



(12) 发明专利

(10) 授权公告号 CN 101636973 B

(45) 授权公告日 2012. 11. 07

(21) 申请号 200880008310. 0

(22) 申请日 2008. 03. 05

(30) 优先权数据

07251059. 7 2007. 03. 14 EP

(85) PCT申请进入国家阶段日

2009. 09. 14

(86) PCT申请的申请数据

PCT/GB2008/000752 2008. 03. 05

(87) PCT申请的公布数据

W02008/110751 EN 2008. 09. 18

(73) 专利权人 英国电讯有限公司

地址 英国伦敦

(72) 发明人 艾伦·麦奎尔 尼尔·哈尔森

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

代理人 李辉

(51) Int. Cl.

H04L 12/46(2006. 01)

H04L 12/56(2006. 01)

(56) 对比文件

CN 1717903 A, 2006. 01. 04,

EP 1134938 A1, 2001. 09. 19,

审查员 聂锦程

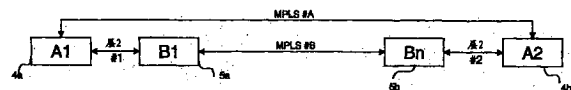
权利要求书 2 页 说明书 15 页 附图 7 页

(54) 发明名称

对业务单元的报头堆栈进行划分的方法、设备

(57) 摘要

一种对数据分组的报头堆栈进行划分的方法,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括可对其指派标签值的标签字段,以使得标签交换节点能够在包括多个网络的通信系统中的标签交换协议的网络路由域中对所述分组进行路由,该方法包括以下步骤:在所述通信系统中的路由域的边缘节点处接收所述分组;以及在所述边缘节点处理所述标签报头堆栈,以增加具有标如下签字段的新的堆栈条目,该标签字段包括一个用于对位于一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目进行定界的值,所述一个堆栈域可以由属于所述通信系统的所述标签交换协议的第一操作域的所述节点来设置,所述至少一个其它堆栈域可以由属于所述通信系统的另一路由域的节点来设置,由此可以实现 MPLS 传输网络方案。



1. 一种对业务单元的报头堆栈进行划分的方法,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括能够对其指派标签值的标签字段,以使得标签交换节点能够在包括多个标签交换协议通信网络的通信系统中的一个标签交换协议通信网络的相应操作域中对所述业务单元进行操作,该方法包括以下步骤:

在所述通信系统中的一个所述标签交换协议通信网络的第一操作域的边缘节点处接收所述业务单元;

在所述边缘节点处理标签报头堆栈,以增加具有标签字段的新的堆栈条目,该标签字段包括一个用于对位于一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目的边界进行标记的值,所述一个堆栈域能够由属于所述通信系统的一个所述标签交换协议通信网络的所述第一操作域的节点设置,所述至少一个其它堆栈域能够由属于所述通信系统的另一标签交换协议通信网络的另一操作域的节点来设置。

2. 根据权利要求 1 所述的方法,其中,所述标签交换协议包括多协议标签交换协议。

3. 根据权利要求 1 所述的方法,其中,所述标签交换协议包括传输多协议标签交换协议。

4. 根据权利要求 1 到 3 中任意一项所述的方法,其中,所述业务单元包括数据分组。

5. 根据权利要求 1 到 3 中任意一项所述的方法,其中,所述操作域包括路由域,并且其中,能够指派所述标签值以使得标签交换节点能够通过所述通信网络的各路由域中对所述业务单元进行路由而对所述业务单元进行操作。

6. 根据权利要求 1 到 3 中任意一项所述的方法,该方法包括以下步骤:对接收所述业务单元的步骤及处理所述报头以增加具有定界值的堆栈条目的步骤重复多于一次,以对所述报头堆栈条目的边界进行标记,从而将所述报头堆栈划分为多个附加堆栈域。

7. 根据权利要求 6 所述的方法,其中,各个附加堆栈域由具有不同定界值的堆栈条目来划分。

8. 根据权利要求 6 所述的方法,其中,各个附加堆栈域由具有相同定界值的堆栈条目来划分。

9. 根据权利要求 1 到 3 中任意一项所述的方法,其中,指派给报头的所述标签字段的并对所述报头堆栈条目进行定界的所述值包括根据标签交换协议的保留值。

10. 根据权利要求 1 到 3 中任意一项所述的方法,其中,具有定界值的所述堆栈条目表示要在如下的点来划分所述报头堆栈,在该点使得与所述边缘节点位于同一路由域中的其它节点仅能够对增加到由增加了所述堆栈定界值的同一路由域所应用的所述报头堆栈中的再多一个附加报头进行操作。

11. 一种被设置为对通信网络中的业务单元的报头堆栈进行划分的设备,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括能够对其指派标签值的标签字段,以使得标签交换节点能够在包括多个网络的通信系统中的所述标签交换协议的网络操作域中对所述业务单元进行操作,该设备包括:

接收机,其被设置为在所述通信系统中的操作域的边缘节点处接收所述业务单元;

一个或更多个处理器,其被设置为在所述边缘节点处理所述标签报头堆栈,以增加具有如下标签字段的新的堆栈条目,该标签字段包括一个用于对一个堆栈域与至少一个其它堆栈域之间的报头堆栈条目进行定界的值,所述一个堆栈域能够由属于所述通信系统的所

述标签交换协议的所述操作域的节点来设置,所述至少一个其它堆栈域能够由属于所述通信系统的另一操作域的节点来设置。

12. 一种移除业务单元的报头堆栈中的边界标记的方法,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括能够对其指派标签值的标签字段,以使得标签交换节点能够在包括多个网络的通信系统中的标签交换协议的网络操作域中对分组进行操作,该方法包括以下步骤:

在所述通信系统的操作域的边缘节点处接收所述分组;

在所述边缘节点处理所述标签报头堆栈,以移除具有如下标签字段的堆栈条目,该标签字段包括一个用于对一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目的边界进行标记的值,所述一个堆栈域能够由属于所述通信系统的所述标签交换协议的第一操作域的所述节点来设置,所述至少一个其它堆栈域能够由属于所述通信系统的另一操作域的节点来设置。

13. 一种包括多个多协议标签交换通信网络的通信系统,各个所述通信网络包括被设置为实现多协议标签交换操作域的多个网络节点,该系统包括:

位于两个所述通信网络之间的边界节点,该边界节点包括:

接收机,其被设置为接收多协议标签交换业务单元;以及

一个或更多个处理器,其被设置为对所接收到的业务单元的多协议标签交换报头堆栈进行处理,以包含如下的定界标签字段值,该定界标签字段值指示了所述两个通信网络的多协议标签交换操作域的操作边界。

14. 根据权利要求 13 所述的系统,其中,所述节点是到所述两个通信网络中的一个通信网络的入口节点,并且其中,所述节点进一步包括一个或更多个处理器,所述一个或更多个处理器被设置为执行以下处理:

在指派了所述操作边界标签字段值之后处理所述报头堆栈,以包含一个或更多个其它多协议标签交换报头字段值,使得位于所述入口节点的所述通信网络的操作域中的其它节点能够对所述业务单元执行一个或更多个操作。

对业务单元的报头堆栈进行划分的方法、设备

技术领域

[0001] 本发明涉及一种将多协议标签交换 (MPLS) 报头堆栈划分为不同域的方法,并且具体地但并不排除涉及一种通过 MPLS 传输网络协议的一个实例来发送符合 MPLS 传输网络协议的另一实例的业务 (在本领域中称为“通过 MPLS 的 MPLS”)的方法。

