On the asymptotic behavior of end-point-based multicast streaming

György Dán, Viktória Fodor and Gunnar Karlsson
KTH, Royal Institute of Technology
Department of Signals, Sensors and Systems
10044 Stockholm
Email: {gyuri,vfodor,gk}@s3.kth.se

Abstract—**In this paper we propose an analytical model of a resilient, tree-based end-node multicast streaming architecture that employs path diversity and forward error correction for improved resilience to node churns and packet losses. We show that this architecture can distribute data to nodes arbitrarily far away from the root of the trees as long as the loss probability is lower than a certain threshold, but the probability of packet reception suddenly drops to zero once this threshold is exceeded. The value of the threshold depends on the ratio of redundancy and on the number of the distribution trees.**

I. INTRODUCTION

The delivery of streaming media over end-point overlays has received much attention recently ([1], [2] and references therein). Although current commercial content delivery networks are capable of supporting many simultaneous streams, end-node-based multicast could considerably decrease the cost of large scale streaming, while being resilient to sudden surges in the client population, such as flash crowds. 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. The main advantages are that such a system is easy to deploy and it reduces 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. They were based on a mesh [3], [4] or a single distribution tree [5]. 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 [6] by distributing

packets to randomly chosen neighbors outside of the distribution tree. Though this scheme provides some resilience to losses, it is known that repeating information is less efficient than using error correcting codes. SplitStream [7] and CoopNet [1] introduce multiple distribution trees and employ priority encoding transmission (PET) [8] based on forward error correction (FEC) [9] to decrease the effects of node failures and to recover from packet losses. Simulations were used to show the resilience of these schemes under various scenarios showing that increasing the number of trees improves the resilience both to packet losses and node churns.

Feasibility issues of small overlays with less than 100 nodes were discussed in detail in [10] based on experimental broadcasts over the Internet, and showed promising results. The experiments showed that poor performance was due in a large extent to packet losses. The feasibility of larger deployments was studied for a CoopNet like end-node overlay 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.

Albeit there is an extensive literature on end-point-based multicast streaming, previous work on the behavior of these systems was limited to simulations. In this paper we present a simple model for a CoopNet like overlay, and evaluate the performance of such a system for a large number of nodes. Our results show that an arbitrarily high packet reception probability can be achieved independent of the number of nodes in the overlay by adding enough redundancy. The packet reception probability goes however to zero if there is not enough redundancy added. The transition between the stable and non-stable states of the system is ungraceful, which can raise problems in a dynamic environment.

The paper is organized as follows. In Section II we briefly describe the architecture of the considered end-point-based application overlay for multicast. In Section III we present the mathematical model and the main results. In Section IV we discuss the performance of the system based on the analytical model and simulations. In Section V we conclude our work.

II. SYSTEM DESCRIPTION

We consider an application overlay as the one described in [1], [2] consisting of a root node and N peer nodes. Peer nodes are organized in t distribution trees, either by a distributed protocol or 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 and a different child node to which they forward data. Child nodes of the root node can have the same parent (i.e. the root) in more than one tree. Upon construction of the distribution trees each node is at the same distance from the root node in all trees, and we will refer to nodes at distance i nodes from the root as members of layer i. In the presence of group dynamics it is the task of the tree building algorithm to ensure that all parent nodes of a node are in the same or almost the same layer. We denote the number of children of the root node in each tree by m, and we call it the multiplicity of the root node. The number of layers in the distribution tree is N/m. Typically the number of distribution trees is no more than the multiplicity of the root node $m \geq t$; we will consider this case in the analysis. We assume that nodes do not contribute more bandwidth towards their children as they use to download from their parents, so that the multiplicity of the peer nodes is one, i.e. they have one child in each distribution tree (See Fig. 1).
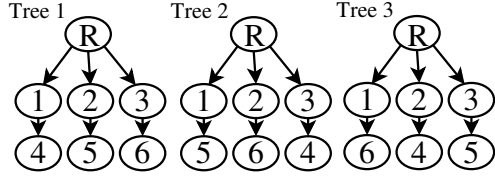

Fig. 1. Multicast tree structure for $t = 3$, $m = 3$ and $N = 6$.

