Proceedings of the Int. Conf. on Computer and Communication Engineering, ICCCE'06
Vol. I, 9-11 May 2006, Kuala Lumpur, Malaysia

Recovery Modeling in MPLS Networks

Wajdi Al-Khateeb[1], Sufyan Al-Irhayim[2], Khalid Al-Khateeb[1]
[1]*Department of Electrical and Computer Engineering, Faculty of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia.*
[2]*Department of Computer Engineering, University of Bahrain, Bahrain*
E-mail: wajdi@iiu.edu.my

Abstract

Transmission of QoS based traffic over packet-switched network typically requires resource reservation or differentiated treatment to guarantee an acceptable level of performance. But it is also essential to bound the disruption caused by failure of nodes or links for a real time traffic to a limit that is acceptable by the application. In this paper, a simulation platform models the impact of the MPLS recovery/protection schemes on the QoS traffic parameters including disruption time and number of out of order packets arriving at the destination. The simulation considers measures to alleviate drawbacks caused by recovery process.

1. Introduction

IP-based Internet is challenged to support multiple classes of service to meet diverse quality-of-service (QoS) requirements. Transmission of QoS based traffic over a packet-switched network typically requires resource reservation or differentiated treatment to guarantee an acceptable level of performance such as throughput, delay, jitter. These requirements imply also strict reliability objectives, such as keeping disruptions for real time traffic caused by failure of nodes or links to a limit that is acceptable by the application.

The Multi-protocol Label Switching (MPLS) brings the Internet backbone one step nearer to that of the PSTN by allowing bandwidth reservation and differentiated treatment of the traffics through its traffic engineering process. It also grants the backbone a high level of availability by adopting protection/switching schemes similar to that of the PSTN backbone [1].

2. Reliability through MPLS

The current Internet inherently has a degree of survivability due to the connection-less IP protocol. Dynamic routing protocols are designed to react to faults by changing routes when routers learn about topology changes, such as congestions or failure, via routing information updates (e.g. link status advertisement). Loss of QoS has not been an issue because current Internet was designed to deliver a best-effort service.

In contrast, the MPLS is connection oriented, which implies greater sensitivity to faults, particularly to interruption of services. Reliability is becoming more important as users expecting a higher level of performance and reliability from the Internet. In practice, fault restoration capabilities are implemented in multiple protocol layers, such as automatic protection switching in physical transmission layers, self-healing in the ATM virtual path layer, and fast rerouting in MPLS.

The ability to protect traffic around a link/node failure is important in mission critical networks. The path recovery is to reroute traffic around a failed path, where packets are redirected to a recovery path in case of working path failure [3]. The traditional recovery in the Internet is based on rerouting. Rerouting is a model that establishes a recovery path after a failure on its working path through recalculating a new "shortest path".

Protection switching, as implemented in MPLS, 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. 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 label switch router (LSR) that has detected a failure.

In MPLS, after a fault is detected, the LSRs will automatically carry out procedures for: fault notifications to other LSRs, search for alternate path, rerouting to the alternate path, and (optional) restoring back to the original path after recovery from failure.

Copyright © 2006 IEEE. All rights reserved.
ISBN 983-43090-0-7

3. MPLS Recovery

MPLS recovery remained for some time a key research issue in the Internet Engineering Task Force (IETF). Several drafts are published proposing options for recovery mechanism [3-6]. A comprehensive framework for MPLS-based recovery is presented in [3]. Well known resilience/recovery concepts from SDH and ATM technologies are mapped to MPLS recovery. Also a number of well established protection schemes from switched circuits backbones are adopted.

3.1 Recovery Schemes

It is important to select appropriate topologies that reflect the needs and practicalities for QoS based Internet backbones. Such topologies are important for both, the analytical and the simulation models. Size, scalability, symmetry, connectivity, and heterogeneity in link capacity are some of the important factors to be considered when selecting topologies. These factors will ensure that the analysis and simulation results are as general as possible.

