MPLS TE 快速重路由技术白皮书

目 录

1 概述 1
1.1 产生背景 1
1.2 技术优点 1
2 MPLS TE 快速重路由技术实现方案 1
2.1 概念介绍 1
2.2 FRR 的基本原理 2
2.2.1 FRR 的保护对象 2
2.2.2 FRR 的实现方式 3
2.3 Bypass 方式的详细介绍 4
2.3.1 主 LSP 的建立 4
2.3.2 Bypass LSP 的建立 4
2.3.3 绑定计算 5
2.3.4 失效检测 5
2.3.5 切换过程 5
2.3.6 切换后 LSP 的维护 5
2.3.7 重优化 6
3 典型应用组网 6
4 参考文献 7

1 概述

1.1 产生背景

MPLS TE 网络中一般都需要实施快速重路由保护，这主要是由 MPLS TE 自身的特点决定的。

对于纯 IP 网络，当局部失效出现的时候，如果到同一个目的地的还有其他路由可以使用，报文会按照这些路由进行转发。在失效引起的路由变化扩散到全网之前，仅靠这种机制就可以比较快速地在局部实现失效保护。

在没有布署 TE 的 MPLS 网络，现在应用比较多的是 LDP 按照 DU 方式建立 LSP。当局部失效出现时，如果还有其他路由可用，LDP 会向上游节点发起 LSP 的建立。由于没有考虑到带宽、优先级和链路属性等 TE 有关的需求，这个 LSP 建立成功的机会相对较大，因此从失效到恢复的过程也相对较短。

在 MPLS TE 网络中，LSP 的建立一般是通过 RSVP 协议按照 DoD 的方式完成的。在头节点，CSPF 利用域内所有路由信息计算出一条满足约束条件的路径，RSVP 按照这个路径建立 LSP。当网络中有局部失效时，需要重建整条 LSP。而在失效引起的路由改变扩散到头节点之前，CSPF 无法算出有效的路径。另外，局部失效可能会引起网络中多条 LSP 的重建。这样一来，利用新计算出来的路径建立 LSP 的过程中，出现带宽不够等问题的机会比较大。因此，与纯 IP 网络和没有布署 TE 的 MPLS 网络比较而言，MPLS TE 网络从局部失效中恢复的时间可能会更长，更需要一种能快速响应失效的机制。

MPLS TE 快速重路由通过预先建立备份路径，实现对 TE 隧道的快速保护倒换，从而减少数据丢失。

1.2 技术优点

MPLS TE 快速重路由是 MPLS TE 中用于保护链路和节点的机制，在 MPLS TE 网络中扮演了重要角色。MPLS TE 快速重路由事先建立本地备份路径，保护 LSP 不会受链路/节点故障的影响。当故障发生时，检测到链路/节点故障的设备就可以快速将业务从故障链路切换到备份路径上，从而减少数据丢失。

快速响应、及时切换是 MPLS TE 快速重路由的特点，它可以将业务中断的时间控制在一个很小的时间段，保证业务数据的平滑过渡。同时，LSP 的头节点会尝试寻找新的路径来重新建立 LSP，并将数据切换到新路径上。在新的 LSP 建立成功之前，业务数据会一直通过保护路径转发。

由于需要预先建立备份路径，MPLS TE 快速重路由会占用额外的带宽。在网络带宽余量不多的情况下，建议只对关键的接口进行快速重路由保护，这一点是部署 MPLS TE 快速重路由时需要注意的。

2 MPLS TE 快速重路由技术实现方案

2.1 概念介绍

- 主 LSP：被保护的 LSP。
- 保护 LSP：用来保护主 LSP 的 LSP。
- Detour LSP：为每一条需要保护的 LSP 创建一条保护路径，该保护路径称为 Detour LSP。

- Bypass LSP：用一条保护路径保护多条 LSP，该保护路径称为 Bypass LSP。
- PLR：Detour LSP 或 Bypass LSP 的头节点，它必须在主 LSP 的路径上，且不能是尾节点。
- MP：Detour LSP 或 Bypass LSP 的尾节点，必须在主 LSP 的路径上，且不能是头节点。
- 链路保护：PLR 和 MP 之间由直接链路连接，主 LSP 经过这条链路。当这条链路失效的时候，可以切换到 Detour LSP 或 Bypass LSP 上。
- 节点保护：PLR 和 MP 之间通过一个 LSR 设备连接，主 LSP 经过这个 LSR 设备。当这个 LSR 设备失效时，可以切换到 Detour LSP 或 Bypass LSP 上。

