but decreases to 0 rapidly if $p > p_{max}$. For $c = 1$ the threshold is $p_{max} = 0.0536$ and for $c = 2$ it is $p_{max} = 0.1292$.

Figure 3 shows $\pi(i)$ as a function of i for $t = 8$, $m = 8$, $n = 8$, $L = 300$ and different ratios of redundancy and packet loss probability p. The figure shows that $\pi(i)$ evolves similarly to $\pi_f(i)$ for $p > p_{max}$. For $p < p_{max}$ it drops even faster than $\pi_f(i)$ since $\pi(i)$ is a function of $\pi_f(i)$ and $\pi_f(L-1)$ as shown by eq. (2). Consequently, we expect that increasing the number of layers in the overlay worsens its performance whenever the overlay is unstable.

Figure 4 shows π as a function of the packet loss probability
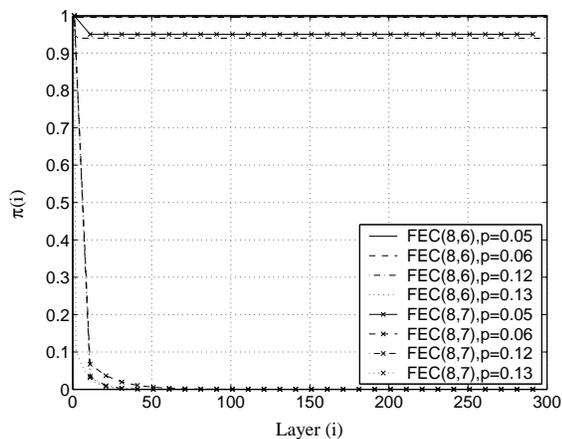
Fig. 3. $\pi(i)$ vs. i for $t = 8$, $m = 8$, $n = 8$. $\pi(i)$ shows similar behavior to $\pi_f(i)$, but the decrease is even faster than in the case of $\pi_f(i)$ for $p > p_{max}$.

Fig. 4. π vs. packet loss probability for $t = 4$, $m = 4$, $n = 4$, $c = 1$. The vertical bars show $\pi(1)$ and $\pi(L-1)$.

for $t = 4$, $m = 4$, $n = 4$, $c = 1$ and various values of L. The vertical bars show the values $\pi(1)$ at the upper end and $\pi(L-1)$ at the lower end. Note, that if $L = 2$ then $L - 1 = 1$, hence there is no vertical bar. The figure shows that the packet possession probability is high as long as the loss probability is below p_{max}. It drops however as the packet loss probability crosses the threshold. The drop of packet possession probability gets worse as the number of layers and hence the number of nodes in the overlay increases. At the same time, for $p > p_{max}$, the difference between $\pi(1)$ and $\pi(L-1)$ (the packet possession probability of nodes that are fertile in the first and the penultimate layers, respectively) increases. The figure shows as well that increasing the number of layers affects the packet possession probability $\pi(1)$ (and $\pi(i)$ in general), that is, the performance experienced by nodes that are already part of the overlay is influenced by arriving nodes. The effect is negligible when the system is stable, but becomes large when it is unstable. The line corresponding to $L = 500$ shows the asymptotic value of the packet possession probability as the number of layers increases. The value of π for large values of L is close to $\pi_f(L-1)$, and it shows the evolution of the fixed point of eq. (4) for $p < p_{max}$. This curve is in accordance with those in Figs. 2 and 3 and in [12], [13].

Figure 5 shows π as a function of the packet loss probability for $t = 8, m = 8, n = 8$ and various values of c and L. Comparing results obtained with different values of c we see that increasing the ratio of redundancy increases the value of the threshold p_{max}. At the same time, increasing the ratio of redundancy makes the drop faster once p_{max} has been exceeded. Comparing Figs. 4 and 5 shows that increasing the FEC block length slightly increases the value of p_{max}, at the price of a faster drop once the threshold in exceeded.