A wide spread topology used in connection with the MPLS analysis and simulation consists of an ingress to egress path representing the main "protected" path, which consists a number of Label-switch Routers (LSR) between the ingress and egress, and one, or more, backup paths around the ingress-egress that would be ready to carry the rerouted traffic following a failure in the protected path. This arrangement resembles the 1+1 and 1:1 backup protection/switching in the infrastructures of conventional telecom backbones. This topology is effective in dealing with:

Ingress-based protection/switching (pre-negotiated)

Fast-rerouting restoration by a specific node detecting a failure

In order to allow such topology to deal with dynamic routing as well, the protected and the backup paths are equipped with equal number of LSRs. The routers in the protected path are connected to their peers in the backup path through cross-links allowing for dynamic rerouting around a failed link. The behavior of the restoration mechanism with regard to packet losses, packets reordering, and resource utilization are detailed in a performance evaluation study [5] based on simulation of MPLS backbone. The simulation supports protection configuration such as pre-negotiated end-to-end, fast rerouting, as well as dynamic protection.

3.1.1 Fast-Rerouting (Haskin's Scheme): In this scheme an alternative Label Switched Path (LSP) route is set to handle fast reroute [6]. A backup route is pre-negotiated in advance, which can be used to carry lower priority traffic that can be preempted by the higher priority protected traffic once switched over to the alternative path. When an established LSP becomes unusable, due to switch or physical link failure, data may need to be rerouted over a backup LSP. The alternative path can be established after the detection of the primary path failure or using the predefined alternative LSP in order to reduce the switchover time.

Haskin's scheme defines a method for setting an alternative path with the objective to provide a quick restoration. Both, one-to-one (1:1) and many-to-one (1: N) protection can be achieved. The main idea in Haskin's scheme is to reverse the traffic at the point of failure back to the ingress, where the traffic flow is redirected via a parallel LSP between ingress and egress switches of the protected path, Figure 1. The main advantages of this scheme are the support of fast rerouting and minimal packet losses. Lost packets are only the transitional ones on the link that experience failure, e.g. link between nodes 5 and 7 in the simulation setup, Figure 4. However packet reordering during switchover from backup path to main after recovering from the failure is its main limitation.

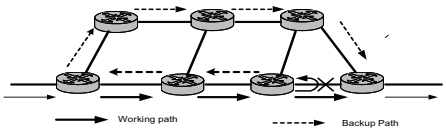

Figure 1: Fast Rerouting Scheme

3.1.2. End-to-End Rerouting (Makam's Scheme): As MPLS-based recovery is expected to become a viable option for obtaining faster restoration than layer 3 rerouting, this scheme [7] sets the procedures for the configuration of working and protection paths. Failure notification information is transmitted to appropriate switching elements which activate appropriate switchover actions. The components for the switching protection are specified as follows:

i. A method for selecting the working and protection path
ii. A method for setup of the working and protection path
iii. A fault detection mechanism
iv. A fault notification mechanism
v. A switchover mechanism
vi. A repair detection mechanism
vii. An (optional) restoration mechanism to the working path after repair

A protection includes both: pre-negotiated and/or dynamic protection mechanism. The dynamic protection requires longer restoration time. The pre-

Copyright © 2006 IEEE. All rights reserved.

negotiated protection assigns a pre-established protection path which is link and node disjoined with the primary working path. The resources such as bandwidth and buffers are predetermined and reserved for the use of the protected path, however, they are either left unused beforehand, or they are allocated with lower priority traffic in the absence of a failure on the protected path. For each protected LSP a protection LSP is established either between the ingress and egress Label Switch Routers LSRs, Figure 2 or between designated recovery switching points. The switching LSR must be notified that an LSP failed in order to switch the LSP to the protection LSP. Once notified, the switching LSR will carry out the switch-over function where the traffic is diverted from the failed LSP to the backup path. For the optional restoration function, a notification message on the repair of the primary path allows the switching LSR to restore back the traffic from the backup path to the primary path. The MPLS signaling protocols CR-LDP and RSVP-TE are extended to support such failure notification.

