Telecommunication Systems 19:3,4, 481–495, 2002

An Efficient Rerouting Scheme for MPLS-Based Recovery and Its Performance Evaluation

GAEIL AHN * and JONGSOO JANG fogone@etri.re.kr
Network Security Department, Electronics and Telecommunication Research Institute (ETRI), 161 Kajong-dong, Yusong-gu,Taejon, 305-350, Korea

WOOJIK CHUN
Department of Computer Engineering, Chungnam National University, 220 Goong-Dong, Yusong-Gu,Taejon, 305-764, Korea

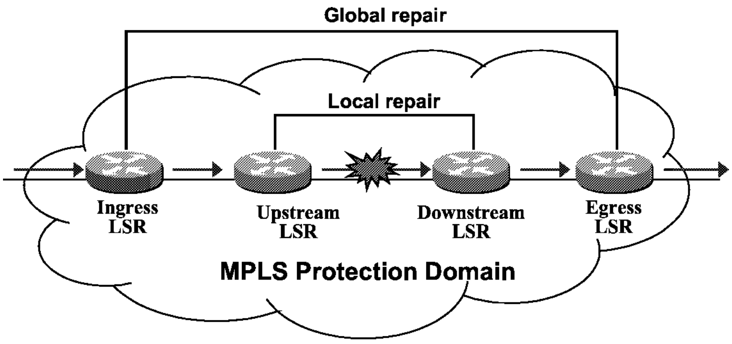
Abstract. The path recovery in MPLS is the technique to reroute traffic around a failure or congestion in a LSP. Currently, there are two kinds of model for path recovery: rerouting and protection switching. The existing schemes based on rerouting model have the disadvantage of more difficulty in handling node failures or concurrent node faults. Similarly, the existing schemes based on protection switching model have some difficulty in solving problem such as resource utilization and protection of recovery path. This paper proposes an efficient rerouting scheme to establish a LSP along the least-cost recovery path of all possible alternative paths that can be found on a working path, which is calculated by the upstream LSR that has detected a failure. The proposed scheme can increase resource utilization, establish a recovery path relatively fast, support almost all failure types such as link failures, node failures, failures on both a working path and its recovery path, and concurrent faults. Through simulation, the performance of the proposed scheme is measured and compared with the existing schemes.

Keywords: MPLS, path recovery, rerouting model, least-cost recovery, path, protection merging LSR, protection switching

1. Introduction

The explosive growth and the advent of sophisticated services of Internet have reduced the actual carried traffic and service quality. So IETF (Internet Engineering Task Force) is developing Multi-Protocol Label Switching (MPLS), which combines flexibility of IP routing and efficiency of link-level switching [Awduche, 5; Rosen et al., 14]. MPLS router, called Label Switching Router (LSR), enables to forward IP packet directly to a next hop by using label inserted in the packet as index into table that specifies the next hop. Label Distribution Protocol (LDP) [Andersson et al., 3] is defined to distribute labels and to establish Label Switched Path (LSP). MPLS also provide capability that can establish a Constraint-based Routed LSP (CR-LSP), which is Explicit-Routed Label Switched Path (ER-LSP) based on QoS. For it, CR-LDP (Constraint-based Routing

* Corresponding author.



	Local repair	Global repair
Preferable Recovery Model	Rerouting	Protection Switching

Figure 1. Global repair vs. local repair.

LDP) [Jamoussi et al., 10] and RSVP (ReSerVation Protocol)-TE [Awduche et al., 6] was proposed.

The ability to protect traffic around a link/node failure or congestion in a LSP can be important in mission critical MPLS networks. The path recovery is to reroute traffic around a failure or congestion in a LSP; that is, packets will be followed along recovery path in case of working path failure [Sharma et al., 15].

There are two basic models for path recovery: rerouting and protection switching. Rerouting is a model that establishes a recovery path after a failure on its working path. Protection switching is a model that establishes a recovery path prior to any failure on the working path. According to how the repairs are affected upon the occurrence of a failure on the working path, there are two ways: global repair and local repair as shown in figure 1. In global repair, protection is always activated on end-to-end basis, irrespective of where a failure occurs. But in local repair, protection is activated by each LSR that has detected a failure.

Rerouting model is very good in resource utilization. But it has the disadvantage that requires longer times or even fails in establishing a recovery path. Because of the problem, most of the existing schemes based on rerouting model considers local repair [Chen and Oh, 7; Yoon et al., 18]. The intent of local repair is to protect against a single link fault. So those schemes has the disadvantage of difficulty in handling node failures or concurrent node faults. This is why popular approach in MPLS fault recovery is global repair, which is used to protect against any link or node fault on the entire path or on a segment of a path.

In global repair, protection switching is the preferred approach because the setup overhead and delay may be greater disadvantages when longer alternative paths are sought. So almost every existing schemes based on global repair adopts protection switching [Haskin and Krishnan, 8; Huang et al., 9]. However, those schemes have some difficulty in solving problem such as resource utilization and both working and recovery path failures. To solve the first problem, [Kini et al., 11] has proposed n-to-1 protection that n working paths are protected using only one recovery path. And also [Mo, 13] has proposed the efficient resource allocation scheme, in which the resource of recovery LSPs is allocated by considering that of working LSPs. Incase of the second and last problem, it is very difficult to find a solution for it because the route of recovery path in protection switching is fixed.

