

# On the Efficiency of Path Diversity for Continuous Media Applications

Bruno Ribeiro<sup>1</sup> Edmundo de Souza e Silva<sup>2</sup> Don Towsley<sup>1</sup> \*

<sup>1</sup>Dept. of Computer Science  
University of Massachusetts  
Amherst, MA 01003  
{ribeiro, towsley}@cs.umass.edu

<sup>2</sup>COPPE/PESC and CS Depart./IM  
Federal Univ. of Rio de Janeiro  
06902 Sophia Antipolis, France  
edmundo@land.ufrj.br

**Computer Science Technical Report 05-19**

April 7, 2005

## Abstract

Streaming media over best-effort networks is still an area rich in challenging problems, and many issues remain to be resolved in order to allow the deployment of large scale continuous media content delivery over the Internet. Recent works have addressed methods to cope with the poor quality of service of the Internet due to its best effort service. One promising technique is the use of distinct paths connecting the source and the receiver to achieve better end-to-end loss behavior. One of the main advantages of this technique (called path diversity) resides in the fact that it is neither based on adding redundancy to the original flow nor introducing changes in the current Internet protocols.

Using analysis we address four main issues that arise with path diversity. First, we study the performance gains and properties as the number of paths is increased beyond two, and when paths are heterogeneous. Second, we compare path diversity against packet interleaving, which is another technique to reduce the impact of packed losses. We also obtain results when different paths share a common segment. Finally, we investigate the performance of the technique when it is used by many applications that collectively consume a large fraction of the bandwidth available in the paths.

*Keywords:* Path diversity, Multi-path streaming, packet interleaving, packet loss, overlay networks

---

\*This work was developed while Bruno Ribeiro was with the Federal Univ. of Rio de Janeiro, COPPE/PESC and CS Depart./IM. This research has been supported in part by the NSF, NFS/CNPq, CAPES (Brasil), CNPq (Brasil) and FAPERJ (Brazil) grants. Bruno Ribeiro had fellowships from CAPES and CNPq. E. de Souza e Silva was supported in part by grants from CNPq, and FAPERJ. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

# 1 Introduction

Path diversity has been proposed as a technique to improve the robustness of applications running over best effort networks. The basic idea of the technique is simple: application packets are sent over two or more distinct paths towards the destination. It was initially proposed as a routing scheme to improve transmission reliability, to smooth traffic, and to handle link failures [16]. It can further increase robustness to streaming applications by reducing the lengths of packet loss bursts [4]. Since the lengths of packet loss bursts have an adverse effect on the overall quality of voice and video applications, reducing them can significantly improve the transmission quality as perceived by the user.

Although path diversity has been shown to significantly benefit applications, there is not yet a good understanding of all of its potential and of its limitations, and how it compares to other schemes that do not rely on multiple paths. For example, the performance of two-path diversity is well understood when the paths are homogeneous [9]. Although in the case of  $N = 3$  homogeneous paths is analyzed in [9], there is no deep analysis in the literature for the impact of increasing the number of paths on the value of path diversity. And when the paths do not exhibit identical loss behavior? One advantage of path diversity is that it reduces loss burst lengths; how does it compare then to a scheme that interleaves packets over a single path for the purpose of reducing loss burstiness? Last, the burstiness reduction benefits of path diversity are greatest when the loss processes on different paths are statistically independent. Hence there is the question of how loss correlation, perhaps due to multiple paths sharing a bottleneck link, affects the benefits of path diversity.

The purpose of this paper is to develop a greater understanding of path diversity by addressing the questions presented above. We develop simple analytical models that describe the loss behavior of path diversity in the presence of bursty loss processes on the different paths. Unlike previous work, we account for the presence of  $N > 3$  non identical paths. These models are then used to identify the benefits of path diversity in the presence of two or more possibly heterogeneous paths. They are also used to compare path diversity to the more traditional approach of packet interleaving for reducing loss burstiness. (For the comparisons, new models for interleaving were also developed.) These models assume that paths exhibit independent loss behavior. We extend the models to account for correlation. The resulting analysis are important contributions of the paper. In addition, from these models we draw the following conclusions:

- When paths exhibit similar loss characteristics, increasing the number of paths beyond  $k$  may increase the probability that a burst of length  $< k$  occurs. It is advantageous to increase the number of paths up to a certain (application-specific) threshold value, but not beyond in order to decrease the probability of a long burst.
- Path diversity and interleaving exhibit similar characteristics. However, for channels with similar loss characteristics, even in the presence of an idealized interleaving scheme, we find that path diversity may provide better performance in the form of a less bursty loss process.
- Path diversity is still beneficial when paths share a common bottleneck. The gains obtained depend heavily on the loss processes of the shared segment and those for the independent segments in the paths.

We also performed some experiments over the Internet in order to measure the statistics of path diversity in a real world environment, with highly heterogeneous paths that also share path segments. We find that the

performance of path diversity is superior when compared to single path. When compared to interleaving its performance depends on the parameters involved.

The remainder of the paper is organized as follows. First in section 2 we develop models to calculate the pmf of the loss burst length for  $N$ -path diversity, both when homogeneous and heterogeneous paths are used by the application. We also study the performance of path diversity when combined with a FEC scheme. In section 3 we show that path diversity and block interleaving share similar properties and develop a new model for block interleaving in order to compare both methods. Our models are extended in section 4 to addresses the performance of path diversity when paths share a common segment. The section also reports results from Internet measurements we perform. In section 5 we study several issues that arise when a large number of applications adopt the technique. For that we developed a simulation model that includes TCP traffic sharing the links with the path diversity traffic. Section 6 comment on related work and our conclusions are presented in section 7.

## 2 Path Diversity with more than two (either homogeneous or heterogeneous) paths

The studies in the literature only consider path diversity over two paths. We refer to the technique when  $N$  paths are used as  $N$ -path diversity, and the objective of this section is to evaluate whether performance improves with increasing number of paths that is, when  $N > 2$ .

The usual model that has been employed in the literature assumes that the end-to-end loss process for a path can be described by a continuous time Gilbert model [4, 9]. Let  $\mathcal{X}_r = \{X_r(t) : t \geq 0\}$  be the loss process of the  $r$ -th path. The state  $X_r(t)$  at time  $t$  assumes one of two values:  $b$  or  $g$ , ( $g$ =good and  $b$ =bad). We assume that the source transmits packets at fixed length intervals (with length  $\tau$ ) and in a round-robin fashion. Therefore, we observe the state of the paths at times  $\tau, 2\tau$  etc., and assume  $\tau = 1$  time unit to simplify the notation. We assume that  $\mathcal{X}_r$  is stationary and independent of  $\mathcal{X}_u$  for all pairs  $r \neq u$ . If the  $n$ -th packet transmitted by the source goes through path  $r$  and  $X_r(n) = b$ , then the packet is lost. Let  $\mu_b^{(r)}$  and  $\mu_g^{(r)}$  the transition rates from the *bad* and *good* states, respectively. Let  $\pi_g^{(r)}$  and  $\pi_b^{(r)}$  be the steady state probabilities that  $\mathcal{X}_r$  is in the “good” or “bad” states, respectively, at times  $\tau, 2\tau, \dots$ , and let  $p_{i,j}^{(r)}(l)$  be the probability of transitioning from state  $i$  to  $j$  for  $\mathcal{X}_r$  in  $l$  time units, i.e.  $p_{i,j}^{(r)}(l) = P[X_r(n+l) = j | X_r(n) = i] = P[X_r(l) = j | X_r(0) = i]$ .

Let  $\mathcal{A}$  be a set of  $N$  paths  $1, \dots, N$ . Let  $\Gamma_N = \{\mathcal{X}_1, \mathcal{X}_2 \dots \mathcal{X}_N\}$  be the set of loss processes associated with these paths and let  $L(\Gamma_N)$  be the random variable that denotes the length of a loss burst using this set of paths. (Note that  $L(\Gamma_N)$  can assume values  $1, 2, \dots$ ) In what follows we obtain  $P[L(\Gamma_N) = m]$  and use it to compare the efficacy of path diversity using two paths with that using three or more paths.

