标签交换路径切换方案研究

何海斌[1, 2]，曹争[1, 2]

（1.东南大学计算机科学与工程学院，南京，211189；2.计算机网络与信息集成教育部重点实验室，南京，211189）

摘　要：为了提高资源利用率，平衡网络负载，流量工程技术收到了推崇和使用。MPLS 的显式路由标签交换路径机制为实施流量工程技术提供了极大的便利。但是，如果原有 LSP 所经路径不能满足需求时，应该为其找到新的路径并进行切换。本文研究并提出了两种在 MPLS 网络中的标签交换路径切换方案，并且通过实验对比了两种方案对现有业务的影响。最后，通过对实验结果的深入分析，找出了影响服务质量的关键因素。

关键词：MPLS；LSP；ER-LSP；标签交换路径

Research on Label Switched Path Switching Schema

He Haibin[1, 2], Cao Zheng[1, 2]

(1. School of Computer Science & Engineering, Southeast University, Nanjing, 211189;

2. Key Laboratory of Computer Network and Information Integration, Ministry of Education, Nanjing, 211189)

Abstract: To facilitate the network resources more effectively and balance the load, traffic engineering (TE) technology is widely applied in network operations. The ER-LSP mechanism in MPLS offers the TE this power more conveniently. This paper focuses on the impacts of the process of switching between two LSPes and proposes two different schema of LSP switching. Experimental results demonstrates the REBUILD_LSP schema is better in the proposed schemata and the final deduction shows why

Key words: MPLS; LSP; ER-LSP; Label Switched Path;

MPLS[1]（Multi-Protocol Label Switching，多协议标签交换）是一项利用绑定在 IP 包中的标签，通过网络进行数据包转发的技术。它把第二层交换技术和第三层路由技术结合在了一起，可以在路由器、支持标签交换的交换机上，仅仅根据一个简单的标签进行数据包的快速转发，而不必执行基于目的 IP 地址的复杂路由查找算法，提高了报文的转发效率。

基于 MPLS 的流量工程可以利用 MPLS 提供的 ER-LSP（Explicit Route Label Switched Path，显式路由标签交换路径）机制创建显式路径，通过控制流量的转发路径，平衡网络资源的利用率。但是，在将一条 LSP 上的流量切换到新的路径上时，需要执行路径切换过程并且降低对现有业务的影响，本文提出了两种在 MPLS 流量工程中进行显式路径切换的方案，并通过实验结果和理论分析找出了在路径切换过程中影响服务质量的关键因素。

作者简介：何海斌，（1988-），男，硕士研究生，E-mail: hbhe@njnet.edu.cn；曹争，（1958-），男，副教授，E-mail: zcao@njnet.edu.cn.

1 相关概念

LSP 切换是指一条 LSP 从一条路径切换到另一条路径，LSP 的切换所引起的路由变化可能会对服务质量造成非常严重的影响。文献[2][3]指出，当网络中的路由发生变化、发生失效恢复等情况时，时延会显著增加，同时丢包、报文失序等情况也会增加。在路由路径发生变化并收敛到新的路由路径过程中，时延增加、丢包量瞬间增大，并伴随着剧烈的时延抖动[4]。

因此，在对 LSP 进行切换的时候，除了要保证网络整体的负载均衡，LSP 切换方案还需要考虑切换的过程尽可能的减少对网络中正在经过的流量造成的影响。具体而言，要减少原有流量因路径切换造成的失序以及报文丢包等现象，即尽可能的保证网络的服务质量。

举例来说，如现有图 1 所示的网络拓扑，当前有一条流 Flow_F 有<A,B,C,D,E> 5 个报文在 S 端的发送队列中等待被依序发送至 R 端，Flow_F 原转发路径为<S,R1,R3,R>，在依次发送了 ABCD 四个