This paper proposes an efficient rerouting mechanism to establish a LSP along the least-cost recovery path of all possible alternative paths that can be found on a working path, which is calculated by a node that has detected a failure. The proposed scheme can establish an alternative LSP quickly without requiring longer times, increase resource utilization, and support almost all failure types such as link failures, node failures, failures on both a working path and its recovery path, and concurrent faults.

The remainder of this paper is organized as follows. The following section overviews and analyzes the existing path recovery schemes. We then explain our scheme and evaluate its performance. The last section contains the concluding remarks.

2. MPLS recovery schemes

In MPLS recovery, working path is the protected path that carries traffic before the occurrence of a fault. And recovery path is the path by which traffic is restored after the occurrence of a fault.

In this paper, we introduce two schemes using rerouting model, which are proposed by Chen and Oh [7] and Yoon et al. [18], respectively, and also two schemes using protection switching, which are proposed by Haskin and Krishnan [8] and Huang et al. [9], respectively. At first, Chen and Oh have proposed a scheme, in which recovery paths are pre-established between links on working path without reserving resources. When a link failure occurs, the downstream LSR that has detected it sends a notification message (e.g., CR-LDP notification message) to its upstream LSR to check and reserve resources and to notify to the upstream LSR. Yoon et al. have proposed an efficient pre-qualified recovery mechanism. In the scheme, the upstream LSR can choose an optimal path up to its downstream LSR by re-calculating the path whenever network status is changed. When a link failure occurs, the upstream LSR that has detected it establishes a recovery path along the pre-calculated optimal path. The two approaches have the advantages of resource utilization and recovery speed because they use rerouting model and local repair. But, they have the disadvantage of big difficulty in handling node failures. This is why schemes that use protection switching model and global repair are popular.

The next is Haskin's scheme and Makam's scheme, which use protection switching model and global repair. Figure 2 shows Haskin's scheme and Makam's scheme. In figure 2, the straight line between LSR1 and LSR9 is the working path. In Haskin scheme, a recovery path is established as follows:

1. The initial segment of the recovery path is established between PEL (Protection Egress LSR) and PIL (Protection Ingress LSR) in the reverse direction of the working

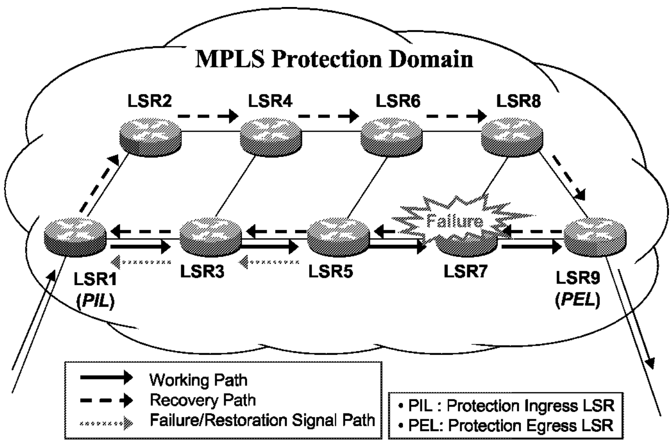


Figure 2. Path recovery: Haskin's scheme and Makam's scheme.

path. In figure 2, the dashed line between LSR9 and LSR1 illustrates such a segment of the recovery path.

2. The second and final segment of the recovery path is established between PIL and PEL along a transmission path that does not utilize any working path.
3. The initial and final segments of the recovery path are linked to form an entire recovery path.
4. When a failure occurs, the node that has detected it reroutes incoming traffic by linking the upstream portion of the working path to the downstream portion of the recovery path. In figure 2, when the node LSR7 fails, the working traffic is rerouted along the recovery path LSR5–3–1–2–4–6–8–9.

The merit of Haskin's scheme is that there is almost no packet loss during link/node failure. However, Haskin's scheme introduces re-ordering of packets in case that traffic is switched back from the recovery path to the working path after a link/node goes up.

In Makam's scheme, a recovery path is established as follows:

1. A recovery path is established between PIL and PEL along a transmission path that does not utilize any working path.
2. When a failure occurs, the node that has detected the failure sends a failure notification message toward its PIL. On receiving the message, PIL reroutes the incoming traffic through the recovery path. In figure 2, when the node LSR7 fails, the working traffic is rerouted along the recovery path LSR1–2–4–6–8–9.

The merit of Makam's scheme is that there is almost no problem in reordering of packets during link/node failure. However, Makam's scheme introduces packet loss because PIL does not execute the protection switching until it receives the failure notification message from a node that has detected a link/node failure.

Both Haskim's and Makam's schemes are poor in resource utilization and cannot support protection in case of link/node failures on both a working path and its recovery path. This is because they use the protection switching model.

