Fast Reroute In MPLS Network And Design MPLS Network For PTIT

Author Tran Cong Hung.Ph.D and Author Nguyen Thi Hoang Ngoc
Dept of Information Technology,
Posts & Telecom Institute of Technology (PTIT), HoChiMinhCity, Vietnam
Tel : +84-8-829-9605 Fax : +84-8-829-5258 E-mail: conghung@ptithcm.edu.vn
Tel : +84-8-829-9605 Fax : +84-8-829-5258 E-mail: hoangngocpk@yahoo.com

This paper presents the ability to quickly reroute traffic around error or congestion in a Label Switched Path (LSP) of MPLS networks. When label switched path established unusable (e.g. due to a physical link or switch failure) data may need to be re-routed over an alternative path. This paper include 5 parts. Part 1 introduce the essential of alternative path in MPLS when nodes or links in network were failed. Part 2 presents the reroute in LSP of MPLS network and the alternative path routing algorithm for Traffic Engineering. Part 3 presents the simulations of the alternative LSP. And part 4 is a MPLS network design for PTIT Network in the future.

I. Introduction

The alternative path can be established after a primary path failure is detected or, alternatively, it can be established beforehand in order to reduce the path switch-over time.

Pre-established alternative paths are essential where packet loss due to an LSP failure is undesirable. Since it may take a significant time for a device on a label switched path to detect a distant link failure, it may continue sending packets along the primary path. As soon as packets reach a switch that is aware of the failure, packets must be immediately rerouted by the switch to an alternative path away from the failure if loss of data is to be avoided. Since it is impossible to predict where failure may occur along an LSP tunnel, it might involve complex computations and establish alternative paths to protect the entire tunnel. In the extreme, to fully protect an LSP tunnel, alternative paths might be established at each intermediate switch along the primary LSP.

This paper consider the way which provides in-band means for quick detection of link and switch failures or congestion along a primary path without resorting to an out of band signaling mechanism.

II. Reroute

Networks using traffic engineering must satisfy network change and maintain stabilize. Any node or link failure which can't destroy the higher priority of network services, special the high service layer. Reroute is a mechanism which make the smallest services failure and reroute is optimised by network topology change.

MPLS Fast Reroute [1], [4] provides a mechanism for automatically rerouting traffic on an LSP if a node or link in the LSP fails. Fast rerouting is accomplished by pre-computing and pre-establishing a number of "protection LSPs" between the source and destination routers. Each link or node in an MPLS network can be protected via a protection LSP. This LSP provides an alternative path for the data being sent through primary LSPs that pass through the link or node should there be a failure. The LSP acts as a temporary tunnel through which all of the affected LSPs can be routed. The fail-over mechanisms are triggered by physical link or

routing events that indicate that the link or node is down. In theory, a router should be able to reroute packets immediately after receiving the event. Ideally there should be no packet loss or interrupted services during the switch-over.

Paths for LSPs are calculated at the LSP headend. Under failure conditions, the headend determines a new route for the LSP. Recovery at the headend provides for the optimal use of resources. However, due to messaging delays, the headend cannot recover as fast as possible by making a repair at the point of failure.

1. Fast reroute operation

[7]The example in Figure 1 i1llustrates how Fast Reroute link protection is used to protect traffic carried in a TE tunnel between devices R1 and R9, as it traverses the mid-point link between devices R2 and R3. [The TE tunnel from R1 to R9 is onsidered to be the primary tunnel and is defined by labels 37, 14, and Pop.] To protect R2–R3 link, you create a backup tunnel that runs from R2 to R3 by way of R6 and R7. This backup tunnel is defined by labels 17, 22, and Pop.

When R2 is notified that the link between it and R3 is no longer available, it simply forwards traffic destined for R3 through the backup tunnel. That is accomplished by pushing label 17 onto packets destined to R3 after the normal swap operation (which replaces label 37 with label 14) has been performed. Pushing label 17 onto packets forwards them along the backup tunnel, thereby routing traffic around the failed link. The decision to reroute packets from the primary tunnel to the backup tunnel is made solely by R2 upon detection of link failure.