背景技术

[0002] MPLS 是最初被构想为用于提高互联网协议 (IP) 路由器的分组转发性能的方法。但是,随后 MPLS 被扩展为使用伪线 (Pseudowire, PW) 封装技术来承载其它层网络技术 (例如,异步传输模式 (ATM)、帧中继 (FR)、准同步数字体系 (PDH) 等)。在由 E. Rosen 等于 2001 年 1 月提交到互联网工程任务组 (IETF) 的标题为“Multiprotocol Label Switching Architecture”的 RFC 3031 中阐述了 MPLS 标准的架构,以引证方式将其内容合并于此。该文档的电子拷贝可从 URL :www.ietf.org/rfc/rfc3031.txt 下载。

[0003] 如本领域内所公知的,MPLS 实质上通过预先设置用于确定一个路由器与下一路由器之间的路径的“标签”,而使得能够进行更快的路由判定。这些“标签”实质上是在短分组报头字段中所承载的标签,这些“标签”由交换 / 转发网络节点 (称为标签交换路由器 (LSR)) 提取出来。LSR 被预先设置为将特定标签与 (多个) 特定出向端口相关联,因此,可以对包含该标签的业务进行路由,而不必详细地检查分组报头。这避免了对 IP 层网络地址进行逐跳路由判定,而是沿着由特定标签集合预先确定的路径来发送业务。

[0004] 在本领域中,已知将 MPLS 标签堆栈化作为用于实现 MPLS 隧道化的手段。为了实现 MPLS 隧道化,通常使用外部传输标签在提供商的网络的多个提供商边缘设备之间建立批量 (bulk) 传输标签交换路径 (LSP) (其充当隧道),并且在各批量 LSP 中,使用内部传输标签来标识各个业务流。各个分组可以承载被组织为后进先出堆栈的许多加标签的报头。在正常的通过 MPLS 网络的转发中,LSR 仅处理顶部标签 (即,最外的标签)。在任意 LSR 处,可以 (通过 LSR 执行“进栈”操作) 将加标签的报头加到堆栈中或 (通过 LSR 执行“出栈”操作) 将加标签的报头从该堆栈移除。标签堆栈化使得能够将多个 LSP 聚合 (aggregation) 为针对一部分路由的单个 LSP,这创建了“隧道”。

[0005] 图 1A 示出了 MPLS 报头。MPLS 报头的长度为 32 位,并包括多个功能字段。MPLS 堆栈包括多个 MPLS 报头,图 1B 示出了 MPLS 标签堆栈在一帧中 (或等价地称为在分组中) 的位置。

[0006] 如图 1A 所示,常规的 MPLS 32 位报头包括一个 20 位的标签字段、一个 3 位的 EXP (试验) 字段、一个 1 位的 S (堆栈) 字段及一个 8 位的 TTL 字段。这里,TTL 字段指示了 MPLS 分组的生存时间。S 字段是堆栈字段的底部,其指示了 MPLS 报头在 MPLS 堆栈中的位置,S = 1 指示 MPLS 报头在 MPLS 堆栈的底部位置 (即,最靠近有效负荷的最内位置),如果 MPLS 报头不在 MPLS 堆栈的底部位置则 S = 0。图 1B 示出了如何将 MPLS 标签堆栈条目定位在层 2 (例如,数据链路层) 的层报头之上且位于网络层或层 3 (例如,IP) 报头之下。可以增加其它报头,例如,如果有效负荷数据不是 IP,则这里还可以存在 PW 控制字报头。

[0007] 图 1C 针对 MPLS 堆栈更清楚地示出了在层 2 (例如,数据链路层) 报头之后 LSR 如何首先接收标签堆栈的顶部,而该标签堆栈的底部最后出现,并最靠近层 3 (例如,网络层) 报头。该底部报头的 S 位被设置为 1。全部其它报头的 S 位被设置为 0。

[0008] 图 2 示出了包括三个 MPLS 报头的 MPLS 堆栈的扩展图。图 2 上部示出了如何将 MPLS 的加标签的报头的三层以堆栈形式置于分组报头内。图 2 下部示出了如何使层 2 报头之后跟随两个 $S = 0$ 的 MPLS 报头以及出现在 MPLS 堆栈的底部的 $S = 1$ 的第三 MPLS 报头。

[0009] MPLS 报头的 20 位标签值字段可以提供的 MPLS 标签值的理论数量为 2^{20} 个。但是在实践中,保留了一些标签值,并且其具有特殊含义。如需了解进一步的信息,可参考由 E. Rosen 等于 2001 年 1 月提交到互联网工程任务组标准组织 (“IETF”) 的标题为 “MPLS Label Stack Encoding” 的请求注解 (Request for Comment) RFC 3032,其指示了值 0 到 15 是保留的。以引证方式将该文档的内容合并于本说明书中。

[0010] 当 LSR 接收到加标签的分组时,首先处理位于标签堆栈顶部 (即,最靠近层 2 报头) 的标签值,以查找转发信息。这使得接收 LSR 能够获知可以将分组转发到哪个端口和 / 或获知在转发之前要对 MPLS 堆栈执行的 (多个) 任意操作。LSR 对 MPLS 堆栈执行的操作的示例包括利用另一值来替换顶部标签堆栈条目、和 / 或使条目从标签堆栈出栈、和 / 或替换顶部标签堆栈条目、和 / 或使一个或多个附加条目进栈到标签堆栈中。LSR 可以执行的其它操作包括获知出向数据链路封装以及正确地转发分组所需的任意其它信息。

[0011] 取决于在 MPLS 报头堆栈中指派给 MPLS 标签字段及其它字段的值,接收 LSR 可以按照多种方式来解释标签字段。标签字段例如可以用作转发标签、源标签,或用作指示了 LSR 必须执行的操作的功能标签。

[0012] 当用作转发标签时,MPLS 报头的标签值字段用作针对 LSP 目的地结束点 “地址” 的代理标识符。在该 MPLS 报头位于 LSP MPLS 报头堆栈顶部的情况下, $S = 0$,然而,如果其是堆栈中的唯一 MPLS 报头,则 $S = 1$ 。它用于基于逐跳方式将数据平面内的 MPLS 业务单元转发到目的地 (除了使用 PHP (penultimate-hop-popping,倒数第二跳出栈) 的情况以外,在 PHP 中,在最后一跳完全移除转发报头)。

[0013] 当用作源标签时,20 位标签值字段用作针对 LSP 源结束点 “地址” 的代理标识符。加标签的报头的这种功能性使用表示它不变地作为 MPLS 报头堆栈的最后条目,因此 S 位将被设置为 1。

[0014] 当用作功能标签时,标签值字段标识出当在 LSR 接收到时需要采取的特定动作。如现有技术中所已知的,MPLS 报头标签字段的保留值规定了 LSR 要对接收到的分组执行的特定功能。例如,标签 14 指示了将业务单元的内容传送到控制 / 管理平面以执行进一步动作的路由器 OAM 告警。在这种加标签的报头中,S 位通常设置为 1 (但是,在其它功能性使用的情况下并不总是如此)。

[0015] S. Bryant 提出了通过服务器 MPLS 网络来传输客户端 MPLS 网络的网络架构,并且在 2006 年 10 月 13 日提交到互联网工程任务组的标题为 “Application of PWE3 to MPLS Transport Networks” 的互联网提案中进行了说明。该文档的拷贝可从 <https://tools.ietf.org/id/draft-bryant-pwe3-mpls-transport-00.txt> 得到。图 3A 示出了 S. Bryant 等所提出的针对如何通过一个运营商的 MPLS 网络 (MPLS 网络 “B”) 从另一个运营商的 MPLS