3. Rerouting scheme using least-cost recovery path

In this paper, we define a term, Candidate-PML, which is LSR on working path that can be used as PML (Protection Merging LSR). For example, in figure 2 Candidate-PMLs of LSR3 are LSR5, LSR7, and LSR9. Similarly, Candidate-PMLs of LSR5 are LSR7 and LSR9. The Candidate-PMLs' address information is easily found in MPLS signaling message such as PV (Path Vector) or ER (Explicit Route) of LDP and RRO (Record Route Object) or ERO (Explicit Route Object) of RSVP when the working path is established.

The proposed scheme uses rerouting model and local repair. So, the recovery path is established by the upstream LSR after a fault occurrence on a working path. If the failure is restored, the established recovery path is released. Figure 3 illustrates the procedure of restoration function briefly. A recovery path is established as follows:

1. The upstream LSR that has detected a failure calculates the least-cost path of all possible alternative paths between itself and each Candidate-PML. As the result, the upstream LSR can know the PML of the least-cost path and the explicit route up to the PML.
2. A recovery path is established along the calculated explicit route from the upstream LSR to the PML. In the recovery path setup, the explicit route is inserted into the ER (Explicit Route) of MPLS signaling message (e.g., CR-LDP, RSVP). And LSPID (LSP Identifier) of the working LSP is used as an ER hop for the purpose of splicing the existing working LSP and its new recovery LSP to be established. The holding priority of the working path may be used as the setup priority of the recovery path.

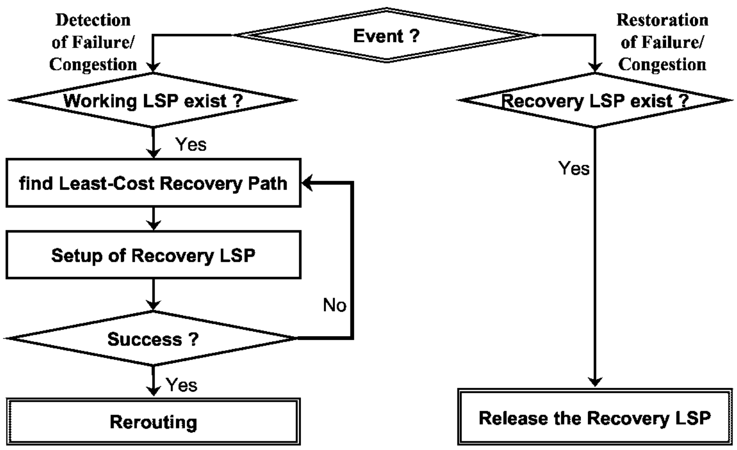


Figure 3. Procedure of the proposed rerouting scheme.

```
procedure Least-Cost-Recovery-Path (UpstreamLSR, CandidatePMLs, TParam)
   // UpstreamLSR is a upstream LSR that has detected a failure //
   // CandidatePMLs is a set of LSRs that can be used as PML //
   // TParam is the traffic parameter value of working LSP //

   least_cost ← ∞
   least_cost_path ← nil

   for ( each PML of CandidatePMLs ) do
        explicit_route ← Get-CR-Path (UpstreamLSR, PML, TParam)
        cost ← Get-Cost (explicit_route)

        if ( cost ≤ least_cost ) then    least_cost_path ← explicit_route
                                         least_cost ← cost
                                  else   break
   end  end

   return least_cost_path
end Least-Cost-Recovery-Path
```

Figure 4. Least-Cost-Recovery-Path algorithm.

3. As soon as the recovery path is established, traffic on the working path is switched over to the recovery path.

4. If the setup of the recovery path fails, go to 1.

Figure 4 shows Least-Cost Recovery-Path algorithm, which is used to calculate the least-cost recovery path of all possible alternative paths between itself and each Candidate-PML. The function *Get-CR-Path*, which is used to find the optimal path between the upstream that has detected a failure and a Candidate-PML, can be easily implemented by using the existing algorithm such as widest-shortest path algorithm [Apostolopoulos et al., 4], shortest-widest path algorithm [Wang and Crowcroft, 17], and shortest-distance path algorithm [Ma and Steenkiste, 12]. *Get-CR-Path* is a function that is used to calculate links delay or hop count of an explicit route.

Figure 5 is an example for the proposed scheme. When LSR11 fails, LSR10 that has detected it calculates an explicit route for new recovery path. The Candidate-PMLs for the working LSP (LSPID 1000) in LSR10 are LSR11, LSR12, LSR13, and LSR14. Assuming that hop count is used in calculating the cost of recovery path, the cost from LSR10 to LSR11 is infinite, the cost to LSR12 is 4, the cost to LSR13 is 3, and the cost to LSR14 is 4. Because the cost from LSR10 to LSR13 is the least value, the explicit route of recovery path becomes LSR10–2–3–13.