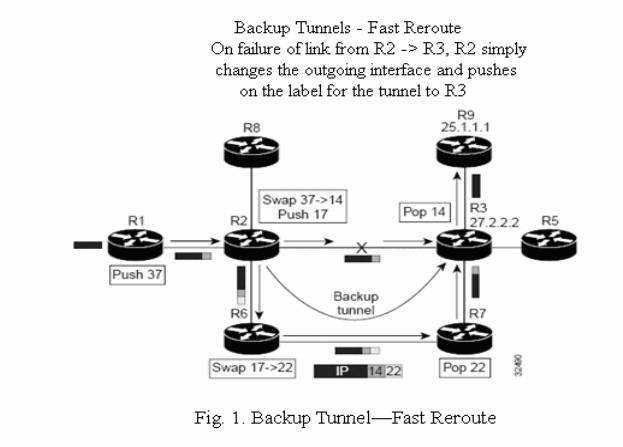

Fig. 1. Backup Tunnel—Fast Reroute

2. Optimal reroute

A feature of layer 2 network called *bridge - and - roll* or *make - before - break [7]*. This is the capability to always set up a new VC while maintaining current VC. Suppose the new and existing paths for a tunnel require resources from common links. However, one or more of those links does not have sufficient capacity to admit the second path. The first, tunnel has to torn down and then re-establish on the new path. However, the links were able to recognize the second path as a replacement for the existing path, the path could be admitted.

3. The alternative path routing algorithm [2]

The problem of routing is to optimize flow of packets through a network (Traffic Engineering) with various constraints such as, optimize network utilization, minimize cost of switching, minimize the number of hop counts, optimize use of link bandwidth, etc. The basic problem of routing is equivalent to the max-flow or shortest path problem and can be solved in polynomial time. A heuristic algorithm for solving the bandwidth constraint routing is proposed. The algorithm consists of two phases a preprocessing phase and a online phase. In the pre-processing phase the critical links in a network is identified and paths between all pairs of edge routers are computed, while in the online phase the routes are calculated based on the information collected during the pre-processing phase.

a) Pre-processing phase

Traffic data from the network is collected for a period of time. Because the measured traffic rate of the links can vary significantly at different time, it must be decided which value to use. One common choice is the 95th-percentile of all rates measured every 5 minutes over a period of time. This value is close to the real peak value as opposed to the traffic spike.

This data is used to identify the critical links. A list of edge routers are collected, path between each pair of edge routers are pre-computed. The shortest path between all the pairs is computed. Let this shortest path have a cost say *shortest path cost* (SPC). If this path contains a critical link then an alternate path with cost less than 110% of the SPC is computed and selected as the primary path and the path which has the critical link is selected as the secondary path. If no such path exists then the shortest path which has the critical link in that is selected as the primary path and the secondary path is left empty.

In this phase the critical links are identified and the paths between every pair of edge router is computed.

b) Online routing phase

In the online phase the shortest distance between the source and the destination is calculated. If there be a critical link in this path then an alternate path which has

a greater hop count (or cost) is sought for. If there exist such a path then that path is named as the primary path and the shortest path with critical link is named as the secondary path else if there exist another longer path, which also contains the critical link then that path is said to be the secondary path, while the shortest path with the critical link is named as the primary path else if no other longer path exist then the shortest path with the critical link in that is named as the primary path and the secondary path is left null.