网络 (MPLS 网络“ A”) 来传输业务的问题的现有技术解决方案。该现有技术解决方案包括形成 MPLS 隧道化。在图 3A 中, 在位于 MPLS 网络“ A”的域中的用户边缘节点 CE1 与 CE2 之间创建 MPLS 标签交换路径 (LSP)。通过以太网承载该 LSP。在 CE1 与 PE2 之间设置以太网, 而在提供商边缘设备 PE1 与 PE2 之间设置以太网伪线 (PW), 以将以太网业务从 PE1 承送到 PE2, 还在 PE2 与 CE2 之间设置以太网。通过 IP 或 MPLS PSN 来承载以太网 PW。如果使用了 MPLS, 则不利用 PHP 来配置隧道 LSP。

[0016] 图 3A 所示的以太网传输 PW 可以支持多个业务 LSP 实例, 例如, 业务 LSP 及主干 (Trunk) LSP (其可以承载业务实例的聚合)。各业务实例在实例聚合体内的识别是基于标签, 并且, 可以通过使标签进栈来将业务实例聚合到主干, 而通过使标签出栈来去除聚合 (这里, 按照本领域中已知的含义来使用术语“进栈”及“出栈”)。

[0017] S. Bryant 提出的现有技术方案通过在两个堆栈之间插入以太网报头来分隔两个 MPLS 报头堆栈, 如图 3B 示意性所示。因此在现有技术中已知的是, 使用插入的以太网层网络报头在功能上分隔两个 MPLS 层网络, 即, 提供如下的网络堆栈, 其中 MPLS 在层次上位于以太网之上而以太网进而又在层次上位于 MPLS 之上。该现有技术解决方案分隔了两个 MPLS 堆栈, 并使得将两个 MPLS 网络在功能上完全断开, 例如, 运营商 A 的数据 / 控制 / 管理平面从运营商 B 的数据 / 控制 / 管理平面隔离出来。但是, 该现有技术解决方案具有特定限制, 例如, 这两个网络都必须都支持相同的缓存技术 (即, 服务器 MPLS 网络及客户端 MPLS 网络必须支持相同的插入以太网技术), 并且这必须在 CE 设备之间按照端到端方式。这是因为 PE 设备会期望以太网呈现, 这会成为通过中间服务器 PSN 网络所承载的客户端 PW。另一限制在于, 通过将层 2 技术 (诸如以太网) 插入到 MPLS 堆栈以实现划分, 这会导致附加的操作开销, 例如, 对退化 (即, 1 跳) 以太网层网络的管理。

[0018] 因此, 在现有技术中已知的是, 以太网层沿用户边缘 (CE1 及 CE2) 设备之间的 LSP 来扩展端对端。

[0019] 此外, 在本领域中所已知的, 通过非 MPLS 网络来支持标签交换路径。在以 Nortel Networks Corporation 的名义提出的欧洲专利申请 EP1 134938 中, Barry Hass 描述了通过非 MPLS 网络来支持标签交换路径的系统、设备及方法。该系统要求通过在 LSP 的非 MPLS 域上创建将 LSP 的第一 MPLS 域与第二 MPLS 域连接在一起的 IP 隧道, 来保留在 LSP 上发送的分组的标签堆栈信息。对分组及标签堆栈信息进行封装并通过 IP 隧道进行发送, 并且将 MPLS 标识符包括在封装后的分组的报头中, 使得可以在第二 MPLS 域中识别出分组及标签堆栈信息。

[0020] 本发明的 MPLS 传输网络方案是为了消除现有方法中使用 MPLS 作为传输网络方面的限制, 现有方法要求沿着在传输提供商的 MPLS 网络上两个客户端 MPLS 设备之间的 LSP 实现一致的技术。本发明是为了提供一种使用 MPLS 作为传输网络的方法, 该方法包括以下步骤: 使用第二 (服务器) MPLS 网络的基础设施来对位于第一 (客户端) MPLS 网络中的两个标签交换路由器 (LSR) 进行互联, 即, 实现通过 MPLS 的 MPLS 客户端 / 服务器关系, 其不需要在 MPLS 报头堆栈中插入不同及固定类型的技术以实现两个 MPLS 网络在功能上的分隔。

发明内容

[0021] 本发明的第一方面包括一种对数据分组的报头堆栈进行划分的方法, 所述报头堆

栈符合标签交换协议并包括多个报头,这些报头中的每一个包括可对其指派标签值的标签字段,以使得标签交换节点能够在包括多个网络的通信系统中的所述标签交换协议的网络路由域中对所述分组进行路由,该方法包括以下步骤:在所述通信系统中的路由域的边缘节点处接收所述分组;以及在所述边缘节点处理所述标签报头堆栈,以增加具有如下标签字段的新的堆栈条目,该标签字段包括一个用于对位于一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目进行定界的值,所述一个堆栈域可以由属于所述通信系统的所述标签交换协议的第一操作域的所述节点来设置,所述至少一个其它堆栈域可以由属于所述通信系统的另一路由域的节点来设置。

[0022] 因此,通过标记出操作边界可以对标签堆栈进行定界,使得位于所述通信系统的操作域的边缘或操作域中的一个或更多个节点仅可以对到边界标记符为止(或位于一对边界标记符中)的堆栈的一部分进行操作。操作域的示例包括路由域。

[0023] 在一个实施方式中,该方法使得能够根据本发明来实现多协议标签交换(MPLS)传输网络方案。

[0024] 在一个实施方式中,对接收所述业务单元的步骤及处理所述报头以增加具有定界值的堆栈条目的步骤重复多于一次,以将所述报头堆栈划分为多个附加堆栈域。

[0025] 在一个实施方式中,各个附加堆栈域由具有不同定界值的堆栈条目来划分。

[0026] 在一个实施方式中,各个附加堆栈域由具有相同定界值的堆栈条目来划分。

[0027] 在一个实施方式中,指派给报头的所述标签字段的并对所述报头堆栈条目进行定界的值包括根据标签交换协议的保留值。

[0028] 在一个实施方式中,具有定界值的所述堆栈条目表示要在如下的点来划分所述报头堆栈,在该点使得与所述边缘节点位于同一路由域中的其它节点仅能够对增加到由增加了所述堆栈定界值的同一路由域所应用的所述报头堆栈中的再多一个附加报头进行操作。

[0029] 在一个实施方式中,所述标签交换协议包括多协议标签交换(MPLS)协议。

[0030] 本发明的另一方面是为了提供一种将多协议标签交换(MPLS)报头堆栈划分为多个堆栈控制域的方法,该方法包括以下步骤:

[0031] 在MPLS节点处接收MPLS加标签的分组;

[0032] 所述MPLS节点将网络界定MPLS报头增加到所述分组的MPLS标签堆栈,其中,对该所述MPLS报头的标签字段值指派如下的预定值,该预定值表示要在如下的点之后对MPLS报头的所述MPLS堆栈进行划分,在该点使得所述MPLS节点的网络仅能够对在网络界定MPLS报头之后并在任意其它网络界定MPLS报头之前增加到所述分组的再多一个附加MPLS报头进行操作。

[0033] 本发明的另一方面是为了提供一种在包括多个MPLS网络的通信系统中对多协议标签交换(MPLS)业务进行路由的方法,所述多个MPLS网络包括MPLS网络的层级,所述层级中的各个MPLS网络包括多个MPLS节点,所述MPLS节点能够与处于所述层级中的不同等级的MPLS节点进行接口连接,使得在所述MPLS网络层级的各个层处,所述客户端MPLS网络能够使用所述层级中的另一MPLS网络作为服务器网络,来将业务从所述客户端MPLS网络中的第一MPLS节点路由到所述客户端MPLS网络中的另一MPLS节点,该方法包括以下步骤:

[0034] 所述服务器网络的入口MPLS边缘节点从所述客户端MPLS网络的所述第一MPLS

节点接收 MPLS 业务,所述 MPLS 业务包括多个分组;以及

[0035] 所述服务器网络的入口 MPLS 边缘节点对各个接收的分组进行处理,以对位于可由所述客户端网络中的节点进行操作的堆栈条目与可由所述服务器网络中的节点进行操作的堆栈条目之间的所述多个分组中每一个的所述 MPLS 堆栈进行划分。