The calculated explicit route is used as ER of CR-LDP Request Message. The LSPID of working path, 1000 is used as ER hop in order to splice the existing working LSP and new recovery LSP. As soon as LSR10 receives a CR-LDP Mapping Message for the recovery LSP, it reroutes traffic on working path to new recovery path. If the link between LSR10 and LSR11 fails, the working traffic is rerouted along the recovery path LSR10–2–11.

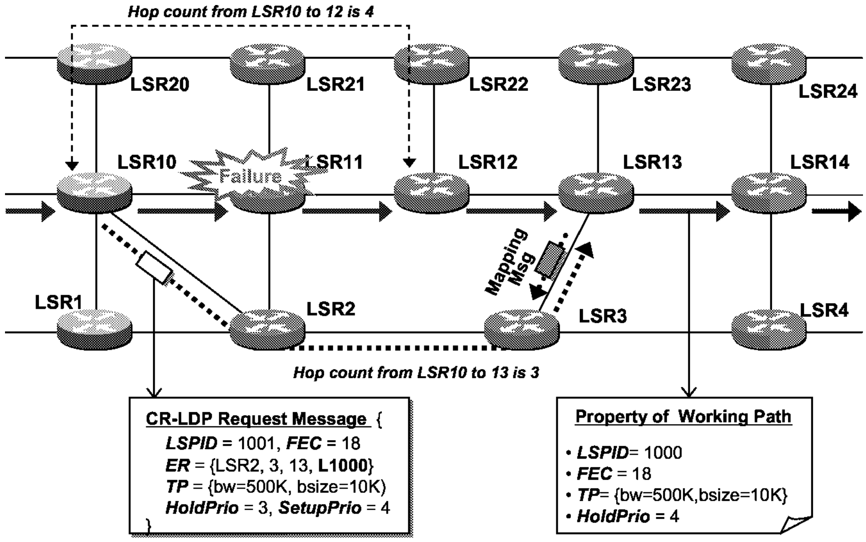


Figure 5. Example for the proposed scheme.

The proposed scheme has the following advantages:

- High resource utilization – because any resource for recovery path is not reserved unless a failure on working path occurs, and because even at the fault occurrence the recovery path with least-cost is chose, thus the least resource is allocated.
- Relatively fast path rerouting – because traffic is rerouted to the recovery path with least cost.
- Protection of recovery path – because our scheme uses rerouting model, thus the recovery path is handled in the same way as the working path.
- No modification of MPLS signaling message.

4. Simulations and performance evaluation

In order to simulate the existing and the proposed schemes, we have extended MPLS Network Simulator [Ahn and Chun, 1], which is operated on Network Simulator (ns) [16]. MPLS Network Simulator supports the setup of CR-LSP based on QoS as well as basic MPLS core functions such as label switching, LDP, CR-LDP, and various options of label distribution.

4.1. Simulation environment

Simulation environment is shown in figure 6. Node0 is IP node. The rest of the nodes are MPLS nodes. Each node is connected with a duplex link with the bandwidth 2 Mbps, a delay of 10 ms, and a CBQ queue.

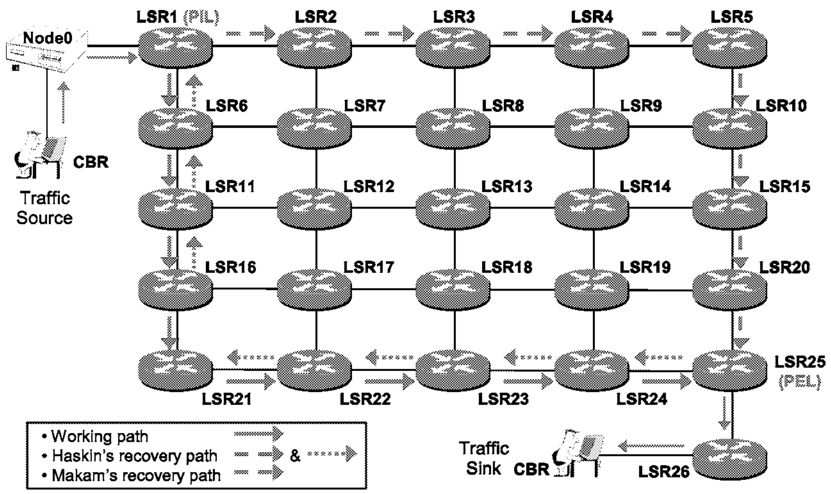


Figure 6. MPLS networks for simulation.

There is one pair of UDP traffic, called CBR. Traffic is injected into Node0 and escape from LSR26. CBR generates 256-byte-long packets at 1 Mbps and at constant bit rate. Each node is monitored every 10 ms to protect the working path.