Figures 6 and 7 show results obtained via simulations for the same scenarios as in Figs. 4 and 5 respectively, and show perfect match with the analytical model. We did not perform simulations for $L = 500$, as even for $t = 4$ an overlay with 500
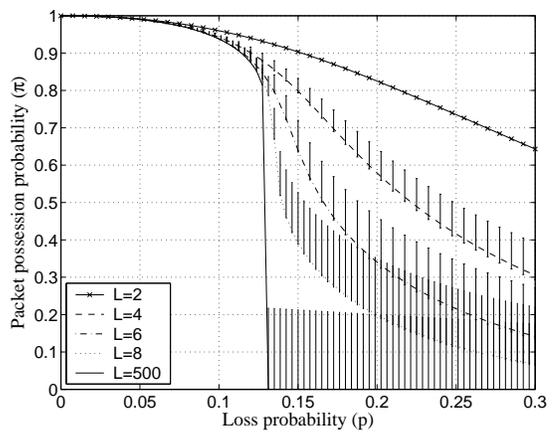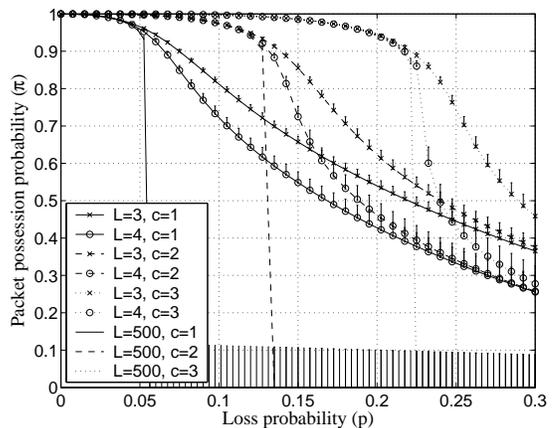
Fig. 5. π vs. packet loss probability for $t = 8$, $m = 8$, $n = 8$. The vertical bars show $\pi(1)$ and $\pi(L-1)$.

layers would consist of approximately 4^{499} nodes.

IV. DISTRIBUTION TREE STABILITY

In this section we analyze the stability of the overlay in the presence of node dynamics. We show a necessary condition for the overlay to be feasible. Then we present an approximate model to describe the evolution of the available capacity in an arbitrary tree of the overlay, and use the model and simulations to study the stability of the overlay.

In the presence of node dynamics, children of the departing nodes have to find a new parent in the fertile tree of the departing node. Finding a parent is however not possible if the number of fertile nodes in the tree is too low, as we will explain later. If some nodes cannot find parents in a tree, the tree becomes disconnected. In [6] a decentralized mechanism was proposed to resolve this problem, in [1] the entity responsible for tree construction was responsible for moving fertile nodes from trees with available capacity to the trees without available capacity. In an environment with high churn rate reallocating fertile nodes between trees can lead
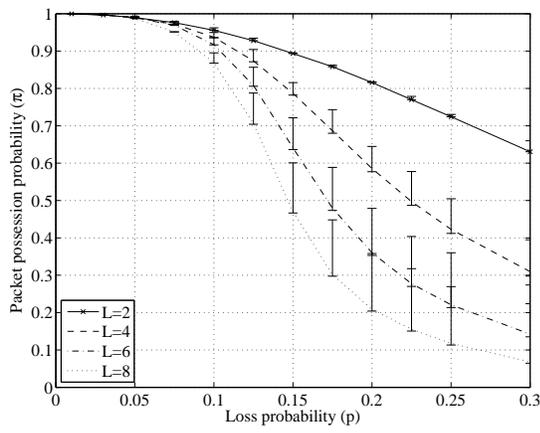
Fig. 6. π vs. packet loss probability for $t = 4$, $m = 4$, $n = 4$, $c = 1$. The vertical bars show $\pi(1)$ and $\pi(L-1)$. Simulations.
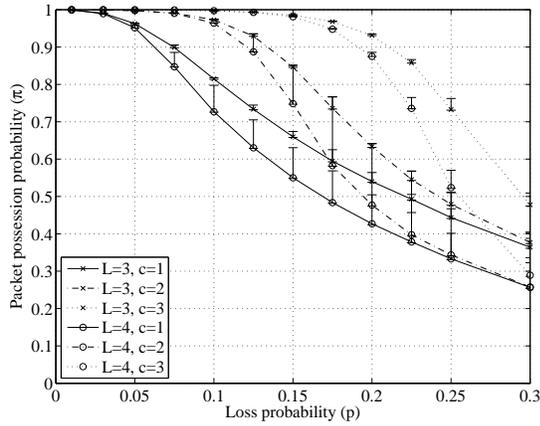
Fig. 7. π vs. packet loss probability for $t = 8$, $m = 8$, $n = 8$. The vertical bars show $\pi(1)$ and $\pi(L-1)$. Simulations.

to high management overhead. Another solution is to allocate arriving nodes to be fertile in the disconnected trees, and hence make the trees connected. This way the failures of the overlay due to node departures can be healed by the arrivals; if a node fails to find a parent, it can reattempt after a reconnection interval T.

