

MPLS 网络中的路径保护和恢复机制^{*})

Path Protection and Recovery Mechanism for MPLS Network

王新红 王光兴

(东北大学网络与通信中心 沈阳110004)

Abstract MPLS is a technology integrating routing with switching and supporting multi-protocol. In order to deliver data reliably, MPLS requires a set of mechanism to provide fast recovery from fault. This paper introduces two path recovery models and compares with them.

Keywords MPLS, Path recovery, Rerouting, Protection switching

1. 引言

随着网络技术的发展, Internet 正在成为人们进行信息交流的重要媒介,除了日常的信息传递外,许多商务活动依赖于它进行重要的通信业务,而且网络中的实时业务日渐增多,因此 Internet 的可靠性变得越来越重要。虽然传统的 IP 路由协议也支持重路由恢复机制,但是连接恢复时间往往需要几秒钟甚至几分钟,这对于很多敏感的实时业务(如 IP 电话)来说是不可接受的。

MPLS 是 IETF 提出的一种路由与交换相结合的技术,它保持了 IP 技术的灵活性、可扩充性和 ATM 的快速交换、流量控制等性能。MPLS 基于标签转发数据流,具有同一标签的数据流遵循相同的路径,可以简单地通过改变标签的方式达到控制数据流向的目的,增强了网络的可控制能力,为实现网络的快速故障恢复提供了可能。

基于 MPLS 的恢复机制与传统的第三层恢复机制相比,对故障具有更加快速的反应能力,同时保持了对第三层网络可见的特点。虽然网络底层也具有恢复机制,但是它对业务提供保护的粒度太粗,而且对于高层操作不可见,这样当提供链路保护时,很难提供节点保护或第三层保护。而基于 MPLS 的恢复机制则具有更大的灵活性,不仅能够提供链路保护、节点保护,还提供对整条路径的保护。目前 IETF 提出了两种基于 MPLS 的路径恢复模型:重新路由(Rerouting)和保护交换(Protection Switching),我们将在后面加以介绍。

2. MPLS 的工作原理

MPLS 网络由标记边缘路由器(LER)和标记交换路由器(LSR)组成,其中 LER 是 MPLS 网络同其它网络的边缘设备,它提供流量分类和标签的映射(作为 Ingress)、标签的移除功能,而 LSR 是 MPLS 网络的核心交换机,它提供标签交换(Label Swapping)、标签分发功能。

MPLS 设备之间运行路由协议(如 OSPF、BGP 等),通过路由协议获取网络拓扑信息,建立路由表,并根据路由表使用标签分发协议(LDP,可使用 RSVP 或 CR-LDP)在节点间分发标签信息,建立数据传输通道——标记交换路径(LSP)。LSR 的功能模块如图1所示。当一个数据流到达入口 LER 后,LER 根据一定的规则为数据流分配一个具有本地意义的标签,在以后的转发过程中,LSR 只根据数据流携带的标签进行查表、交换,不再进行其他处理,直到到达出口 LER,将

标签移去,恢复原始数据流。这种方式在入口和出口节点间数据流的传输是透明的,因此数据流的通路可以看成隧道(Tunnel)。

MPLS 这种基于标签的转发方式,很容易实现对数据流向的控制,分配了相同标签的数据流将走相同的路径,如果要改变其流向,只需要改变其标签映射就可以了。另外,分组头部可以携带多个标签,形成标签堆栈,这时标签交换基于最顶部的标签进行。标签堆栈的方式允许多个控制平面同时作用于一个分组,简化了操作。在路径恢复中,利用标签堆栈的方式,可以方便地建立保护路径、旁路故障位置。

MPLS 中支持两种选路方式:逐级跳路由(Hop-by-hop Route)和显式路由(Explicit Route)。逐级跳路由允许每个节点独立地选择下一跳,与 Internet 当前使用的路由方式一样;显式路由类似于源路由,它在网络边缘节点(如入口或出口 LSR)规定 LSP 中的部分或全部 LSR,路径根据路由表或网络和业务的限制条件(如带宽、时延等),由网络管理指定或约束路由由动态计算得到。显式路由的使用,使网络的业务流控制更易实现,并且显式路由信息只在 LSP 建立的时候传递,可以减少开销。

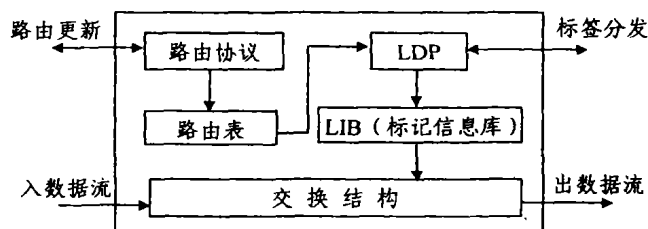


图1 LSR 功能模块图

3. 两种 MPLS 路径恢复模型

目前, IETF 已经提出了两种 MPLS 的路径恢复模型:重新路由和保护交换,下面首先给出 MPLS 路径恢复中常用的几个术语,然后对两种恢复模型分别加以介绍。