Throughout this sub-section, all path loss processes  $\mathcal{X}_r$  are assumed to be mutually independent and in steady state. Let  $\Theta(r, \Gamma_N)$  be the fraction of loss bursts in which the first packet lost was sent via path  $r$ . Without loss of generality we assume that the application selects paths in a round robin manner so, if a packet is sent via path  $r$ , the next packet in sequence is sent via path  $(r+1)|_N$  where “ $|$ ” is the modulo operator defined as follows:  $(r)|_N = \zeta + 1$ , and  $\zeta$  is the remainder of  $(r)/N$ .

We begin by deriving  $p_{i,j}^{(r)}(l)$ . (In equation (1) below we omit the superscript  $r$  to simplify notation.) It is well known that, for a two state Markov model:

$$\begin{bmatrix} p_{g,g}(l) & p_{g,b}(l) \\ p_{b,g}(l) & p_{b,b}(l) \end{bmatrix} = \begin{bmatrix} \pi_g & \pi_b \\ \pi_g & \pi_b \end{bmatrix} + \begin{bmatrix} \pi_b & -\pi_b \\ -\pi_g & \pi_g \end{bmatrix} e^{-l(\mu_g + \mu_b)} \quad (1)$$

where  $l$  is the time variable. The following corollary calculates the probability that a loss burst has length  $m$  when multiple paths are used. It extends the results in [9] for two homogeneous paths to also more than two and even heterogeneous paths.

**Lemma 1**

$$\Theta(r, \Gamma_N) = \frac{\pi_g^{(r-1)} \pi_b^{((r)|_N)}}{\sum_i^N \pi_g^{(i-1)} \pi_b^{((i)|_N)}} \quad (2)$$

*Proof:* Construct an overall (discrete time) loss model that represents the loss process that the application sees. This model includes the variable that indicates the path currently in use and  $N$  variables, each representing the state of the corresponding path (good or bad). (Therefore we have  $N2^N$  states.) Let  $B_r$  be the event that a burst starts in path  $r$ . This occurs if: (a) the loss process of path  $(r-1)|_N$  is in a “good” state when a packet is sent via path  $(r-1)|_N$  and; (b) the loss process on path  $r$  transits to a “bad” state while it is in use, independent of the other paths. In steady state, and since all loss processes are independent processes, after summing over all possibilities, we obtain:  $P[B_r] = 1/N \pi_g^{(r-1)|_N} \pi_b^{(r)}$ . Equation (2) is obtained after normalizing by  $\sum_{r=1}^N P[B_r]$ . ■

Using the above lemma we can calculate  $P[L(\Gamma_N) = m]$ , i.e. the probability that the length of a burst is  $m$  when we have  $N$  independent and possibly heterogeneous paths.

**Corollary 1** For  $N \geq 2$ ,

$$P[L(\Gamma_N) = m] = \sum_{r=1}^N \Theta(r, \Gamma_N) \times \begin{cases} \left[ \prod_{i=1}^{m-1} \pi_b^{((r+i)|_N)} \right] \pi_g^{((r+m)|_N)} & \text{for } m < N-1, \\ \left[ \prod_{i=1}^{N-2} \pi_b^{((r+i)|_N)} \right] p_{g,g}^{((r+N-1)|_N)}(N) & \text{for } m = N-1, \\ \sigma(r, \Gamma_N) & \text{for } m \geq N \end{cases} \quad (3)$$

$$\begin{aligned} \text{where, } \sigma(r, \Gamma_N) &= \left[ \prod_{i=1}^{N-2} \pi_b^{((r+i)|_N)} \right] p_{g,b}^{((r+N-1)|_N)}(N) \\ &\times \left[ \prod_{i=1}^{(m-N)} p_{b,b}^{(r-1+i)|_N}(N) \right] p_{b,g}^{((r+m)|_N)}(N) \end{aligned}$$

*Proof:* We condition on a burst starting on path  $r$ .

For  $m < N-1$ . Given that a loss occurs, each of the following  $m-1$  packets will each be lost with probability  $\pi_b^{((r+i)|_N)}$  due to the assumption that the loss processes are stationary and independent. Therefore,

a burst of length  $m$  occurs with probability  $\prod_{i=1}^{m-1} \pi_b^{((r+i)|N)} \pi_g^{((r+m)|N)}$  since it must end when the path is in the good state.

For  $m = N - 1$ : given that a loss occurs, the probability of losing all remaining  $N - 2$  packets is  $\prod_{i=1}^{N-2} \pi_b^{((r+i)|N)}$ . Note that both the first packet after the burst and the packet immediately preceding the burst must be received correctly, and these packets are sent via the same path. This event occurs with probability  $p_{g,g}^{((r+N-1)|N)}(N)$  and, since it is independent of the others events, the second equality in equation (3) follows.

For  $m \geq N$ : after the first loss, the following  $N - 2$  packets are lost with probability  $\prod_{i=1}^{N-2} \pi_b^{((r+i)|N)}$  and the  $(N - 1)$ -st packet is lost with probability  $p_{g,b}^{((r+N-1)|N)}(N)$ . Subsequent packets are sent via the paths that already lost a packet and they are all lost with probability  $\prod_{i=1}^{(m-N)} p_{b,b}^{((r-1+i)|N)}(N)$ . The burst ends with a correctly received packet and this event occurs with probability  $p_{b,g}^{((r+m)|N)}(N)$ . ■

From Corollary 1 it is possible to draw some conclusions about path diversity when all paths are homogeneous (i.i.d.  $\mathcal{X}_r$ ). In this case, let  $\mathcal{H}_N$  denote the set of loss processes for  $N$  paths. Corollary 2 below compares the probability that a loss burst is smaller than  $k$  when the number of paths is  $k$  against that when the number of paths is larger. For the comparison, the following lemma is necessary.

**Lemma 2** For  $N$  finite,  $p_{g,g}(N) > \pi_g$ .

*Proof:* From equation (1)  $p_{g,g}(N) = \pi_g + \pi_b e^{-N(\mu_g + \mu_b)}$ . The lemma immediately follows. ■

**Corollary 2**  $P[L(\mathcal{H}_k) < k] \geq P[L(\mathcal{H}_N) < k]$  for  $1 < k < N$ ,  $N$  finite.

*Proof:* After simplifying equation (3) for homogeneous paths we obtain  $P[L(\mathcal{H}_k) = m] = P[L(\mathcal{H}_{k+1}) = m]$  for  $m < k - 1$ . Also from the same equation:  $P[L(\mathcal{H}_k) = k - 1] = \pi_b^{k-2} p_{g,g}(k)$  and  $P[L(\mathcal{H}_{k+1}) = k - 1] = \pi_b^{k-2} \pi_g$ . From Lemma 2 we have  $p_{g,g}(k) > \pi_g$  and then  $P[L(\mathcal{H}_k) = k - 1] > P[L(\mathcal{H}_{k+1}) = k - 1]$ . The corollary follows by summing the appropriate terms. ■

We can use Corollary 2 to elaborate on the following question: “Does the efficacy of path diversity improve with an increasing number of paths?”. The answer depends on the application’s sensitivity to small and long loss bursts, i.e. increasing the number of paths may not necessarily improve the quality perceived by the application. If we increase the number of paths, the probability of a long burst decreases, but the probability of a small burst may increase. Since the probability of a very long burst is generally very small, it may not be efficient to perform path diversity over many paths.

Figure 1 plots the probability that a loss burst of length greater or equal than  $m$  occurs as  $m$  increases and the number of independent paths varies from 1 to 6. All curves in Figure 1 were obtained using equation (3) and parameter values taken from [20]. From the figure it is evident that, beyond four paths, the improvements in the burst probabilities are negligible.