A. Overlay feasibility

We say that the overlay is feasible for given m, t and N if the overlay can be constructed. A necessary and a sufficient condition for the tree to be feasible was shown in [6]. Those conditions were based on the number of cogs that nodes want to use and are willing to offer. The condition shown here extends those conditions and is a condition that relates the parameters of the overlay to each other.

Proposition 1: If the overlay is feasible for arbitrary N, then $m \geq t - 1$.

Proof: We show that this condition is necessary in a well maintained overlay. To prove the proposition we calculate N_s, the number of nodes that do not have to forward data in any

of the t distribution trees and show that N_s is negative for $m < t - 1$ for some N. We use the term drain for such nodes to distinguish them from freeriders that do not want to offer any cogs. The number of drains is maximal if there is at most one fertile node in each tree with less than t children. Using the notations introduced in Section II the number of nodes in layer i in a tree is $N_i = mt^{i-1}$. In a well maintained tree with $L \geq 2$ sterile nodes are located in layers $L-1$ and L. The number of drains is equal to the number of nodes minus the number of nodes that are fertile in any tree, i.e., the number of nodes in layers 1 to $L-1$ minus the number of nodes without children in layer $L-1$, and has to satisfy $N_s \geq 0$.

$$N_s = N - t \left\{ m \frac{t^{L-1} - 1}{t - 1} - \left[mt^{L-2} - \lceil \frac{N_L}{t} \rceil \right] \right\} \geq 0, \quad (5)$$

where N_L is the number of nodes in the last layer given as $N_L = N - m\frac{t^{L-1}-1}{t-1}$. Rewriting ineq. (5) we get that

$$m \geq t \lceil \frac{N_L}{t} \rceil - N_L. \quad (6)$$

If this condition is not satisfied then $N_s < 0$, and hence the overlay cannot be constructed. The condition can be satisfied independent of N_L by choosing $m \geq t - 1$ since $N_L \geq 1$. In this case the overlay is feasible for any N. ∎

The condition is not sufficient for feasibility since the overlay cannot always be kept well maintained in the presence of node departures and it cannot be ensured that there be at most one fertile node in each tree with less than t children. A consequence of Proposition 1 is that if the number of freeriders in the overlay is more than $m - t + 1$ then the overlay might become infeasible. Furthermore, if the number of freeriders in the overlay is more than m then the overlay is infeasible for any N.

B. Evolution of the available capacity

We define the available capacity in the overlay as the sum of the unused offered cogs of fertile nodes. For example, if there are f drains and no fertile nodes with available cogs then the available capacity is ft. To calculate the available capacity in the overlay we use induction. Initially, the available capacity in the overlay is mt, since the root node can support m nodes in each tree. Upon arrival of an arbitrary node the available capacity does not change, since the node consumes one available cog in each of the t trees and adds t available cogs in its fertile tree. Hence the available capacity remains mt. Similarly, a departure does not change the available capacity as long as the overlay remains feasible. Since the available capacity in the overlay is mt, the available capacity per tree is m on average.

Trees of the overlay can however become disconnected after the departure of a node. Upon departure of a node the available capacity decreases by $t - 1$ in the departing node's fertile tree and increases by one in its $t - 1$ sterile trees. The available capacity in the departing node's fertile tree can decrease below zero, in which case that tree becomes disconnected. In the following we show how the probability of disconnection depends on the parameters t and m of the overlay.

We consider the stationary state of the system, when the arrival and departure rates are equal. We assume that the interarrival times of nodes are exponentially distributed, this assumption is supported by several measurement studies [15], [16]. We approximate the distribution of the session holding times by an exponential distribution. The distribution of the session holding times was shown to fit the log-normal distribution [15], however, using the exponential distribution makes modeling easier and as we will see, the model gives a good match with the results of simulations where we use the log-normal distribution. For a given arrival intensity λ the mean number of nodes in the overlay is $\overline{N} = \lambda/\mu$, where $1/\mu$ is the mean session holding time.

