The Effect of Delay Mismatch in MPLS Networks Using 1+1 Protection

David Griffith, SuKyoung Lee
National Institute of Standards and Technology (NIST)
100 Bureau Drive, Stop 8920
Gaithersburg, MD 20899-8920
Email: {david.griffith, sukyoung}@nist.gov

Liliya Krivulina
Santa Monica College
1900 Pico Boulevard
Santa Monica, CA 90405-1628
Email: liliya@smc.edu

***Abstract*— High-capacity optical-fiber backbone networks protect information flows belonging to their premium customers by routing two copies of the customer's data over disjoint paths. This scheme, known as 1+1 protection, ensures that the customer will experience no service interruptions even if a fiber cut occurs somewhere in the network. A protection scheme based on this concept was recently proposed for Multi-Protocol Label Switched (MPLS) packet flows. This proposal requires the MPLS routers located at the ingress and egress edges of the MPLS network to protect certain data flows by creating two disjoint label switched paths (LSPs). This scheme allows data to flow even if a link failure occurs on one of the LSPs. There is a design issue related to the delays associated with the two LSP; a sufficiently large difference in the propagation delays can cause performance degradations that may result in an unsatisfactory quality of service (QoS) on the protected flow. In this paper we examine the impact of delay mismatch on restoration performance, probability of packet loss, and packet jitter, and we show that these metrics are adversely affected by large LSP delay differences.**

***Index Terms*—MPLS, 1+1 Protection , Quality of Service (QoS)**

I. Introduction

MPLS and Generalized MPLS (GMPLS) provide a common control plane over many types of transport networks, including optical networks, which creates new requirements for protection architectures [1]. During the past several years proposals have been made in [2], [3], [4] and [5] to incorporate optical restoration mechanisms into MPLS. These mechanisms are an extension of automatic protection switching (APS) principles from Synchronous Optical Network (SONET) ring networks to the more general mesh topologies that are being deployed in current-generation optical transport networks (OTNs). They have been discussed extensively in the literature; a summary appears in [6]. Optical protection mechanisms create dedicated backup lightpaths that are disjoint from the working lightpath that normally carries the protected data flow. Optical 1+1 protection reserves resources on two disjoint lightpaths and sends duplicate data streams over both lightpaths. These lightpaths share common termination nodes. The node closest to the data source, from which the two lightpaths diverge, is known as the ingress node, while the node where the two lightpaths merge is the egress node. The egress node selects one lightpath's data to forward to the destination based on the measured optical signal to noise ratio. If one lightpath fails, the egress node is able to switch over to the other lightpath nearly instantaneously.

This research was partially supported by the program, "SURFing the Information Technology Lab: A NIST-NSF Partnership," under Agreement #EIA-0097873.

A variation on the 1+1 concept for optical networks has recently been proposed for networks using MPLS [7]. Unlike the traditional 1+1 concept in the transport world, the approach uses duplicate paths at the MPLS layer, rather than at the optical layer. In terms of restoration time, the 1+1 scheme has a significant advantage [8] over soft protection reservation schemes. However, this method requires the network operator to resolve performance degradation issues due to variations in the delay between the two paths. The challenge is to design a appropriate restoration strategy that synchronizes the two paths. In this paper, we investigate the effect of the delay mismatch between the two paths and discuss mechanisms for improving the performance of networks using MPLS 1+1 protection.

The remainder of this paper is organized as follows. In Section II we review the existing work on 1+1 protection at the MPLS layer. In Section III we discuss some of the effects of delay mismatch on the restoration performance of the MPLS 1+1 system, and describe some of the resulting design issues. In Section IV we examine the effect of delay mismatch on QoS parameters, specifically packet jitter and the probability of packet loss. We also examine simulation results that illustrate the impact of delay mismatch on jitter and packet loss rate. We summarize the discussion in this paper in Section V.