Consider a sequence  $\{X_i(1), \dots, X_i(n)\}$  obtained from the loss process for  $i$ -path diversity ( $i = 1$  indicates single path) such that  $X_i(j) = 1$  if the  $j$ -th packet sent is lost and  $X_i(j) = 0$  otherwise. Let  $\lambda_{ik}$  denote

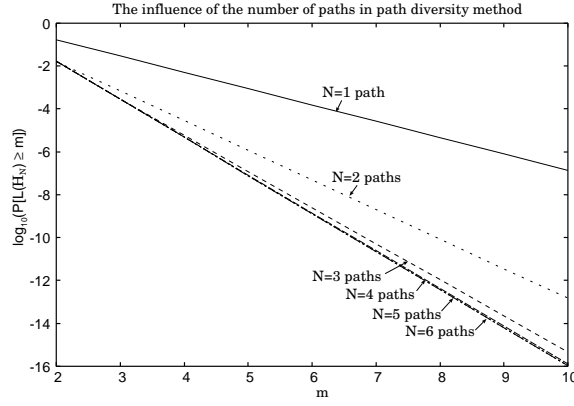


Figure 1: Model parameters:  $p_{g,b}(1) = 0.014$  and  $p_{b,b}(1) = 0.173$ .

the rate at which loss bursts of length  $k$  or greater occur, i.e.  $\lambda_{ik} = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbf{1}\{X_i(j) = 0, X_i(j+1) = 1, \dots, X_i(j+k) = 1\}$ . Another way to show the efficacy of path diversity is to calculate  $\phi_{N,m} = 1 - \lambda_{N,m}/\lambda_{1,m}$ . (Similar definition can be used for interleaving.) Note that if  $\phi_{N,m}$  is equal to  $a$  then  $N$ -path diversity has  $a \times 100\%$  less bursts of length  $\geq m$  than the single path scheme. The quantity  $\phi_{N,m}$  can be easily obtained from  $P[L(\Gamma_N) \geq m]$  and  $P[B_r]$  above. Figure 2 plots  $\phi_{N,m}$ . From the figure it is easy to see that the total number of bursts increases with path diversity w.r.t. single path, since  $\phi_{N,1}$  is negative. However, 2-path diversity has nearly 90% less bursts of length  $\geq 2$  than single path. One should also note that increasing the number of paths from 2 to 3 or even 4 is less efficient than 2-path diversity to lower the number of bursts of length 2 or more. On the other hand, 3 or 4-path diversity is more efficient than 2-path diversity to lower the number of bursts of length 3 or more. This result is counter intuitive, since one might expect that, if the application uses more paths, the probability of bursts of length two or longer would decrease. This behavior is more pronounced if the loss processes are such that long bursts are likely to occur (that is, if we increase the expected time in the “bad” states of the loss models). Thus, increasing the number of paths may have a negative impact on the quality perceived by the application. Consequently, the application must carefully evaluate the effect of adding paths taking into account the loss processes involved.

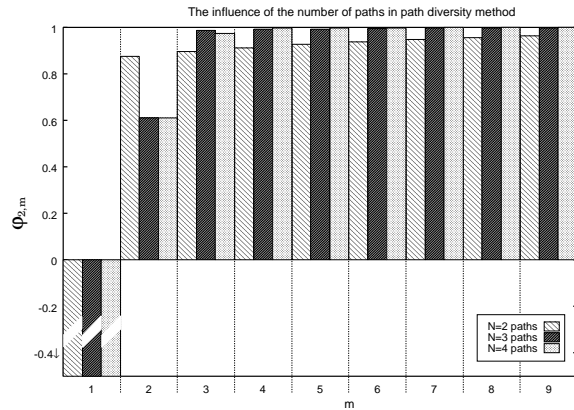


Figure 2: Model parameters:  $p_{g,b}(1) = 0.01$  and  $p_{b,b}(1) = 0.84$ .

The next example considers heterogeneous paths. Figure 3 illustrates an example in which an application has up to four available paths to use: a relatively *good* path (with  $\mathcal{X}_1$  such that  $p_{g,b}^{(1)}(1) = 0.0141$  and  $p_{b,b}^{(1)}(1) = 0.1732$ ) and three relatively *bad* paths (with  $\mathcal{X}_i$  such that  $p_{g,b}^{(i)}(1) = 0.028$  and  $p_{b,b}^{(i)}(1) = 0.837$ ,  $i = 2 \dots 4$ ). The figure plots  $\varphi_{N,m}$  for different values of  $N$ . When  $N = 1$  it is assumed that the application send packets via the good path. When  $N > 1$  the application uses the good path and the remaining  $N - 1$  bad paths. Furthermore, the measure  $\varphi_{N,m}$  is calculated w.r.t. the good path.

Figure 3 shows that, depending on the value of the burst length being considered, the number of bursts of length greater or equal than  $m$  using  $k > 1$  paths may be larger or smaller than using the single best path. Using two or three paths drastically reduces the number of  $\geq 4$  burst lengths. However, note that, 3-path diversity may increase the number of  $\geq 5$  burst lengths w.r.t. the single best path bursts. Note that one may also need to consider the probability of such bursts to occur in order to evaluate the overall performance of the method. (In this case, although not shown, this probability is small.)

All the values of  $\varphi_{N,m}$  in Figure 3 are calculated w.r.t. the best path. However, in practice, one may not know the best path in advance. Consequently, if we calculate  $\varphi_{N,m}$  w.r.t. (for instance) a uniformly weighted sum of the number of burst lengths in all paths, the advantage of path diversity would increase when compared to the results of Figure 3. Overall, the efficacy of path diversity is better than using only a single path, even if the additional path is worse than the first.

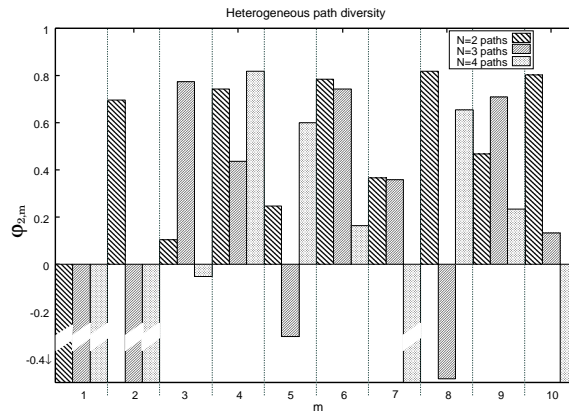


Figure 3: Graph of  $\varphi_{N,m}$  for various values of  $N$ .

It is well known that multimedia applications are affected by the length of a loss burst [14]. It is also clear that we can combine a FEC scheme with path diversity to reduce the length of a burst [17]. We apply our model to analyze the effectiveness of FEC schemes coupled with path diversity and compare efficacy of the technique when two and three paths are used. For this comparison additional metrics are required. Suppose that our objective is to maximize the number of packets corrected by the FEC scheme of [6], combined with the path diversity. This FEC scheme can correct a single packet in a loss burst of any length. Define the function

$$\phi(\mathcal{H}_N) = \sum_{m=1}^{\infty} P[L(\mathcal{H}_N) = m]/m$$

which is the fraction of packets corrected by the scheme of [6]. (Other, FEC schemes can be used as well,

such as the one in [7], but we choose the scheme of [6] in this example since the function  $\phi(\mathcal{H}_N)$  is very simple in this case.)

Figure 4 shows  $\phi(\mathcal{H}_2) - \phi(\mathcal{H}_3)$ , as a function of the two parameters,  $p_{g,b}(1)$  and  $p_{b,g}(1)$  of the loss model. Since  $\phi(\mathcal{H}_2) - \phi(\mathcal{H}_3)$  is always positive, it is clear that two paths is better than three paths for this FEC scheme. This conclusion is also valid if one increases the number of paths further. This peculiar behavior follows directly from Corollary 2.