Advantage of this scheme is that it requires almost no packet reordering. However, the notification message delivery time results in packet loss. As the

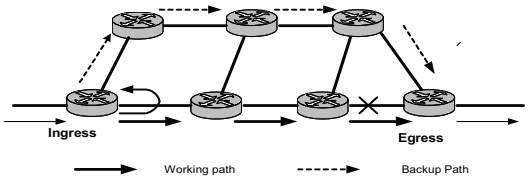

Figure 1.2: End-to-end Protection Scheme

pre-negotiated routing in Haskin and Makam schemes selects a backup path only once at the LSP time, it may not reflect the exact status of network resources at the time of fault [5] proposes an approach where exact status of network status information are exchanged among LSRs so that the backup path selection engine use up-to-date information and decide an optimal backup path for a possible failure.

4. Disruption Time

Service disruption time is inevitable with any protection scheme. It depends on the topology and how fast fault detection/reporting proceeds in end-to-end protection or how far the node located from the ingress in the fast rerouting schema is. The service disruption time is generally determined by the delay difference encountered by the packets traveling from the ingress to egress over the working path and that encountered over the protected path.

The difference increases with failure location being close to the egress node. Since delays over the two paths may differ considerably with larger network size, the effect of such difference is to be considered carefully and solution needs to be found to alleviate the consequences. Delay between Ingress and Egress over the working path is the sum of the individual delays of links located along the protected route, while the delay over the backup route is the sum of the delays of links between ingress and egress located along the backup route plus the delays of links between the node sensing the failure on the working path and the ingress.

5. Adaptive Delay

An "adaptive delay" is proposed in this paper to alleviate the negative effect of the recovery action. It is an additional adjustable delay added to the delay of each link to reduce the disruption time found between the stream of packets arriving at the egress over the main route and the packets arriving over the backup path following the rerouting as a result of the failure in the main route. The rules applicable to the adaptive delay are as follows:

- The added delays are maximized when added to the standard links delay at the main route. The adaptive delay is adjusted such that the standard delays of the main route plus the added adaptive delays to each link between the ingress and egress are within the time limit acceptable by the given application.
- The adaptive delays added to the links crossed by the packets over the backup routes are adjusted such that the difference between the total delay over the main route and the backup route is minimized.
- For favorable system conditions, e.g. maximum allowable delay between ingress and egress and location of failure from ingress node, delays over main route and backup route could become equal resulting in zero disruption time
- The further the location of the failure from the ingress node, the shorter the added adaptive delays to the standard delays of the links across the backup routes. A limit is set when the duration of the adaptive delay added to the link delays of the backup route is reduced to zero when the length of the backup route exceeds that of the main route by a given limit making the disruption time inevitable.

6. Out-of-order Packets

Out-of-order packets are the result of the traffic restored back from the backup path to the main route

Copyright © 2006 IEEE. All rights reserved.

following the repair of the main path. The number of the out-of-order packets, that is undesirable in real time traffic, appears in both schemes, e.g. the "end-to-end rerouting as well as the "fast rerouting. Although destination nodes can be equipped with means to reorder the out-of-order packets in non-realtime applications, such reordering may takes time that is beyond the acceptable limits of a real-time application hence they are dropped in the later applications.

While the service disruption occurs generally after a failure in the main route, the out-of-order packets occur after the repair of the main route as a result of restoring the traffic from the backup path back to the main path.

The number of the out-of-order packets is generally small in the end-to-end rerouting, negligible in the dynamic rerouting, it may be quite high in the fast rerouting scheme depending on how far is the down stream failure located from the ingress LSR.