II. MPLS 1+1 Protection

The proposal in [7] discusses the basic design principles of MPLS 1+1 protection. 1+1 protection in the transport layer duplicates traffic on two label switched paths that respectively split and merge at ingress and egress Label Switching Routers (LSRs), as shown in Fig. 1. The ingress node is responsible for duplicating packets that are received from the flow source, assigning sequence numbers to them, and sending one copy downstream on each of the two protection LSPs. The egress node is responsible for filtering the two received streams so that only one copy of each packet is forwarded to the flow's destination. This approach is simple to manage and provides fast end-to-end protection. Furthermore, it fills a gap that cannot be covered by either Interior Gateway Protocol (IGP)-rerouting which is very slow or MPLS Fast Rerouting (FRR) which does not provide end-to-end protection.

The MPLS 1+1 scheme proposed in [7] treats both LSPs (e.g. LSP_0 and LSP_1 in Fig. 1) as working paths while traditional MPLS protection [2] designates LSP_0 (or LSP_1) and LSP_1 (or LSP_0) as working and protection paths, respectively. Because the MPLS 1+1 scheme provides a packet level protection service, packets should be buffered to temporally align

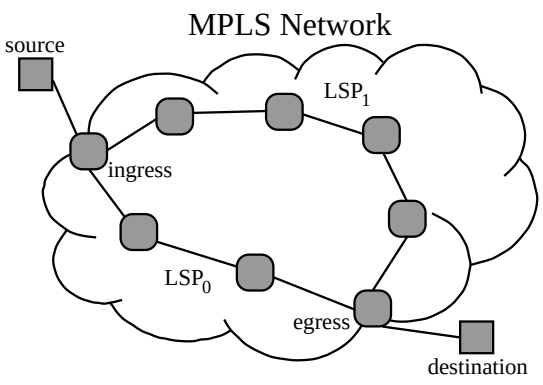

Fig. 1. MPLS 1+1 protection across an MPLS cloud. This provides for continued service in the event of a failure on one of the LSPs.

the two LSPs and compensate for variations in delay between the two paths. The goal is to buffer both paths such that the path that is leading (i.e. whose packets tend to be received by the egress first as defined in [7]) has the same delay as the path that is lagging. In addition, routing algorithms for choosing multiple non-overlapping paths (e.g. the heuristic created by Bhandari [9]) can be modified to choose two LSPs so that the expected propagation delays on the two paths are as close as possible.

The packet selection scheme at the egress is carried out based on the packet sequence number, which is contained in the MPLS shim header, and on the status of a sliding receive window maintained at the egress. Packets are accepted or rejected by the egress LSR based on whether their sequence numbers fall within the range defined by the window at the time of their arrival. If a packet is accepted, the window is adjusted so that its lower limit is one greater than the sequence number of the accepted packet. The operation of the window can be seen in Fig. 2 for the case where the first packet has sequence number 1. In [7], the authors describe several constraints on the range, L, of the window. For instance, L must be large enough so that it is greater than the longest likely burst of lost packets on either LSP, so that the packet sequence numbers do not fall outside the window range and result in all data being lost until the sequence numbers wrap and reenter the window's range from below.

III. Effect of Delay Mismatch on Restoration Performance

The behavior of MPLS 1+1 protection during LSP outage events was discussed in [7], which noted that there will be a delay in new packet arrivals if a failure occurs on the leading (i.e. less delayed) LSP. In this section we quantify this behavior using a simple, deterministic system model. In this model, we assume that the packet transmission rate g is the same on the two LSPs. Packets are uniformly spaced along the time axis on each LSP, with an inter-packet spacing of $1/g$ seconds. In addition, we assume without loss of generality that the fixed propagation delays D_0 and D_1 on LSP_0 and LSP_1, respectively, are related as $D_0 < D_1$. The arrivals of packets on the two LSPs are shown in Fig. 3. The arrival time of the n^{th} packet on LSP_i is $D_i + n/g$. From the figure, we see that D_1 must fall