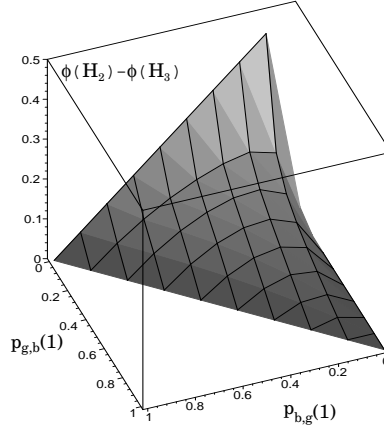


Figure 4: Difference between the fraction of packets recovered by a chosen FEC scheme with two and three paths.

### 3 Path Diversity and Packet Interleaving

Several other methods exist that can mitigate the negative impact of loss bursts in the application. One of these is *packet interleaving*, which has the advantage of not making use of multiple paths and, similar to path diversity, not increasing throughput by adding redundancy to the original stream. The most common interleaver is that exemplified in Figure 5, which is called a *block interleaver*. Throughout this section, we refer to the indices of the packets that indicate the order in which they are generated by the application as the *original index (OI)*. Clearly this differs from the order packets are transmitted. In the example of Figure 5, packets with *OI* 1, 5, 9 and 2 are transmitted in sequence. There are two ways to order packets belonging to two consecutive blocks. Let  $j$  be the *OI* index of the last packet in a block. Then, in one scheme the first packet in the next block has *OI* index  $j + 1$ , and the block follows the same pattern as its predecessor. The second scheme is similar to the first, but *OI* indices are mirrored every other block. In this paper we study the first approach.

With reference to Figure 5, let  $\eta$  and  $\Delta$  be the number of columns and lines in a block, respectively. Let  $Y(\eta, \Delta, \mathcal{X})$  be a random variable equal to the length of a loss burst when an  $\eta \times \Delta$  interleaving scheme is used, and the loss process of the transmission path is  $\mathcal{X}$ . Since all the variables and measures of interest depend on  $\eta$ ,  $\Delta$  and  $\mathcal{X}$ , we omit this dependency to simplify notation.

Path diversity and packet block interleaving have some similarities. Both methods try to reduce the





*Proof:* Choose a packet at a random position in a block, and refer to the received (re-ordered) packet stream. Let  $\Phi_k$  be the event that a loss burst starts at the  $(k+1)$ -th position, conditioned that position  $k$  was selected. Clearly, the event  $\Phi_k$  occurs if the selected packet  $k$  is received and packet  $k+1$  is lost. Let  $u = f(k+1)$ , i.e.  $u$  is the order of transmission of packet  $k+1$ . We consider three cases:

- (i)  $\Delta < u \leq \Delta\eta$ : The  $k$ -th packet is received with probability  $\pi_g$  and packet  $k+1$  belongs to the next column with respect to that of packet  $k$  (see Figure 5), and so it is lost with probability  $p_{g,b}(\Delta)$ . Then a burst starts with packet  $k+1$  with probability  $\pi_g p_{g,b}(\Delta)$ .
- (ii)  $1 < u \leq \Delta$ : Note that, in this case, packet  $k+1$  is transmitted before packet  $k$ . Since packet  $k+1$  is lost and packet  $k$  is not,  $P[\Phi_k] = \pi_b p_{b,g}((\eta-1)\Delta-1)$ .
- (iii)  $u = 1$ : In this case packet  $k$  is in the block preceding that of packet  $k+1$ , and note that packets  $k$  and  $k+1$  are transmitted in sequence. Therefore, a burst starts with packet  $k+1$  with probability  $\pi_g p_{g,b}(1)$ .

Since each packet in a block is equally likely to be selected, after unconditioning on the position of the selected packet we have:

$$P[\Phi] = \pi_g p_{g,b}(\Delta) \frac{\eta-1}{\eta} + \pi_b p_{b,g}((\eta-1)\Delta-1) \frac{\Delta-1}{\eta\Delta} + \frac{\pi_g p_{g,b}(1)}{\eta\Delta}. \quad (6)$$

Equation (5) is obtained after rearranging the terms in (6). ■

**Corollary 3** *The maximum value of  $P[\Phi]$  occurs when  $\Delta \rightarrow \infty$ .*

*Proof:* From equation (1) it is clear that  $p_{g,b}(l)$  and  $p_{b,g}(l)$  monotonically increase with  $l$ . Therefore, the first two terms in equation (5) monotonically increase with  $\Delta$ .

Consider the last term in (5). Since  $p_{b,g}(l)$  increases with  $l$ , the term in brackets in (5) monotonically decreases with  $\Delta$ , and its maximum value occurs when  $\Delta = \eta = 2$ . In this case, it is easy to see (using (1)) that the term in brackets is zero, and so it is always negative for  $\Delta \geq 2$  and  $\eta \geq 2$ . Furthermore, as  $\Delta \rightarrow \infty$ , this last term is equal to zero. Since the first two terms in (5) are monotonically increasing, the maximum value of  $P[\Phi]$  occurs when  $\Delta \rightarrow \infty$ . ■

As a consequence of Corollary 3 and equation (4),  $E[\Upsilon(\eta, \Delta, \mathcal{X})]$  is minimized when  $\Delta \rightarrow \infty$ , for any value of  $\eta \geq 2$ , independent of the loss process. This is an idealized scenario, since latency is equal to infinity, but the results in this case serves as a bound to evaluate the performance of interleaving.

Let  $\Upsilon(\eta, \mathcal{X}) = \lim_{\Delta \rightarrow \infty} \Upsilon(\eta, \Delta, \mathcal{X})$ . We focus on its probability mass function (pmf).

**Theorem 1**

$$P[\Upsilon(\eta, \mathcal{X}) = m] = \begin{cases} \pi_b^{m-1} \pi_g & \text{for } m < \eta - 1, \\ \pi_b^{\eta-2} p_{g,g}(1) & \text{for } m = \eta - 1, \\ \pi_b^{\eta-2} p_{g,b}(1) [p_{b,b}(1)]^{m-\eta} p_{b,g}(1) & \text{for } m \geq \eta \end{cases} \quad (7)$$

*Proof:* Refer to Figure 5. Packet loss events in different columns are independent of each other in the limit as  $\Delta \rightarrow \infty$ . As a consequence, only packets from the same column have correlated loss processes.

One can then compare this scheme with  $\eta$ -path diversity as follows. Packets from different columns in interleaving have sequential  $OI$  indexes and see independent loss processes, like packets sent via different paths in path diversity. Packets from the same column are sent one unit of time apart and see the same loss process. This is similar to packets being sent using the same path in the path diversity method. However, in this last case, packets are spread  $\eta$  time units apart. Equation (7) can then be obtained following the same steps used to derive (3). ■

It is interesting to compare Theorem 1 against Corollary 1 to observe that in the idealized interleaving scenario, interleaving and packet diversity methods have similar characteristics. A result similar to Corollary 2 can also be stated. Define a function  $\phi'(\eta, \mathcal{X})$  similar to  $\phi(\mathcal{H}_N)$  (see Section 2):  $\phi'(\eta, \mathcal{X},) = \sum_{m=1}^{\infty} P[Y(\eta, \mathcal{X}) = m]/m$ . Figure 6 shows  $\phi'(2, \mathcal{X}_1) - \phi'(3, \mathcal{X}_1)$  as a function of parameters  $p_{b,g}(1)$  and  $p_{g,b}(1)$  of the loss model. It is clear from Figures 6 and 4 that similar conclusions can be drawn from the two methods.

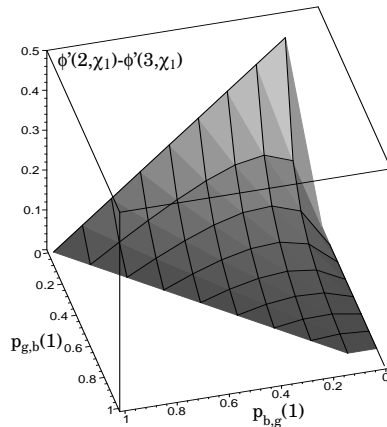


Figure 6: Difference between the fraction of packets recovered by FEC with  $\eta \in \{2, 3\}$ .