Here again, the proposed "adaptive delay" considered earlier to minimize the duration of the disruption time, can help keeping the number of the out of order packets to a minimum. The value of the adaptive delay added to the standard delay of the links in the backup route are set such that the packets will travel through this path faster than those traveling through the main route. As a result, larger number of these packets is cleared from the backup path before packets restored back to the main route, following the repair of the main route, start arriving at the egress.

7. Simulation

The simulator, MNS-2 [9] is used to analyze the recovery behavior in an MPLS domain without and with the proposed "adaptive delay. MNS-2 is a patch developed as an extension to the network simulator NS-2. It extends the IP protocols to those of the MPLS. Out of the many features in the MPLS, this paper focuses on the recovery scheme of the fast recovery

7.1 Simulation scenario and results

The selected MPLS domain is made of a main path 1-3-5-7-9, of which node 1 is the Ingress and node 9 is the egress. The backup path 1-2-4-6-8-9 shares its ingress and egress with the main path. The recovery scheme is that of the fast routing. The failure of link 5-7 affects the traffic forwarding activities as follows:

a) During normal operation the traffic between source, node 0, and destination, node 10, proceeds between ingress, node 1, and egress, node 10, via the working path 1-3-5-7-9, Figure 3.

b) Following link 5-7 failure, all packets traveling the link between nodes 5 and 7 are lost. Although practically part of these packets may survive, e.g. those past the point of failure, the simulator is not capable of such treatment

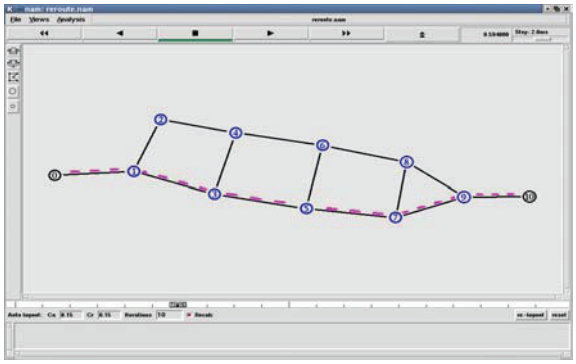

Figure 3: Regular traffic via main path

c) The node immediately sensing a link failure in the main path, node 5 in this scenario, implements the fast rerouting of all packets arriving at this node back to the ingress, which on its part reroutes them to the destination via the backup path, Figure 4.

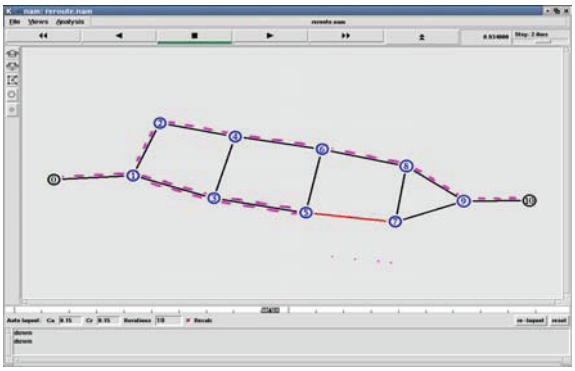

Figure 4: Fast rerouting

d) A disruption time follows the link failure caused by discontinuity of the packet flow. The disruption slot is the time gap between the arrival of the last packet (found at the head of node 7 prior to the link 5-7 failure) at the destination, node 10, which is the last packet traveling through the main route, and the last packet arriving at node 5 after this node is notified of the failure ahead of it, Figure 5. This packet leads the stream of subsequent packets traveling along the backup route 5-3-1-2-4-6-8-9. While the delay encountered by the last packet over the main route mounts to the delay of two links, e.g. 7-9 and 9-10, it takes the first packet over the backup routes a delay of 8 links. The simulation using standard delay results in a disruption time of $t_{disr.}$= 895-822=73 ms, Figure 5.

Copyright © 2006 IEEE. All rights reserved.