报文之后（此时报文 A 已经到达 R 端，B 报文处于 R3-R 之间，C 报文处于 R1-R3 之间，报文 D 在 S-R1 之间，报文 E 正等待被发送），原转发路径被调整，从<S,R1,R3,R>变成了新的<S,R1,R2,R3,R>。

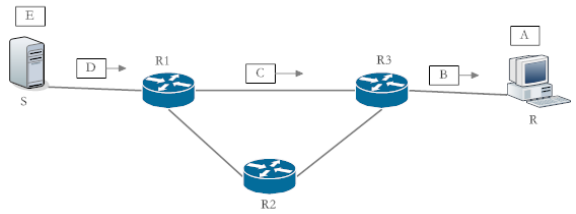

图 1 报文在路径调整前的瞬时状态图

报文失序，是指在接收端接收的报文顺序与发送端发送的顺序不一致。例如报文 D 在报文 C 之前到达接收端，使接收端接收到的报文顺序变为<A,B,D,C,E>。如果报文不能按照发送顺序到达目的地，那么可能导致该报文被丢弃，从而降低服务质量，文献[5]总结了报文失序对 TCP 和 UDP 性能的影响。导致端到端报文失序的原因主要有丢包重传以及网络内部由于并行路径、路由器并行处理等。描述报文失序的测度有失序比率、失序程度、滞后时间、失序字节数等[6]，其中失序比率的定义为：失序比率=发生失序的报文总数/接收端收到的报文总数。

丢包，即报文丢失，是指报文从发送端发出，在网络转发的过程中由于网络拥塞、路由改变等原因被丢弃，最终未被接收端收到的现象。如在上例中，假设报文 C 在到达 R3 的时候由于某种原因被丢弃，导致报文 C 没有被 R 端收到，R 最终收到的报文为<A,B,D,E>，则说明 Flow_F 发生了丢包问题。对于基于 UDP 的应用来说，报文丢失意味着收到的数据不完整，对应用造成影响。对于基于 TCP 的应用来说，报文丢失意味着需要重传丢失的报文，既影响了应用，又加重了网络的负载。

综上，由于报文失序以及丢包等问题对视频会议、VoIP 等以多媒体实时应用为代表的网络应用具有较大的影响[7][8][9]，因此一个好的路径切换过程应该尽可能的减少网络丢包和报文失序问题。

2 LSP 切换方案

在 MPLS 网络中，一条 ER-LSP（显式路由标签交换路径）从 LSR 的角度而言有 LSP 的资源属性、显式路径表项、转发策略三部分信息描述。ER-LSP 资源属性主要包括 LSP 末端路由器 ID、该路径需保留的带宽、LSP 建立/保持优先级、是否允许 IGP 协议在执行路由计算时使用该隧道等；显式路径表项则是一条用来构造 ERO（显式路由对象）的 EXP-PATH 资源，需要时从 EXP-PATH 资源池分配；转发策略则指可以在该 ER-LSP 上转发的转发等价类 FEC。

当准备从一条 OLD-ER-LSP 切换到另外一条 NEW-ER-LSP 时，在完成显式路径计算并在 LSR 上生成了代表该路径的 NEW-EXP-PATH 表项之后，对两条 ER-LSP 切换的方式可以有两种：一种方式是在原来的基础上调整 EXP-PATH，即将 OLD-EXP-PATH 替换成新的 NEW-EXP-PATH；第二种方式则是新建一条 ER-LSP，然后替换 OLD-ER-LSP，最后再将原流经 OLD-ER-LSP 的流量切换到 NEW-ER-LSP 上。下面详细介绍这两种方案。

方案 1-ADJUST_EXP，只调整 OLD-ER-LSP 的显式路径：

1) 向 LSR 申请一个 EXP-PATH 表项，根据指定路径<Hop1, Hop2, …, Hopn>，建立 NEW-EXP-PATH；
2) 将 LSP 隧道的显式路径由 OLD-EXP-PATH 替换成 NEW-EXP-PATH，其余设置不需要改变；
3) 由于只是 ER-LSP 内部属性变动，对外界而言该 ER-LSP 没有发生变化，因此在这种方案中转发策略也不需要改变；
4) 最后回收原有资源，在这种方案中只需回收 OLD-EXP-PATH 资源。