Figure 7 compares the efficiency of each method to reduce loss bursts when both the number of paths in path diversity and the interleaving parameter  $\eta$  vary, in the idealized interleaving case. One observes that, even in the idealized case, path diversity is better than packet interleaving when  $\eta = N$  in this example. From equations (7) and (3) it is clear that, for  $m < \eta - 1$  and  $\eta = N$ , path diversity and packet interleaving have the same loss burst length statistics.

In what follows we present a method to calculate  $P[Y(\eta, \Delta, \mathcal{X}) = m]$  for any value of  $\Delta$  and  $\eta$ . Unlike the idealized interleaving case no closed form was found. Instead, we devise a simple algorithm to calculate  $P[Y(\eta, \Delta, \mathcal{X}) = m]$ . The results will also be useful to check the convergence of  $P[Y(\eta, \Delta, \mathcal{X}) = m]$  as  $\Delta \rightarrow \infty$ , and check if the results of Theorem 1 can be used as an approximation for small interleaving latencies.

Similar to the definition of  $\Phi_k$ , choose a packet at random in a block and let  $\Psi_k(\eta, \Delta, \mathcal{X}) = m$  be the event that a loss burst of length  $m$  starts just after the selected packet, conditioned on the successful packet transmission in position  $k$ . For the event to occur, packet  $k$  must be correctly received, packets  $k+1, \dots, k+m$  lost and packet  $(k+m+1)$  also correctly received. Let  $\mathcal{S}(k, m)$  be the set of indexes  $\{f(k), f(k+1), \dots, f(k+m+1)\}$ , and let  $\gamma_1(\mathcal{S}(k, m))$  be the smallest index in  $(\mathcal{S}(k, m))$ ,  $\gamma_2(\mathcal{S}(k, m))$  the second smallest index in  $(\mathcal{S}(k, m))$ , etc. We also define two other functions:  $f^{-1}(a)$  and  $c(v, m)$ . The first function maps the transmission index  $a$  into the original packet position in a block.  $c(v, m) = g$  if  $v = k$ ,

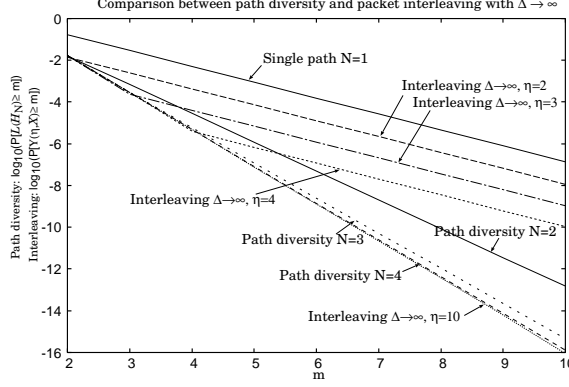


Figure 7: Path diversity versus packet interleaving, graph of  $\log_{10} P\{Y(\eta, \mathcal{X}) \geq m\}$  and  $\log_{10} P\{L(\mathcal{H}_N) \geq m\}$  with parameters  $p_{g,b}(1) = 0.0141$ ,  $p_{b,b}(1) = 0.1732$ , infinite latency ( $\Delta \rightarrow \infty$ ),  $\eta \in \{2, 3, 4, 10\}$ , and  $N \in \{1, 2, 3, 4\}$ .

where  $k$  is the index of the selected packet, and  $g$  indicates that packet  $k$  is correctly received;  $c(v, m) = b$  if  $k + 1 \leq v \leq k + m$ , where  $b$  indicates that packet  $v$  is lost. Finally,  $c(v, m) = g$  if  $v = k + m + 1$ .

## Theorem 2

$$P[Y(\eta, \Delta, \mathcal{X}) = m] = \frac{1}{P[\Phi]\Delta\eta} \sum_{k=1}^{\Delta\eta} P[\Psi_k(\eta, \Delta, \mathcal{X}) = m] \quad (8)$$

and

$$P[\Psi_k(\eta, \Delta, \mathcal{X}) = m] = \pi_{\zeta(i)} \times \prod_{i=1}^{m+1} p_{\zeta(i), \zeta(i+1)}(\gamma_{i+1}(\cdot) - \gamma_i(\cdot)) \quad (9)$$

where  $\zeta(i) = c(f^{-1}(\gamma_i(\mathcal{S}(k, m))), m)$ , and the dependence of  $\zeta(i)$  and  $\gamma_i$  on  $\mathcal{S}(k, m)$  is omitted to clarify notation.

*Proof:* The proof parallels that for  $P[\Phi]$ . In this case however, for each selected position  $k$  in a block, we have to analyze the transmission position of each packet in a burst. Assume packet  $k$  is selected. Then for the event  $\Psi_k(\eta, \Delta, \mathcal{X}) = m$  to occur, packet  $\gamma_1(\mathcal{S}(k, m))$  (the first packet transmitted in set  $\mathcal{S}(k, m)$ ) must be lost or correctly received according to its position in the burst which is given by  $f^{-1}(\gamma_1(\mathcal{S}(k, m)))$ . The value of the function  $\zeta(i) = c(f^{-1}(\gamma_1(\mathcal{S}(k, m))), m)$  indicates if the packet is lost or not ( $b$  or  $g$ ). The product in equation (9) is obtained by observing which are the loss process transition probabilities that must take place in order to trigger the event  $\Psi_k(\eta, \Delta, \mathcal{X})$ . Each packet in a block is equally likely to be selected, and so the sum in equation (8) follows from unconditioning on the selected event. Since  $Y(\eta, \Delta, \mathcal{X})$  is the random variable that gives the length of a burst, the result is obtained after normalizing by  $P[\Phi]$ , i.e. the probability of a burst. ■

We first compare  $P[Y(\eta, \Delta, \mathcal{X}) = m]$  and  $P[Y(\eta, \mathcal{X}) = m]$  to see if the latter produces reasonable approximations to the former. Figure 8 shows that, for latencies greater or equal to 0.5 (recall that latency is equal to  $(\eta - 1)(\Delta - 1)$ ),  $P[Y(\eta, \Delta, \mathcal{X}) \geq m] \approx P[Y(\eta, \mathcal{X}) \geq m]$  when  $\eta$  is “small”. As we increase  $\eta$  maintaining a constant latency,  $P[Y(\eta, \Delta, \mathcal{X}) \geq m]$  also increases, reducing the efficacy of the approach.

We also note that  $P[Y(\eta, \Delta, \mathcal{X}) \geq m]$  decays with  $m$  at a high exponential rate when  $m < \eta - 1$ , and decays

at a much slower rate when  $m \geq \eta - 1$ . It is not immediately clear that increasing  $\eta$  while maintaining a constant latency (and so decreasing  $\Delta$  proportionally to  $\eta$ ), improves the efficiency of interleaving. For instance, consider the curves when the latency is equal to 0.5s and  $\eta$  changes from 7 to 10. Note that  $P[\Upsilon(7, 5, \mathcal{X}) \geq m] < P[\Upsilon(10, 3, \mathcal{X}) \geq m]$  for  $m \leq 8$  but  $P[\Upsilon(7, 5, \mathcal{X}) \geq m] > P[\Upsilon(10, 3, \mathcal{X}) \geq m]$  when  $m > 8$ .

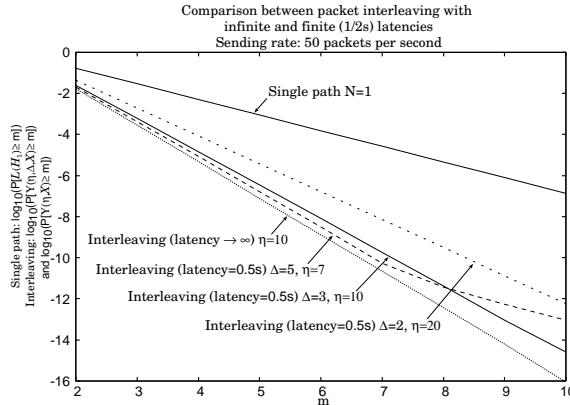


Figure 8: Comparing latencies in packet interleaving, graph of  $\log_{10}P\{\Upsilon(\eta, \Delta, \mathcal{X}) \geq m\}$  and  $\log_{10}P\{L(\mathcal{H}_1) \geq m\}$  with parameters  $p_{g,b}(1) = 0.0141$ ,  $p_{b,b}(1) = 0.1732$ .

