Method for Paths' Optimization during Path Recovery in MPLS Network

Veneta Aleksieva[1]

Abstract – **MPLS(Multi Protocol Label Switching) is being used in many corporate networks and public infrastructures and as a backbone technology of many Autonomous Systems. Many mission critical applications require better resilience than that provided by the current Internet routing convergence process. During path recovery in MPLS networks large numbers of packets may be dropped. This paper presents a method, which overcomes part of this problem by optimizing paths during path recovery in MPLS network.**

Keywords – **MPLS, LSP, backbone networks**

I. INTRODUCTION

MPLS(Multi Protocol Label Switching) networks are currently evolving towards an universal and convergent network, capable of flowing multiservice traffic as voice, data and video over the same IP based infrastructure. In a real situation in most of MPLS networks is used a physical trace of fiber optic. But sometimes, this fiber cut can cause all the traffic in the fiber to totally interrupt, which is equal to at least tens of Gbps capacity or sometimes even up to hundreds of Gbps capacity. The loss of this huge amount of traffic can bring a significant impact on our economy. Thus, network protection and survivability is of paramount to today's telecommunication networks. This is the main reason for applying of MPLS conception of LSP priorities.

MPLS uses Label Switching Paths (LSPs) priorities. The purpose of them is to mark some traffic as more important than others and allow them to use resources from less important LSPs (pre-empt the less important LSPs). This makes it possible for an important LSP to be established along the most optimal path for this LSP, regardless of existing reservations, if those reservations have a lower priority than this LSP. When LSPs need to reroute, important LSPs have a better chance of finding an alternate path the lower priority LSPs. Best effort traffic that does not need the same treatment in the network, can be mapped to low priority LSPs and higher priority LSP can pre-empt those low priority LSPs if it becomes necessary.

II. RELATED WORKS

The recovery of the MPLS network is based on the algorithm that is applied in order to detect the faults and to route the data flow in an alternative path. For a MPLS based backbone network, the fault-tolerant issue focuses on how to protect the traffic of LSP against node and link failures. In IETF, two well-known recovery mechanisms (protection switching and rerouting) have been proposed, but many researchers create every year some better suggestions, which have different advantages and disadvantages [1].

When an IP packet travel on a MPLS domain, it follows a predetermined path depending on the Forwarding Equivalence Classes (FEC) [2] to which it was assigned by the ingress router. The two main approaches to determine the desired granularity for FEC and determining the paths for the Label Switching Paths (LSP) are:

- **Offline path calculation** - This way of doing path calculations can lead to optimal resource usage, predictable routing and stable network configurations, because determined paths with an off line tool without the LSRs directly participating in the process.
- **Constraint based routing** - Each LSR determines an explicit route for each traffic trunk (aggregation of traffic flows) originating from that LSR based on the bandwidth and the cost of the links and other topology state information [2].

In practice, the traffic engineer will specify the endpoints of a traffic trunk and assign a set of attributes to the trunk about the performance' expectations and behavioral characteristics of the traffic trunk, but there are two main categories of how to set up a LSP:

- **Static LSP**
- **Signaled LSP**.

Static LSP is a LSP that is manually configured via CLI or SNMP. Visiting each LSR and using network management to set the label and interface typically create this kind of LSP.

Dynamic signaling protocols have been designed to allow single routers to request the establishment and label binding to FEC for an end-to-end path. The router that needs to setup an LSP simply determines the best path through the network according to the local constraints and requests the routers in the path to establish a LSP and distribute the label binding to FEC. Configuring a new LSP, over a domain that is MPLS and signaling enabled, does not require anything beyond the configuration in the instantiating router. Signaling is a way in which routers exchange relevant information. In an MPLS network, the type of information exchanged between routers depends on the signaling protocol which is being used. At a base level, labels must be distributed to all MPLS enabled routers that are expected to forward data for a specific FEC and LSPs created. The MPLS architecture does not assume any single signaling protocol [3] and so four methods have been specified for label distribution:

• Label Distribution Protocol (LDP)[4]

[1] Veneta P. Aleksieva is with the Department of Computer Science and Engineering, Technical University of Varna, str."Studentska "1, 9010 Varna, Bulgaria, e-mail: ven7066@abv.bg

• Resource Reservation Protocol extension for MPLS (RSVP-TE)[5,6,7]

• Constrained Routing with LDP (CR-LDP)[8,9,10]

• Distributing labels with BGP-4[11]

Multiple protection routing schemes are possible. To minimize disruption and control overhead, it is used protection routing schemes that change the route from the origin LSP when it traverses a failed link before it fails. Among this class of routing reconfiguration techniques, link-based protection is the most widely used. Thus, link-based protection is good decision, but this scheme can extend to path-based protection, which can be viewed as a special case of link-based protection in an overlay topology.