2.2 FRR的基本原理

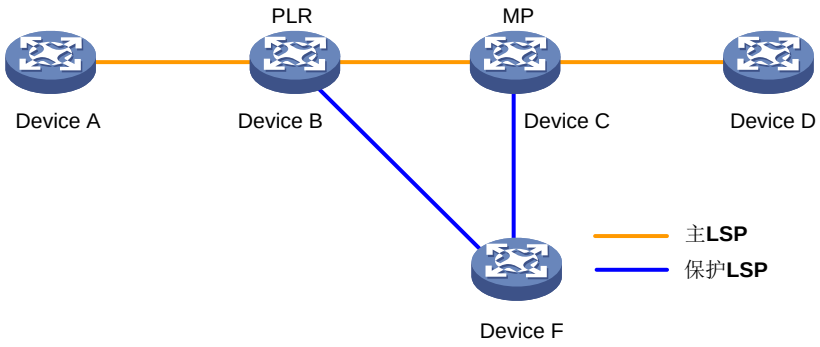
MPLS TE 快速重路由的基本原理是用一条预先建立的 LSP 来保护一条或多条 LSP。预先建立的 LSP 称为保护 LSP，被保护的 LSP 称为主 LSP。MPLS TE 快速重路由的最终目的就是利用保护 LSP 绕过故障的链路或者节点，从而达到保护主 LSP 的目的。

2.2.1 FRR 的保护对象

根据保护的对象不同，FRR 分为两类：

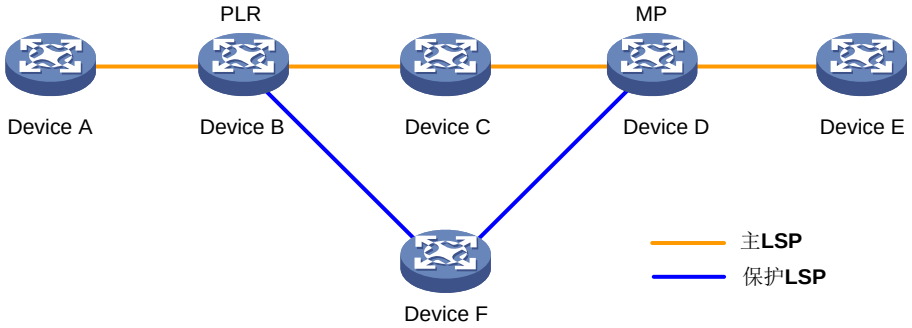
- 链路保护：PLR 和 MP 之间由直连链路连接，主 LSP 经过这条链路。当这条链路失效时，流量可以切换到保护 LSP 上。如图 1 所示，主 LSP 是 Device A→Device B→Device C→Device D，保护 LSP 是 Device B→Device F→Device C。

  图1 FRR 链路保护示意图



- 节点保护：PLR 和 MP 之间通过一台设备连接，主 LSP 经过这台设备。当这台设备失效时，流量可以切换到保护 LSP 上。如图 2 所示，主 LSP 是 Device A→Device B→Device C→Device D→Device E，保护 LSP 是 Device B→Device F→Device D，Device C 是被保护的设备。