[0036] 按照这种方式,通过将特殊保留的标签插入到所述报头堆栈中,在一个 MPLS 网络上透明地传输另一 MPLS 网络。在本发明的一个实施方式中,MPLS 网络的层级包括客户端/服务器网络层级。

[0037] 在一个实施方式中,对 MPLS 业务的所述路由包括由 MPLS 节点对 MPLS 业务执行交换操作。

[0038] 在一个实施方式中,对 MPLS 业务的所述路由包括由 MPLS 节点对 MPLS 业务执行转发操作。

[0039] 在一个实施方式中,对 MPLS 业务的所述路由包括由 MPLS 节点对 MPLS 业务执行路由操作。

[0040] 在一个实施方式中,所述客户端 MPLS 网络能够使用所述层级中的另一所述 MPLS 网络作为服务器网络,在所述客户端网络中将业务从所述第一 MPLS 节点透明地路由到第二 MPLS 节点。

[0041] 在一个实施方式中,上述方法进一步包括以下步骤:

[0042] 所述服务器 MPLS 网络的出口 MPLS 边缘节点从所述入口 MPLS 边缘节点接收所述 MPLS 业务;

[0043] 所述出口 MPLS 边缘节点对各个接收的分组进行处理,以移除由所述入口 MPLS 边缘节点所增加的划分;以及

[0044] 所述出口 MPLS 边缘节点执行路由操作,以将所接收到的 MPLS 业务转发到所述客户端 MPLS 网络的入口 MPLS 边缘节点。

[0045] 在一个实施方式中,在包括多个 MPLS 网络的通信系统中执行该方法,其中,所述多个 MPLS 网络的第一 MPLS 网络中的第一 MPLS 节点能够连接到所述多个 MPLS 网络的另一 MPLS 网络中的入口 MPLS 边缘节点,其中,所述另一网络的入口 MPLS 边缘节点能够连接到另一 MPLS 网络的出口 MPLS 边缘节点,并且其中,所述另一 MPLS 网络的出口 MPLS 边缘节点能够连接到所述第一 MPLS 网络中的第二 MPLS 节点,该对业务进行路由的方法包括以下步骤:

[0046] 使用另一 MPLS 网络作为传输网络在所述第一 MPLS 网络中将业务从所述第一 MPLS 节点路由到第二 MPLS 节点;

[0047] 所述第二网络的所述入口 MPLS 边缘节点从所述第一 MPLS 网络的所述第一 MPLS 节点接收 MPLS 业务,所述 MPLS 业务包括多个分组;以及

[0048] 所述入口 MPLS 边缘节点对各个接收到的分组进行处理,以对位于可由所述第一网络中的节点操作的堆栈条目与可由所述第二网络中的节点操作的堆栈条目之间的所述多个分组中每一个的 MPLS 堆栈进行划分。

[0049] 在一个实施方式中,对各个接收的分组进行处理以对所述 MPLS 堆栈进行划分的步骤包括由入口 MPLS 边缘节点执行以下步骤:

[0050] 增加包括 MPLS 报头的网络定界符,该 MPLS 报头具有指派给标签字段的堆栈划分

值,该网络定界符作为堆栈条目增加到所接收到的分组的 MPLS 堆栈;

[0051] 在所述网络定界符堆栈条目之后将其它 MPLS 报头增加到所述 MPLS 标签堆栈,以将接收到的分组通过另一 MPLS 网络转发到所述出口 MPLS 边缘节点。

[0052] 本发明的另一方面是为了提供一种包括多个 MPLS 网络的通信系统,该多个 MPLS 网络包括:

[0053] 客户端 MPLS 网络及服务器 MPLS 网络的层级,所述层级中的各个 MPLS 网络包括能够与所述客户端及服务器层级中不同等级的节点进行接口连接的多个节点,使得在 MPLS 网络客户端及服务器层级的层中,所述客户端 MPLS 网络能够使用所述层级中的另一 MPLS 网络作为服务器网络,在客户端 MPLS 网络中将业务从第一节点路由到第二节点;

[0054] 入口服务器 MPLS 节点,其被设置为从所述客户端 MPLS 网络的第一 MPLS 节点接收 MPLS 业务,所述 MPLS 业务包括多个分组,所述入口 MPLS 节点包括:

[0055] 处理器装置,其被设置为对各个接收到的分组进行处理,以对可由所述客户端网络中的节点操作的堆栈条目与可由所述服务器网络中的节点操作的堆栈条目之间的所述多个分组中每一个的所述 MPLS 堆栈进行划分。

[0056] 在一个实施方式中,所述多个 MPLS 网络的第一 MPLS 网络中的第一 MPLS 节点能够连接到所述多个 MPLS 网络的另一 MPLS 网络中的入口 MPLS 节点,其中,所述另一网络的入口 MPLS 节点能够连接到所述另一 MPLS 网络的出口 MPLS 节点,并且其中,所述另一 MPLS 网络的出口 MPLS 节点能够连接到所述第一 MPLS 网络中的第二 MPLS 节点,

[0057] 一种使用所述另一 MPLS 网络作为传输网络在所述第一 MPLS 网络中将业务从第一节点交换/转发到第二节点的方法,该方法包括以下步骤:

[0058] 所述第二网络的所述入口 MPLS 节点从所述第一 MPLS 网络的所述第一 MPLS 节点接收 MPLS 业务,所述 MPLS 业务包括多个分组;以及

[0059] 所述入口 MPLS 节点对各个接收的分组进行处理,以对可由所述第一网络中的节点操作的堆栈条目与可由所述第二网络中的节点操作的之间的所述多个分组中每一个的所述 MPLS 堆栈进行划分。

[0060] 本发明的另一方面是为了提供一种对数据分组的报头堆栈去除划分的方法,所述报头堆栈符合标签交换协议并包括多个报头,所述各个报头包括可对其指派标签值的标签字段,以使得标签交换节点能够在通信系统的网络路由域中对所述分组进行路由,所述通信系统包括多个网络,该方法包括以下步骤:

[0061] 在所述通信系统的路由域的边缘节点处接收所述分组;

[0062] 在所述边缘节点处理标签报头堆栈,以移除具有如下标签字段的堆栈条目,该标签字段包括一个用于对一个堆栈域与至少一个其它堆栈域之间的报头堆栈条目进行定界的值,所述一个堆栈域可由属于所述通信系统的所述标签交换协议的第一路由域的节点来设置,所述至少一个其它堆栈域可由属于所述通信系统的另一路由域的节点来设置。

[0063] 通过移除所述报头堆栈中的划分来对所述报头堆栈去除划分。标签堆栈的各个部分返回到其不相交(disjoint)的形式。按照这种方式,在去除划分之后(即,在移除所述堆栈中的边界标记符之后),这两个标签堆栈在物理上或在逻辑上被分开。例如,这两个标签堆栈当在它们的父 MPLS 网络中时在物理上/空间上不相交,和/或当一个堆栈承载于另一堆栈上时在逻辑上不相交。

[0064] 在任意给出的数据分组中,根据一个或更多个协议来附接 (append) 多个报头 (例如,参照图 2),例如,可以存在层 2 和 / 或层 3 报头。符合标签交换协议的报头堆栈的示例是多协议标签交换报头堆栈。

[0065] 可对其指派标签值的标签字段使得标签交换节点能够对网络路由域中的分组执行特定操作,例如,路由、转发等。在本发明的一些实施方式中,并不是全部标签都具有路由 (例如,目的地址 (DA) 代理转发) 语义。一些标签具有源语义,一些标签具有本地动作语义,等等。因此,标签字段使得标签交换节点能够对假设的标签语义采取合适的动作。