The root uses block based FEC, e.g. Reed-Solomon codes, 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 $n = k + c$. We denote this FEC scheme by FEC(n,k). FEC can be used to implement MDC as described in [1] or PET (UXP) as described in [8] if layered coding is supported by the coder. In case of losses the lost packets can be reconstructed as long as no more than c packets are lost out of n packets. The root 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, and once they received at least k packets of a block of n packets they recover the remaining c packets and send them to the child nodes in the corresponding distribution trees. A packet received from the parent node after it has been decoded based on other packets in the block will be discarded.

III. MATHEMATICAL MODEL

Our goal is to calculate the probability $\pi(i)$ that a node in layer i of the distribution tree receives or can reconstruct an arbitrary packet. We denote the probability that a packet is lost on the path between two adjacent peer nodes by p ($0 < p < 1$), and assume that the probability of losses is independent. We assume that the probability that a node is in possession of a packet is independent of the probability that another node in the same layer possesses a packet from the same block of packets. We will comment on the validity and possible effects of these assumptions later.

In the following we give a nonlinear recurrence equation to calculate the evolution of $\pi(i)$. As the root node is assumed to possess all packets, we have the initial condition

$$\pi(0) = 1. \tag{1}$$

Let us define the probability $\pi_a(i)$ as the probability that a node in level i receives an arbitrary packet from its parent node. With the above assumptions this probability can be expressed as

$$\pi_a(i) = (1 - p)\pi(i - 1), \quad i > 0 \tag{2}$$

A node in layer i will possess a packet if it receives it, or if it can reconstruct it using FEC, which is reflected in the following equation for $i \geq 1$

$$\pi(i) = \pi_a(i) + (1 - \pi_a(i)) \tag{3}$$
$$\sum_{j=k}^{n-1} \binom{n-1}{j} \pi_a(i)^j (1 - \pi_a(i))^{n-1-j}.$$

By substituting (2) into (3) we get the nonlinear recurrence equation

$$\pi(i+1) = \pi(i)(1-p) + \sum_{j=k}^{n-1} \binom{n-1}{j} \tag{4}$$
$$(\pi(i)(1-p))^j (1 - \pi(i)(1-p))^{n-j},$$

which describes the evolution of $\pi(i)$. If (4) has a fixed point, then in that fixed point $\pi(i+1) = \pi(i)$.

We can rewrite (4) by subtracting $\pi(i)$ from both sides and omitting the indices to

$$f(\pi) = -\pi p + \tag{5}$$
$$\sum_{j=k}^{n-1} \binom{n-1}{j} (\pi(1-p))^j (1 - \pi(1-p))^{n-j}.$$

Since $\pi(i+1) - \pi(i) = f(\pi(i))$ we have that for any layer i if $f(\pi(i)) < 0$ then $\pi(i+1) < \pi(i)$, if $f(\pi(i)) > 0$ then $\pi(i+1) > \pi(i)$ and if $f(\pi(i)) = 0$ then $\pi(i+1) = \pi(i)$ and $\pi(i)$ is a fixed point. Starting with $\pi(0) = 1$ as in eq. (1) the value of $\pi(i)$ will decrease as long as $f(\pi(i)) < 0$. The roots of $f(\pi)$ correspond to the fixed points of (4), and if $f(\pi)$ has a real root r in the interval $(0,1)$ and the derivative $f^{(1)}(r) = \frac{d}{d\pi}f(\pi)|_{\pi=r} < 0$ then $\pi(\infty) = \lim_{i \to \infty} \pi(i) = r$. A root r with $f^{(1)}(r) > 0$ is not stable on the other hand. We will call the system stable if $\pi(\infty) > 0$ and unstable otherwise. To check the existence and the number of real roots of $f(\pi)$ in $(0,1)$ we investigate the signs of $f(\pi)$ at the endpoints of the interval (see Fig. 2 for some examples). For any $p > 0$ the ratio of successfully received or recovered packets has to be less than 1, so that $f(1) < 0$. Since $\pi = 0$ is
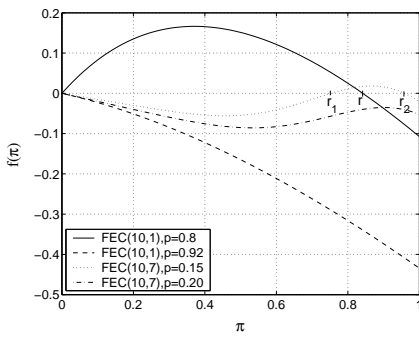
Fig. 2. $f(\pi)$ vs. π for different ratios of redundancy and loss probabilities. At the root closest to 1 (if it exists) the derivative is negative and hence the corresponding fixed point is stable.

