

MPLS 上游标签分配方式在 PW 保护倒换中的应用

陈春玲,张锦跃

(南京邮电大学 计算机学院,江苏 南京 210003)

摘要:为提高 PW 保护倒换的速度,降低倒换的丢包率,采用 MPLS 上游标签分配的方式来分配冗余 PW 的标签。通过对下游标签分配方式和上游标签分配方式两者间 PW 倒换的对比,来详细说明上游标签分配方式的优点,并通过在标签封装结构中占用一个比特位来标识标签的分配方式。研究和实验表明,在 PW 保护倒换中应用上游标签分配方式来分配 PW 标签,可以提高 PW 切换的速度,降低 PW 切换的丢包率。上游分配 PW 标签时,为冗余的 PW 分配相同的标签,PW 倒换时就可以省去切换 PW 标签的时间,因此可以实现冗余 PW 间快速、平滑的切换。

关键词:MPLS;PW;上游标签分配;保护倒换

中图分类号:TP311.5

文献标识码:A

文章编号:1673-629X(2011)09-0055-04

Application of MPLS Upstream Assigned Label in PW Protect Switchover

CHEN Chun-ling, ZHANG Jin-yue

(College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract:The application of MPLS upstream assigned label in PW protect switchover can improve the speed of switchover. Illustrate the upstream assigned label by the comparison between downstream assigned label and upstream assigned label. And take a bit in label encapsulation structure to indicate label distributed by which mode. Research and experimental results show that the application of MPLS upstream assigned label in PW protect switchover can well reduce handoff delay and packet drop rate. Assigned the same label for redundancy PW when use the mode of upstream assigned label, so that can save the time for the switch for PW label when PW protect switchover.

Key words:MPLS;PW;upstream assigned label;protect switchover

0 引言

MPLS^[1](Multi-Protocol Label Switching)作为一种性价比高、业务能力强的骨干网交换技术,用短而定长的标签来封装网络层分组,直接根据标签进行转发,而必像 IP 一样进行复杂的路由查找和转发^[2]。MPLS 技术易于实现基于 IP 技术的虚拟专用网 VPN^[3-6]业务,满足 VPN 管理和扩展性的需求。PW^[7-9]是运营商网络为用户提供的伪线路服务。MPLS PW 由一对方向相反的 VC-LSP(Virtual Circuit Label Switched Path)组成,并更低一层的隧道承载。一条隧道可以承载成千上万条 PW,甚至更多。PW 转发表分为 PW 发送标签和接收标签。PW 发送流量时,先打上 PW 发送标签,再加上承载 PW 的隧道封装,然后在

指定的连路上发送。当 PW 收到流量时,先剥离隧道封装,露出的分组上带了 PW 标签,根据这个标签查找本地的 PW receive label 表项,在相关表项关联的接口上进行发送。

为了提高网络的可靠性,通常为主 PW 部署一条或多条冗余^[10]的 PW 作为备用 PW。正常情况下,流量只在主用 PW 上转发,当主用 PW 失效时,将流量切换到备用 PW 上进行传送,即任何时候只有一条 PW 被下发到转发平面用于数据转发。若承载的隧道发生异常,则需要将该隧道所承载的 PW 都要进行切换,并切换的速度越快越好。MPLS PW 的建立和维护是通过扩展 LDP 协议实现的。LDP^[11]协议采用下游分配(Downstream assigned)的方式来分配标签,即流量转发的下游设备为特定的转发等价类(Forwarding equivalence class, FEC)分配标签,并将标签与 FEC 的绑定信息通告给上游,其标签空间是“每平台标签空间”(Per platform label space),就是说设备分配的标签在

收稿日期:2011-01-17;修回日期:2011-05-07

作者简介:陈春玲(1961-),男,江西兴国人,教授,硕士,研究方向为软件技术及其在通信中的应用。

该设备中是唯一的。下游设备分配标签建立 LSP 后,在其标签转发表中增加了相应的标签转发表项。对下游设备而言,其为 FEC 分配的标签都是唯一的,当收到从上游发送来的 MPLS 的分组时,根据分组上携带的标签查询其标签转发表得到一个唯一的转发表项,然后根据该转发表项将数据包转发出去。标签由下游设备分配,主备 PW 的 PW 标签由于下游不相同而分配到不同的标签。当发生主备 PW 切换时,不仅要切换其隧道封装,还要切换 PW 私网标签。