方案 2-REBUILD_LSP，新建一条 ER-LSP 并应用 OLD-ER-LSP 的所有配置：

1) 向 LSR 申请一个 EXP-PATH 表项，根据指定路径<Hop1, Hop2, …, Hopn>，建立 NEW-EXP-PATH；
2) 向 LSR 申请一条新的 LSP 资源，并将 OLD-ER-LSP 上的所有配置都应用到 NEW-LSP；
3) 将 NEW-ER-LSP 的显式路径设置为 NEW-EXP-PATH。此时 LSR 会根据该路径，通过 RSVP 协议为 NEW-ER-LSP 预留指定资源；

4) 调整原有流量导入策略，将 OLD-ER-LSP 上的 FEC 转发接口变为 NEW-ER-LSP；
5) 回收原有资源：删除 OLD-ER-LSP，删除 OLD-EXP-PATH。

以上两种方案的差别在于一个是显式路径建立好后才开始切换，另一个则是直接替换 LSP 的显式路径。假定两种方案都能成功建立新的路径，则在方案 ADJUST_EXP 中，经过第 2 个步骤，新的路径建立完成，流量就开始从新的路径上通过；在方案 REBUILD_LSP 中，流量在第 4 步后开始从 NEW-ER-LSP 路径转发。

3 实验评估

为了对比以上两种路径切换方案对当前流经该 ER-LSP 的网络流所造成的服务质量影响，本文使用 GNS3[10]搭建了一个实验网络，其拓扑结构如图 2 所示。在该 MPLS 网络中，所有路由器都开启 Telnet 远程操作路由器的功能及权限；所有的链路速率均为 10Mbps，每条链路上的可保留带宽设置为 7Mbps；C 端为控制服务器，其职责是通过 Telnet 方式与 LSR 交互，动态建立与撤销 LSP；S 端运行流量产生器，并按照指定的发送速率向 R 端发送测量报文；R 端为一台运行 Linux 操作系统的测量点，它的职责是接受测量任务，并反馈测量结果。

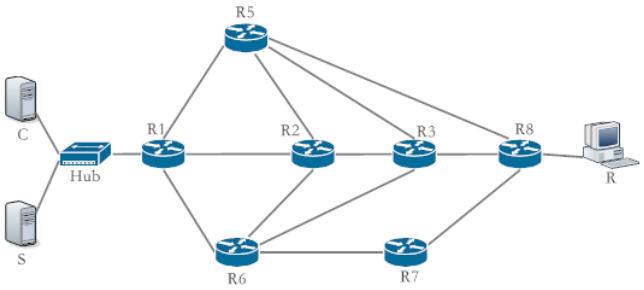

图 2 评估两个方案的实验拓扑

在实验过程中，每次进行路径选择时，会先随机从由若干条有效显式路径组成的集合中选取与上一条路径不同的路径，然后依据指定的路径在入口 LSR 上建立该 EXP-PATH 表项。每次创建的 LSP 的保留带宽为 3Mbps。本次实验选取的显示路径集合如下所示，集合中的元素<i,j,k>代表该条路径依次经过拓扑图中的 R_i、R_j、R_k 三个节点，路径的起点 LSR 为 R_i，R_j 为中间 LSR，目的 LSR 为 R_k。

EXP-PATH-SET = {
<1,2,3,8>，<1,2,5,8>，<1,2,5,3,6,7,8>，<1,2,5,3,8>，
<1,2,6,7,8>，<1,2,6,3,8>，<1,2,6,3,5,8>，<1,5,2,3,8>，
<1,5,2,6,3,8>，<1,5,2,6,7,8>，<1,5,3,8>，<1,5,3,6,7,8>，
<1,6,2,3,8>，<1,6,2,3,5,8>，<1,6,2,5,3,8>，<1,6,2,5,8>，
<1,6,3,5,8>，<1,6,3,8>，<1,6,7,8> }