a zero of $f(\pi)$ we have to calculate $f^{(1)}(0)$ to see the sign of $f(0+) = \lim_{\pi \to 0+} f(\pi)$.

If $k = 1$ then we get that $f^{(1)}(0) = n(1-p) - k$, and thus if $p < (n-k)/n$ then $f^{(1)}(0) > 0$ and consequently $f(0+) > 0$. Hence there has to be at least one root r in $(0,1)$ for which $f^{(1)}(r) < 0$ resulting in a asymptotically stable fixed point (the solid line in Fig. 2). It follows that for any loss probability p there is a ratio of redundancy c/k above which $\lim_{i \to \infty} \pi(i) > 0$. Otherwise, if $p \geq (n-k)/n$, then the number of real roots in $(0,1)$ is either zero or an even number, and using Sturm's theorem [11] we find that the number of roots in $(0,1)$ is 0 for any p (the dashed line in Fig. 2).

If $k > 1$ then we have $f^{(1)}(0) = -p$, which is always negative, and thus the number of real roots in $(0,1)$ is either zero or an even number. By using Sturm's theorem we find for any p that the number of real roots in $(0,1)$ is no more than two (counting their multiplicity). If they exist, denoted by r_1 and r_2 ($r_1 \leq r_2$), then $f^{(1)}(r_2) < 0$, and $\pi(\infty) = r_2$ (the dotted line in Fig. 2). Since $f(\pi) > 0$ for $r_1 < \pi < r_2$, the above result holds for any $r_1 < \pi(0) \leq 1$ as initial condition. Similarly, even if $r_1 < \pi(i) < r_2$ for some i, we have $\pi(\infty) = r_2$. With other words, the system can recover from disturbances, as long as $\pi(i) > r_1$. Without a proof we state that $r_1 < r_2$ for $0 < p < p_{max}$, where for p_{max} $r_1 = r_2$. However, $f(\pi)$ has no roots in $(0,1)$ for $p > p_{max}$ (the dash-dotted line in Fig. 2).

In the following we discuss the validity and effects of certain assumptions made in the model. Losses in the Internet are known to be correlated [12], which might influence the results obtained with the model. In particular, if losses occur in bursts at the output links of the nodes, the burstiness influences the results if packets from the same block are distributed over the same distribution tree, i.e. $t < n$, but does not influence them otherwise. Bursty losses that occur in the backbone will influence the results if packets from different distribution trees traverse the same bottleneck link in the backbone. Worse performance has to be expected however if bursty losses occur on the input links of the nodes, in which case the potential of FEC to recover from losses decreases. End-nodes with asymmetric links (e.g. ADSL) will have bottlenecks at the outgoing links, hence correlated losses are more likely to occur

there. The assumptions $n \leq t$ and $m \geq t$ are made to ensure independence of the losses of packets in the same block and to ensure that each node has different parents in all of the trees respectively. Removing these assumptions will make losses more correlated, and hence worsen the performance of the distribution tree.

The model does not take into account node departures, which have to be dealt with in the considered scenario. We argue however that node departures can be incorporated in the model as an increase of the packet loss probability as $p_d = N_d/N * \overline{T}$, where N_d is the mean number of nodes departing per time unit and $\overline{T}$ is the mean of the time nodes need to recover from the departure of a parent node. The rationale for this hypothesis is that node departures can be treated as bursty losses on the output link of the departing node, and can be modeled as independent if $n \leq t$ and $m \geq t$. An evaluation of this hypothesis is subject of future work.