III. Analysis

To recover a failure, protection of end-to-end connectivity does not need to know where the failure is. Once the two end nodes of the working path detect a network failure, they just perform the switching-over actions. Protection is carried out at the two end nodes of a working LSP. Thus, for a specific link failure, only those unaffected protection LSP can be used to protect the working path. LSP restoration also allows spare capacity sharing among different protection LSPs. The key condition to ensure full failure recovery is that a fiber link should reserve an amount of protection capacity that is maximal among all the link failure situations.

In Fig.1 is presented one example of MPLS network with LSP and its LSP restoration when link failure is occurred. Primary LSP start from LSR1 and follows LSR1-LSR2-LSR3-LSR4, but when link failure between LSR3 and LSR4 arise, packets, which travel on this path, will switch on the backup (alternative) LSP: LSR1-LSR2-LSR3-LSR7-LSR4. This backup path is created before link failure in off-line phase on the protocol, but on this link travel primary traffic.

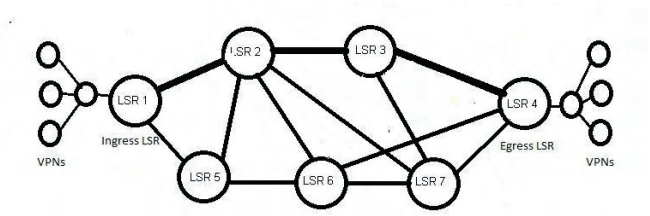

Fig. 1. MPLS network

This means that this link may be responsible for two LSPs – basic path and protection path, and they are sharing protection capacity on their overlaying links. The same links protect more than one LSP. Therefore, the condition of spare capacity sharing is that each time there is only one single network failure and only one of the working paths is recovered, because if more than one link failure arises, it leaves all the traffic on the other working path totally lost.

According to some recent studies on node failure protection with path restoration or shared backup path protection, it is found that a network that supports only single-link failure protection can essentially provide a high percentage, for example, more than 90%, of single-node failure protection without bringing in any extra protection capacity, i.e. it may use the protection capacity that is specially planned for the single-link failure restoration to recover single-node failures and more than 90% node-failure traffic flows can be recovered[13]. Typically, for a mesh network, as a backbone MPLS network, a span restorable network can have a spare capacity efficient around 50-70%, while a shared backup path protection network can achieve spare capacity efficiency around 30-40%[14].

IV. Problem Formulation

In link-based protection, the source node of a failed link reroutes the traffic originally passing through the failed link along a detour route to reach the tail node of the link. Thus, the protection routing only needs to be defined for each link that requires protection; in contrast, the base routing defines routing for each primary LSP.

If as the alternative path is on not empty link and it has own traffic, which uses temporary close to full bandwidth of link, available bandwidth may be not enough for this new repaired traffic. For example, in fig.1 primary path on link LSR3-LSR4 will switch on LSR3-LSR7_LSR4, but on link LSR7-LSR4 traffic exists in the same moment. This traffic, for which this link is the primary path, has higher priority than the new added traffic. This means, that if bandwidth is less than all traffic, some part of rerouting traffic will be lost before repairing of the original link and restoring on this link this traffic.

It is difficult to find optimal routing for alternative link, because optimal routing depends on each interface. For example, if the protocol on layer 3 is IS-IS, this route will be optimal, but if the protocol on layer 3 is OSPF- would not.

This means that for optimizing of network performance and minimizing of packet losses, when MPLS recovery occurs, must be found new algorithm, which will evaluate the behavior of MPLS recovery mechanism.

V. Solution strategy

Based on analysis, the new method for recovery must consist of two phases:

- **Off-line phase** – In this time there are optimized both routing and protection routing in the same time, using original RSVP-TE protocol. The main goal of this stage is to minimize the overload, when failure occurs.
- **On-line phase** – In this time, after failure occurs and traffic is sending on alternative (recovery) path, LSR applies protection routing as fast reroute. This gives to the MPLS network advantage, because if in the moment this alternative link has not enough bandwidth for both traffics, instead of packet losses, these packets will reroute on second alternative link, which is temporary and dynamically created in the moment.

To discover information for MPLS traffic, including Virtual Private LAN Service (VPLS) information, must enable the appropriate agents. They are different for different vendors – Cisco, Huawei, Juniper, Laurel etc. The agents that retrieve

MPLS data use either Telnet or SNMP to retrieve the data. Before enabling the MPLS agents, it must configure Telnet and SNMP access on these devices. Agents that retrieve VPLS information can retrieve large amounts of data. Enabling these agents can add significant processing time to the discovery process.