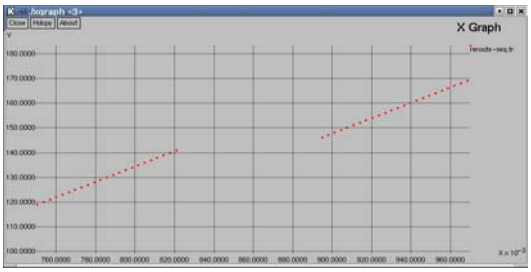

Figure 5: Disruption time, standard delay

e) Following the repair of the main route, node 5 will be notified of this repair and it responds by stopping the rerouting back the packet stream over the backup route. The result of this is a flow of two streams of packets. Old packets already found on the backup route will continue flowing over the backup route and new stream of packets over the main route, Figure 6.

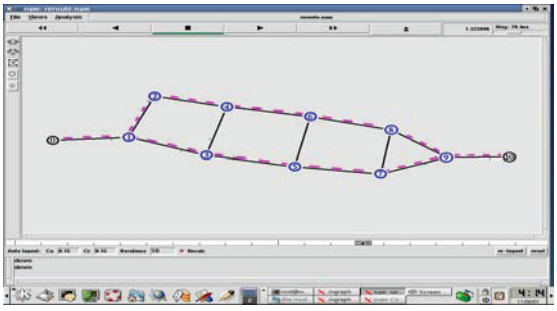

Figure 6: Two streams of packets

From the instant when the egress starts receiving fresh packets over the shorter segment of the network, e.g. the main route, older batch of packets that are located over backup route are considered as out-of-order-packets, Figure 7. The number of out-of-order packets using standard delay is 18 packets.

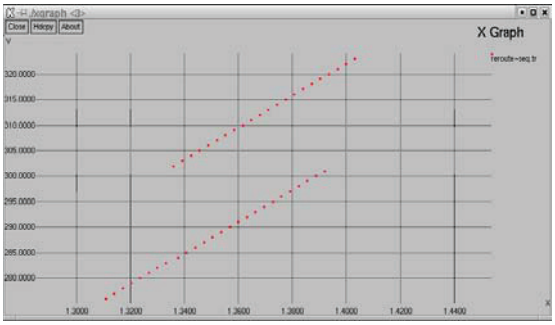

Figure 7: Out-of-order packets, standard delay

7.2 Simulation with adaptive delay

The disruption time is minimized by applying the "adaptive delay". The adaptive delay for the links are selected such that the difference between the total delays through the main route and the backup route, e.g. from the node before the failure to the egress via the ingress, is minimized. In the example of the simulation, Delay (1-3-5-7-9) – Delay(1-3-5-3-1-2-4-6-8-9) = Minimum. The simulation using the adaptive delay results in an improved disruption time of $t_{disr.}$= 895-840=55 ms, Figure 8, as compared to 73 ms with standard delay, which mounts to an improvement of 25%.

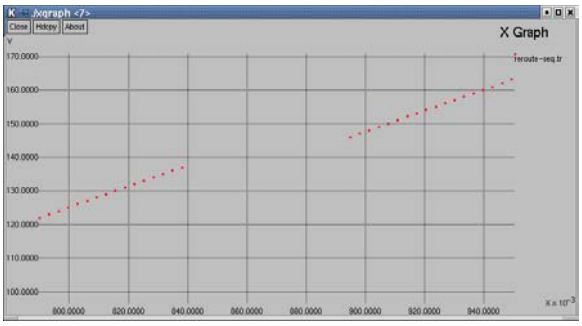

Figure 8: Disruption time using adaptive delay

Number of out-of-order packets are reduced in the simulation with adaptive delay setting. The numbers drops to 8 packets, Figure 9, as compared to 18 in the standard setting.