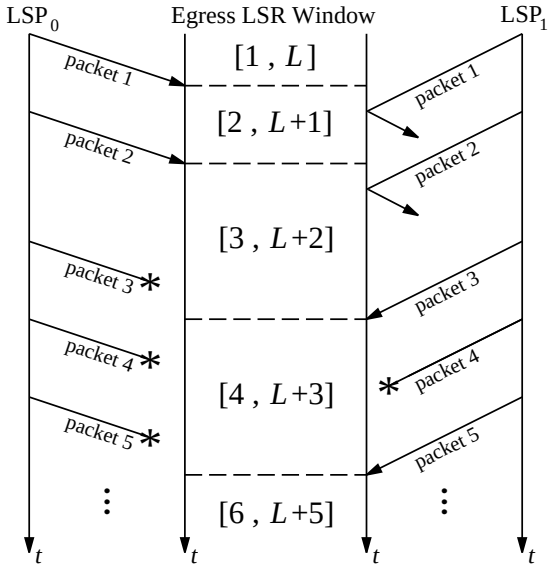

Fig. 2. An illustration of the windowing function used at the egress LSR in a MPLS 1+1 protection system. The window has length L and is adjusted upon each receipt of a packet whose sequence number lies within its range.

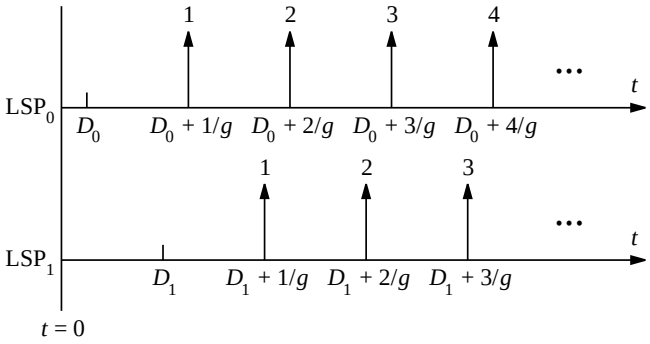

Fig. 3. Deterministic MPLS 1+1 system with equal line rates on the two LSPs. Sequence numbers start at 1 in this example.

between the arrival times of two packets on LSP_0. This gives us

$$D_0 + \frac{k}{g} < D_1 < D_0 + \frac{k+1}{g} \tag{1}$$

for some non-negative integer k, where $k = \lfloor g(D_1 - D_0) \rfloor$. Under these assumptions, the n^{th} packet arrival on LSP_1 at the egress occurs between the arrival times for the $(k+n)^{\text{th}}$ and $(k+n+1)^{\text{th}}$ packets on LSP_0.

If LSP_1 fails, there is no effect on the packet stream at the egress, because every packet passed downstream by the egress node is pulled from LSP_0 in this model. If a failure occurs on LSP_0 at time $t = T$, there will be a delay between the last packet received on LSP_0 before the failure event (call this packet n, where $n = \lfloor g(T - D_0) \rfloor$) and the first packet received on LSP_1 and forwarded downstream by the egress. Assuming that no packets are lost on LSP_1 after the failure event, the first packet received by the egress LSR from LSP_1 after the failure of LSP_0 is packet $n - k$. This packet is discarded by the egress LSR because the egress sequence number window cov-

ers the range $[n+1, n+L]$ after the receipt of packet n from LSP_0. The egress continues to discard packets until it receives packet $n+1$ from LSP_1; the egress will discard a total of $k+1$ packets from LSP_1 between the failure of LSP_0 and the resumption of traffic using packets from LSP_1. The time between the arrival of packet n on LSP_0 and the arrival of packet $n-k$ on LSP_1 is $D_1 - D_0 - k/g$. The time gap between the arrival of packet $n-k$ on LSP_1 and the arrival of packet $n+1$ on LSP_1 is $(k+1)/g$. Thus the total time lag between the arrival of the last packet on LSP_0 and the arrival of a packet on LSP_1 that is forwarded downstream is $(D_1 - D_0) + 1/g$.