如图 1 所示,PW1、PW2 为 PE1 的主备 PW,其隧道是 MPLS 方式的 LSP。业务的转发过程^[9,10]为:在 PE 入口处将收到的原始数据进行封装,先封装 PW 标签 1030,然后封装隧道标签 1024。当承载 PW1 的隧道发生异常时,要将 PW1 切换到 PW2 上。这时不仅要隧道标签从 1024 切换到 1025,还要将 PW 标签从 1030 切换到 1031。切换后再进行数据传送。也就是说,PW 切换时要进行两次封装切换。如果发生异常的隧道承载了多条 PW,切换这些 PW 都要做两次封装的切换,从开始切换到切换完成就会有相当一部分的时延,切换时产生的丢包率就会相对高点。

文中采用 MPLS 上游分配^[12](Upstream assigned) 标签的方式来分配 PW 标签,以提高 PW 保护倒换的速度。

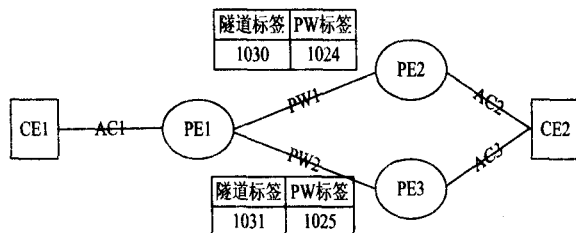


图 1 PW 切换图

1 上游标签分配

上游标签分配是指流量转发的上游设备为特定的转发等价类分配标签,并将标签与 FEC 的绑定信息通告给下游。该标签分配模式的标签空间为“上下文指定标签空间”(Context-Specific Label Space),也就是分配的标签只在其“上下文指定的标签空间”中是唯一的。一个 LSR 维护一个或多个上下文指定标签空间。一般情况下,在“每平台标签空间”中查找该标签,但是上游分配标签时,必须在其“上下文指定标签空间”中查找该标签。“上下文指定标签空间”的一个典型例子就是“每接口标签空间”。当从一个特定的接口收到一个 MPLS 时,其最外层标签必须在该接口的“每接口标签空间”中进行查找,接收该 MPLS 包的接口就是这个包的上下文。

上游分配标签模式与下游分配标签模式二者之间

并不需要协调。一条跨越多个 LSR (Label Switching Router) 的 LSP,可能其中一些 LSR 用下游分配的标签来绑定 FEC,并通告其上游 LSR;而另外一些 LSR 则是用上游分配的标签来绑定 FEC,并通告其下游 LSR;甚至,其中一个 LSR 可能用下游分配的标签来绑定 FEC 并通告其上游 LSR,而用上游指配标签来绑定 FEC 并通告其下游 LSR。如果两个邻接 LSR 为一条 LSP 的两端,则要么上游 LSR 向下游发布一个上游分配的标签,要么下游向上游发布一个下游分配的标签,但不可能是两者都存在。具体使用哪种标签分配模式取决于所用 LSP 的应用。一个特定的标签值可能既是上游分配给一个 FEC 进行绑定,还是下游分配的另一个不相同 FEC 绑定。上游分配标签的模式是可选的:当下游 LSR 不支持时,上游不能使用该模式。

任何要求使用上游分配标签模式的应用都必须明确说明,否则将系统默认采用下游分配标签模式来分配标签。LSR 通过上、下游之间的标签分配协议来通知其对端使用上游或下游分配标签。