在每次实验中，由 S 向 R 方向按指定的平均速率发送 UDP 报文，每个测量报文含有一个序号字段以及一定大小的数据，序号字段用来标记报文的发送顺序以便 R 端测量和计算报文的失序比率。为降低由偶然因素对实验结果造成的偏差，每次实验持续时间需要足够长。在每次的实验中，C 端运行的测试控制程序按设定的 LSP 切换频率在一个随机时间段内切换一次路径，路径的切换方式按照前一节中提出的切换过程进行。

在本文的实验中，选取了不同的切换间隔和不同的流量发送速率，每组实验均持续 600 秒，进行了多组实验。表 1 列出了本文其他的实验参数以及相应的实验结果。其中（1,2,3）组有较短的切换时间间隔，（4,5,6）组有着较长的切换时间间隔。

表 1 实验参数以及实验结果

实验组	切换间隔（秒）	发送速率(Mb)	丢包率 1	失序比率 1	丢包率 2	失序比率 2
1	0.2–1.5	2.50	8.31%	0.00%	0.07%	0.38%
2	0.2–1.5	2.75	8.34%	0.00%	0.09%	0.47%
3	0.2–1.5	3.00	8.35%	0.00%	0.07%	0.53%
4	2 – 11	2.50	1.94%	0.00%	0.15%	0.16%
5	2 – 11	2.75	2.20%	0.00%	0.24%	0.18%
6	2 – 11	3.00	2.08%	0.00%	0.39%	0.25%

注：表中丢包率 1、失序比率 1 为 ADJUST_EXP 方案的结果；丢包率 2、失序比率 2 为 REBUILD_LSP 方案的结果。

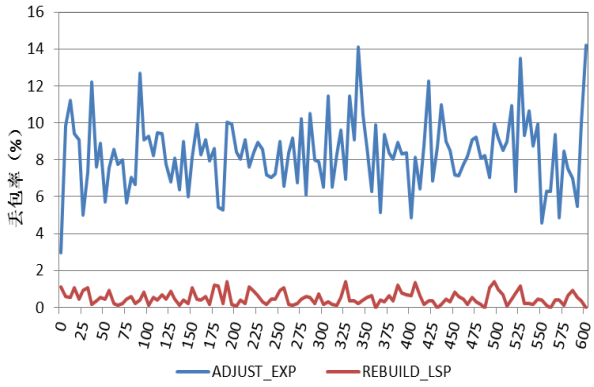

图 3 实验组 2 的丢包率分布图

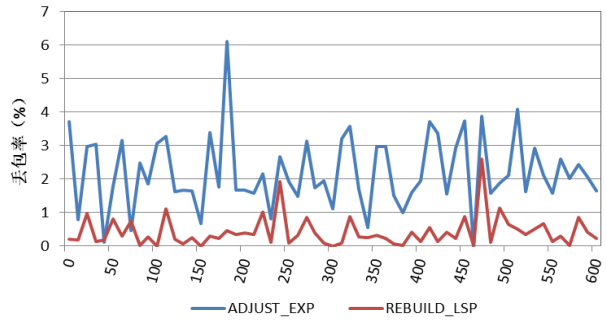

图 4 实验组 5 的丢包率分布图

图 3 和图 4 给出了实验过程中 6 组实验其中的两组实验在其 600 秒的实验周期内的丢包率情况趋势图，数据均由实验过程中产生和保存的 trace 记录文件分析统计而得，并在计算时将失序的报文也算作丢失报文。