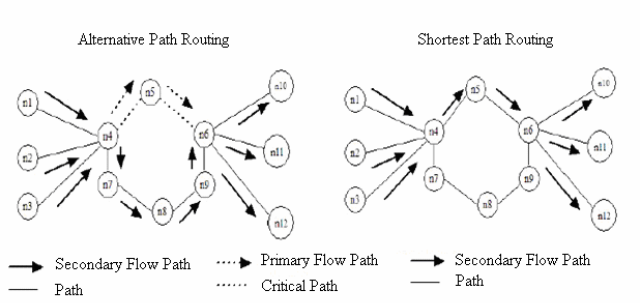

Fig. 2. A Network consisting of 12 nodes

An example network with link bandwidths is shown in Table 1. There are 4 paths being setup two from n1, one each from n2 and n3. The path from n1 to n3 requests a bandwidth of 0.25Mb, while that from n1 to n10 requires a bandwidth of 1.5Mb. The traffic from n2 to n11 requires a bandwidth of 2.5. The traffic path from n3 to n12 requires 3.5Mb bandwidth.

Table 1. An example network with link bandwidth

ID link	Start Node	End Node	Link band width (Mbit)	95% peak rate (Mbit)	Critical link	Edge Node
1	n1	n4	2	1.75	No	Yes
2	n2	n4	3	2.5	No	Yes
3	n3	n4	4	3.75	Yes	No
4	n4	n5	5	8	Yes	No
5	n5	n6	5	8	No	Yes
6	n4	n7	4	0	No	Yes
7	n7	n8	4	0	No	Yes
8	n8	n9	4	0	No	Yes
9	n9	n6	4	0	No	Yes
10	n6	n10	2	1.5	No	Yes
11	n6	n11	3	2.5	No	Yes
12	n6	n12	4	3.5	No	Yes

- Shortest path first algorithm

When the shortest path first algorithm is run to setup the 4 paths. The traffic from n1, n2, n3 to n10, n11, n12 flows via the links n4_n5_n6 as it is the shortest path to the destination. This leads to congestion hence packet loss and delay. Though the alternate path to the destinations exist it is not used by the SPF algorithm.

Table 2. The shortest path routing algorithm

Path ID	Start Node	End Node	Link bandwidth	Path
1	n1	n10	1.5	n1_n4_n5_n6_n10
2	n2	n11	2.5	n2_n4_n5_n6_n11
3	n3	n12	3.5	n3_n4_n5_n6_n12
4	n1	n3	0.25	n1_n4_n3

- Alternative path routing algorithm

In pre-processing phase, network statistics are collected and the critical links are identified. From the statistics of this particular network the critical links are identified as the links from n4 to n5 and n5 to n6.

In the online phase, the shortest path between n1 to n10 is n1_n4_n5_n6_n10. Here in this path links n4_n5_n6 is congested so the alternate path is sought for. There exists an alternate path from n1 to n2, which is n1_n4_n7_n8_n9_n6_n10. This path has a hop count of 2 greater than the shortest path, but this path doesn't have a critical link, so it is named as the primary path while the shortest path with the critical link is named as the secondary path. This same process is done for setting up the paths between n2, n3 to n11 and n12 respectively. For the path between n1 to n3 no alternate path is sought for as the shortest path between n1 to n3 n1_n4_n3 dose not have any critical link.

Table 3. The alternate path routing algorithm

Path ID	SN	EN	BW	Primary path	CL	Secondary path	CL
1	n1	n10	1.5	n1_n4_n7_n8_n9_n6_n10	no	n1_n4_n5_n6_n10	yes
2	n2	n11	2.5	n2_n4_n7_n8_n9_n6_n11	no	n2_n4_n5_n6_n11	yes
3	n3	n12	3.5	n3_n4_n7_n8_n9_n6_n12	no	n3_n4_n5_n6_n12	yes
4	n1	n3	0.25	n1_n4_n3	no	NULL	no

4. Bandwidth reservation considerations

Generally there is no need to exclusively allocate bandwidth resources to the alternate LSP. The holding

priority of the primary LSP can be used as traffic-triggered resource preemption priority for the alternate LSP in case the primary LSP fails and traffic is switched to the alternate LSP as described in this paper. When there is no traffic, other LSPs sharing the interface should get full access to bandwidth and other system resources. Consequently, if the traffic-triggered priority of the alternative LSP is greater than the holding priorities of the other LSPs using an interface in the alternate path, the alternate LSP can preempt bandwidth and other system resources as soon as traffic gets rerouted via the alternate LSP. This enables high-priority LSPs, which are being rerouted, to preempt resources from lower priority LSPs without explicit bandwidth reservation for the alternate path.