图2 FRR 节点保护示意图



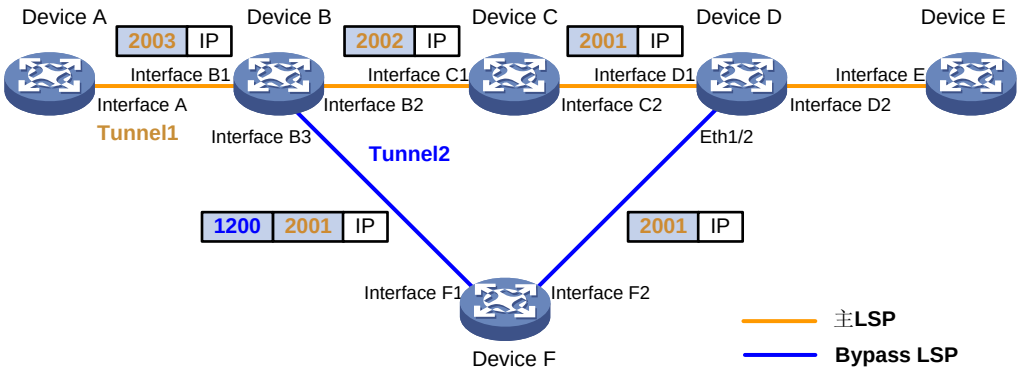
2.2.2 FRR 的实现方式

实现快速重路由有两种方式：

- Detour 方式：One-to-one Backup，分别为每一条被保护 LSP 提供保护，即为每一条被保护 LSP 创建一条保护路径，该保护路径称为 Detour LSP。
- Bypass 方式：Facility Backup，用一条保护路径保护多条 LSP，该保护路径称为 Bypass LSP。

Detour 方式实现了每条 LSP 的单独保护，但需要的开销相对较大。在实际使用中，Bypass 方式被更广泛使用。目前，Comware 只支持 Bypass 方式。

图3 Bypass 方式快速重路由



Bypass 方式快速重路由如图 3 所示，Device A→Device B→Device C→Device D→Device E 为主 LSP，Device B→Device F→Device D 为 Bypass LSP。当 Device B 到 Device C 的链路失效或节点 Device C 失效时，主 LSP 上的数据会切换到 Bypass LSP 上。Device B 发送报文时先压入 Device D 为 Device C 分配的标签，再压入 Bypass LSP 的标签，即 Device F 为 Device B 分配的标签。也就是，在 Device B→Device F→Device D 这条路径上，LSP 使用两层标签。Device D 收到报文后，弹出 Device D 为 Device F 分配的标签以后（如果 Bypass LSP 的 Device D 为 Device F 分配的标签是隐式空标签，那么 Device F 弹出标签，Device D 收到的报文只有 Device D 为 Device C 分配的标签），继续用 Device D 为 Device C 分配的标签进行转发。

2.3 Bypass方式的详细介绍

MPLS TE 快速重路由基于 RSVP-TE 建立保护 LSP。为了实现快速重路由功能，需要扩展 RSVP 消息中 SESSION_ATTRIBUT 和 RECORD_ROUTE 对象的几个标志位：

- PATH 消息的 SESSION_ATTRIBUT 对象中，使用标志位指明该 LSP 是否需要局部保护、是否记录标签、是否为 SE 风格、是否有要保护带宽。
- RESV 消息的 RECORD_ROUTE 对象中，使用标志位指明该 LSP 是否已经被保护、是否已经切换、是否被保护了带宽、是否是被节点保护。

被保护 LSP 建立与普通 LSP 建立的区别就在于这几个标志位的处理。

下面将结合图 3 从以下几个方面详细介绍 Bypass 方式快速重路由：

- 主 LSP 的建立
- Bypass LSP 的建立
- 绑定计算
- 失效检测
- 切换过程
- 切换后 LSP 的维护
- 重优化

2.3.1 主 LSP 的建立

主 LSP 的建立过程与普通 LSP 基本相同，只是增加了绑定计算，并在 PATH 和 RESV 消息中增加了几个相关标记和子对象。

RSVP 从头节点（图 3 中的 Device A）逐跳向下游发送 PATH 消息（经过 Device A→Device B→Device C→Device D→Device E），从尾节点（图 3 中的 Device E）逐跳向上游发送 RESV 消息。在处理 RESV 消息时分配标签，预留资源，建立 LSP。

主 LSP 的建立是通过在头节点（Device A）手工配置隧道来触发的。在建立主 LSP 前，如果通过命令指定该 LSP 具有快速重路由属性，RSVP 就会在 PATH 消息的 SESSION_ATTRIBUTE 对象中增加局部保护标记、记录标签标记、SE 风格标记。如果还为该 LSP 指定了带宽，则还会增加带宽保护的标记。下游节点在收到 PATH 消息以后，通过局部保护标记，就能分辨出该 LSP 是一条需要快速重路由保护的 LSP。