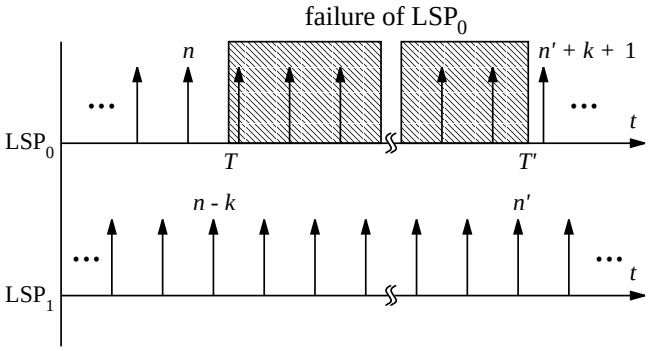

Fig. 4. Illustration of failure event on LSP_0, beginning at time $t = T$ and ending at time $t = T'$.

When service on LSP_0 is restored, a packet will arrive at the egress from LSP_0 after the receipt of packet n' on LSP_1. From (1), the first packet received on LSP_0 after restoration of service will be packet $n' + k + 1$. When this packet is received, the window range is $[n'+1, n'+L]$. There are two possible outcomes. If $k \leq L-1$, then the packet from LSP_0 is accepted and the window advances so that it covers the range $[n'+k+2, n'+k+L+1]$. Subsequently received packets from LSP_1 will be dropped by the egress; thus, k packets have been lost. If $k \geq L$, then packet $n'+k+1$ from LSP_0 is rejected by the egress and packet $n'+1$ is received from LSP_1 and passed downstream. The egress will continue to forward packets from LSP_1 and reject packets from LSP_0, even though the packets from LSP_1 are arriving later than their copies that were forwarded over LSP_0. This has serious consequences in the event that a failure occurs on LSP_1, for in that case no packets will be received from LSP_1 while the egress continues to reject packets from LSP_0 because their sequence numbers lie outside the range of the egress' receive window. If the transmission continues for a sufficiently long period of time, the situation will be resolved by the wrapping of the sequence numbers so that packets are accepted from LSP_0, but this may involve the loss of a considerable amount of data, unless some recovery management scheme is used, as noted in [7].

IV. Effect of Delay Mismatch on QoS

A. Effect on Packet Jitter

Instantaneous packet jitter is defined in [10] using the difference in the delay times of two sequentially received packets as measured by the receiving node. If S_i is the time at which packet i was sent and R_i is the time when packet i is received, then the delay difference between packets i and j is

$$\begin{aligned} D_{i,j} &= (R_j - R_i) - (S_j - S_i) \\ &= (R_j - S_j) - (R_i - S_i). \end{aligned} \tag{2}$$

Jitter is measured using an adaptive process in which the measured delay difference between sequentially received packets is the forcing function. The adaptation function for the jitter measurement is

$$J_n = J_{n-1} + \frac{|D_{n-1,n}| - J_{n-1}}{16}. \tag{3}$$

If the jitter process $\{J_n\}$ and the delay difference process $\{D_{n-1,n}\}$ are stationary, then the expected jitter can be found using the expected delay difference between sequentially received packets at the egress node. In the deterministic model, the delay difference between packets received on a given LSP is zero. One obtains non-zero jitter measurements in the deterministic system due to random packet losses on each LSP or due to LSP failures, in which case the changes in jitter are transient in nature.

In the case of random packet losses, the expected delay variation between packets is

$$\mathsf{E}\{D_{n-1,n}\} = (D_1 - D_0)p_{\text{diff}}, \tag{4}$$

where p_{diff} is the probability that the $(n-1)^{\text{th}}$ and n^{th} packets forwarded by the egress LSR were received from different LSPs.

To compute p_{diff}, we condition on which LSP produced the $(n-1)^{\text{th}}$ packet forwarded by the egress LSP, giving

$$p_{\text{diff}} = p_{n-1,n}(0,1)p_{n-1}(0) + p_{n-1,n}(1,0)p_{n-1}(1), \tag{5}$$

where $p_{m,n}(i,j)$ is the conditional probability that the n^{th} packet forwarded by the egress LSR came from LSP_i given that the m^{th} packet came from LSP_j, and $p_n(i)$ is the probability that the n^{th} packet forwarded by the egress LSR came from LSP_i.