Of course, if bandwidth efficiency is not an issue, bandwidth resources can be explicitly reserved for the alternate LSP also.

III. Advantage

The presented method of setting the alternative label switched path has the following benefits [4]:

- Path computation complexity is greatly reduced. Only a single additional path between the source and destination switches of the protected path segment needs to be calculated. Moreover, both primary and alternative path computations can be localized at a single switch avoiding problems that can arise when computations are distributed among multiple switches.

- The amount of LSP setup signaling is minimized. With small extensions to RSVP or LDP, a single switch at ingress of the protected path can initiate label allocations for both primary and alternative paths.

- Optionally, presence of traffic on the alternative path segment that runs in the reverse direction of the primary path can be used as an indication of a failure or congestion of a downstream link along the primary path. As soon as the source switch detects the reverse traffic flow, it may stop sending traffic downstream of the primary path and start sending data traffic directly along the final alternative path segment.

It is fair to note that this technique increases the likelihood of data packet reordering during the path rerouting process. Therefore benefits of the reducing the alternative path latency should be weighed against possible problems associated with short term packet reordering.

IV. Simulations of MPLS backbone network

1. Network simulation [3]

To simulate MPLS Network, there is a lot of software, but in this paper, we introduce Network Simulation (NS). This is free software and it supports various features for MPLS.

NS is developed by VINT (Virtual Internet Testbed), with jointed USB, … NS simulates events that happen in network and uses optimal resource [8].

2. Simulations

There are some simulations of MPLS backbone network. These include two cases, data are transmited in the alternative path when the primary path is failed and data are transmited in two path: primary and alternative path [6], [9].

With NS, we can create some LSR node, like as follow:

```
set LSR2   [$ns mpls-node]
set LSR3   [$ns mpls-node]
set LSR4   [$ns mpls-node]
```

Create link between Node0 and Node1 and bandwith reservation is 1Mb, and queue: droptail

```
$ns duplex-link $Node0 $LSR2   1Mb   10ms DropTail
$ns duplex-link $Node1 $LSR2   1Mb   10ms DropTail
```

Establish LDP Protocol config in LSRs

```
$ns configure-ldp-on-all-mpls-nodes
```

Establish constraint LSP in 2_4_6_8 path

```
$ns at 0.7 "$LSRmpls2 setup-crlsp  8  2_4_6_8 1200  500K  1000 100 7 2"
…
```

In the first simulation, we can see easily:

- Network topo includes 11 nodes: 2 Node0 and Node10 are CE customer, 9 nodes from Node1 to Node9 are LSRs in MPLS network backbone.

Node0 is UDP data source which is passed in network to destination- Node10.

- Data is transmited in the predicted path base Shosted dynamic algorithm which based the shorted path: 0 - 1 - 3 - 5 - 7 - 9 - 10.

- And data will be transmited in backup path when link 5 - 7 is failed (figure 3): 0 - 1 - 3 - 5 - 6 - 8 - 7 - 9 - 10.

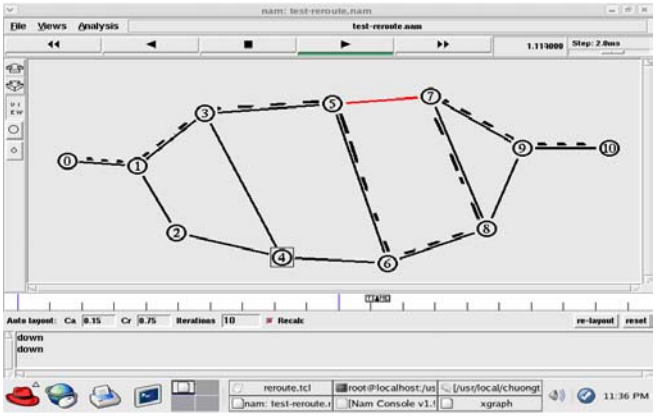

Fig. 3. Traffic is transmited in different paths (the alternative path)

The second simulations, traffic can share in two paths:

All two TCP sources are transmited in a path: 2 - 4 - 6 - 8, while there isn't any source which transmited in 3 - 5 - 7 - 8. Therefore, the congestion in this path is surely happen, due to packets loss is to be avoided.