4.2. Simulation results and evaluation

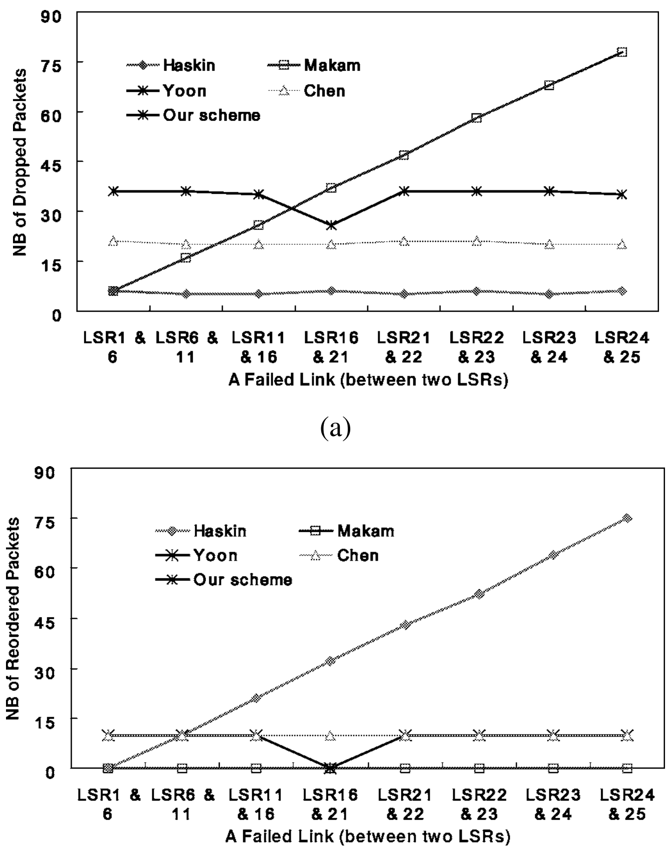
Several metrics are proposed for performance evaluation of path recovery schemes by [Andersson et al., 2]. Each scheme is measured and evaluated in terms of packet loss, reordering of packets, and resource utilization.

In this simulation, two kinds of service, best-effort service and guaranteed service, are used as service for traffic to be protected. In best-effort service, a working LSP and its recovery LSP are setup by using ER-LSP, which does not consider resource reservation but can utilize the rest bandwidth of the link. In guaranteed service, a working LSP and its recovery LSP are setup by using CR-LSP, which is guaranteed the bandwidth required but cannot utilize the rest bandwidth of the link. In this paper, we considered the following failure types:

- link failure on working path,
- node failure on working path,
- concurrent faults on working path,
- node failures on both working and recovery paths.

Figure 7 shows the performance comparison of each scheme by the location of a failed link in best-effort service. Figures 7(a) and (b) show the performance for packet loss and reordering of packets, respectively.

Haskin scheme has almost no packet loss during link failure as shown in figure 7(a). However, it introduces re-ordering of packets in case that traffic is switched back from



(a)

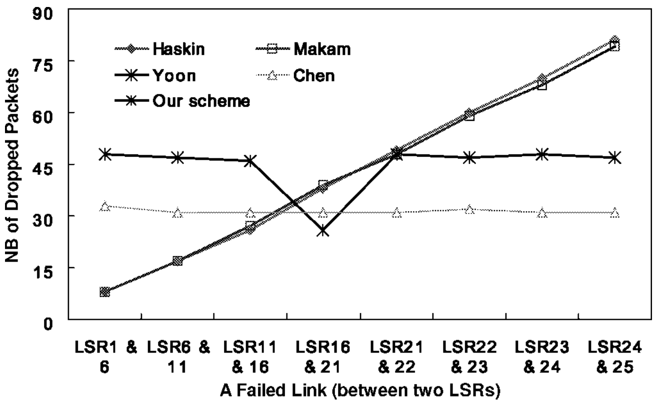
(b)

Figure 7. A link failure in best-effort service: (a) packet loss; (b) reordering of packets.

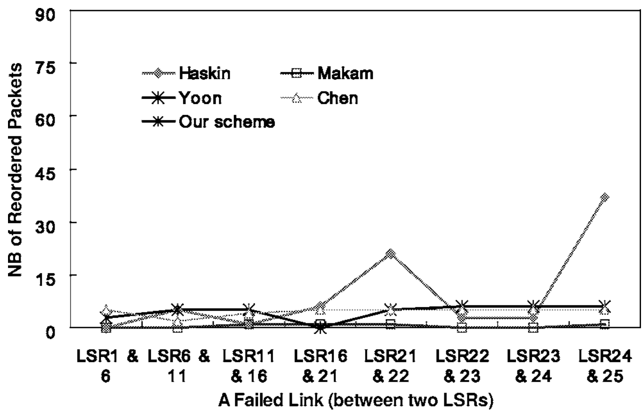
the recovery path to the working path after a link goes up. As shown in figure 7(b), the number of the reordered packets increases more in proportion to the distance between PIL (i.e., LSR1) and the LSR that has detected the link failure.

Makam's scheme is almost no problem in reordering of packets during link failure as shown in figure 7(b). However, Makam's scheme introduces packet loss because PIL cannot execute the protection switching until it receives a failure notification message from the LSR that has detected a link failure. As shown in figure 7(a), the number of the dropped packets increases more in proportion to the distance between PIL and the LSR that has detected a link failure.