1.1 上游的处理

上游的 LSR 设备负责分配与 FEC 绑定的标签。由于上游分配标签不是其入标签,因此不会在上游 LSR 的 ILM 中创建表项。相反,该标签作为上游 LSR 的出标签,会在上游 LSR 的 NHLFE 中添加相应表项,通过查找 NHLFE 表项可将这个标签压入 MPLS 包的标签栈中。如果上游的 LSR 为 LSP 的 Transit 部分,则其 NHLFE 与 ILM 相关联。如果上游的 LSR 是 LSP 的 ingress 部分,则其 NHLFE 直接与 FEC 相关联。举一个简单的例子来说,一条 LSP 经过三个 LSR:LSR1 为 INGRESS,LSR2 为 TRANSIT,LSR3 为 EGRESS。转发数据包时,LSR1 收到数据包后,将包的地址作为 FEC 查找其 FTN (FEC TO NHLFE) 表,得到 FEC 所映射的一个 NHLFE 所对的表项,然后查找该表项,得到相应出接口、下一跳、标签及标签操作 PUSH,将这个标签压到数据包后从对应的出接口转发出去。LSR2 收到这个 MPLS 包后,根据上下文查找所“上下文指定标签空间”相应的 ILM,得到标签操作 SWAP 及一组相应的 NHLFE,然后根据 NHLFE 表项进行相应的操作。LSR3 收 MPLS 包后,处理过程同 LSR2,查 ILM 表时查标签操作 POP,将标签弹掉。

1.2 下游的处理

如果下游 LSR 收到的 MPLS 数据包的最外层标签是上游指配的,则在其“上下文指定标签空间”中查找这个标签,而不是在“平台空间”中查找。下游 LSR 为其上游 LSR 维护一个“上下文指定标签空间”来“保存”上游指配的标签空间。如果上游发布一个绑定 FEC 的标签给下游,下游根据这个标签在“上下文指

定标签空间”所组织的 ILM 表创建一个表项。这个标签可能是其收到的包的最外层标签,也可能是弹掉一层或更多层标签后才露出来的标签。这就意味着下游 LSR 必须能区分一个标签是上游分配还是下游分配的。

如果上游通过 IP 或 MPLS 隧道来转发这个包,则下游通过接收包的隧道确定上游指配标签,也就是说,这个隧道就是该 MPLS 包的“上下文”,在这个隧道的基础上确定其“标签空间”。

2 PW 保护倒换

倒换即切换,将流量由故障的链路切换到可用的链路上^[13]。PW 保护倒换^[10]是指需要进行保护倒换时,若备用的 PW 是可用的,则将正进行流量转发 PW 切换到备用的 PW 上。切换的速度越快越好。

2.1 保护倒换的触发因素

倒换的触发因素^[10]主要分为:外部倒换和信令倒换两部分。

外部倒换是指通过手工配置的命令触发的保护倒换。外部倒换的优先级由高到低为:

- (1)清除倒换(Clear):清除所有的外部倒换命令。
- (2)锁定倒换(Lockout of Protection):数据流锁定在 PW 上传输。
- (3)强制倒换(Forced Switch):强制数据流在备用 PW 上传输。

(4)手工倒换(Manual Switch):手动将数据流从 UP 状态的 PW 切换到 BLOCKED 状态的 PW 上,更新相互切换的 PW 状态。

信令倒换是指通过协议信令触发的保护倒换。可以触发冗余 PW 切换的信令包括 VCCV LSP PING、VC-CV BFD、MPLS OAM(Operation, Administration and Maintenance,操作、管理和维护)和链路失效所产生的接口事件。

2.2 保护倒换的过程

当本地要进行切换,必须先向对端发送切换请求消息,然后等待对端回复的确认信息。只有收到确认信息后才能进行 PW 切换。当对端节点收到 PW 切换请求消息时,先检查要切换的 PW 是否满足激活的条件,如果不满足则丢弃这个切换请求消息;若满足激活条件,则先去激活原处于 UP 状态的 PW 后再激活新的 PW,然后向切换请求端回复一个确认消息。切换时,首先要切换 PW 私标签,然后再切换其公网隧道标签。

任何时候,只有一条 PW 处于激活状态,其被下发到 MPLS 的转发平面用于数据转发。保护倒换时先更新切换的 PW 的状态,通知 MPLS 转发面删除旧的 UP 状态的 PW 表项,再把新的 UP 状态的 PW 下发到转发