3.1 MPLS 路径恢复中的几个术语

①工作路径(Primary Path):故障发生前承载数据流的路径,是被保护路径,也称为主路径、活动路径。

②恢复路径(Recovery Path):发生故障后,数据流采用的路径,该路径根据一定的 MPLS 恢复机制建立,也称为备份路径、可选路径或保护路径。

^{*})国家自然科学基金项目(69973011),王新红 博士研究生,主要研究方向为 MPLS、QoS 保证等。王光兴 教授,博士生导师,主要研究方向为计算机网络和多媒体通信。

③PSL(Path Switch LSR):负责在工作路径和恢复路径间切换业务流的LSR.PSL是恢复路径的起点,但不一定是工作路径的起点,即它可能是入口LSR,也可能是LSP上其它的中间LSR。

④PML(Path Merge LSR):负责接收恢复路径上业务流的LSR,它将业务流重新合并到工作路径上,如果它本身是目的节点,则将业务流传递到高层协议作进一步处理.PML一定是恢复路径的终点,但不一定是工作路径的终点。

⑤FIS(Fault Indication Signal):表示发生故障的信号。它向上游或下游相邻节点传递,直到到达负责启动恢复机制的LSR.MPLS的信令协议中包含有这样的通知消息,如RSVP-TE中的PathErr消息,CR-LDP中的Withdraw消息。在CR-LDP中,随着Withdraw消息到达每个LSR,LSP被拆掉。在RSVP-TE中,PathErr消息只用于故障通知,只有随后的PathTear或ResvTear到达时,LSP才被拆掉。这些消息最终到达PSL,PSL终止FIS消息的传递。在1+1保护模式中,FIS也应该向下传递到PML。另外,有些网络也采用数据链路层来通知故障信息,如ATM网络中的OAM信元。

3.2 重新路由恢复模型

重新路由是在检测到工作路径发生故障后,动态地建立一条恢复路径,并将数据流转移到恢复路径上传输。恢复路径的建立基于故障位置、路由策略和拓扑信息等。按照路径的建立方式,重新路由恢复模型又可分为两类:按需建立(established-on-demand)和预验合格(pre-qualified)。

按需建立是在故障发生后重新选路以建立恢复路径。预验合格是在建立工作路径的同时,考察一条已存在的路径是否能够合格地承载工作路径上的业务流,如果合格,将其指定为工作路径的备份,一旦发生故障,直接将业务流转移到该备份路径上传输。预验合格方法不需要重新选路,节省了选路时间,故障恢复比按需建立方式快。

在MPLS网络中,LSP的建立不仅依赖于路由表中的路由信息,同时也必须使用信令协议进行标签分配和映射。这两个过程都要耗费时间,而且重新选路也并不一定总能够成功。如果在路径上的每个LSR都能够重新选路,又恰巧都没有成功,那么,当故障信息传递到PSL时,可能已经过了很长时间。所以,在信令协议中做了一些限制,例如,CR-LDP不支持在故障检测点重新路由,只能在入口节点重新选路;RSVP-TE允许根据拓扑结构的变化重新选路,但是也只限于故障检测点和入口节点。

3.3 保护交换

保护交换是在建立工作LSP的同时也建立一条备份LSP。备份LSP可以和、也可以不和工作LSP使用相同的链路或节点。没有故障时,业务流由工作LSP承载。当工作LSP上的LSR检测到有故障后,发送FIS消息通知PSL,然后PSL将业务流转移到备份LSP上。显然,采用这种方法,已经预先建立了恢复路径,工作路径失效后,不用再分发标记建立恢复LSP,因而路径恢复很快,但是由于提前为恢复路径预留了资源,在资源利用方面效率不高。

在保护交换中,备份LSP一般是预先配置好的,其备份模式有以下几种:

(1)1+1保护

工作LSP有特定的备份LSP,备份LSP的资源只能由工作LSP使用,即使备份LSP上没有业务传输,它的资源也不能被其他业务流使用。通常,数据同时沿着工作LSP和备份LSP传递,而PML只从工作LSP上读数据。出现故障后,失效点发送消息给PML,通知PML改为由备份LSP上读数

据。这种方式是最快的一种路径恢复方式,但是它所占用的资源也最多。

(2)1:1保护

每条工作LSP也有特定的备份LSP供其使用,但是备份LSP是空闲的,没有故障时数据流不使用它。发生故障后,失效点发送消息给PSL,通知它将数据转移到备份LSP上传输。这种方式与1+1保护不同的是,没有故障时备份LSP上的资源可以被其他低优先级的业务流使用,提高了备份资源的使用效率。但是发生故障后,被保护的業務流将抢占备份LSP,原来备份LSP上传输的低优先级业务流会被丢弃。