其中图 3 给出了切换间隔在 0.2 秒至 1.5 秒之间，测量报文平均发送速率为 2.75Mbps 的实验组 2 在实验过程中的丢包率分布情况。该图的以每 5 秒为一个统计周期，共计算得出 120 个数据记录。从图中可以看出，REBUILD_LSP 方案的丢包率远远低于 ADJUST_EXP 方案。

图 4 给出了切换间隔在 2 秒至 11 秒之间，测量报文平均发送速率为 2.75Mbps 的实验组 5 在实验过程中的丢包率分布情况。由于该组实验的切换间隔较长，因此在计算该组实验的丢包情况时，采用以 10 秒为间隔时间进行计算，共计算得到 60 个数据记录。从图中也可以看出两种方案中 REBUILD_LSP 方案丢包率较低。

4 实验结果分析

从总体上来看，实验结果显示，在两种方案中，ADJUST_EXP 方案丢包率高出 REBUILD_LSP 方案很多倍。但是在 ADJUST_EXP 方案中，几乎没有报文失序问题；而 REBUILD_LSP 方案在频繁的路径切换中的失序比率比较大。

对比实验组（1,2,3）和实验组（4,5,6），可以发现，在 ADJUST_EXP 方案中，LSP 切换时间间隔越小，即切换频率越大，则导致的丢包率越大；但是 REBUILD_LSP 方案由切换频率导致的丢包问题没有前一种方案明显。

对于在 ADJUST_EXP 方案中丢包严重而几乎没有失序问题这个现象，本文猜想是 LSR 在改变显式路径时的做法是先拆除原有路径，然后再建立新的显式路径，在拆除原有路径和建立新的路径之间有时间窗口，从而导致丢包严重。为了验证这种猜想，本文使用 Wireshark[11]捕获实验网络中的报文，用以观察 ADJUST_EXP 方案中路径的拆除和建立顺序以及所需的大致时间，图 5 是在一次 ADJUST_EXP 方案的实验过程中捕获的 RSVP 信令报文（过滤了其他无关报文）。

No.	Time	Source	Destination	Protocol	Info
162	150.791656	192.168.254.1	192.168.254.8	RSVP	PATH Message.
163	150.851133	192.168.16.6	192.168.16.1	RSVP	RESV Message.
170	156.588806	192.168.254.1	192.168.254.8	RSVP	PATH TEAR Mess
171	156.603311	192.168.254.1	192.168.254.8	RSVP	PATH Message.
172	156.679065	192.168.16.6	192.168.16.1	RSVP	RESV Message.

图 5 被 Wireshark 捕获的 RSVP 报文

从图中可以看到，在被捕获报文序号为 170 的 RSVP PATH TEAR 消息发送后，经过约 14.5 毫秒之后请求建立新路径的 PATH 消息才被发出；在经过约 75.8 毫秒的时延后，该 PATH 请求收到了被捕获报文序号为 172 的 RESV 响应消息，至此新路径才完成建立。因此，从路径拆除到新路径建立成功总共花了约 90.3 毫秒的时间，而这段时间内到达 LSR 的报文由于找不到其相应的路由，因此被 LSR 丢弃，造成了较大的丢包。

另外，报文的发送速率与丢包率之间的关联没有预想中的大，为探究报文的发送速率与丢包率之间的联系，本文进行了如下推导分析。

为便于讨论，假定：

1）每秒发送 p 个探测报文，即报文发送速率为 p 报文/秒；

2）整个测量过程报文发送的持续时间为 D 秒；

3）两次路径切换的时间间隔为 T 秒，即切换频率为 1/T；

4）在 ER-LSP 上成功建立一条显式路径所需时间为 Tb，路由表项重建时延为 Tr；

通过上文的分析得出，在切换过程中，新的路径在建立成功之前，报文将被丢弃，因此单次切换导致的丢包数为：

$$Loss = p \cdot (Tb + Tr) \quad (1)$$

故在单次实验中，总丢包数为：

$$Loss_{total} = \frac{D}{T} \cdot Loss = \frac{D}{T} \cdot p \cdot (Tb + Tr)$$

(2)