We assume packets are dropped independently on LSP_i with probability p_i. Because $D_0 < D_1$, every packet that appears on LSP_0 is forwarded by the egress; thus $p_{n-1}(0) = 1 - p_0$. Suppose that the $(n-1)^{\text{th}}$ packet forwarded by the egress LSR is the m^{th} packet transmitted from the ingress LSR over LSP_0. This packet arrived at the egress LSR at time $t = D_0 + m/g$. The n^{th} packet transmitted downstream by the egress LSR arrived at the egress from LSP_1 at time $t = D_1 + (m+\ell)/g$, where $\ell = 1, 2, \ldots$. From (1), it follows that if no packets are lost, packet $k+m+\ell$ will arrive at the egress from LSP_0 before packet $m+\ell$ arrives on LSP_1. So for packet $m+\ell$ to be selected by the egress from LSP_0, that packet must not be lost, while packets $m+1, m+2, \ldots, m+k+\ell$ must be lost on LSP_0 and packets $m+1, m+2, \ldots, m+\ell-1$ must be lost on LSP_1. The probability of this event is $(1-p_1)p_0^{k+\ell}p_1^{\ell-1}$. The total probability $p_{n-1,n}(0,1)$ is therefore

$$\begin{aligned} p_{n-1,n}(0,1) &= \sum_{\ell=1}^{\infty}(1-p_1)p_0^{k+\ell}p_1^{\ell-1} \\ &= \frac{(1-p_1)p_0^{k+1}}{1-p_0p_1}. \end{aligned} \tag{6}$$

The probability that a packet with sequence number n is accepted from LSP_1 is the probability that packet n was lost on LSP_0, along with the packets $n+1, n+2, \ldots, n+k$. This event occurs with probability $p_{n-1}(1) = p_0^{k+1}(1-p_1)$. Given that packet n was accepted from LSP_1, the probability that the next accepted packet is from $\text{LSP}_0$and has sequence number $n+k+\ell$ is the probability that packets $n+1, n+2, \ldots, n+\ell-1$ are lost on LSP_1 and packets $n+k+1, n+k+2, \ldots, n+k+\ell-1$ are lost on LSP_0, while packet $n+k+\ell$ is not lost. For a particular value of ℓ, this event occurs with probability $(1-p_0)(p_0p_1)^{\ell-1}$. If $\ell > L-k$, packet $n+k+\ell$ from LSP_0 will fall outside the range of the sliding window and be rejected, along with any subsequent packets from LSP_0 (until the window values wrap). The conditional probability therefore is

$$\begin{aligned} p_{n-1,n}(1,0) &= \sum_{\ell=1}^{L-k}(1-p_0)(p_0p_1)^{\ell-1} \\ &= \frac{(1-p_0)(1-(p_0p_1)^{L-k})}{1-p_0p_1}. \end{aligned} \quad (7)$$

If $L \gg k$, this can be approximated as $(1-p_0)/(1-p_0p_1)$. Thus, we have

$$\mathsf{E}\{D_{n-1,n}\} = \frac{2(1-p_1)(1-p_0)}{1-p_0p_1}(D_1-D_0)p_0^{\lceil g(D_1-D_0)\rceil}. \quad (8)$$

A plot of the normalized expected jitter $g\mathsf{E}\{J\}$, which is measured in packet intervals, that is introduced into the deterministic arrivals system by packet error is given in Fig. 5 for the case where $p_0 = p_1$. The jitter is plotted versus packet loss probability for three values of normalized delay offset, $g(D_1 - D_0)$. The peak jitter occurs when $p = 0.4$, and does not exceed a single packet period. Given that the jitter on each LSP is zero, reflecting the behavior of an idealized constant bit rate stream, the average jitter may be unacceptable, even for relatively low packet loss probabilities. In addition, these curves depict only average jitter values. Localized events can cause large increases in jitter, which can result in the loss of packets if they are judged to be in violation of their flow's traffic contract by an admission controller in the network.