To model the evolution of the available capacity we use a two-dimensional Markov process with state (ν, ι), corresponding to the number of nodes in the overlay and the free capacity in an arbitrary tree respectively. The state space of the process is $\{N_l \ldots N_u\} \times \{c_l \ldots c_u\}$. The parameters N_l and N_u are the lower and the upper bounds on the number of nodes in the overlay that the model considers. Similarly, c_l and c_u are the lower and the upper bounds on the free capacity that the model considers. We set $N_l = 0.9N$, $N_u = 1.1N$, $c_l = -(m-1)t$ and $c_u = mt$, so that the model is computationally feasible but the probability of $\nu \notin \{N_l \ldots N_u\}$ and $\iota \notin \{c_l \ldots c_u\}$ is negligible. The model is approximate, since the available capacity in an arbitrary tree is not independent of the available capacity in the other trees (since their sum is constant). A model that considers the evolution of all trees would be $t+1$ dimensional, and hence computationally not feasible. Another approximation is the use of a limited state space.

We denote the arrival rate of the nodes by λ and the mean session holding time by $1/\mu$. We denote by $q_{i,j}^{k,l}$ the transition intensity from state (i,j) to state (k,l) and $a(j)$ is the probability of that an arriving node is assigned to be fertile in the chosen tree given that the available capacity is j in that tree. The transition intensities are then given as ($N_l \leq i \leq N_u$ and $c_l \leq j \leq c_u$)

$$q_{i,j}^{max(i-1,N_l),min(j+1,c_u)} = (t-1)\mu/t$$
$$q_{i,j}^{max(i-1,N_l),max(j-t+1,c_l)} = \mu/t$$
$$q_{i,j}^{min(i+1,N_u),max(j-1,c_l)} = (1-a(j))\lambda$$
$$q_{i,j}^{min(i+1,N_u),min(j+t-1,c_u)} = a(j)\lambda$$
$$q_{i,j}^{i,j} = -\sum_{k\neq i,l\neq j} q_{i,j}^{k,l}.$$

The above intensities correspond to the departure of a sterile node, the departure of a fertile node, the arrival of a sterile node, and the arrival of a fertile node respectively. The last line corresponds to the diagonal of the transition intensity matrix. To calculate the steady state distribution of the Markov process we use the following distribution of $a(j)$

$$a(j) = \begin{cases} 0 & j > m \\ (1-F(\frac{j-m}{t/2-1}))^{t-1} & j \leq m, \end{cases} \quad (7)$$
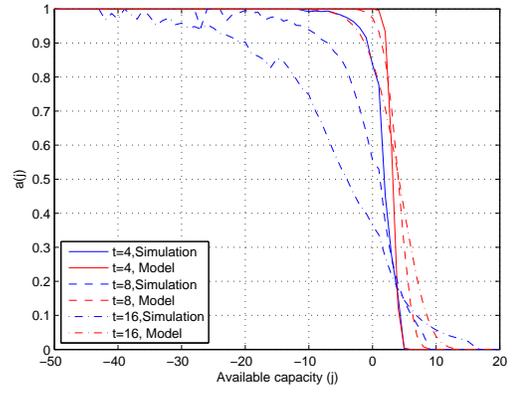
Fig. 8. $a(j)$ vs. available capacity for $t = 4, 8, 16$ and $m = t$.

where $F()$ is the standard normal distribution function. The rationale behind the distribution of $a(j)$ is the following. An arriving node is chosen to be fertile in the tree with the lowest available capacity among all t trees. This happens with probability 0 if $j > m$, since there has to be at least one tree with available capacity below m so that the total available capacity in the overlay can be mt. For $j \leq m$ we assume independence of the available capacity in the other trees and model their distribution by a normal random variable with mean m and standard deviation $t/2 - 1$. Hence, in this case the probability that the arriving node is assigned to be fertile is the probability that the available capacity in all other trees is higher than j.