Figure 9 presents results with smaller latencies than those shown in Figure 8. As expected, the efficacy of interleaving is reduced when latency decreases from 0.5 secs to 0.1 secs, but the method is still efficient compared to the case where no interleaving or path diversity is employed (the single path case). Furthermore, the result with  $\eta = 2$  outperforms that with  $\eta = 4$ . Note also that the results when latency is equal to 0.1 secs are very different from those that use a latency value of 0.5 secs.

In Figure 9 we also compare  $P[\Upsilon(\eta, \Delta, \mathcal{X}) \geq m]$  and  $P[\Upsilon(\eta, \mathcal{X}) \geq m]$ , with  $P[L(\mathcal{H}_1) \geq m]$ , using the same loss process  $\mathcal{X}$  for all paths. Homogeneous paths are chosen to emphasize the differences in the two methods, independent of the path loss statistics. It is interesting to observe that 2-path diversity performs better than interleaving with  $\eta = 10$  when  $m \leq 3$ .

## 4 Path diversity with shared bottlenecks

In the previous sections we studied path diversity when none of the paths share bottlenecks. In practice, the probability that an additional path will share a bottleneck with some of the other paths increases with the number of paths [18]. In this section we will account for shared bottlenecks and determine if they affect previous conclusions. One may expect that if all the used paths share a bottleneck then there may be no significant advantages of using multiple paths. Furthermore, packets sent via different paths may experience different delays and the difference between these delays values may be non-negligible. As a consequence, loss correlations between adjacent packets may decrease lowering the probability of long loss bursts. Note that this effect is similar to that of packet interleaving. The issue is to quantify the influence of the bottleneck(s) and different delays in the efficacy of path diversity.

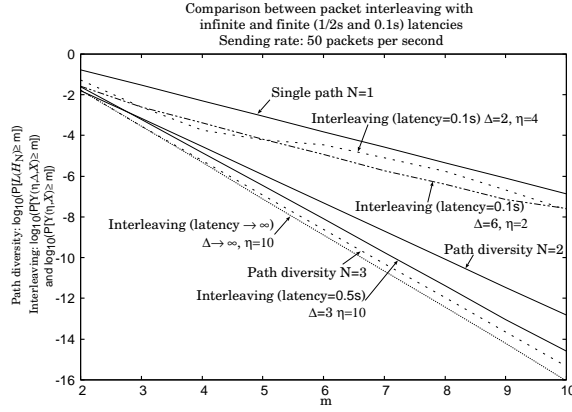


Figure 9: Packet interleaving with low latency. Graph of  $\log_{10} P\{Y(\eta, \Delta, X) \geq m\}$  and  $\log_{10} P\{L(\mathcal{H}_N) \geq m\}$  with parameters  $p_{g,b}(1) = 0.0141$ ,  $p_{b,b}(1) = 0.1732$ .

In what follows we will study the influence of a bottleneck shared by all paths over the loss burst length through analysis. The model we use is as follows. Each path from a source to a destination is partitioned into two segments, each associated with an independent Gilbert loss process. In the independent path case neither path shares a segment. In the shared bottleneck case all paths share one segment. The same methodology used in previous sections to obtain the measures of interest for the independent path case can be also employed for the shared bottleneck case. The resulting equations are slightly more complex and we omit the development for conciseness.

Figures 10 and 11 presents a sample result using the analytical model. In Figure 10 the parameter values for all Gilbert models associated to the path segments are identical. In Figure 11 the Gilbert models associated with the independent segments are parameterized such that the mean loss burst length is increased by ten times with respect to the models in Figure 10 while preserving the loss rate.

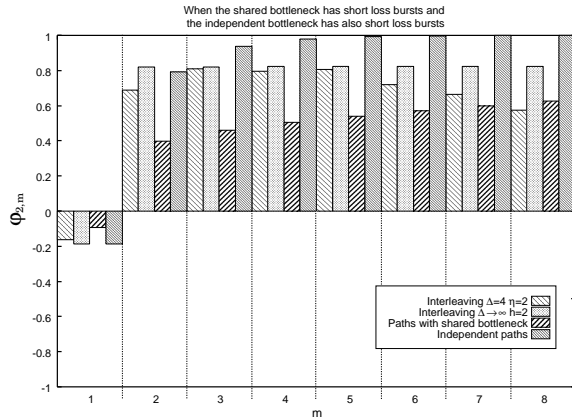


Figure 10:  $\phi_{2,m}$  when all segments are associated with the same loss process model.

In Figure 10 we observe that 2-path diversity performs worse than interleaving, but the gains are significant with respect to single path. In Figure 11 even though one may expect that the efficacy of 2-path

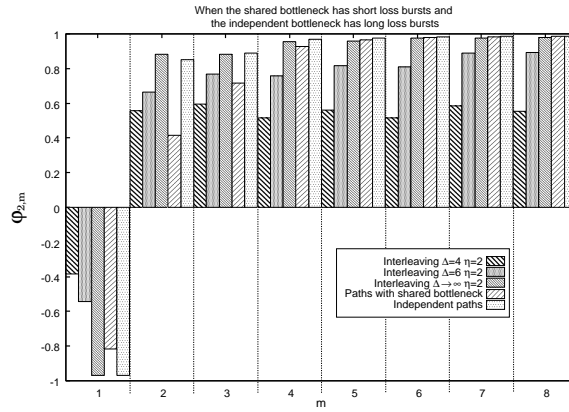


Figure 11: Graph of  $\phi_{2,m}$  when the shared segment is associated with a different loss process model.

diversity would not change much since the loss rate was preserved, its performance significantly increases when compared to Figure 10. This occurs because, in this example, long loss bursts are more likely to occur in the independent path segments, instead of the shared segment. In summary, the effect of a shared bottleneck link in the overall performance of the method may vary drastically, and the method may be either worse or better than interleaving, but it is still superior when compared to the single path case.

In what follows we report on some experiments done over the Internet in order to access the effectiveness of path diversity. The topology of all Internet experiments is given in Figure 12. (UFRJ, UFMG, UFJF are universities in Brazil.) In our experiments we sent regularly spaced packets from UFRJ at a rate of 119

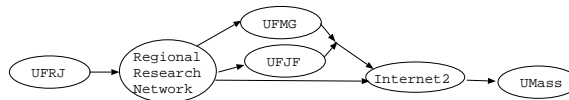


Figure 12: Experiments topology.

Kbps. Three paths from UFRJ to UMass are possible. In this scenario, all paths share a segment at UFRJ and at UMass, i.e. all packets sent traverse the same route while in UFRJ (UMass) and go to the same output (input) router. The segment inside UFRJ is the bottleneck (has the highest loss rate).

One should note that there may be non-negligible differences in the delays experienced by packets that follow different paths. When the difference in latencies is high and this difference is due to the delays in the segments from the source till the shared segment, the loss burst statistics may be significantly affected. This leads to temporal independence in the loss processes, similarly to interleaving. In our scenario, since one shared segment is at the source of packets, and this segment has the worst loss statistics, one should expect that the different path latencies will not favor loss statistics. However, the gains due to the difference in latencies can be easily observed in our analytical models, though the results are not presented here due to space limitations.