(3)1(m):n保护

n条工作LSP共用1(m)条备份LSP。如果备份LSP被占用后,其它的工作LSP将失去保护。这种方式通常用于多个工作LSP同时发生故障可能性不大的情况,它的优点是节省备用资源的使用费用。

3.4 两种恢复模型的比较

重新路由的方式在故障发生后建立恢复路径,故障恢复速度相对较慢,但是在网络资源利用方面具有优势。保护交换在故障前已经建立了备份路径,故障恢复速度快,但是资源利用率不高。表1是对两种恢复模型的比较。

表1 两种恢复模型的比较

恢复模型	重新路由		保护交换	
	按需建立	预验合格	1+1	1:1,1:n,m:n
恢复时间	故障后		故障前	
资源占用	不用预留资源	不用预留资源	每条工作路径都预留一条备份路径	1或n条工作路径共享1或m条预留的备份路径
恢复速度	慢,取决于路由表更新、故障信息传递和恢复LSP建立时的信令传递时间	慢,取决于故障信息传递和恢复LSP建立时的信令传递时间	快,取决于故障到修复点的传输时间	快,取决于故障到修复点的传输时间
支持的服务	尽力而为服务(TCP,UDP)	高优先级服务(VPN,实时业务)	严格QoS服务(EF)	松散QoS服务(AF)
恢复路径长度	是当前可用的最短路径	不一定是当前可用的最短路径	路径是由配置完成的,不一定是最短路径	

4. MPLS 恢复周期

MPLS定义了三种恢复周期:MPLS恢复周期、MPLS复原周期和MPLS动态重路由周期。MPLS恢复周期负责检测故障,并将数据流转移到恢复路径上传输;复原周期是在故障修复后,将数据流重新搬回原来的工作路径上;最后一个周期是网络状态稳定后,重新寻找最优路径。复原周期适用于显式路由的情况,而动态重路由周期适用于逐级跳路由。本文主要介绍MPLS恢复周期。

MPLS恢复周期如图2所示,由以下几步组成:

(1)故障检测时间T1——从发生故障到检测到故障的时间。

(2)延缓时间T2——从检测到故障到开始启动恢复动作间的等待时间。设置延缓时间是为了使网络底层的保护措施能够生效,它可以设为0。如果配置成在PSL而不是在检测到故障的LSR等待,那么延缓时间就在通知时间之后。

(3)通知时间 T3——开始传递 FIS 消息到 PSL 恢复操作开始的时间。如果是 PSL 自身检测到故障,那么这个时间为 0。

(4)恢复操作时间 T4——从第一个到最后一个恢复动作的时间。这包括 PSL 和 PML 间的消息交换。显然,重新路由方式要比保护交换方式所需的 T4 时间长。

(5)业务流恢复时间 T5——从最后一个恢复动作到第一个被恢复的数据包到达 PML 的时间。这一时间取决于故障位置、采用的恢复机制以及恢复路径上的传输时延。



图2 MPLS 恢复周期

5. 恢复范围

5.1 本地修复

在网络中并非所有的链路和节点都容易失效,实际上只有某些节点或链路比较脆弱,容易受到攻击,本地修复的目的就是保护某一条链路或某一个节点。它在发生故障相邻最近的上游节点启动恢复动作,这样,故障检测点就是修复点,不需要再采用信令协议将故障信息发送到修复点,缩短了故障传输时间,路径恢复得更快。本地修复分为两种:

(1)链路保护 链路保护是为了保护某一条链路。图3所示为保护 LSR1 和 LSR2 间链路的示意图。备份路径的实现可以采用 LSP 隧道的形式。当数据流转到备份路径时,在原来的标签上压入一个标签放在顶层,备份路径上的数据转发根据顶层标签进行,到达备份路径出口后,将顶层标签弹出。例如在图3中,LSR1 和 LSR2 间的链路使用标签 15,失效后,数据流从 LSR1 转向 LSR3,并被标记为 (21, 15),到达 LSR2 后,弹出顶部的标签,LSR2 仍然使用标签 15 转发。但是数据转发使用的是 {入接口,入标签} 到 {出接口,出标签} 的映射,虽然保留了入标签,但是入接口可能改变了。为了解决这一问题,提出了两种方案。一种是将原始的入接口映射为一个虚拟接口,虚拟接口由配置完成;另一种方法是使用全局标记空间。