Figure 8 shows $a(j)$ from eq. 7 and obtained via simulations for $t = 4, 8, 12$ and $m = t$. For $j < 0$ the model assumes $a(j)$ to be higher, while for j close to m to be lower than it is according to the simulations. The probability of $j < 0$ is however small, and hence eq. 7 is a pessimistic estimate. It will be subject of future work to derive a more precise distribution for $a(j)$. We can calculate the steady state distribution $\psi(i,j)$ of the Markov process using the transition intensity matrix $Q = (q_{i,j}^{k,l})$ [17]. The probability that the tree is disconnected is then $p_d = \sum_{N_l \leq i \leq N_u, j < 0} \psi(i,j)$.

C. Performance evaluation

For the evaluation we consider a mean session duration of $1/\mu = 306$ s as it was measured in [15]. We use exponential session length distribution in the model and log-normal distribution in the simulations. The reconnection interval is set to $T = 1$ s in the simulations unless otherwise stated. The first measure we consider is the blocking probability p_b^a, the probability that an arriving node can not join the overlay because it finds it in a state in which there are at least two disconnected trees. Due to the PASTA [17] property, this is the same as the probability that at least two trees are disconnected. Based on the model we can calculate p_b^a assuming independence of the trees as $p_b^a = 1 - (1-p_d)^t - tp_d(1-p_d)^{t-1}$. Figure 9 shows p_b^a as a function of the root multiplicity m for $t = 4$, $t = 8$ and $t = 16$ as obtained with the approximate mathematical model.

Figure 10 shows results for the same parameters obtained via simulations. Both figures show that for a given t increasing m decreases the probability of blocking. But increasing the value of t for a given m increases p_b^a. Figure 11 obtained using the model and showing p_b^a as a function of t for different m/t ratios leads to the same conclusion. For $m = t$ the blocking probability increases sharply as the number of trees increases. For $m = 2t$ the blocking probability remains very low however. The same results were obtained via simulations and are shown in Fig. 12. These results suggest that an overlay with high churn rate is only feasible for high values of m. For low values of m the trees get disconnected with a high probability. Comparing the results obtained with the model and the simulations shows that the approximate model describes the behavior of the overlay with good accuracy.
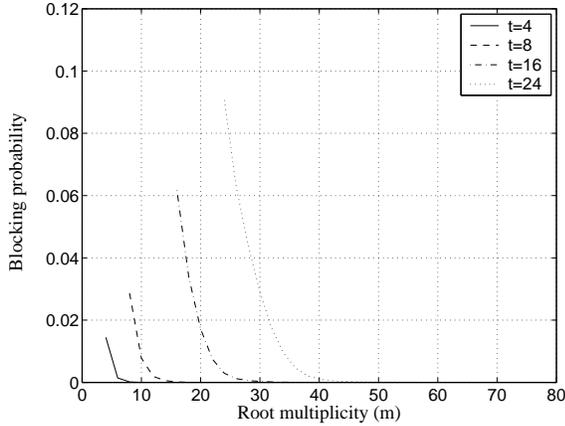
Fig. 9. Blocking probability (p_b^a) vs. root multiplicity for $t = 4$, $t = 8$, $t = 16$, $\lambda = 16.7/s$.

Fig. 10. Blocking probability (p_b^a) vs. root multiplicity for $t = 4$, $t = 8$, $t = 16$, $\lambda = 16.7/s$. Simulation results.

The next measure we study is the reconnection failure probability p_f^r, the ratio of the number of failed attempts to find a parent node and the total number of attempts to find a parent node. Figure 13 shows simulation results for the reconnection

Fig. 11. Blocking probability (p_b^a) vs. number of trees for $m = t$, $m = 1.5t$, $m = 2t$, $\lambda = 16.7/s$.

Fig. 12. Blocking probability (p_b^a) vs. number of trees for $m = t$, $m = 1.5t$, $m = 2t$, $\lambda = 16.7/s$. Simulation results.

failure probability as a function of the root multiplicity for $t = 4$, $t = 8$, $t = 16$. The figure shows that the reconnection failure probability decreases as m increases similar to the blocking probability. Figure 14 shows simulations results for the reconnection failure probability as a function of the number of trees for $m = t$, $m = 1.5t$, $m = 2t$. The conclusions are similar to those regarding the blocking probability; the reconnection failure probability slightly decreases however for large values of t.