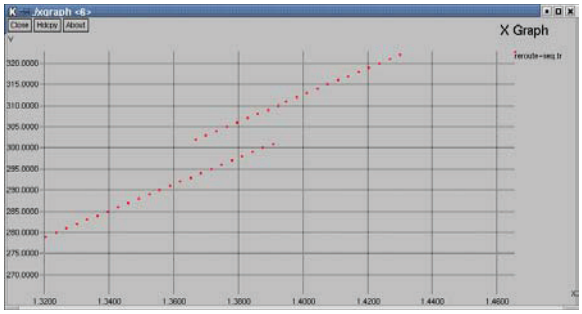

Figure 9: Out-of-order packets using adaptive delay

Finally, disruption time and out-of-order packets resulting from recovery are shown together in Figure 10 for simulation with standard delay and by using adaptive delay, Figure 11.

Copyright © 2006 IEEE. All rights reserved.

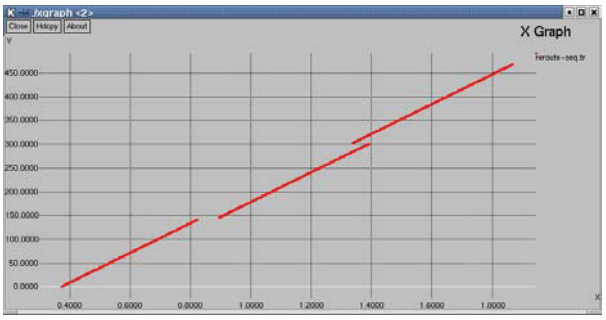

Figure 10: Recovery with standard delay

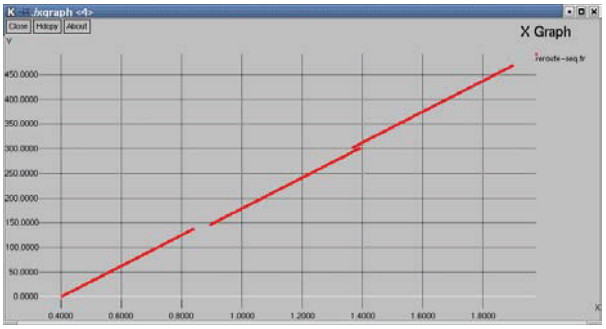

Figure 11: Recovery with adaptive delay

8. Conclusion

In this paper, recovery behavior of MPLS networks is modeled. Negative impact of the recovery on "disruption time" and "out-of-order packets" is analyzed using recovery simulation on an MPLS domain platform. A traffic parameter, the "adaptive time delay" is proposed that can alleviate the effect of the recovery on these two traffic parameters as a result of link failure in the main route, which affects the disruption time, and the restoration of the traffic from the backup route to the main route, which affects the out-of-order packets.

References

[1] W. F. Al-Khateeb et al, "Reliability Objectives inNext-generation Internet", *9th Asia Pacific Conference on Communication,* 2003, pp 192-197.

[2] T. M. Chen, T. H. Oh, "Reliable services in MPLS" *IEEE Communications Magazine*, 1999, pp 58-62.

[3] V. Sharma et al, "Framework for MPLS-Based Recovery", *IETF Internet Draft*, July 2001.

[4] K. Owens et al, "A Path Protection/Restoration Mechanism for MPLS Networks", *IETF Internet Draft*, July 2001.

[5] G. Ahn, J. Jang, "An Efficient Rerouting Scheme for MPLS-Based Recovery and Its Performance Evaluation", *Kluwer Academic Publishers*, 2002.

[6] D. Haskin, R. Krishnan, "A Method for Setting an Alternative Label Switched Paths to Handle Fast Reroute", *IETF Internet Draft*, Nov. 2000.

[7] S. Makam et al, "Protection/Restoration of MPLSNetworks", *IETF Internet Draft*, Oct. 1999.

[8] S. Yoon et al, "An Efficient Recovery Mechanism for MPLS-based Protection LSP", *ATM and High Speed Intelligent Internet Symposium*, 2001, pp 75–79.

[9] G. Ahn, W. Chun, "Design and implementation of MPLS network simulator (MNS) supporting QoS", *15th International Conference on Information Networking*, 2001, pp 694-699.

Copyright © 2006 IEEE. All rights reserved.