另外，总发包数：

$$P_{total} = D \cdot p \tag{3}$$

从而，当 p > 0 时，单次实验总丢包率为：

$$Loss_{rate} = \frac{Loss_{total}}{P_{total}} = \frac{\frac{D}{T} \cdot p \cdot (Tb + Tr)}{D \cdot p} = \frac{(Tb + Tr)}{T} \tag{4}$$

当 p=0 时，显然有：

$$Loss_{rate} = 0 \tag{5}$$

综合公式 4、公式 5，得出：

$$Loss_{rate} = \begin{cases} 0 & p = 0 \\ \frac{(Tb + Tr)}{T} & p > 0 \end{cases} \tag{6}$$

由公式 6 可知，每次实验中因路径导致的丢包由新的路径建立时延、路由表的重建时延以及切换时间间隔三部分决定，与数据的发送速率、持续的时长没有理论上的联系；但数据的发送速率大可能导致路由器用于处理报文转发的 CPU 消耗变多，从而延长了路由表项的重建时间，即在发送速率较大时 Tr 也变大。在 ADJUST_EXP 方案中，由于拆除原有路径和建立好新的路径间的时间窗口内无法路由报文，因此报文在到达路由器之后由于没有相应的转发路径从而导致报文被丢弃，在这种方式下 Tb 和 Tr 都存在；而在 REBUILD_LSP 方案中，由于在拆除路径之前建立好了新的路径，只需要重建路由表，因此没有 Tb 部分的时延，而路由表的重建只在入口 LSR 内部执行，没有网络的延迟，加之路由表重建时延 Tr 通常要比 Tb 小很多，因此丢包率比 ADJUST_EXP 方案低了很多。

5 结束语

从上述分析可以得出，要减少由于路径切换而导致的网络丢包，可以从显式路径的重建时延、路由表项的重建时延、路径的切换频率三方面考虑。但是显式路径的重建时延以及路由表项的重建时延由当前网络的负载等因素决定，人为可控制性很小。因此从实验结果看来，采用预先建立 LSP 然后再对流量进行切换的 REBUILD_LSP 方案，加上降低 LSP 的切换频率，可以最大程度减少因路径的切换导致的网络丢包等问题。

参考文献

[1] Rosen E, Viswa Nathan A. Multiprotocol Label Switching Architecture[S]. RFC 3031, IETF, 2001

[2] Wang F, Mao Z M, Wang J, et al. A measurement study on the impact of routing events on end-to-end Internet path performance[J]. ACM SIGCOMM Computer Communication Review 36.4 (2006): 375-386

[3] Wang F, Gao L, Wang J, et al. On understanding of transient inter-domain routing failures[J]. 13th IEEE International Conference on Network Protocols. IEEE, 2005

[4] Pucha H, Zhang Y, Mao Z M, et al. Understanding network delay changes caused by routing events[J]. ACM SIGMETRICS Performance Evaluation Review 35.1 (2007): 73-84.

[5] 陈均华.多路径传输中乱序与负载均衡研究[D]:[硕士学位论文]. 浙江大学.2010

[6] Morton A, Ciavattone L, Ramachandran G, et al. Packet reordering metrics[J]. IETF internet-standard: RFC 4737, 2006

[7] Sun L, Wade G, Lines B, et al. Impact of packet loss location on perceived speech quality[J]. In 2nd IP-Telephony Workshop, pp. 114-122. 2001

[8] Arthur C M, Girma D, Harle D, et al. The effects of packet reordering in a wireless multimedia environment[C]. Wireless Communication Systems, 2004, 1st International Symposium on. IEEE, 2004: 453-457.

[9] Design V C. Video codec design[J]. Wiley. 2002.

[10] Graphical Network Simulator - GNS3[EB/OL]. http://www.gns3.net/ 2014.5

[11] Wireshark[EB/OL]. http://www.wireshark.org/ 2014.5