We show the impact of delay difference on jitter in non-deterministic MPLS 1+1 systems in Fig. 6. We simulated a MPLS 1+1 system in which the packet streams arriving at the egress LSR were Poisson processes with the same mean packet interarrival time. The packet loss probability was $p = 10^{-6}$ on both LSPs in all three cases. The jitter was computed over runs of 5000 packets each, and the curves are ensemble averages taken over 100 runs. The sliding window had length 100. In Fig. 6(a), the delay difference is on the order of a single packet period, and the average jitter is very close to the mean packet interarrival time, as we would expect. Increasing the delay difference to 10 average packet intervals produced a greater initial overshoot, slower convergence of $\{gJ_n\}$, and a long-term offset of approximately 10% from the average packet interarrival time, as shown in Fig. 6(b). In Fig. 6(c), the delay difference is 100 packet intervals. Because of the long delay and the low packet loss probability, we do not see the impact of the delay difference until the 500^{th} packet. After this point, the jitter

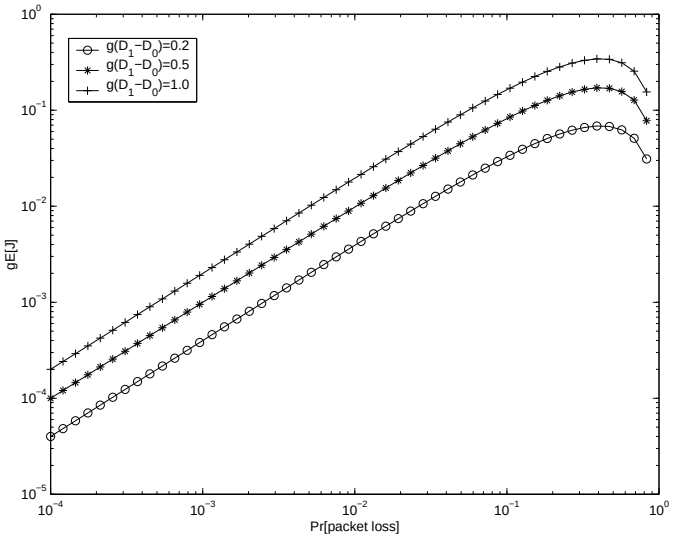

Fig. 5. Normalized jitter in a MPLS 1+1 system with deterministic packet arrival times.

curve becomes very noisy, approaching three times the mean packet interarrival time at places. Reducing the size of the sliding window helps only in cases where the delay difference is very large. For $g(D_1 - D_0) = 200$, we found that if $L < 40$ the jitter curve is well behaved, but setting $L = 4$ for the case where $g(D_1 - D_0) = 100$ did not eliminate the noise. Thus buffering at the ingress seems to be the best solution.

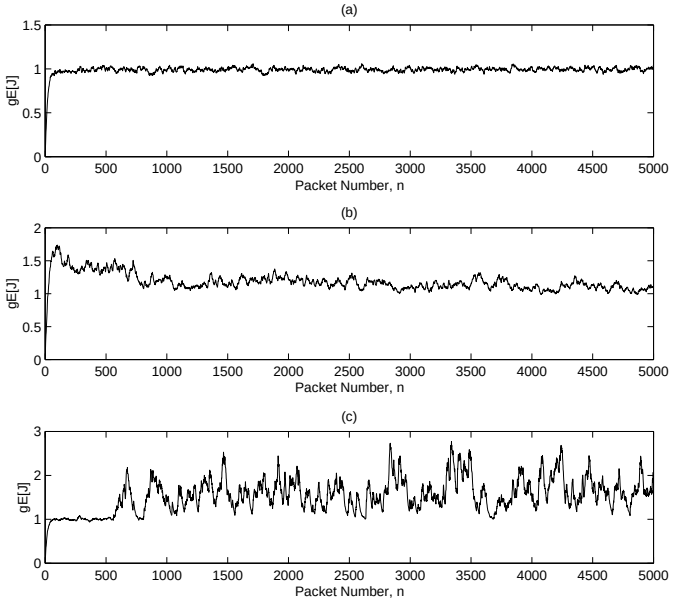

Fig. 6. Plots of average normalized jitter in a MPLS 1+1 system with random packet loss and exponential packet interarrival times. **(a)**: $g(D_1 - D_0) = 1$, **(b)**: $g(D_1 - D_0) = 10$ **(a)**: $g(D_1 - D_0) = 100$