IV. PERFORMANCE EVALUATION

In this section we show results obtained with the model presented in the previous section and simulations. In all scenarios we set $t = n$ and $m = 50$ for easy comparison. For the simulations we considered the streaming of a 128 kbps stream to 50000 nodes organized in 1000 layers, which results in $m = 50$. The root node sends at a bitrate of 6400 kbps and could be hosted on a 10 Mbps link, while the peer nodes have 128 kbps connections both uplink and downlink. Nodes choose their parent nodes at random, so that with a small probability nodes can have the same parents in different trees. In each run of the simulation the root node sends $10000 \times m$ packets,

Figure 3 shows $\pi(1000)$ as a function of p for $k = 10$ and $k = 20$ and different values of c. The figure shows that for every (n,k) pair there is a loss probability p_{max} above which the reception probability in nodes far from the root node suddenly becomes 0. Below p_{max} the reception probability is close to 1 and is slowly decreasing. This stepwise, ungraceful decrease of the reception probability is an undesired feature for systems working in a dynamic environment such as the Internet. The figure shows that increasing the number of trees, i.e. the FEC block length, slightly improves the resilience of the distribution tree to losses, which is in accordance with [1], [9].
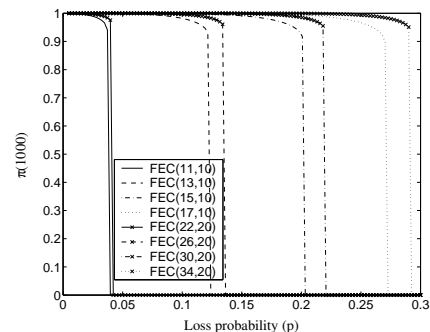

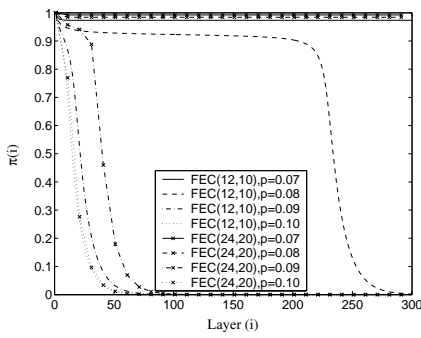Fig. 3. $\pi(1000)$ vs p for different ratios of redundancy.

Fig. 4. $\pi(i)$ vs. i for different ratios of redundancy and loss probabilities.
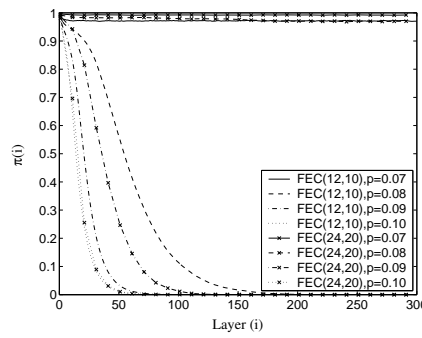

Fig. 5. $\pi(i)$ vs. i for different ratios of redundancy and loss probabilities, simulation results.
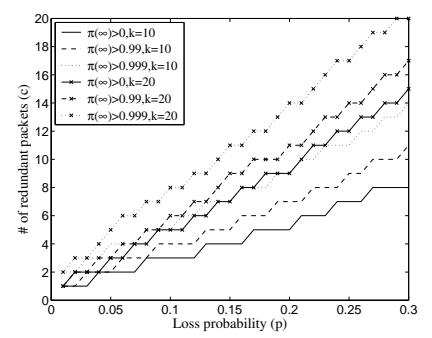

Fig. 6. Number of redundant packets vs. p for different performance objectives.