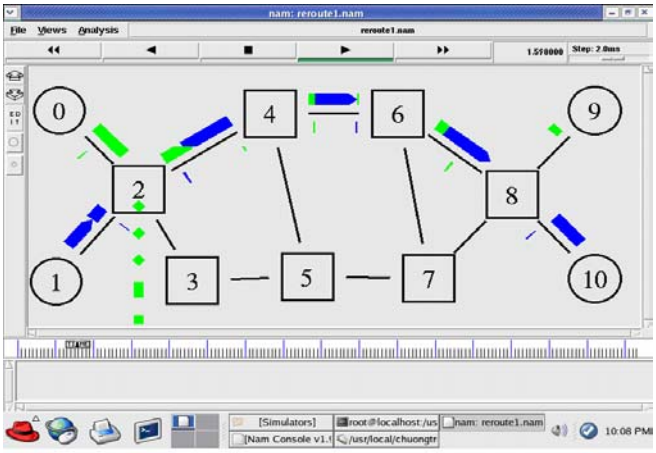

Fig 4. Routing based IP traditional

So, this problem is why don't we use resource in the alternative path? And MPLS backbone is used to solve this problem.

Data from two different sources are distributed in two path to destination. However, there is the packet loss, but a little. The routing based IP traditional problem is solved by using the MPLS backbone network (figure 5).

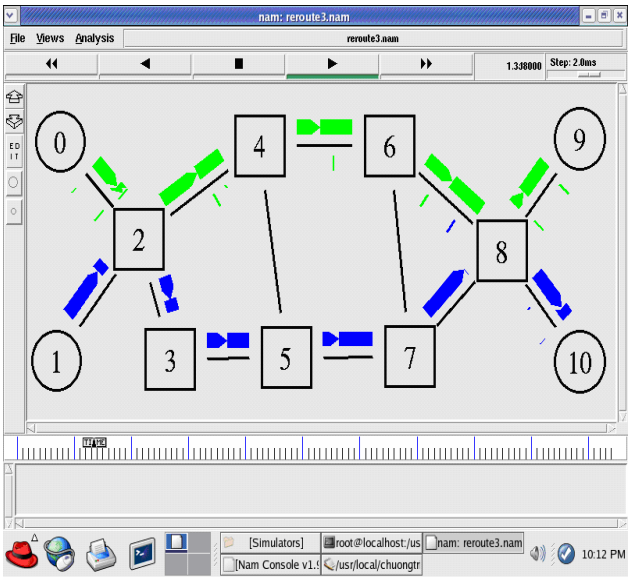

Fig. 5. Network using MPLS backbone

V. Design MPLS backbone network for PTIT

With the extension of PTIT campus for Da Nang City besides the big campuses at Ha Noi capital and Ho Chi Minh City, the network is interested issue. And there is a MPLS network topology which was designed that used for PTIT in the future.

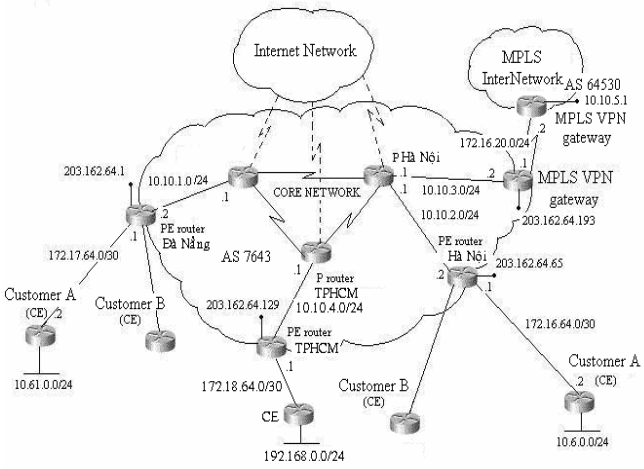

Fig. 6. MPLS Network topology of PTIT in the future

Besides that, the PTIT network has a gateway, MPLS VPN gateway, to connect to MPLS Internetwork.