Basic algorithm of this suggestion is presented below:

When a packet arrives at a router, its next-hop is computed using the network map minus the failed links. If this next-hop would send the packet out an interface that has a failed link, then the router follows next steps for each packet:

1. to remember the failed link
2. to recompute the route using this new failure information
3. to return to step one if the new next-hop also incurs a failure or, if not, forwards the packet to its next- hop

Algorithm for each packet is:

```
Initialization:
        packet.failed links = NULL
Packet Forwarding:
        while (TRUE)
    path = ComputePath(M − packet.failed links)
    if (path == NULL)
        abort("Path is absent")
    elseif (path.next hop == FAILED)
        packet.failed links != path.next hop
    else
        Forward(packet, path.next hop)
    Return
```

Moreover, short explanation of mathematical model of this suggestion is presented below.

There are LSRs $X = \{x_i\}, i = \overline{1,n}$, which are connected with links with bandwidths $D = \{d_j\}, j = \overline{1,k}$ and cost of link $C = \{c_j\}, j = \overline{1,k}$, and $H(k) = \| h_{ij} \| i, j = \overline{1,n}, k = \overline{1,K} - classes, h_{ij}$ – intensity of K-class, which is sent from LSR_i to LSR_j in KBps. Algorithm will found LSP as queue from links $E = \{(r,s)\}$ with throughput $\{\mu_{rs}\}$ and dispersion of flows for all classes $F(k) = [f_{rs}(k)]$ that cover all traffic from each class $H(k)$ and minimize number of packets' losses CLP_k, used minimal cost of network. The main goal is:

$$\min_{E\{\mu_{rs}\}} C_{\Sigma}(M) = \Sigma \sum_{(r,s)\in E} C_{rs}(\{\mu_{rs}\}) \quad (1)$$

And main condition is:

$$CLP(\{\mu_{rs}\};\{f_{rs}\}) \le CLP_k \quad (2)$$

In [15] authors found mathematical expression for CLP_{rs} :

$$CLP_{rs} = P_k = \left[\sum_{k=0}^{n_{rs}} \left(\frac{f_{rs}}{\mu}\right)\frac{1}{k!} + \left(\frac{f_{rs}}{\mu}\right)^{n_{rs}} \frac{1}{n_{rs}} \sum_{K=1}^{N} \left(\frac{f_{rs}}{n_{rs}\mu}\right)^{k}\right]^{-1} \left(\frac{f_{rs}}{\mu}\right)^{n_{rs}} \frac{1}{n_{rs}!} \sum_{K=1}^{N} \left(\frac{f_{rs}}{n_{rs}\mu}\right)^{N_{rs}} \quad (3)$$

Based on (3) average probability of packets' losses in entire MPLS network from K-class is:

$$CLP_k = 1 - \prod_{(r,s)} \left(1 - CLP_{rs}(\mu_{rs}; f_{rs})\right) \quad (4)$$

The main goal of this algorithm is to optimize LSP, in order to minimize packet losses during path recovery process, but it does not reduce time for LSP recovery.

Administrator may affect on the choice of primary LSP, when use bandwidth, priority, administrative weight and attributes and affinity. Configurations on primary and backup paths are presented below:

```
ip rsvp bandwidth <B>                                                    (5)
        where B=75% by default and it is bandwidth of interface
tunnel mpls traffic-eng <s> {H}                                          (6)
        where s=[0;7] setup priority, H=[0;7] holding priority, 0 –high
        priority , by default s=7,H=7
mpls traffic-eng administrative-weight <M>                               (7)
        where M=[0;2^32-1]- metric, which overwrite IGP metric
tunnel mpls traffic-eng path-selection metric {te|igp}                   (8)
        where igp is by default and it is used when channel has delay
mpls traffic-eng attribute-flags <0x0-0xffffffff>{ mask<0x0-0xffffffff>} (9)
tunnel mpls traffic-eng path-option 1 explicit name straight
tunnel mpls traffic-eng oath-option 2 dynamic                            (10)
```

This configuration shows that in LSRs is possible to define proper static path, but when link failure occurs, LSR will find dynamically new one. This will work, if in ingress router is configured fast reroute with command:

```
tunnel mpls traffic-eng fast-reroute                                     (7)
```

But on the protected link:

```
tunnel mpls traffic-eng backup-path <backup-tunnel>                      (8)
```