[0066] 本发明的另一方面是为了提供一种被设置为实现对数据分组的报头堆栈进行划分的方法的标签交换节点,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括可对其指派标签值的标签字段,以使得节点能够在包括多个网络的通信系统中的所述标签交换协议的网络路由域中对所述分组进行路由,该方法该方法包括以下步骤:

[0067] 在所述节点接收所述分组;

[0068] 在所述节点对标签报头堆栈进行处理,以增加具有如下标签字段的堆栈条目,该标签字段包括一个用于对一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目进行定界的值,所述一个堆栈域可由属于所述通信系统的标签交换协议的第一路由域的节点来设置,所述至少一个其它堆栈域可由属于所述通信系统的另一路由域的节点来设置。

[0069] 本发明的另一方面是为了提供一种包括多个数据分组的通信信号,各个分组包括报头及有效负荷,所述报头包含报头堆栈,该报头堆栈包括全部符合相同标签交换协议的多个报头字段,其中,为至少一个报头字段指派如下的值,该值表示在所述报头之后,可由与先前对所述报头字段的堆栈条目进行操作的符合所述标签交换协议的节点的操作域不同的所述标签交换协议的操作域中的节点,来对符合所述标签交换协议的所述报头的堆栈条目进行操作。

[0070] 所述操作域可以包括路由域、或转发域、或节点可以执行动作的任意其它的合适域。

[0071] 本发明的另一方面是为了提供一种包括多个数据分组的通信信号,各个分组包括报头字段及有效负荷,所述报头字段包含具有多个报头字段的报头堆栈,由与所述多个报头字段中的另一报头字段的标签分配协议不同的标签分配协议来增加所述多个报头字段中的至少一个报头字段,其中,为至少一个报头字段指派如下的值,该值表示在所述报头之后,可由所述标签交换协议的与先前对所述报头字段的堆栈条目进行操作的符合所述标签交换协议的节点的路由域不同的路由域中的节点,来对符合所述标签交换协议的所述报头的堆栈条目进行操作。

[0072] 在本发明的一些实施方式中,所述报头堆栈中的标签字段值由不同类型的标签分配协议来增加。

[0073] 可以由与不同控制域相关联的网络来提供 MPLS 报头字段。在本发明的一些实施方式中,所述 MPLS 报头的堆栈全部属于同一网络方。

[0074] 本发明的另一方面是为了提供一种包括多个多协议标签交换通信网络的通信系统,各个所述通信网络包括被设置为实现多协议标签交换操作域的多个网络节点,该通信系统包括:位于两个所述通信网络之间的边界节点,该边界节点包括:接收机,其被设置为

接收多协议标签交换业务单元；以及一个或多个处理器，其被设置为对所接收到的业务单元的多协议标签交换报头堆栈进行处理，以包含以下定界标签字段值，该定界标签字段值指示了所述两个通信网络的多协议标签交换操作域的操作边界。

[0075] 在一个实施方式中，所述节点是到所述两个通信网络中的一个通信网络的入口节点，并且其中，所述节点进一步包括一个或多个处理器，所述一个或多个处理器被设置为执行以下处理：在指派了所述操作边界标签字段值之后处理所述报头堆栈，以包含一个或多个其它多协议标签交换报头字段值，使得位于所述入口节点的通信网络的操作域中的其它节点能够对所述业务单元执行一个或多个操作。

[0076] 在一个实施方式中，由所述通信网络中的另一节点处理所述定界标签字段值，并且不对由位于所述两个通信网络中的另一通信网络的操作域中的节点指派给所述报头堆栈的标签字段值进行操作。

[0077] 在一个实施方式中，在所述通信网络中接收所述业务单元的节点仅对所述报头堆栈处理到所述定界标签字段值为止，并且忽略由位于所述两个通信网络中的另一通信网络的操作域中的节点指派给所述报头堆栈的任意标签字段值。

[0078] 在一个实施方式中，所述多协议标签交换协议的类型在所述多个通信网络中的至少两个通信网络之间不同。

[0079] 根据本发明的划分方案将所述客户端 MPLS 层网络中的 IP 地址视为无关，并且在所述服务器层 MPLS 网络中并不需要知道所述客户端 IP 地址。相反，本发明在服务器层 MPLS 网络上透明地承载客户端层 MPLS 网络。

[0080] 以上阐述了本发明的各个方面，并且通过所附的独立权利要求、优选实施方式及从属权利要求给出。对于本领域普通技术人员而言显而易见的是，可以按照任意合适的方式来组合各个方面及优选实施方式。

附图说明

[0081] 下面将参照仅作为示例的附图来说明本发明的优选实施方式，在附图中：

[0082] 图 1A 到图 1C 示出了常规 MPLS 报头的各种图；

[0083] 图 2 示出了在 MPLS 报头中如何以堆栈形式设置 MPLS 字段的扩展图；

[0084] 图 3A 示出了现有技术的 MPLS 传输网络方案；

[0085] 图 3B 示出了在图 3A 所示的现有技术方案中，中间技术（以太网）如何分隔两个单独的 MPLS 传输堆栈；

[0086] 图 4A 示出了其中可以根据本发明的一个实施方式实现 MPLS 传输方案的通信系统；

[0087] 图 4B 示出了根据本发明的一个实施方式的 MPLS 传输网络方案的更多细节；

[0088] 图 4C 示出了根据本发明的一个实施方式的划分后的 MPLS 堆栈；

[0089] 图 5A 及图 5B 示出了根据本发明的一个实施方式的划分后的 MPLS 堆栈，其中划分 MPLS 堆栈条目中 $S = 0$ ；

[0090] 图 6A 及图 6B 示出了根据本发明的一个另选实施方式的划分后的 MPLS 堆栈，其中划分 MPLS 堆栈条目中 $S = 1$ ；

[0091] 图 7A 示出了为其提供了根据本发明的一个实施方式的递归堆栈 MPLS 传输方案的

通信系统；以及

[0092] 图 7B 示出了如何多次划分 MPLS 堆栈以针对图 7A 所示的通信系统来实现本发明的一个实施方式。

具体实施方式

[0093] 下面描述发明人当前所构想的本发明的最佳方式及各个优选实施方式。本领域普通技术人员可以理解，明显的是，特定的部件或步骤包含有其存在对于实现本发明而言必要的特定特征，即使并未明确地列出这些特征（这是为了在说明书中出于清楚和简洁的目的，并且认为隐含地包括这些特征，除非明确说明了相反的含义）。所述这里的全部技术术语均保留 RFC 3031 及 RFC 3032 指派给这些技术术语的含义（如上所述），除非另有明确相反的指示。

[0094] 术语“业务单元”表示分组、信元、帧和 / 或任意其它合适的业务成分。这里各种类型的业务单元用作其它类型业务单元的同义概念 (synecdoche)，并且为了本发明的目的而将其视为等同物。为了与使用术语 MPLS “分组”的 RFC 3031 及 3032 保持一致，这里应当一致地使用该术语，但是，本领域普通技术人员可以理解的是，对“分组”一词的引用是对任意适当的等价类型业务单元的引用，除非“分组”一词的上下文明确地指示了其它含义。

[0095] 图 4A 示出了根据本发明的包括多个网络 2 及 3（网络 2 及 3 分别包括多个节点 4a、4b、5a、5b）的通信系统 1。

[0096] 各个网络的节点 4a、4b、5a、5b 被设置为用于实现对接收到的业务分组的标签交换或转发，并且在一个实施方式中，这些节点包括根据标准多协议标签交换 (MPLS) 通信协议而实现的标签交换路由器 (LSR)。