对需要快速重路由保护的 LSP（根据先前的 PATH 消息中的标记判断），各个节点向上游发送 RESV 消息时，会在 RRO 中记录 RESV 消息的出接口、LSR ID 和标签。这些信息被逐跳累计传递到各个上游节点。

各节点第一次收到 RESV 消息时，根据 RRO 中记录的这些信息，为该 LSP 选择合适的 Bypass LSP。为主 LSP 选择合适的 Bypass LSP 的过程称为绑定，绑定的具体算法请参见“2.3.3 绑定计算”。

2.3.2 Bypass LSP 的建立

当一个没有快速重路由属性的隧道被指定保护一个接口以后，它所对应的 LSP 就成为 Bypass LSP。

Bypass LSP（Device B 上的 Tunnel2）的建立是通过在 PLR（Device B）手工配置触发的。Bypass LSP 可以被指定保护多个接口，但不能保护它自己的出接口。它的配置与普通 LSP 的区别为：

- 不能为 Bypass LSP 配置快速重路由属性。也就是说，Bypass LSP 不能同时是主 LSP，LSP 不能被嵌套保护。
- 除了 Bypass LSP 的带宽之外，还需要配置保护带宽。

快速重路由只能进行链路保护或节点保护。在配置建立一条 Bypass LSP 时就应该规划好它所保护的链路或节点，并且要确保该 Bypass LSP 不会经过它所保护的链路或节点。否则，即使 Bypass LSP 建立成功，主 LSP 与它绑定计算成功，也不能真正起到保护作用。

Bypass LSP 的带宽一般是用于保护主 LSP 的，Bypass LSP 上所有资源仅为切换后使用。用户在配置时需要保证 Bypass LSP 的带宽大于等于被保护的所有 LSP 所需的带宽和，否则会有主 LSP 不能绑定到 Bypass LSP 上。

2.3.3 绑定计算

绑定计算是为一条主 LSP 绑定 Bypass LSP 的过程。绑定计算是为了获取切换时转发所需要的必要数据，如 Bypass 隧道接口、Bypass LSP 的 NHLFE、MP 分配的标签等。计算结果将保存下来，当发生局部失效的时候可以立即使用，这也是 MPLS TE 快速重路由可以迅速响应失效的原因。绑定计算必须在切换之前完成，如果绑定计算成功，RESV 会向上游节点通告该主 LSP 已经被保护。

当有多条 Bypass LSP 保护同一条主 LSP 时，按下面的规则选择 Bypass LSP：

- 节点保护优先于链路保护。
- 优先选择剩余带宽大于等于主 LSP 带宽的 Bypass LSP。

2.3.4 失效检测

失效检测的目标是尽快发现链路和节点失效并触发切换，缩短流量中断的时间。

可以通过三种方法检测链路或节点失效：

- 链路层协议检测：链路层发现失效的速度跟接口类型直接相关。
- RSVP 的 hello 机制：为每个需要保护的物理接口使能 hello，当对端接口也使能了 hello，就会在两个 LSR 设备之间定时发送 hello 消息和回应。当链路或节点失效的时候，hello 消息或回应消息会丢失，如果连续三次丢失消息，认为链路或节点失效。hello 机制检测失效的速度相对比较慢。
- 利用 BFD 检测邻居连通性：BFD 是一种快速检测机制，利用 BFD 可以迅速检测到链路或节点失效。

2.3.5 切换过程

切换是指启用 Bypass LSP，主 LSP 的流量和 RSVP 协议消息都不再通过失效链路或节点转发。

转发的数据首先切换到 Bypass LSP。在进行绑定计算时，数据转发所需要的内层标签 2001（MP 分配的标签）已经存放在 NHLFE 中，这时只要标记该 LSP 已经切换，数据就可以通过 Bypass 隧道进行转发了。

2.3.6 切换后 LSP 的维护

切换以后，原有链路不再可用。为使 LSP 不被超时删掉，RSVP 需要在 PLR（Device B）和 MP（Device D）之间维持消息刷新。

PATH 消息经过修改以后通过 Bypass 隧道（Device B 的 Tunnel2）发给 MP。MP 收到 PATH 消息，确认自己是 MP 节点，RESV 消息也经过修改以后经过多跳 IP 转发（经过 Device D→Device F→Device B），发送给 PLR 节点。