We collected 12 sets of one hour traces (during 8 days at difference time intervals) using the topology described in Figure 12, by sending a continuous flow of packets in all the three different paths at rate equal to 119 Kbps in each path. We show the results for a set that presented large loss rates over the paths so that

the benefits of reducing loss burst lengths were more evident. The expected difference in latencies between the UFRJ-UFJF-UMass and UFRJ-UMass paths is approximately 67 msec and 95% of the values were less than 81 msec. Figure 13 shows  $\phi_{2,m}$  for paths UFRJ-UFJF-UMass, UFRJ-UMass and UFRJ-UFGM-UMass w.r.t. the single path which can be either the UFRJ-UMass path or UFRJ-UFJF-UMass path. Also included are results for packet interleaving over the best path, which is the UFRJ-UMass path in our experiments. We chose 80 milliseconds for our packet interleaving latency, to avoid adding too much delay over long paths, which otherwise would jeopardize the use of interactive streaming applications. The appropriate value of  $\eta$  was chosen using Theorem 1, as described in section 3.

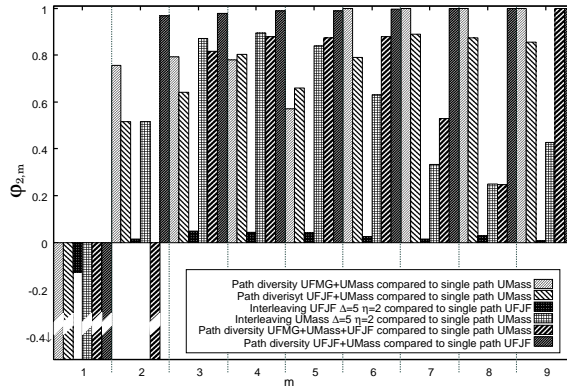


Figure 13:  $\phi_{N,m}$  using interleaving with  $\Delta = 5$   $\eta = 2$  and  $N$ -path diversity for  $N = 2, 3$ .

The paths are highly heterogeneous, since the loss rates were 0.036, 0.10 and 0.25 for the UFRJ-UMass and UFRJ-UFGM-UMass and UFRJ-UFJF-UMass, respectively. The figure illustrates the large benefits that can be obtained with packet diversity with respect to single path. The performance w.r.t. interleaving depends on the interleaving parameters and the number of paths as shown in the figure.

## 5 Path Diversity Scalability

Multimedia content delivery networks (CDNs) are planned to serve large populations of users. One may consider the use of path diversity to reduce traffic loss and improve audio and video quality as perceived by the users for such applications. In this case, the amount of bandwidth used by applications employing path diversity may become a large fraction of all bandwidth usage. This could introduce undesirable correlations among the path loss processes and call into question assumptions made in previous sections. On the other hand, spreading traffic over multiple paths lowers the traffic burstiness, which in turns reduces loss probabilities [10]. In this section we analyze 2-path diversity when the applications that use this method consume a large fraction of the available network bandwidth and address the above issues.

Figure 14 shows an example with two CDNs, each with a number of video and audio servers. In this example there are two possible ways to serve client requests. In the first case, a server in one of the CDNs is chosen to stream the media requested by a client. The second possibility is to use two servers, each located at a different CDN, to jointly serve the client. For instance, one could stream odd numbered blocks of the requested stream and the other the even numbered blocks.



This last approach corresponds to path diversity. As mentioned in previous sections, the efficacy of path diversity in shortening loss bursts is influenced by: (1) the degree of correlation among the loss processes on different paths and; (2) the degree of traffic smoothness that occurs as a consequence of spreading traffic through two or more paths.

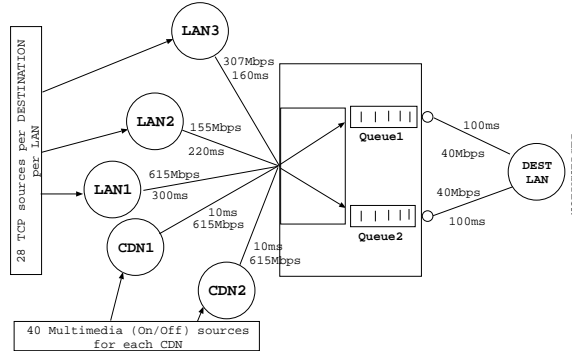


Figure 14: NS topology, path diversity sources (On/Off CBR) with TCP sources.

In what follows we assume that the CDNs are responsible for more than half of the traffic in the routers and the remaining load is generated by greedy TCP sources. The model is shown in Figure 14. It is a simulation model and details concerning the behavior of the sources can be included. The speed of the links and propagation delays are shown in Figure 14. Multimedia sources are modeled as Markovian On/Off sources with mean On and Off times equal to  $10^3$  seconds. Each continuous media source transmits CBR traffic at a rate of 125 packets/sec during the On period and packets are 1,500 bytes long. All TCP sources have a maximum window size equal to 50 packets and *MSS* equal to 1,500 bytes. The buffer size of each router is 50 packets and both use the drop-tail policy. 80 multimedia sources and 168 TCP sources are used in the experiment.

Four to eight runs Each consisting of 83 simulated time hours were used to obtain one measure. (Since the confidence intervals are very tight, they are not shown in the figure.) Three cases were studied: (1) all continuous media CDNs sources use path diversity; (2) a single CDN source using path diversity while the others do not; (3) all sources not using path diversity; (4) all except one of the CDNs sources use path diversity (in this case we are interested in the CDN source that does not use path diversity).

Figure 15 plots the results of our simulation. In this figure we plot  $P[L \geq m]$  for the CDN source of interest. From the graph it is clear that, when only a single source uses path diversity, (“Only One Using PD”) this source benefits the most. The second best result occurs when all sources use path diversity (“All Using PD”). Even when a source does not use path diversity while others do, the first source can benefit as shown by the curve labeled “All But One Using PD”. This can be inferred by comparing the results of this case with those when none of the sources use path diversity.

Comparing the curves “All But One Using PD” and “None Using PD” shows (which was already known) that path diversity smooths the traffic and benefits the sources that do not use this technique. From Figure 15 we also conclude that, even when correlation is introduced between the loss processes when all sources use path diversity, the negative impact in  $P[L \geq m]$  is not sufficiently large to hurt the overall benefit of the technique. This is true since  $P[L \geq m]$  for the “All Using PD” curve is lower than that for the “None Using

PD” curve in the range of values plotted. The best case for a CDN source is achieved when only one source uses path diversity. We then conclude that the benefit of traffic smoothing is not as strong as the benefit obtained by introducing independence between the loss processes on the different paths. Note that traffic smoothing occurs when most of the CDNs sources use path diversity, but this is also the case where the correlation between the path loss processes has the highest value. When only one source uses path diversity the correlation between the loss process of the two paths is negligible and the source can fully benefit from the technique.

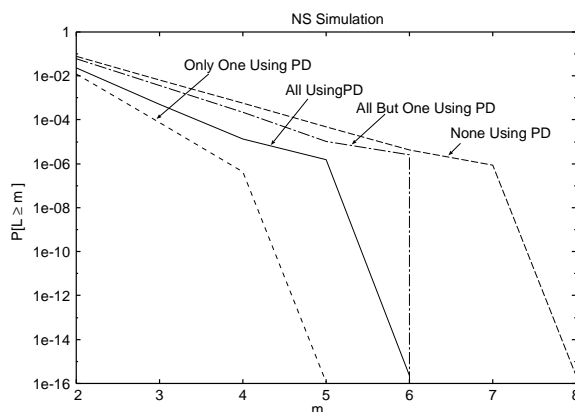


Figure 15: Loss burst distribution for the model of Figure 14.

## 6 Related Work

Path diversity was first proposed by Maxemchuk [16]. Banerjea [5] also studied path diversity in the context of IP/ATM routing (see Gustafsson and Karlsson [10] for a survey on the subject). One of their conclusions is that path diversity performs load balancing among routers as it sends traffic through more than one path and, as a consequence, smooths the traffic in the links.

Gogate and Panwar [8] were the first to apply path diversity to multimedia traffic. In that work, mobile multi-hop radio environments were studied, and in [15] *ad-hoc* networks were considered. Apostolopoulos [1] and also Girod, Steinbach and Liang [12], studied path diversity in the Internet. Following these works, many studies have been published on the technique: [2, 3, 4, 9, 11, 13, 19, 17].