This is some config command [1] for MPLS - VPN Network:

! Config iBGP

router bgp 7000 // AS number of Vietnam give to PTIT

no bgp default ipv4-unicast

```
        neighbor PTIT-VPN peer-group
        neighbor PTIT-VPN remote-as 7000
        neighbor PTIT-VPN update-source loopback0
    !
    ! Interface with P router DaNang
       interface Serial abc/0/0
        ip address 10.10.1.2 255.255.255.252
        tag-switching mtu 1536
        tag-switching ip
    !
    ! CustomerA connect to P router DaNang
           ip vrf KH_A
            rd 7000:100
            route-target export 7000:100
            route-targer import 7000:100
      !
           interface Serial 5/0/0
            ip vrf forwarding KH_A
            ip address 172.17.64.1 255.255.255.252
      !
           router bgp 7000
```

VI. Conclusion

MPLS has some advantages , special traffic engineering in MPLS. With ability reroute-base constraint, explicit routing traffic share and reroute in network which solved the congestion and using the best reserved resource in network.

In Vietnam, MPLS is developing in VDC, … and it will be deployed in some company and PTIT in the near future.

VII. References

A. Books

[1]. "Advanced MPLS Design and Implementation Vivek Alwayn", CCIE #2995 Cisco Press 201 West 103rd Street, Indianapolis, IN 46290 USA

B. Dissertations or Theses

[2]. Shyam Subramanian and Venkatesan Muthukumar, "Alternative Path Routing Algorithm for Traffic Engineering", Department of Electrical and Computer Engineering, Las Vegas.

[3]. Kevin Fall and kannan Varadhan, "Network Simulation", April 14, 2002

C. Articles in Conference Proceedings

[4]. "draft-haskin-mpls-fast-reroute-05.txt", November 2000

[5]. K. Kompella and Y. Rekhter, "draft-ietf-ccamp-gmpls-routing-09.txt", Juniper Networks, October 2003.

[6]. Gaeil Ahn and Woojik Chun,"Overview of MPLS Network Simulator: Design and Implementation", Department of Computer Engineering, Chungnam National University, Korea

D. Standards

[7]. "MPLS Traffic Engineering Fast Reroute — Link Protection", Release Number 12.0(16)ST

[8]. "Network Simulator" (ns_doc.pdf)

[9]. http://www.isi.edu/nsnam/ns/ns-documentation.html

Author's Profile

TRAN CONG HUNG was born in VietNam in 1961

He received the B.E in electronic and Telecommunication engineering with first class honors from HOCHIMINH university of technology in VietNam, 1987.

He received the B.E in informatics and computer engineering from HOCHIMINH university of technology in VietNam, 1995.

He received the master of engineering degree in telecommunications engineering course from postgraduate department HaNoi university of technology in VietNam, 1998.

He received Ph.D of engineering degree at HaNoi university of technology in VietNam, 2004.

His main reseach areas are B – ISDN performance parameters and measuring methods, and QoS in high speed network technologies, MPLS.

Currenthy, he is a lecturer, deputy head of Faculty of Information Technology II and head of section Network & Data Transmission in Post and Telecom Institute of Technology (PTIT), in HOCHIMINH City, VietNam.

Author's Profile

NGUYEN THI HOANG NGOC was born in VietNam in 1981.

She is the fifth year student of Posts & Telecommunications Institute of Technology.

Her main research areas are MPLS VPN based ATM.