平面生成新的转发表项,这样使流量转移到 UP 的 PW 上。

3 方案实现

在现在的方案中,PW 标签与公网 LSP 的标签都是通过 LDP 协议分配得到的,即都是由下游 LSR 分配的。下游分配与 FEC 绑定的标签,并将绑定信息通告给其上游。该标签作为下游 LSR 的入标签,会在其全局 ILM 表中添加相应的表项。由于该标签在下游设备中是唯一的,下游收到上游发来的 MPLS 包后根据最外层标签查找全局 ILM 表能找到一条唯一的表项,然后根据表项进行相应的操作。由于 PW 标签由下游分配的,因此主备 PW 一般是不相同的。PW 保护倒换时更两次封装切换操作,切换的时延随着切换 PW 条数的增多而越来越大。

本方案在原有方案的基础上进行了一些改进。其公网隧道标签还是采用下游分配方式,而对于 PW 标签,文中则采用上游分配标签的方式为 PW 分配标签。通过手工指配为主备 PW 指定相同的标签。下游 LSR 为其上游 LSR 维护一个“上下文指定标签空间”来“保存”上游指配的标签空间,并在“标签空间”对应的 ILM 表中为该标签添加相应的表项。上游选择一条公网 LSP 隧道来转发封装上游分配标签的 MPLS 包给下游。这条 LSP 隧道就是这个上游标签的“上下文”。

当下游 LSR 从 LSP1 隧道收到一个 MPLS 包后,根据“标签上游分配标记”判断包的最外层是否为上游分配。若非上游分配,则根据标签查找“每平台标签空间”对应的 ILM 表,再进行相应的操作。如果为上游分配的,则先根据该 MPLS 报文的“上下文”确定其“上下文指定标签空间”,根据标签查找该“上下文指定标签空间”所对应的 ILM 表进行相应的操作。为了使下游 LSR 能区分出 MPLS 标签栈的最上层标签是否为上游指配的标签,文中在标签头的 EXP 域中占用一个位来标识该标签为上游分配。如图 2 所示。



图2 标签封装结构

(1) Label Value: 长度为 20bits, 用来标识一个 FEC。

(2) Exp: 3bits, 保留, 协议中没有明确规定, 通常用作服务等级。文中用该域的最低位来标识标签是否为上游分配。定义这个位为“标签上游分配标记”。

(3) S: 1bit, MPLS 支持多重标签。值为 1 时表示为最底层标签。

(4) TTL: 8bits, 和 IP 报文中的 TTL 意义相同, 可以用来防止环路。

以图 1 所示的组网为例,隧道入口 PE1 收到发往 CE2 的数据包,根据接收包的接口得到接口绑定的 PW 标签,将 PW 标签压入包中;并根据 PW 标签查找 PW 转发表项得到转发的隧道,由隧道得到公网隧道的标签,将该标签封装到 MPLS 中能过隧道将包透传到隧道的出口 PE2。PE2 收到 MPLS 数据包后,剥离公网标签后,此时 MPLS 包的最外层标签是为上游分配的。PE2 根据接收数据包的隧道确定包的“上下文”,由“上下文”确定“上下文指定的标签空间”;查找该标签空间对应的 ILM 表,得到相应对项,根据表项执行相应操作,并把数据包从 PW 标签关联的出接口转发出去。

4 实验仿真

仿真的硬件环境为:两个 CE 设备、两个 PE 设备。

软件环境为:模拟网络操作系统、通用测试平台软件。

仿真测试的目标:测试本方案是否可行性及保护倒换的时间。

测试组网图:如图 1 所示的组网图。

正确组建 VPWS 方式的 L2VPN 后,显示 VC 信息如图 3 所示,上游 LSR 为主备 PW 分配相同的 PW 标签。

在 PE1 中手动切换 PW:

如图 4 所示,PW 切换成功,说明方案是可行的。

```

[PE1]display mpls l2vc
Total ldp vc : 2      1 up      0 down      1 blocked
Transport  Client      Service  VC      Local      Remote
VC ID     Intf         ID       State   VC Label   VC Label
100       Vlan100     --       up      1024       1025
200       Vlan100     --       blocked 1024       1025
[PE1]
  
```

图 3 VPWS 方式 VC 信息

```

[PE1-Vlan-interface100]mpls l2vc switchover
[PE1-Vlan-interface100]dis mpls l2vc
Total ldp vc : 2      1 up      0 down      1 blocked
Transport  Client      Service  VC      Local      Remote
VC ID     Intf         ID       State   VC Label   VC Label
100       Vlan100     --       blocked 1024       1025
200       Vlan100     --       up      1024       1025
[PE1-Vlan-interface100]
  