Figure 15 shows p_f^r as a function of $\overline{N}$ for $t = 4$, $t = 8$ and $t = 16$ and $m = t$. The figure shows that the failure probability slowly decreases as the number of nodes in the overlay increases, hence a large overlay is more resilient to node departures than a small one. Figure 16 shows the reconnection failure probability as a function of T, the reconnection interval. The figure shows that increasing the reconnection interval decreases the failure probability. The reason for this phenomenon is that the longer a node waits the higher the probability that a fertile node arrives to the disconnected tree by the time it tries to reconnect. But the decreased failure probability comes at
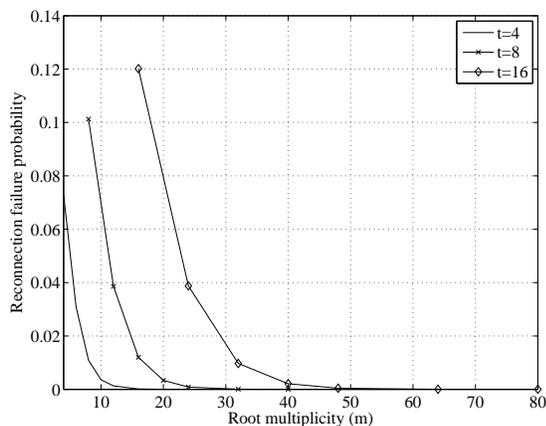
Fig. 13. Reconnection failure probability vs. root multiplicity for $t = 4$, $t = 8$, $t = 16$ and $\lambda = 16.7/s$. Simulation results.
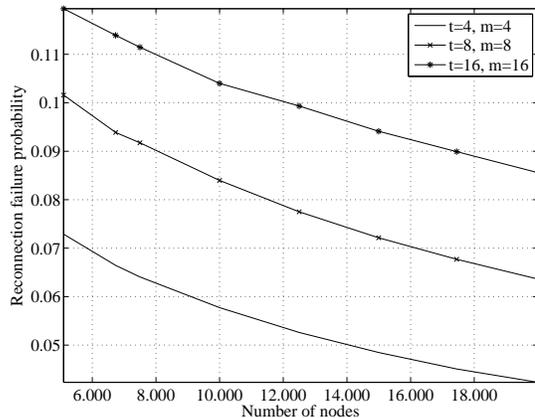
Fig. 15. Reconnection failure probability vs. mean number of nodes for $t = 4$, $t = 8$, $t = 16$ and $m = t$. Simulation results.
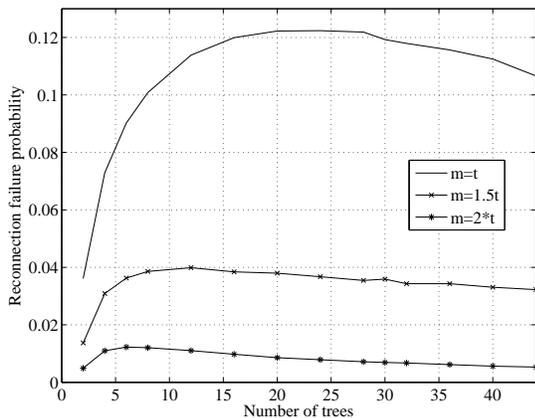
Fig. 14. Reconnection failure probability vs. number of trees for $m = t$, $m = 1.5t$, $m = 2t$ and $\lambda = 16.7/s$. Simulation results.
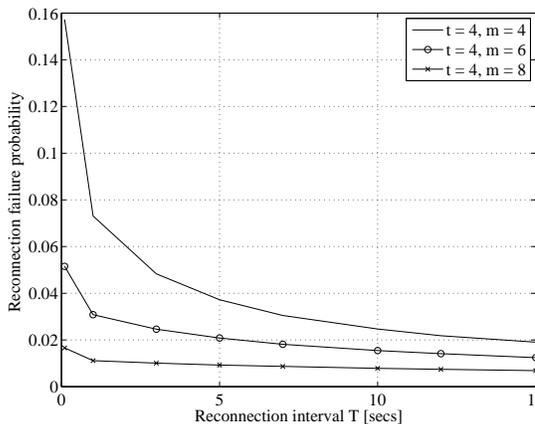
Fig. 16. Reconnection failure probability vs. reconnection interval for $\lambda = 16.7/s$. Simulation results.

the price that nodes have to wait longer between reconnection attempts and hence loose more packets. This result suggests that there is an optimal value of T for given t, m and FEC parameters, which has to be set dynamically to achieve best performance.