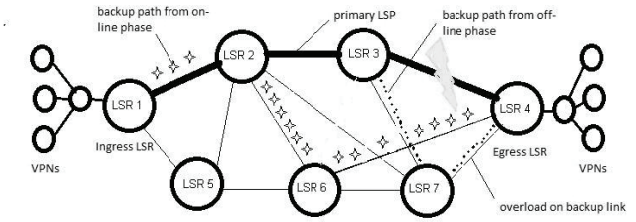

Fig. 2. Example of LSPs in MPLS network

On Fig.2 is presented one example of this algorithm. When link between LSR3 and LSR4 is failed, primary LSP is failed to, and recovery mechanism will switch traffic on backup path. Then it is calculated throughput on basic traffic on this link. For example, if there are two traffics, each of them with average capacity 4.92GB and free bandwidth on reserved link is 10GB, but 4.92GB from this link are already used for high priority traffic. This means, that 4.92GB basic+4.92GB new= 9.84GB, but only 7GB (70% from 10GB) are borrowed about this class of traffic. This is with 2.84GB more than the capacity of the link and they must be rerouted dynamically on different link, because otherwise they would be lost. To overcome this problem, part of the traffic with low priority is

sent back to the LSR2, and LSR2 recalculate new path to the LSR4, without link LSR7-LSR4, where is the problem with overflow. The new dynamic path for this traffic will be LSR1-LSR2-LSR6-LSR4. When the link between LSR3 and LSR4 is repaired, traffic will travel on the primary path LSR1-LSR2-LSR3-LSR4.

VI. Conclusion and Future work

MPLS (Multi protocol Label Switching) is being used in many corporate networks and public infrastructures and as a backbone technology of many Autonomous Systems. This is a connection oriented technology that arises to palliate the problems that current networks have related to speed, scalability and traffic engineering.

A traditional traffic engineering algorithm computes an effective base routing that optimizes a network metric, such as minimizing congestion cost or maximum link utilization. Then, a protection routing is derived from this method, for example, through fast rerouting. While simple and well studied, this traditional approach can easily result in serious network congestion and performance unpredictability under failures.

In this paper first it is formally defined the problem of overflow after the LSP recovery process and then is explained reasons for its challenging. After that it is introduced the key ideas of algorithm, which overcomes packets' losses during the LSP recovery and finally it is given one example of this suggestion.

For the future work this algorithm will be implemented in MPLS module in Network Simulator 2 to be possible to compare with well known recovery schemes in MPLS networks all qualitative and quantitative parameters, because it is important to find in each of them relationship among failure rate of routing paths, the repaired time for finding of protection/alternative path and number of availably alternative paths, number of packets' losses etc.

Acknowledgement

The work presented in this paper was supported within the project BG 051PO001- 3.3.04/13 of the HR Development OP of the European Social Fund 2007-2013

References

[1] Aleksieva V. P., Comparison Studies on Path Recovery Schemes in MPLS Network, ICEST'10, Macedonia, Bitola 2010, vol.2, p.497-500
[2] RFC 3031
[3] L. Andersson, P. Doolan, N. Feldman, A. Fredette, B. Thomas
[4] "LDP Specification (RFC 3036)",2001, http://rfc-3036.rfc-list.net/
[5] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, A. Malis "Constraint-Based LSP Setup using LDP (RFC 3212)",2002, http://rfc-3212.rfc-list.net/
[6] L. Andersson, G. Swallow
[7] "The MPLS Working group decision on MPLS signaling protocols"RFC 3468, 2003
[8] L. Berger, Y. Rekhter "Generalized MPLS Signaling – Implementation Survey", draft-ietf-ccamp-gmpls-signaling-survey-preview-3a,2002
[9] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow,"RSVP-TE: Extensions to RSVP for LSP Tunnels (RFC 3209)",http://rfc-3209.rfc-list.net/, 2001
[10] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin,"Resource ReSerVation Protocol (RSVP) (RFC2205)",http://rfc-2205.rfc-list.net/, 1997
[11] D.Awduche, A.Hannan, X.Xiao,"Applicability Statement for Extensions to RSVP for LSP-Tunnels (RFC 3210)", http://www.ietf.org/rfc/rcf3210.txt, 2001
[12] Y. Rekhter, E. Rosen, "Carrying Label Information in BGP-4 (RFC 3107)", http://rfc-3107.rfc-list.net/, 2001
[13] http://www.network-protection.net/node-failure-protection/
[14] http://www.network-protection.net/shared-backup-path-protection-sbpp/
[15] www.foibg.com/ibs_isc/ibs-11/ibs-11-p12.pdf