```

图 4 手动 PW 切换

分别对两种标签分配模式下的 PW 保护倒换的时间进行测试,得到一系统的数据,对比这些数据,得到

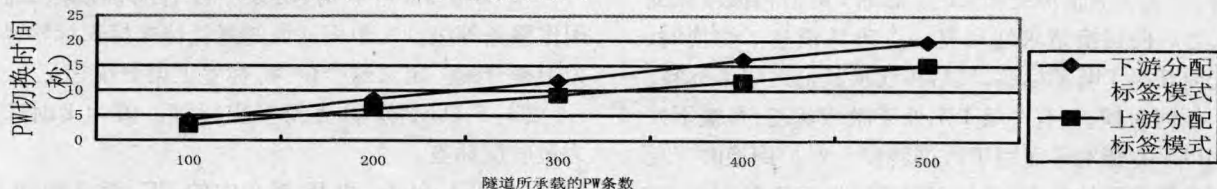


图 5 切换时间对比图

(下转第 62 页)

PW 切换时间对比图(见图 5)。由于 PW 保护倒换的时间是毫秒级别,因此不易对单少 PW 切换进行测试。为便于对倒换时间的测试,使一条公网 LSP 隧道承载多条 PW。

从实验可以得出:MPLS 上游标签分配方式在 PW 保护倒换中的应用是完全可行的,且能提高 PW 保护倒换的速度,从而降低 PW 保护倒换产生的丢包率。

5 结束语

实验表明,PW 保护倒换中应用上游分配标签模式是完全可行的。由于上游可以为多备 PW 分配相同的标签,PW 切换时,只要切换 MPLS 包的公网隧道封装,而不用对 PW 的标签封装进行切换,因此加快了 PW 保护倒换的速度,从而也就降低了由于 PW 保护倒换所产生的丢包率。

参考文献:

- [1] RFC3031 Multiprotocol Label Switching Architecture[S]. IETF,2001.
- [2] 陈业芳. MPLS 技术及发展前景分析[C]//海南省通信学会学术年会论文集. 海南:[出版者不详],2006.
- [3] 奚小获,张 蓉. PWE3 技术概述和参考模型[J]. 飞行器测控学报,2005(3):69-74.
- [4] 李 缓,安 宁. MPLS VPN 的关键技术及其应用[J]. 科技信息,2008(35):497-498.
- [5] 袁 嘉. 基于 MPLS 的 PWE3 实现研究[J]. 现代有线传输,2005(6):51-56.
- [6] Zhou Weihua, Ni Xianle, Ding Wei. A New Architecture of Converged Networks [C]//Proceedings of ICCT. [s. l.]:[s. n.],2003:316-319.
- [7] RFC4096 Transport of Layer 2 Frames Over MPLS[S]. IETF,2007.
- [8] RFC3985 Pseudo Wire Emulation Edge - to - Edge (PWE3) Architecture[S]. IETF,2005.
- [9] RFC 4447 Pseudowire Setup and Maintenance Using the Label Distribution Protocol[S]. IETF,2006.

建筑结构项目,实际需要处理的约束点以及直线数量较大,最后划分的网格一个比较大的问题就是有非常密集的点之间也进行了三角形连接以及四边形划分。这样在实际的图形显示中,原本的三角形在实际的划分效果图上展示的时候就像是一条直线。于是综合实际项目的计算精度和实际考虑,最终评估后得出结论:相邻 10 厘米之内的两个约束点的情况下,在一开始加入 map 处理的时候就将剔除其中一个,以保留线约束为优先。不考虑其对整个建筑结构所产生的影响,是由于其相对于整个建筑结构平面来说是微乎其微的。

另外一个实际应用过程中问题就是三角划分后的网格有些会出现大钝角,为了避免这种情况,在划分三角形的时候对生成的三角形进行评级。我们知道三角形的角度与其三条边相关,如果最大边的平方等于其余两边的平方和,则这个三角形的最大角为直角;如果最大边的平方小于其余两边的平方和,则为锐角;如果最大边的平方大于其余两边的平方和,则为钝角。所以在生成三角形的同时,对三条边进行排序,然后计算出最大角的角度值,如果符合 60 度左右的锐角,就保存此三角形,反之,则排除该三角形。在实际的计算过程中,考虑到程序运行时用到的正弦和余弦计算的费时和费内存的问题,最终采取较小的两边的平方和除以较大边的平方的比值来作为评级的依据,并且通过边长来共同约束大钝角三角形的出现。