[0097] 根据本发明，在各个网络 2、3 中，至少多个节点 4a、4b、5a、5b 适于与所述多个网络的至少一个其它节点进行接口连接。这里将包括与另一网络的朝向外部的接口的这些节点称为边缘节点。取决于边缘节点对业务流进行标签交换 / 转发操作的方向，边缘节点用作业务流进入到该特定网络中的入口 (ingress) 节点或用作业务流离开该特定网络的出口 (egress) 节点。

[0098] 各个网络 2、3 可以具有与该通信系统中的另一网络相同或不同类型的控制和 / 或管理平面技术。但是在各个网络 2、3 中，全部节点都位于同一操作（例如，路由）中，即，在图 4A 中，节点 4a 及 4b 属于一个操作（例如，路由）域，而节点 5a 及 5b 属于一个不同的操作（例如，路由）域。在本发明的一个实施方式中，形成该通信系统的多个网络中的两个或更多个（或全部）可以具有相同的管理域。在本发明的另一实施方式中，形成该通信系统的多个网络中的两个或更多个（或全部）可以具有不同的管理域。

[0099] 因此，通信系统 1 包括两个或更多个不同的 MPLS 网络，各个 MPLS 网络至少包括一个不同的控制 - 平面实例，并且在一些实施方式中，在各个 MPLS 网络中全部功能部件（诸如控制 - 平面、管理平面及数据平面）是独立的。在本发明的一个实施方式中，这两个或更多个 MPLS 网络属于同一运营方（例如，可能属于同一运营商的不同商业用户线），但是，在另选实施方式中，它们属于不同的运营方（例如，不同的网络运营商）。

[0100] 如上所述，在图 4A 中示出了本发明的一个实施方式，其中，多个网络包括表示为 MPLS 网络 A 及 MPLS 网络 B 的两个 MPLS 网络 2、3。在 MPLS 网络 A 中有两个边缘节点 A1

及 A2,而在 MPLS 网络 B 有两个边缘节点 B1 及 Bn,并且由虚曲线箭头表示的标签交换路径 (LSP)6 从 A1 经由 B1 及 Bn 到达 A2。节点 A1 及 A2 位于网络 A 的操作 (例如,路由) 域中,而节点 B1 及 Bn 位于网络 B 的操作 (例如,路由) 域中。

[0101] 将术语“MPLS 节点”定义为与 RFC 3031 保持一致,RFC 3031 将 MPLS 节点定义为能够进行以下处理的节点:能够运行 MPLS 并且知道 MPLS 控制协议,能够运行一个或更多个路由协议,并且能够基于标签来转发分组。在图 4A 中,节点 4a、4b 及 5a、5b 形成了单独的 MPLS 网络域的一部分。根据 RFC 3031,MPLS 网络域包括执行 MPLS 操作 (诸如路由及转发) 并且还位于一个路由域中的节点的连续 (contiguous) 集合。

[0102] 出于简明的目的,在这两个 MPLS 网络 2、3 的每一个中,在图 4A 中示出了位于 MPLS 网络中的仅两个边缘节点 A1 及 A2 (例如,LSR),以及仅两个边缘节点 B1 及 Bn (例如,也是 LSR)。但是,本领域普通技术人员可以理解的是,在任意实际部署中,各 MPLS 网络很可能包括多个节点,并且若干个节点可以提供核心节点及边缘节点功能。

[0103] 在本发明的一个实施方式中,对网络 A 的管理控制与对网络 B 的管理控制不同,例如,网络 A 处于第一网络运营商“A”所控制的域中,而对网络 B 的管理控制处于另一网络运营商“B”所控制的域中。但是在另选实施方式中,运营商实体“A”及“B”可以相同。

[0104] 考虑以下这种情况:其中,运营商 A 会发现通过运营商 B 的 MPLS 网络沿 LSP 6 从节点 A1 经由节点 B1 及 Bn 到另一节点 A2 来发送业务 (MPLS 分组后的数据) 是有利的。这要求分组经由入口边缘节点 B1 进入到 MPLS 网络 B 中并在出口边缘节点 Bn 处离开 MPLS 网络 B。对于本领域普通技术人员而言显而易见的是,在实践中 LSP 6 在节点 B1 与 Bn 之间可以穿过运营商 B 的网络中的多个节点。

[0105] 常规上,当在 MPLS 网络中的 LSR 之间发送 MPLS 业务时,各个 LSR 对包括 MPLS 标签信息的路由表 (该路由表称为标签信息库 (LIB)) 进行查询。各个 LSR 使用接收到的 MPLS 分组入口接口及入口标签来对 LIB 执行查找操作,这使得能够确定出口接口及出口 MPLS 标签,然后,LSR 在经由 LIB 所指示的出口接口来转发接收到的分组之前,用出向标签来替换入向标签。虽然各个标签仅与相邻 LSR 之间的特定链路有关,但是集中起来 LSR 表将创建 MPLS 网络上的标签交换路径。MPLS 堆栈包括多个 MPLS 标签,常规上仅有最外 (最后) 的标签用于转发。通过在 LSR 标签表中指示合适的条目来使最外的 MPLS 标签“出栈”(即,移除)。这使得通过在出口 LSR 处的 LIB 中提供合适的指示而在该出口 LSR 移除一个或更多个 MPLS 标签。但是,针对使运营商 A 按照在这两个 MPLS 网络中独立地使用 S 位 (并且实际上全部其它功能) 的方式将运营商 B 的 MPLS 网络用作传输网络的情况,要求某些手段来防止与网络 A 相关联的堆栈条目被与网络 B 相关联的 LSR 操作。这使得能够在用作客户端的网络 A MPLS 节点与用作服务器的网络 B MPLS 节点之间建立客户端/服务器关系。如图 4A 所示,这使得 MPLS 节点 A1 与 MPLS 节点 A2 直接地成为对端,而 MPLS 节点 B1 及 Bn 建立针对运营商 A 的业务的服务器传输能力。

[0106] 这要求在 MPLS 堆栈中存在划分机制,使得运营商 B 的边缘及核心 MPLS 节点 (B1、B2、...、Bn) 并不具有对由运营商 A 的客户端 MPLS 网络中的 MPLS 节点 (A1、A2) 所指派的标签进行控制的能力。

[0107] 本发明的这种划分机制使得在这两个 MPLS 网络 2、3 之间能够建立这种客户端/服务器关系,现在简要地参照图 4C,并且能够解决与将来自网络 B 的新的 MPLS 报头 8b 增

加到包括已由网络 A 所应用的那些 MPLS 报头 8a 的 MPLS 堆栈 11 中相关联的问题。本发明所提出的堆栈划分 9 用作这两个 MPLS 堆栈控制域之间的缓冲区或网络定界符,即,堆栈划分 9 对图 4C 所示的网络堆栈条目进行划分,并且使得能够针对这两个 MPLS 网络中的每一个在 MPLS 堆栈条目中独立地设置 S 位。

[0108] 本发明的潜在优点在于,设置了 MPLS 传输网络方案,其中本发明在关于将客户端设备与服务器设备互联的段层 (section layer) 技术的类型方面不同于习惯上的做法。虽然在一个实施方式中,使用同一通信技术来支持作为 LSP 的各端的客户端设备与服务器之间的通信,但是在另选实施方式中并不一定如此。