V. CONCLUSION

In this paper we presented an analytical model of a resilient end-point-based overlay for multicast streaming based on multiple minimum-depth distribution trees that employs path diversity and forward error correction for improved resilience to node churns and packet losses. We showed that the probability that an arbitrary node in the overlay possesses a packet is high while the loss probability in the network is lower than a certain threshold. After reaching the threshold this probability suddenly drops. The decrease rate gets faster as the number of nodes in the overlay grows. The value of the threshold depends on the ratio of redundancy and on the number of the distribution trees. We derived a necessary condition for the feasibility of the overlay and developed an approximate

model to study the stability of the overlay in the presence of node dynamics. Simulations show that the model is rather accurate and helps to understand the effect of the parameters of the overlay on its stability. It is subject of our future work to improve the approximate model by considering the correlations between the available capacity in different trees and by finding a more precise distribution for $a(j)$. Another open issue is how ungraceful departures (nodes that depart without sending a notification about the departure) can be incorporated in the model. The results presented here serve to improve our understanding of the behavior of the overlay in a dynamic environment and help us to design an overlay with improved stability properties.

REFERENCES

[1] V. Padmanabhan, H. Wang, and P. Chou, "Resilient peer-to-peer streaming," in *Proc. of IEEE ICNP*, pp. 16–27, 2003.
[2] K. Sripanidkulchai, A. Ganjam, B. Maggs, and H. Zhang, "The feasibility of supporting large-scale live streaming applications with dynamic application end-points," in *Proc. of ACM SIGCOMM*, pp. 107–120, 2004.

[3] Y. Chu, S. Rao, S. Seshan, and H. Zhang, "A case for end system multicast," *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, 2002.

[4] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast," in *Proc. of ACM SIGCOMM*, 2002.

[5] S. Banerjee, S. Lee, R. Braud, B. Bhattacharjee, and A. Srinivasan, "Scalable resilient media streaming," in *Proc. of NOSSDAV*, 2004.

[6] M. Castro, P. Druschel, A. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "SplitStream: High-bandwidth multicast in a cooperative environment," in *Proc. of ACM SOSP*, 2003.

[7] B. Lamparter, A. Albanese, M. Kalfane, and M. Luby, "PET - priority encoding transmission: A new, robust and efficient video broadcast technology," in *Proc. of ACM Multimedia*, 1995.

[8] K. Kawahara, K. Kumazoe, T. Takine, and Y. Oie, "Forward error correction in ATM networks: An analysis of cell loss distribution in a block," in *Proc. of IEEE INFOCOM*, pp. 1150–1159, June 1994.

[9] M. Guo and M. Ammar, "Scalable live video streaming to cooperative clients using time shifting and video patching," in *Proc. of IEEE INFOCOM*, 2004.

[10] E. Setton, J. Noh, and B. Girod, "Rate-distortion optimized video peer-to-peer multicast streaming," in *Proc. of ACM APPMS*, pp. 39–48, 2005.

[11] X. Zhang, J. Liu, B. Li, and T. Yum, "Coolstreaming/donet: A data-driven overlay network for peer-to-peer live media streaming," in *Proc. of IEEE INFOCOM*, 2005.

[12] G. Dán, V. Fodor, and G. Karlsson, "On the asymptotic behavior of end-point-based multimedia streaming," in *Proc. of International Zürich Seminar on Communication*, 2006.

[13] G. Dán, V. Fodor, and G. Karlsson, "On the stability of end-point-based multimedia streaming," in *Proc. of IFIP Networking 2006*, May 2006.

[14] I. Reed and G. Solomon, "Polynomial codes over certain finite fields," *SIAM J. Appl. Math.*, vol. 8, no. 2, pp. 300–304, 1960.

[15] E. Veloso, V. Almeida, W. Meira, A. Bestavros, and S. Jin, "A hierarchical characterization of a live streaming media workload," in *Proc. of ACM IMC*, pp. 117–130, 2002.

[16] K. Sripanidkulchai, B. Maggs, and H. Zhang, "An analysis of live streaming workloads on the Internet," in *Proc. of ACM IMC*, pp. 41–54, 2004.

[17] L. Kleinrock, *Queueing Systems*, vol. I. Wiley, New York, 1975.