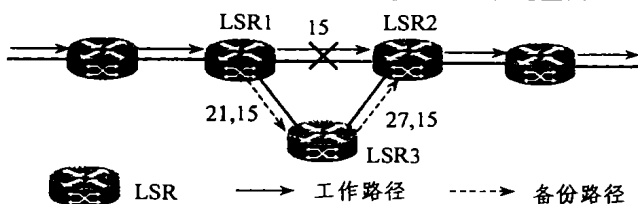


图3 链路保护

(2)节点保护 节点保护是为了保护某些脆弱的节点,当其失效后,整条路径能够迅速地由故障中恢复出来,图4为保护节点 LSR2 的示意图。节点保护仍然使用标签堆栈的方式。但是由于备份路径入口和出口的 LSR 不是相邻 LSR (LSR1 和 LSR3),因此标签映射时会出现问题。如图4中,LSR1 和 LSR2 间使用标签 15,LSR2 和 LSR3 间使用标签 17,LSR2 失效后,LSR1 使用 (23, 15) 将数据流转向 LSR4,到达 LSR3 后,去掉顶部标签,剩下标签 15,而 LSR3 不能识别标签 15。为解决这一问题,IETF 草稿^[8]提出了一种方案。建立工作 LSP 时,在 RESV 消息中包含 Record Route 对象,而 Record Route 对象中有 LSP 经过的每个 LSR ID 及其使用的标签,这样,路径

上的每个节点都知道每条链路所使用的标签。那么,当 LSR1 要将数据流转向备份路径时,它将使用标签 (23, 17)。这种方法依然存在标签使用范围的问题,当然最简单的解决办法就是使用全局标记空间。

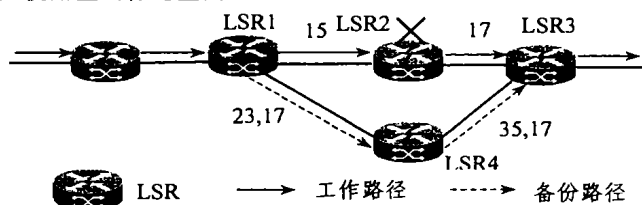


图4 节点保护

5.2 全局修复

全局修复的目的是为了保护整条路径,它通常用于不能预测哪条链路或节点会失效的情况。恢复动作由入口 LSR 启动,通常恢复路径不使用和工作路径相同的链路和节点,这样可以对工作路径中的所有链路和节点提供保护,如图5所示。全局修复中,需要使用 FIS 通知入口 LSR,而故障点一般距离入口 LSR 较远,所以 FIS 消息传递需要较长时间,路径恢复比较慢。

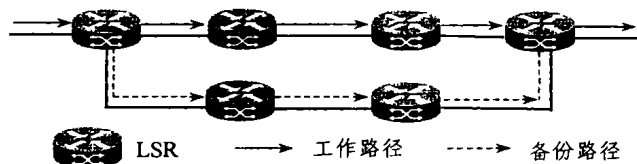


图5 全局修复

结束语 MPLS 在面向无连接的网络中引入了面向连接的技术,它采用简单的标签转发机制,实现了数据的快速、高效传输,被认为是未来骨干网络的主要传输技术之一。MPLS 控制与转发相分离的结构以及对显式路由的支持,都增强了它的路由能力,并且基于标签的转发方式,提高了对数据流控制的灵活性。这些特点对于网络的快速故障恢复都具有很大的优势。而且,MPLS 路径恢复机制具有很大的灵活性,网管员可以根据业务的不同服务类型提供不同级别的保护。例如,可以对实时业务提供预先建立、资源预留的路径保护,而对尽力而为的业务只需采用按需建立的重新路由方式。但是,MPLS 网络的路径保护和恢复机制,目前还处于草案阶段,还有很多问题需要进一步讨论。

参考文献

- 1 Rosen E, Viswanathan A, Callon R. Multi-Protocol Label Switching Architecture. RFC 3031, Jan. 2001
- 2 Andersson L, et al. LDP Specification. RFC 3036, Jan. 2001
- 3 Rosen E, et al. MPLS Label Stack Encoding. RFC 3032, Jan. 2001
- 4 Jamoussi B, et al. Constraint-based LSP Setup using LDP. Internet Draft draft-ietf-mpls-cr-ldp-05. txt, work in progress, Feb. 2002
- 5 Sharma V, et al. Framework for MPLS-based Recovery. Internet Draft draft-ietf-mpls-recovery-frmwk-03. txt, work in progress, July 2001
- 6 Owens K, et al. A Path Protection/Restoration Mechanism for MPLS Networks. Internet Draft draft-chang-mpls-path-protection-03. txt, work in progress, July 2001
- 7 Harrison R. Protection and Restoration in MPLS Networks. Data Connection White Paper, Oct. 2001
- 8 Goguen R, Swallow G. RSVP Label Allocation for Backup Tunnels. Internet Draft, draft-swallow-rsvp-bypass-label-01. txt, work in progress, Nov2000