[0109] 在图 4B 中,该附图被设置为使得与图 3A 所示的现有技术方案的差别更明显,但是该通信系统的元件被标记为与图 4A 所示的元件保持一致。附图中的图 4B 示出了在本发明的一个实施方式中,支持边缘节点 A1 与 B1 (图 3A 中的 CE1 与 PE1) 之间通信的技术 (即,通信协议) 如何与在 B1 (PE1) 与 Bn (图 3A 中的供应商设备 PE2) 之间所使用的技术不同,并且示出了该技术可以与在 Bn (PE2) 与 A2 (图 3A 中的用户设备 CE2) 之间所使用的技术不同。因此在图 4B 中,在 A1 与及 B1 之间使用第一通信协议 (与在图 4B 中示为层 2#1 的第一类型段层技术相对应),而在 Bn 与 A2 之间使用不同的段层通信协议 (在图 4B 中示为层 2#2)。在每种情况下,在 B1 与 Bn 之间并不存在中间的端到端的层网络技术——网络 A 的 MPLS 业务并不被来自网络 B 的 MPLS 业务的任意中间技术报头所分隔。虽然在图 4A 及图 4B 中将其称为层 2,但是“层 2#1”及“层 2#2”不一定总是表示 OSI 协议层级。在本发明的一些实施方式中,取决于本发明的具体实施方式,根据实际情况,段层表示任意合适的面向连接的电路交换 (CO-CS)、面向连接的分组交换 (CO-PS) 或无连接的分组交换 (CL-PS)。例如,本领域普通技术人员可以使用术语“层 2”来表示 CO-PS 层,而术语“层 3”可以用于表示 CL-PS 层。

[0110] 在本发明的一些实施方式的上下文中,术语“段层”表示网络堆栈的最底部 (即,实际层或物理层),其中,需要将信息 (来自高层的二进制信号) 编码 (字典映射) 为适于调制到电磁波上的形式 (该电磁波可以在诸如铜或光纤的固体物理介质上引导,或可以包括例如作为无线电波而传播的自由空间电磁波)。因为 MPLS 并不具有段层规范,因此 MPLS 依赖于在 MPLS 层的下方具有一些其它层网络技术 (即使是诸如点到点单跳网络的劣化网络)。在本发明的一些实施方式中,在这些低层网络自身创建了多跳交换连接的意义上来讲,这些低层网络是“真实的”网络。

[0111] 在图 4C 中,在其标签已被网络 A 中的设备指派了值的堆栈条目与其标签已被网络 B 中的设备指派了值的堆栈条目之间划分单个 MPLS 堆栈。所示的 MPLS 分组的报头 11 包括层 2 报头元素 7、层 3 报头元素 10 及 MPLS 报头堆栈 12。在这个实施方式中,MPLS 报头堆栈 12 包括两个 MPLS 网络报头 (8a 及 8b) 及网络定界符堆栈条目 9,该网络定界符堆栈条目 9 在分别属于 MPLS 网络 A 和 MPLS 网络 B 的不同路由域的堆栈条目集合 8a 与 8b 之间提供了功能划分。网络定界符堆栈条目 9 使得这两个层 MPLS 网络能够例如通过按照客户端 / 服务器模式工作而实现网络层级。

[0112] 按照这种方式,本发明使得单个 MPLS 堆栈能够承载符合在各个网络中所指派的不同值的 MPLS 报头,并且使得 MPLS 网络能够在功能上保持独立。这里,术语“层 2”及“层 3”可以表示 OSI 层,和 / 或,在一些实施方式中,“层 2”仅表示任意合适的面向连接的分组

交换层技术,而“层 3”表示任意合适的无连接的分组交换层技术。

[0113] 本发明寻求对预定值进行限定,该预定值被指派给该 20 位的 MPLS 标签字段并且单独或与其它 MPLS 报头字段相结合地指示了 MPLS 报头堆栈被划分为子划分(sub-partition),这些子划分分别包含由位于特定 MPLS 网络域中的设备独立地指派的报头值。通过确保各个 MPLS 网络域并不试图将这些值重新指派为在另一 MPLS 网络域中设置的报头字段,减轻了和 / 或消除了现有技术中针对 MPLS 传输网络的上述限制。有利的是,通过将 MPLS 报头堆栈划分为不同的网络域,能够递归使用 MPLS 报头堆栈,使得可以分层级地提供多深度传输网络,如附图中的图 7A 及图 7B 示意性所示并如稍后所述。

[0114] 本发明包括将堆栈划分值插入到 MPLS 堆栈条目的标签字段中,该堆栈划分值具有将 MPLS 堆栈分隔为两个或更多个单独的堆栈控制域的功能,各个堆栈控制域与单独的 MPLS 网络相关联。按照这种方式,该堆栈划分值用作网络定界符。通过具有堆栈划分标签值的网络定界符堆栈条目,将其值已被一个网络运营商的控制或管理平面所指派的堆栈条目与其值已被另一网络运营商所指派的那些条目分隔开。因此,根据本发明的网络定界符包括其标签字段具有已经由服务器 MPLS 网络域所指派的堆栈划分值的 MPLS 堆栈条目。

[0115] 通过客户端 MPLS 网络与服务器 MPLS 网络之间的自适应功能而插入网络定界符,该自适应功能由服务器层网络实现。将来自客户端 MPLS 网络(再参照图 4A,即网络 A)的 MPLS 分组提交给服务器 MPLS 网络(图 4A 中的网络 B),而位于服务器层 MPLS 边缘处的入口节点 B1 将两个新的 MPLS 报头附接到 MPLS 堆栈。第一 MPLS 堆栈条目包含网络定界符(网络定界符包括具有预定堆栈划分值的 MPLS 报头),而第二报头是用于通过服务器层网络进行转发的“正常”MPLS 报头。在服务器 MPLS 网络 B 的出口边缘节点(如图 4A 中示出为 Bn)处,移除这些堆栈条目,并且将客户端 MPLS 分组提交给客户端 MPLS 网络 A。在出口边缘节点 Bn 使它的最终转发报头出栈之后,出口边缘节点 Bn 仅能够“看到”“网络定界符”堆栈条目。展现出的网络定界符标签告知边缘节点 Bn 下一个报头是属于一个不同 MPLS 网络的 MPLS 分组。

[0116] 因为在正常的转发中沿 LSP 的各个节点并不增加新的加标签的报头,所以如果节点 B1 对这两个报头应用了“堆栈定界符”+“正常转发”,则在 B2 与 B3 之间、B3 与 B4 之间交换“正常转发”报头,以此类推,直到节点 Bn 接收到该分组为止。这表示沿着从 B1 到 Bn 的 LSP,堆栈深度保持为 2,而仅在 B1(增加“堆栈定界符”的节点)及 Bn(移除“堆栈定界符”的节点)观察到“堆栈定界符”报头。在一些实施方式中,将更多的报头增加到 B1 与 Bn 之间的分组路由(en-route),但是,这会在增加(多个)报头的节点之间(例如,在位于 B1 与 Bn 之间的一些任意中间节点 Bk 与 Bm 之间)得到更低的 LSP。中间节点 Bk 及 Bm 分别增加 / 移除(多个)这些附加报头,使得例如当分组抵达 Bn 时这些附加报头并不出现。

[0117] 在 5A 图及图 5B 中示出了上述示例。在图 5A 及图 5B 中,示出了两层 MPLS 网络报头堆栈的示例。图 5A 及图 5B 示出了在堆栈条目的标签字段中存在预定“划分”值以及 $S = 0$ 位,这指示了网络定界符以上的 MPLS 报头与网络定界符以下的 MPLS 报头属于不同的 MPLS 网络。这些报头中的各个报头通常受到不同的管理控制。

[0118] 在图 5A 中,在 MPLS 堆栈中仅有一个 $S = 1$ 堆栈条目,这与诸如在 RFC 3031、RFC 3032 等中总结的当前 MPLS 标准一致,并且在网络 B 的划分中,针对标签堆栈条目(#B1 及 #B2) 而将 S 位设置为 0。图 5B、图 6A 及图 6B 示出了允许在与特定 MPLS 网络相关联的各个