5 结束语

文中所述的四边形有限元网格划分法是一种边生成边调整的自动生成方法,在程序执行初期会根据实际点的密集度和相邻距离计算出一个合理的划分间距,随后通过约束数据的预处理和不合理点的剔除,根据实际需要选择三角法、组合法以及矩形法中的一种合适的算法来进行网格划分,最后再对划分后的数据进行优化处理方便后续计算。文中的三种算法是在实际的项目中摸索得出的,具有较高的实用价值,在实际的建筑结构的分析中,近万个约束点以及直线的网格划分只需几分钟即可完成,并且有效地应用于实际的软件中。

文中的算法还是有值得提高的地方,如何更进一步提高划分速度、划分出更加规整的四边形网格、智能地整合三种划分算法将是未来的主要研究方向。

参考文献:

- [1] 胡恩球,张新访,向文,等.有限元网格生成方法发展综述[J].计算机辅助设计与图形学学报,1997,9(4):378-383.
- [2] 关振群,宋超.有限元网格生成方法研究的新进展[J].计算机辅助设计与图形学学报,2003,15(1):1-14.
- [3] 张玉峰,朱以文.有限元网格自动生成的典型方法与研究前瞻[J].武汉大学学报:工学版,2005,38(2):54-59.
- [4] 李娅,赵文.有限元网格自动剖分及优化方法[J].四川工业学院学报,2003,22(3):104-106.
- [5] 王明强,朱永梅,刘文欣,等.有限元网格划分方法应用研究[J].机械设计与制造,2004(1):22-24.
- [6] George P L, Hermeline F. Delaunay's mesh of a convex polyhedron in dimension d. Application to arbitrary polyhedra[J]. International Journal for Numerical Methods in Engineering, 1992(33):975-995.
- [7] Blacker T D, Stephenson M B. Paving: a new approach to automated quadrilateral mesh generation[J]. International Journal for Numerical Methods in Engineering, 1991, 31(4):811-847.
- [8] Lo S H. A new mesh generation scheme for arbitrary planar domains[J]. International Journal for Numerical Methods in Engineering, 1985, 21:1403-1426.
- [9] Zhu J Z, Zienkiewicz O C, Hinton E, et al. A new approach to the development of automatic quadrilateral mesh generation[J]. International Journal for Numerical Methods in Engineering, 1991, 32:849-866.
- [10] Lo S H. Automatic mesh generation and adaptation by using contours[J]. International Journal for Numerical Methods in Engineering, 1991, 31:689-707.
- [11] Lo S H, Lau T S. Generation of hybrid finite element mesh[J]. Microcomputers in Civil Engineering, 1992(7):235-241.
- [12] Knuth D E. The Art of Computer Programming[M]. 2nd ed. Reading, MA: Addison-Wesley, 1975.
- [13] Rivara M C. New longest-edge algorithms for the refinement and/or improvement of unstructured triangulations[J]. International Journal for Numerical Methods in Engineering, 1997, 40:3313-3324.
- [14] 赵熠,张新访,赵建军.前沿法生成四边形网格的改进算法[J].计算机工程与应用,2002,38(9):64-66.
- [15] 成建梅.具特征约束的有限元三角网格自动生成方法[J].长江科学院院报,2003,20(4):39-43.

(上接第 58 页)

- [10] Draft-ietf-pwe3-redundancy-02.txt. Pseudowire (PW) Redundancy[S]. IETF, 2009.
- [11] RFC5036 LDP Specification[S]. IETF, 2007.
- [12] RFC5331 MPLS Upstream Label Assignment and Context-

Specific Label Space[S]. IETF, 2008.

- [13] 陈春玲,张锦跃,陈云.基于 Mobile IP 的 WiMax 网关切换方法的研究与实现[J].计算机技术与发展,2011,21(4):115-119.