The works of Apostolopoulos and his co-authors [2, 3, 1, 4] study the effect of path diversity on video distortion considering two paths. Their focus is to combine path diversity and *multiple description coding* (MDC) for video applications. Video distortion was due to packet losses and the loss process was modeled using the Gilbert model in their most recent works [2, 4]. They studied the performance of path diversity when joint or disjoint links were used, and also considered balanced and unbalanced paths. They found four main benefits of using multiple paths over a single path: better network behavior, shorter burst losses, reduced probability of connection outages and better MDC video quality.

Liang *et al* [12, 13] also studied path diversity with two paths. In [13] MDC was considered and in [12] FEC was used to recover from losses. Both simulation and experiments using the Internet were performed.

They also measured the impact of shared links on path diversity.

Golubchick et al. [9] analyzed the use of path diversity for pre-stored video using two paths, and modeled the loss process of each path using the Gilbert model. They concluded that, in general, when compared to single-path streaming, 2-path diversity improves the loss characteristics of the connection independently of the use of an erasure code, reduces the data loss rate, improves the error burst length distribution conditioned that an error occurs and has zero lag1-autocorrelation. There is also an analysis for the 3-path diversity case for some set of Gilbert model parameters. Although the authors studied the case  $N = 3$ , their result in the  $N = 3$  case is neither general nor conclusive.

Wang, Panwar and co-workers [19, 15] studied multiple description coding with two-path diversity as a delivery mechanism for image and video over ad-hoc/wireless networks. Nguyen and Zakhor [17] proposed a FEC scheme designed specifically for path diversity and delay sensitive applications (two paths were considered). This work studied the existence of disjoint paths in the Internet and proposes a heuristic to find disjoint paths between two end points. Issues concerning the existence of multiple disjoint paths in the Internet are also discussed in [18].

## 7 Conclusions and future work

In this work we address several issues concerning the path diversity technique. First we developed a model to obtain the pmf of the length of a loss burst when an application employs path diversity over  $N$  either homogeneous or heterogeneous paths. The model reveals interesting anomalies when the number of paths increases. The analysis shows that increasing the number of paths does not necessarily improve the efficacy of the method, even when the loss processes of all paths are identical and independent processes. Similar conclusions are reached when path diversity is combined with the FEC scheme of [6]. Equation (3) (for heterogeneous paths) can be employed by the analyst to decide, based on path statistics, the number of paths to be used by an application.

We then compare path diversity and packet block interleaving. We show that both methods have similar characteristics. Using the same model parameters as in [20] (obtained from real traces) we concluded that path diversity outperforms interleaving, though the difference may not be noticeable depending on the model parameters.

We extended the analytical models we develop to study the impact of a shared bottleneck on the loss burst statistics. The results show that the efficacy of path diversity to reduce long loss bursts depends both on the loss rate as well as on the loss burst lengths of the shared path segment. We also perform a few Internet experiments and compared path diversity with the single path case and with interleaving.

Finally we address the scalability of the method, that is, the resulting effect when a large fraction of the available bandwidth is consumed by servers that use path diversity to stream media to clients. The efficacy of path diversity relies on the loss process of the paths being independent. However, increasing the number of applications that employ the technique introduces correlation among the loss processes, which in turns lower the efficiency of the method. On the other hand, distributing the load over many paths smooth the overall traffic, which causes a reduction on the packet loss probability. We investigate the existing trade-offs using a simulation model. We conclude that the best scenario for an application is that when only a few of

them use path diversity.

Path diversity can be employed jointly with packet interleaving and other techniques for reducing the impact of losses. Studying the efficacy when a mixture of these techniques are used will be the subject of future work. Furthermore, Although not addressed in this paper, our equations can be extended to handle more complex Markovian loss models.

## References

- [1] John Apostolopoulos. Reliable video communication over lossy packet networks using multiple state encoding and path diversity. *Visual Communications and Image Processing (VCIP)*, pages 392–405, January 2001.
- [2] John Apostolopoulos, W. Tan, Susie Wee, and G. W. Wornell. Modeling path diversity for multiple description video communication. *ICASSP*, May 2002.
- [3] John Apostolopoulos and Susie Wee. Unbalanced multiple description video communication using path diversity. *ICPI*, October 2001.
- [4] John Apostolopoulos, Tina Wong, Wai tian Tan, and Susie Wee. On multiple description streaming with content delivery networks. *Proceedings of the IEEE INFOCOM*, 2002.
- [5] Anindo Banerjea. Simulation study of the capacity effects of dispersity routing for fault tolerant real-time channels. *Proceedings of the ACM SIGCOMM*, 26(4):194–205, Oct. 1996.
- [6] Jean-Chrysostome Bolot and Andrs Vega-Garca. Control mechanisms for packet audio in the internet. In *Proceedings of INFOCOM*, pages 232–239, 1996.
- [7] D.R. Figueiredo and E. de Souza e Silva. Efficient mechanisms for recovering voice packets in the internet. In *Proceedings of IEEE/Globecom, Global Internet: Application and Technology Symposium*, pages 1830–1837, December 1999.
- [8] N. Gogate and S. S. Panwar. Supporting applications in a mobile multihop radio environment using route diversity part I: Non-real time data. *Proceedings of the ICC*, 1998.
- [9] L. Golubchik, J. C.S. Lui, T. F. Tung, A. L.H. Chow, W.-J. Lee, G. Franceschinis, and C. Anglano. Multi-path continuous media streaming: What are the benefits? In *Performance Evaluation*, volume 49, pages 429–449. Kluwer, September 2002.
- [10] E. Gustafsson and G. Karlsson. A literature survey of traffic dispersion. *IEEE Network Magazine*, 11(2):28–36, Mar./Apr. 1997.
- [11] Y. J. Liang and B. Girod. Rate-distortion low-latency video streaming using channel-adaptative bistream assembly. *IEEE International Conference on Multimedia and Expo (ICME)*, Aug. 2002.
- [12] Y. J. Liang, E. G. Steinbach, and B. Girod. Real-time voice communications over the internet using packet path diversity. *Proc. ACM Multimedia*, pages 431–440, Sept/Oct 2001.

- [13] Y. J. Liang, E. G. Steinbach, and B. Girod. Channel-adaptative video streaming using packet path diversity and rate-distortion optimized reference picture selection. *IEEE Fifth Workshop on Multimedia Signal Processing*, Dec. 2002.
- [14] Yi J. Liang, John G. Apostolopoulos, and Bernd Girod. Analysis of packet loss for compressed video: Does burst-length matter? *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-2003*, Apr. 2003.
- [15] S. Lin, S. Mao, Y. Wang, and S. S. Panwar. A reference picture selection scheme for video transmission over ad-hoc networks using multiple paths. *IEEE International Conference on Multimedia and Expo (ICME)*, Aug. 2001.
- [16] N. F. Maxemchuk. *Diversity routing in store and forward networks*. PhD thesis, University of Pennsylvania, 1975.
- [17] Thinh Nguyen and Avideh Zakhor. Path Diversity with Forward Error Correction (PDF) system for delay sensitive applications over the internet. In *Proc. of the IEEE INFOCOM*, San Francisco, CA, Apr. 2003.
- [18] Renata Teixeira, Keith Marzullo, Stefan Savage, and Geoffrey M. Voelker. Characterizing and measuring path diversity of internet topologies (extended abstract). In *Proceedings of the ACM SIGMETRICS*, pages 304–305, 2003.
- [19] Yao Wang, Shivendra Panwar, Shunan Lin, and Shiwen Mao. Wireless video transport using path diversity: Multiple description vs. layered coding. *Proc. International Conference on Image Processing*, 2002.
- [20] M. Yajnik, S. Moon, Jim Kurose, and Donald Towsley. Measurement and modelling of the temporal dependence in packet loss. *Proc. of the IEEE INFOCOM*, 1999.