The performance of the proposed scheme is better than that of Makam's scheme in packet loss and better than that of Haskin's scheme in reordering of packets as shown in figure 7. And also the proposed scheme has the advantage of almost uniform performance without connecting with failure location. Chen's scheme has better performance than the proposed scheme and Yoon's scheme has the same performance to the proposed scheme as shown in figure 7.



(a)

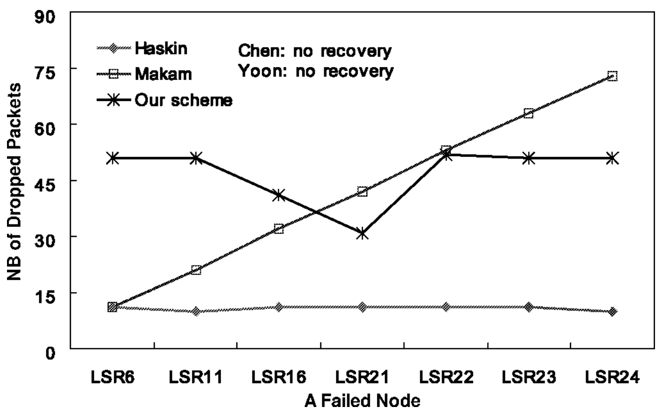


(b)

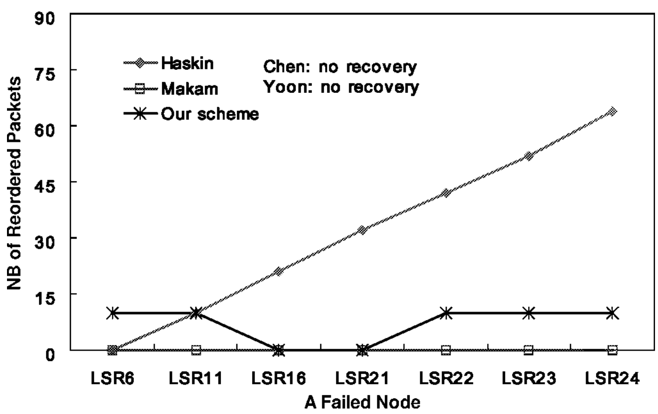
Figure 8. A link failure in guaranteed service: (a) packet loss; (b) reordering of packets.

Figure 8 shows the performance comparison of each scheme by the location of a failed link in guaranteed service. Haskin scheme introduces a certain amount of packet loss, which was not shown in the simulation of the best-effort traffic. This is because PEL (or PML) receives more packets at the same time from working CR-LSP and recovery CR-LSP during the switchover time after the node goes up, thus the packets exceeding the reserved bandwidth is discarded. That means the reordering of packet problem may result in the packet loss problem incase of guaranteed service. In all schemes except Haskin's scheme, the simulation results are similar to those of best-effort service.

Figures 9 and 10 show the performance comparison of each scheme by the location of a failed node in best-effort service and guaranteed service, respectively. In all schemes, the simulation results are similar to those of a link failure except Chen's and Yoon's schemes. Those two schemes cannot support the fault recovery incase of node failures.



(a)

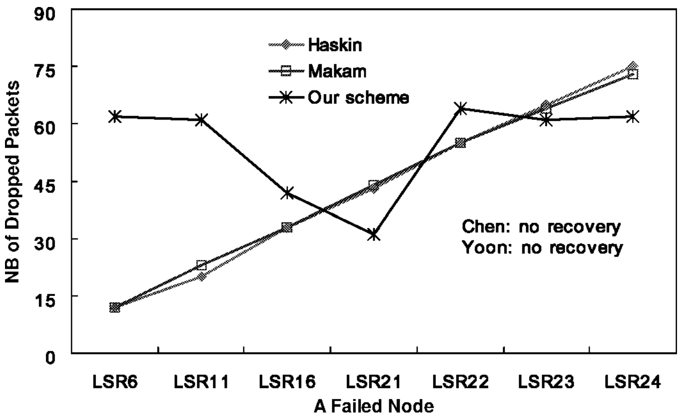


(b)

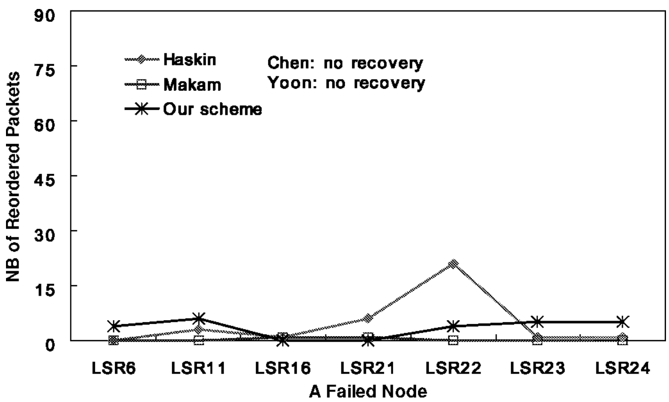
Figure 9. A node failure in best-effort service: (a) packet loss; (b) reordering of packets.