B. Effect on Packet Loss Probability

By sending duplicate copies of each packet over disjoint paths, MPLS 1+1 protection allows for a considerable reduction in the packet loss rate, in addition to providing a method

for rapidly recovering from failure on either of the LSPs. In situations where the delay and transmission rates of the two paths are closely matched, the net packet loss rate can trivially be shown to be $p_0 p_1$, where p_0 and p_1 are the loss rates on LSP_0 and LSP_1, respectively. However, if there is a significant difference in the propagation delays associated with the two LSPs, then the probability of packet loss can actually be higher, due to the existence of an additional packet loss mechanism that we analyze here. A packet will be lost if each copy of it is dropped in transit. Packets can also be lost if one copy is dropped in transit and the other copy is rejected by the egress LSR because its sequence number lies outside the range defined by the sliding window. This will happen if additional packets arrive on the LSP that dropped the packet, advancing the window, before the undropped copy arrives from the other LSP.

Using the deterministic model that we introduced in Section III with independent packet losses on each LSP, we find by conditioning on the LSP packet loss events that

$$\begin{aligned}\mathsf{Pr}\{\text{loss}\} &= p_1 p_0 + p_0(1-p_1)\mathsf{Pr}\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\} \\ &\quad + p_1(1-p_0)\mathsf{Pr}\left\{\text{loss}|\overline{\mathcal{L}_0} \cap \mathcal{L}_1\right\} \qquad (9)\end{aligned}$$

where the events $\mathcal{L}_0$ and $\mathcal{L}_1$ occur when a packet is lost on LSP_0 and LSP_1, respectively. Now, $\mathsf{Pr}\left\{\text{loss}|\overline{\mathcal{L}_0} \cap \mathcal{L}_1\right\} = 0$ because any packet that arrives from LSP_0 in this model will not be discarded; it always appears before its counterpart arriving from LSP_1. Letting the packet of interest have sequence number n, we see that $\mathsf{Pr}\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\}$ is the probability that there is at least one successfully received packet from LSP_0 before time $t = D_1 + n/g$, the time when the copy of packet n arrives from LSP_1. From (1), we see that

$$D_0 + \frac{k+n}{g} < D_1 + \frac{n}{g} < D_0 + \frac{k+n+1}{g}. \qquad (10)$$

The only way for the window not to advance so that sequence number n is out of range is for the packets with sequence numbers $n+1, n+2, \ldots, n+k$ to be dropped by LSP_0. This will occur with probability $\mathsf{Pr}\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\} = 1 - p_0^k$. Thus the probability of packet loss at the egress LSR is

$$\mathsf{Pr}\{\text{loss}\} = p_0 - (1-p_1)p_0^{k+1}, \qquad (11)$$

where $k = \lfloor g(D_1 - D_0) \rfloor$ is the relative offset in packets of the streams on the two LSPs.

In Fig. 7 we plot $\mathsf{Pr}\{\text{loss}\}$ for the cases $k = 0$ and $k = 1$, for the case where the loss probabilities on the two LSPs are equal. For $k = 0$, which occurs when the delay difference between LSPs is less than one packet interval, the loss probability is just p^2, the probability that both copies of a packet are lost. Once the delay difference exceeds one packet period, the loss behavior of the system approaches that of the leading LSP, which is LSP_0 in this case. This demonstrates the importance of using buffering to control packet loss. In addition, it is clear that the two LSPs should be routed, if possible, so that the LSP with the higher loss rate is also the one with the longer average packet delay.