Figure 4 shows $\pi(i)$ as a function of i for different block lengths n and loss probabilities and a ratio of redundancy of $c/k = 0.2$. We see that $\pi(i)$ is close to one in the cases when $p < p_{max}$, while it becomes almost 0 after some i otherwise. The value of i at which $\pi(i)$ breaks down depends on how far p is from p_{max}. For $k = 10$ $p_{max} = 0.0799$ and for $k = 20$ $p_{max} = 0.0885$. The positive effects of the increased block length can be seen by comparing results at $p = 0.08$, where for $k = 20$ the system is stable, whereas for $k = 10$ it is unstable. Figure 5 shows simulation results for the same scenarios as Fig. 4. The comparison shows perfect match for $p = 0.07$ and $p = 0.10$, while the simulation results show worse behavior than the analytical ones when p is close to p_{max}. The difference is rather big for $k = 10$ and $p = 0.08$, when $p - p_{max} = 10^{-4}$, in which case deviations from the mean loss probability in the individual layers make the deterioration faster than that predicted by the model. For higher values of m the simulation gives more accurate results as the probability of deviation is lower due to the central limit theorem.

Figure 6 shows c, the number of redundant packets needed to ensure $\pi(\infty) > 0$, $\pi(\infty) > 0.99$ and $\pi(\infty) > 0.999$ for $k = 10$ and $k = 20$. The figure shows a closely linear relationship between the number of redundant packets needed and the loss probability. For low values of p the number of redundant packets needed to ensure $\pi(\infty) > 0.999$ is close to the number of redundant packets needed for $\pi(\infty) > 0$, and hence in a dynamic environment the ratio of redundancy has to be set higher to prevent a severe decrease of $\pi(i)$ due to a sudden increase of the loss probability. A similar problem can occur towards the end of a broadcast, when departures cause the number of nodes in the overlay to decrease. The increased loss probability due to node churn might exceed the level of stable operation and can lead to $\pi(\infty) = 0$. The root node can prevent this from happening by increasing the ratio of redundancy c/k towards the end of the broadcast.

V. CONCLUSION

In this paper we presented a mathematical model for the analysis of an end-point overlay for multicast based on multiple distribution trees. We showed that for any loss probability there is a ratio of redundancy which ensures that even nodes far away from the root of the trees receive a non-zero ratio of the information. We showed that this multicast scheme shows a non-graceful performance degradation once the loss probability exceeds a certain threshold. The threshold depends on the number of distribution trees and the ratio of redundancy used. Albeit the model does not consider certain characteristics of end-point overlays, it helps to analyze the stability criteria of these systems. A model including correlated losses, inhomogeneous packet loss probabilities, the case $n > t$ and dynamic user population can be built based on e.g. the Gilbert model [13], and is subject of future work.

REFERENCES

[1] V. N. Padmanabhan, H.J. Wang, and P.A Chou, "Resilient peer-to-peer streaming," in *Proc. of IEEE ICNP*, 2003, pp. 16–27.
[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*, 2004, pp. 107–120.
[3] Y. Chu, S.G. Rao, S. Seshan, and H. Zhang, "A case for end system multicast," *IEEE J. Select. Areas Commun.*, vol. 20, no. 8, 2002.
[4] Y. Chawathe, "Scattercast: an adaptable broadcast distribution framework," *Multimedia Systems*, vol. 9, no. 1, pp. 104–118, 2003.
[5] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast," in *Proc. of ACM SIGCOMM*, 2002.
[6] S. Banerjee, S. Lee, R. Braud, B. Bhattacharjee, and A. Srinivasan, "Scalable resilient media streaming," in *Proc. of NOSSDAV*, 2004.
[7] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "Splitstream: High bandwidth content distribution in a cooperative environment," in *Proc. of IPTPS*, 2003.
[8] B. Lamparter, A. Albanese, M. Kalfane, and Luby M., "PET - priority encoding transmission: A new, robust and efficient video broadcast technology," in *Proc. of ACM Multimedia*, 1995.
[9] 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*, June 1994, pp. 1150–1159.
[10] Y. Chu, A. Ganjam, T.S.E Ng, S.G. Rao, K. Sripanidkulchai, J. Zhan, and H. Zhang, "Early experience with an internet broadcast system based on overlay multicast," in *Proc. of USENIX*, 2004.
[11] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in real algebraic geometry*, Springer Verlag, 2003.
[12] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Meaurement and modelling of the temporal dependence in packet loss," in *IEEE INFOCOM*, 1999, pp. 345–352.
[13] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 69, pp. 1253–1265, September 1960.