Figure 11 shows a comparison of each scheme by the number of sequential and concurrent faults. In this simulation, node failures occur from PEL toward PIL concurrently and sequentially. As shown in figure 11, both Haskin's scheme and Makam's have no connection with concurrent faults. The proposed scheme shows poor performance in concurrent faults. The reason is that concurrent faults may make networks unstable, thus wrong ER information may be used in establishing a recovery path. In that case, the try and error operation for establishing a LSP may be repeated until a recovery path is established.

Figure 12 shows resource utilization comparison of each scheme by the number of failed nodes. In this simulation, resource means label, bandwidth, and buffer reserved for recovery path. The proposed scheme shows the best performance than any other schemes in resource utilization.



(a)



(b)

Figure 10. A node failure in guaranteed service: (a) packet loss; (b) reordering of packets.

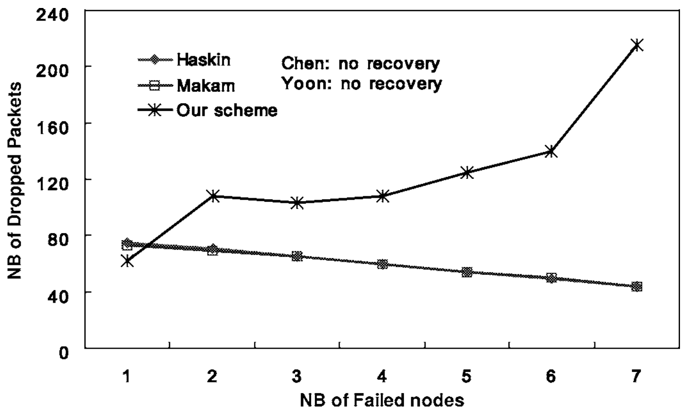


Figure 11. Concurrent faults in guaranteed service: packet loss.

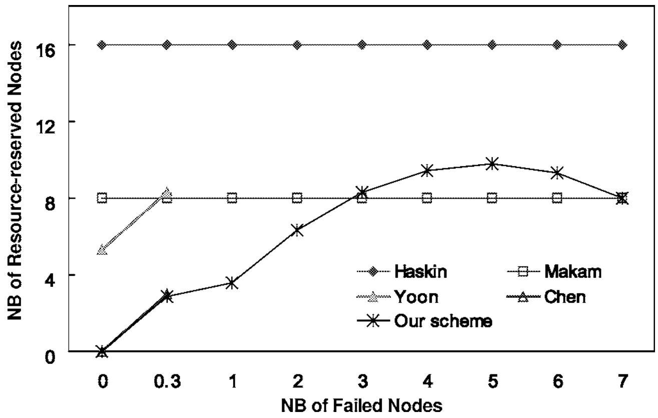


Figure 12. Resource utilization: 0.3 in X-axis means link failure.

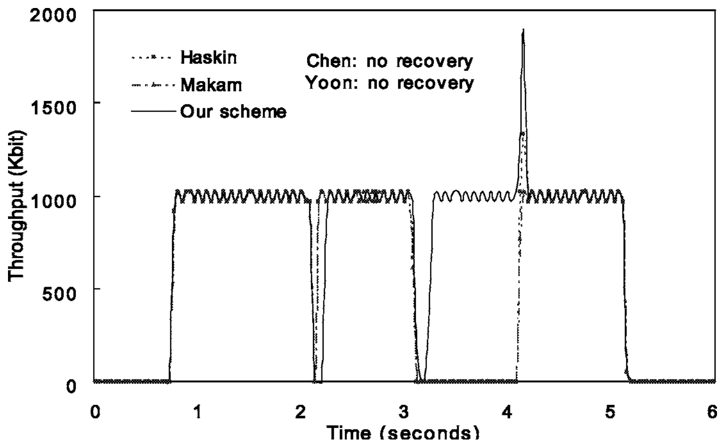


Figure 13. Node failure on both working path and recovery path.

Figure 13 shows throughput of each scheme in case of node failures on both a working path and its recovery path. In this simulation, a node failure on a working path occurs at 2 seconds and a node failure on its recovery path occurs at 3 seconds. Those failures are restored at 4 seconds. As shown in figure 13, only our scheme protects the recovery path as well as the working path.

Table 1 shows the performance comparison of each MPLS recovery scheme based on the simulation results. Both Chen's and Yoon's schemes have very good performance in case of a link failure. But they has disadvantage that cannot recover node failures and concurrent faults. Table 1 shows that our scheme has better performance than the existing schemes in link failures, in node failures, in failures on both a working path and its recovery path, and in resource utilization, even if it has a weakness in concurrent faults.

Table 1
Performance evaluation and comparison of each MPLS recovery scheme.