切换以后，主 LSP 的 PathTEAR、ResvERR、RescTEAR 和 PathERR 消息的发送路径也相应变化。

在节点保护切换以后，被保护的节点（Device C）可能会因为 PATH 消息超时而向下游发送 PATHTEAR 消息，MP（Device D）节点会忽略这个消息。另外，MP 切换的时候会在原来的 LSP 入接口（Device D 的 Interface D1）上发 ResvTear 消息，这样可以让被保护的节点（Device C）尽快释放相应的资源。

2.3.7 重优化

重优化是指按照配置的时间间隔定时对已经建成的 LSP 进行路径计算，LSR 设备按照计算出来的路径发起新的 LSP 建立过程。新的 LSP 建立成功以后会删除原来的 LSP，LSP 隧道的转发切换到新的 LSP。

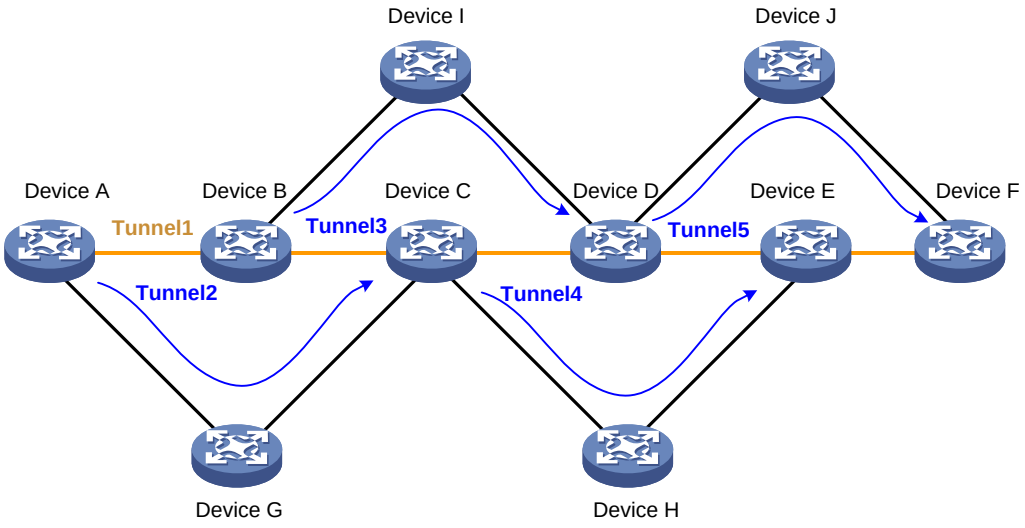
每条 LSP 隧道都可以配置重优化，当 LSP 建成以后，重优化就会启动。

对快速重路由来说，重优化的另一个作用是让被 Bypass 保护的主隧道（Device A 的 Tunnel1）恢复到正常状态。因为快速重路由保护主要用于临时性保护，所以一般需要为有快速重路由属性的隧道配置重优化。当主 LSP 还没有切换时，只有重优化计算出来的路径与原有路径不同时才会建立新的 LSP；当主 LSP 发生了切换，即使重优化计算出来的路径与原有路径相同也会建立新的 LSP。

3 典型应用组网

在网络中的关键节点上配置 FRR 保护，可以实现对通过该节点流量的保护。

图4 MPLS FRR 组网



如图 4 所示，运营商对用户提供了带宽批发业务，使用 MPLS TE 隧道接入用户的连接，使异地的用户网络通过运营商网络连接。因为承载业务为重要的业务，因此需要对主 LSP Tunnel1 经过的路

径进行保护。利用 MPLS TE FRR，可以实现通过保护路径 Tunnel2、Tunnel3、Tunnel4 和 Tunnel5 分别保护主 LSP 路径上的 Device B、Device C、Device D 和 Device E。

4 参考文献

- RFC 3209：RSVP-TE: Extension to RSVP for LSP Tunnels
- RFC 4090：Fast Reroute Extensions to RSVP-TE for LSP Tunnels
- Internet Draft, “draft-ietf-mpls-nodeid-subobject-01”