V. Summary

In this paper we examined the MPLS 1+1 protection scheme that was introduced in [7] and examined the impact of mismatch between the propagation delays of the two LSPs used in the scheme. Using the qualitative discussion some of the effects of delay mismatch on restoration performance in [7], we computed the gap length and number of packets lost during the failure of the leading LSP in a simple system. We examined the effect of delay mismatch on jitter and demonstrated that delay mismatch can introduce considerable levels of noise into the measured packet jitter, even when the jitter on the individual LSPs is small. We also developed a theoretical model of packet loss performance and showed that even small delay offsets can eliminate any packet loss probability reduction that the system gains from using duplicated packets. The best solution to these problems appears to be using constrained routing to reduce mismatch and buffering the leading path at the ingress, rather than shortening the sliding window at the egress.

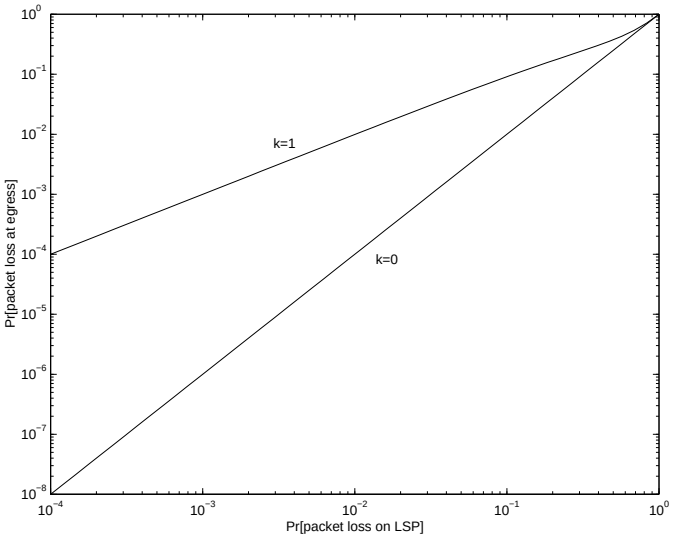

Fig. 7. Packet loss probability for a MPLS 1+1 system with deterministic packet arrival times, where $k = \lfloor g(D_1 - D_0) \rfloor$ is a measure of the relative temporal offset of the data streams on the two LSPs.

References

[1] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y. Rekhter, "Generalized multiprotocol label switching: An overview of routing and management enhancements," *IEEE Communications Magazine*, vol. 39, no. 1, pp. 144–150, January 2001.

[2] V. Sharma and F. Hellstrand, "Framework for MPLS-based recovery," IETF Internet Draft, 2002.

[3] E. Mannie et al., "Recovery (protection and restoration) terminology for GMPLS," IETF Internet Draft, 2002.

[4] C. Huang, V. Sharma, K. Owens, and S. Makam, "Building reliable MPLS networks using a path protection mechanism," *IEEE Communications Magazine*, vol. 40, no. 3, pp. 156 – 162, 2002.

[5] R. Bartos and M. Raman, "A heuristic approach to service restoration in MPLS networks," in *Proceedings of the IEEE International Conference on Communications 2001 (ICC 2001)*, 2001, pp. 117–121.

[6] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE Network*, vol. 14, no. 6, pp. 16–23, Nov–Dec 2000.

[7] R. Nagarajan, M. A. Qureshi, and Y. T. Wang, "A packet 1+1 path protection service for MPLS networks," IETF Internet Draft, March 2002.

[8] Y.A. Kim, J.K. Choi, and S.G. Jong, "Analysis of the end-to-end recovery algorithm in the IP over WDM network," in *Proceedings of the International Conference on Optical Internet (COIN)*, 2002.

[9] R. Bhandari, "Optimal physical diversity algorithms and survivable networks," in *Proceedings of the Second IEEE Symposium on Computers and Communications, 1997*, 1997, pp. 433–441.

[10] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," IETF RFC 1889, January 1996.