<table>
<tr><th colspan="2">Schemes / Characteristic</th><th>Haskin</th><th>Makam</th><th>Chen</th><th>Yoon</th><th>Our scheme</th></tr>
<tr><td colspan="2">Recovery model</td><td>Protection switching</td><td>Protection switching</td><td>Rerouting</td><td>Rerouting</td><td>Rerouting</td></tr>
<tr><td colspan="2">Repair mechanism</td><td>Global</td><td>Global</td><td>Local</td><td>Local</td><td>Local</td></tr>
<tr><td colspan="2">Recovery support</td><td>End-to-end</td><td>End-to-end</td><td>Link</td><td>Link</td><td>End-to-end</td></tr>
<tr><td rowspan="2">Best-effort service</td><td>Packet loss</td><td>Good</td><td>Bad</td><td rowspan="5" colspan="2">* Link failures: Normal/Good

* Node failures: Not support</td><td>Normal</td></tr>
<tr><td>Reordering of packets</td><td>Bad</td><td>Good</td><td>Normal</td></tr>
<tr><td rowspan="2">Guaranteed service</td><td>Packet loss</td><td>Bad</td><td>Bad</td><td>Normal</td></tr>
<tr><td>Reordering of packets</td><td>Good</td><td>Good</td><td>Good</td></tr>
<tr><td colspan="2">Resource utilization</td><td>Bad</td><td>Normal</td><td>Good</td></tr>
<tr><td colspan="2">Concurrent faults</td><td>Good</td><td>Good</td><td>Not support</td><td>Not support</td><td>Bad</td></tr>
<tr><td colspan="2">Failure on both working and recovery paths</td><td>Not support</td><td>Not support</td><td>Not support</td><td>Not support</td><td>Good</td></tr>
</table>

5. Conclusion

In mission critical MPLS networks, it is very important to quickly reroute traffic around a failure or congestion in a LSP. This paper proposed a rerouting scheme using the least-cost based recovery path, which can increase resource utilization, establish a recovery path relatively fast, support all failure types such as link failures, node failures, failures on both a working path and its recovery path, and concurrent faults.

In this paper, two ideas in establishing a recovery path were proposed. One is Candidate-PMLs that can be used as an PML of a recovery path. The other is Least-Cost Recovery-Path algorithm used to calculate the least-cost path of all possible alternative paths between itself and each Candidate-PML. The simulation results show that our scheme has better results in link and node failure, in the failure of both a primary and its alternative path, and in resource utilization than the existing schemes, even if it has a weakness in concurrent faults.

We are planning to simulate and evaluate the proposed scheme by using more comparison criteria such as $1:n$ protection in large MPLS networks.

References

[1] G. Ahn and W. Chun, Design and Implementation of MPLS network simulator supporting LDP and CR-LDP, in: *IEEE Internat. Conf. on Networks (ICON2000)*, Singapore, September 2000, pp. 441–446.

[2] L. Andersson, B. Cain and B. Jamoussi, Requirement framework for fast re-route with MPLS, Internet draft (June 1999).

[3] L. Andersson, P. Doolan, N. Feldman, A. Fredette and B. Thomas, LDP specification, RFC 3036 (January 2001).

[4] G. Apostolopoulos et al., QoS routing mechanisms and OSPF extensions, RFC2676 (August 1999).

[5] D.O. Awduche, MPLS and traffic engineering in IP networks, IEEE Communications Magazine (December 1999) 42–47.

[6] D.O. Awduche et al., RSVP-TE: Extensions to RSVP for LSP tunnels, Internet draft (August 2001).

[7] T.M. Chen and T.H. Oh, Reliable services in MPLS, IEEE Communications Magazine (December 1999) 58–62.

[8] D. Haskin and R. Krishnan, A method for setting an alternative label switched paths to handle fast reroute, Internet draft (May 2000).

[9] C. Huang, V. Sharma, S. Makam and K. Owens, A path protection/restoration mechanism for MPLS networks, Internet draft (July 2000).

[10] B. Jamoussi et al., Constraint-based LSP setup using LDP, Internet draft (February 2001).

[11] S. Kini, M. Kodialam, T.V. Lakshman and C. Villamizar, Shared backup label switched path restoration, Internet draft (November 2000).

[12] Q. Ma and P. Steenkiste, On path selection for traffic with bandwidth guarantees, in: *Proc. of IEEE Internat. Conf. on Network Protocols (ICNP)*, October 1997.

[13] L. Mo, General considerations for bandwidth reservation in protection, Internet draft (July 2000).

[14] E. Rosen, A. Viswanathan and R. Callon, Multiprotocol label switching architecture, RFC 3031 (January 2001).

[15] V. Sharma et al., Framework for MPLS-based recovery, Internet draft (July 2001).

[16] UCB/LBNL/VINT, ns notes and documentation, URL: `http://www-mash.cs.berkeley.edu/ns` (July 1998).

[17] Z. Wang and J. Crowcroft, Quality of service routing for supporting multimedia applications, IEEE Journal of Selected Areas in Communications (JSAC) 14(7) (September 1996) 1228–1234.

[18] S. Yoon et al., An efficient recovery mechanism for MPLS-based protection LSP, in: *Joint 4th IEEE Internat. Conf. on ATM (ICATM 2001)*, Seoul, Korea, April 2001, pp. 75–79.