堆栈控制域中设置一个 $S = 1$ 位的另选实施方式。

[0119] 在图 5B 中,在 MPLS 堆栈的各个划分中,对最靠近 MPLS 堆栈的层 3 报头结束位置的堆栈条目指派 $S = 1$ 位,并且对网络定界符指派中间 $S = 0$ 值。图 6A 及图 6B 示出了其中网络定界符不是由 $S = 0$ 位表示而是与 $S = 1$ 位相关联的另选实施方式,但是,本发明的不同实施方式可以对网络定界符的 S 位指派不同的值。

[0120] 再参照图 4A,现在针对从左到右的业务流(如该附图中的带箭头点线所示),更详细地说明根据本发明的一个实施方式的 MPLS 传输方案。对于双向业务,可以在相反方向提供类似的功能。为了实现根据本发明的传输方案,运营商 B 在节点 B1 与 Bn 之间建立(通过任意机制)标准的 MPLS LSP(这里称为 LSP_X)。在正常情况下 LSP_X 具有加单个 MPLS 标签的报头,并且转发标签通常(但不一定总是如此)在各个跳被交换。除了包含转发标签之外,将另一个 MPLS 报头字段编码为: $S = 1$, $EXP =$ 被认为为了该 LSP_X 最终要承载的来自网络 A 的 MPLS 客户端分组的目的而提供足够的性能的任意值, $TTL =$ 足够用于从 B1 到达 Bn 的任意长度。在一个实施方式中,PHP 用在 LSP_X 上,但是,这并不是发明人所构想的本发明的最佳模式。这是为了确保正常的 MPLS 转发报头,具体地说其中最后一跳转发标签在节点 Bn 可见。这提供了对节点 Bn 接收到的哪些分组与 LSP_X(而不是在节点 Bn 终止的任意其它 LSP) 相关联的清楚指示。

[0121] 按照本领域普通技术人员已知的任意常规方式来设置 B1,以期望在节点 B1 的某被输入端口处接收来自运营商 A 的节点 A1 的 MPLS 分组。来自节点 A1 的 MPLS 分组可以具有允许的任意标签值、 S 值、 EXP 值及 TTL 值。由运营商 B 的 MPLS 网络通过使用本发明的这个实施方式的 MPLS 传输方案,按照层级的方式(例如,客户端/服务器)来透明地承载这些 MPLS 分组。连接节点 A1 及 B1 的服务器段层是能够按照层级的方式(例如,客户端/服务器)来透明地承载 MPLS 分组的任意合适的通信技术,例如,为本领域普通技术人员已知的以太网。用于 A1 与及 B1 之间的通信技术的该“连接”在功能上在 A1 及 B1 处完全终止,即具体地说,该“连接”并不通过运营商 B 的 MPLS 网络延伸。

[0122] 在节点 Bn 与 A2 之间还建立另一服务器段层技术“连接”,并且其满足与上述 A1 与 B1 之间的服务器层连接相同的要求。虽然在本发明的一些实施方式中, A1 与 B1 之间的服务器段层技术以及 Bn 与 A2 之间的服务器段层技术不同,但是在其它实施方式中,在 A1 与 B1 之间以及在 Bn 与 A2 之间提供相同的服务器段层技术。通常,在 A1 与 B1 之间以及在 Bn 与 A2 之间建立服务器段层技术连接的方式对根据本发明的 MPLS 传输网络方案的设置没有影响。

[0123] 再参照附图中的图 4A,当 MPLS 分组从节点 A1 到达节点 B1 时,节点 B1 对接收到的分组的 MPLS 堆栈至少执行以下操作:

[0124] (i) 将包括网络定界符的 MPLS 报头增加到来自节点 A1 的 MPLS 分组;并且

[0125] (ii) 在网络定界符报头之后增加具有 B1 在正常情况下将指派的类型的另一 MPLS 报头,以沿上述 LSP_X 将分组从 B1 转发到 Bn。

[0126] B1 所增加的网络定界符包含被指派了堆栈划分值的标签字段。这有效地指示了 MPLS 网络堆栈的该部分已经终止而在网络定界符之后是属于不同 MPLS 网络的 MPLS 堆栈条目。由网络 B 的入口节点 B1 来为网络定界符堆栈条目的其它字段指派合适的值,诸如 S 位(例如, $S = 0$ (推荐),虽然也可以使用 $S = 1$)、 EXP 字段(例如, $EXP =$ 任意)及 TTL 字段

(TTL = 1(推荐))。指派 TTL 字段的值 (TTL = 1(推荐), 虽然可以使用其它 TTL 值), 以防止在分组被错误传送的情况下不正确地转发该分组。

[0127] 还针对网络 B 中的 MPLS 报头按照正常的方式来为由 B1 增加到 MPLS 堆栈的其它 MPLS 报头来指派值, 例如, S 位被设置为 $S = 0$ (推荐, 虽然在理论上如果这是 B1 与 Bn 之间的网络 B 中的正常转发报头的堆栈底部, 则可以使用 $S = 1$), EXP 字段值被设置为被认为为了该 LSP_X 要承载的来自网络 A 的 MPLS 客户端分组的目的是提供足够的性能的任意值, 而 TTL 字段值被设置为足够用于分组从 B1 到达 Bn 的任意值。

[0128] 然后, 通过运营商 B 的网络 (在 LSP_X 中) 透明地传输来自节点 A1 如上所述被封装的 MPLS 分组, 直到它们抵达节点 Bn 为止。当节点 Bn 移除 LSP_X 的正常 MPLS 转发报头时, 这会暴露出网络定界符堆栈条目。这指示了在 MPLS 堆栈中在网络定界符堆栈条目之后还存在属于另一 MPLS 网络 (即, 在这种情况下为运营商 A 的 MPLS 网络) 的一个或更多个 MPLS 报头的其它集合。然后, 节点 Bn 移除网络定界符堆栈条目, 并且因为之前的设置, 所以能够获知可以从哪个端口将其发送出节点 Bn, 以到达运营商 A 的网络中的节点 A2。

[0129] 本发明的一个实施方式使用标签信令协议 (例如, LDP 或 RSVP-TE), 来在节点之间 (例如, 从节点 Bn 到节点 B1) 分配标签并设置这些节点。可以在相关 MPLS 网络中按照这种方式来分配全部标签 (包括堆栈划分标签)。在其它实施方式中, 可以通过管理和 / 或 OSS 技术执行标签分配及设置。这些是本领域普通技术人员已知的常规技术。

[0130] 附图中的图 7A 及图 7B 示出了可以如何扩展本发明、以表示出包括客户端 / 服务器关系层级的 MPLS 网络的多层情况。如图 7A 所示, 来自运营商 A 的网络的 MPLS 业务使用运营商 B 的 MPLS 网络作为传输网络, 运营商 B 的 MPLS 网络进而使用运营商 C 的 MPLS 网络作为传输网络, 运营商 C 的 MPLS 网络进而使用运营商 D 的 MPLS 网络作为传输网络。图 7B 示出了多次划分后的 MPLS 堆栈, 其中示出了多个堆栈划分 #1、#2、#3。这通过本发明所提出的针对 MPLS 业务的数据 - 平面封装技术而使得各个网络运营商的 MPLS 堆栈条目在 MPLS 堆栈中保持为串接在一起。

[0131] 参照本发明的两层实施方式, 使用对本领域普通技术人员已知的任意常规技术 (诸如以上简要说明的), 通过由网络管理和 / 或 MPLS 信令在各个网络域中恰当地设置入口 LSR 及出口 LSR, 而将 MPLS 堆栈划分值附接为网络定界符堆栈条目。

[0132] 本发明的一个实施方式并不强制要求在网络 A 中使用的堆栈划分值与在网络 B 中使用的任意堆栈划分值之间的关系。在针对三层或更多层 MPLS 传输网络而实现的本发明的各个实施方式中, 也不要求在用于网络定界符的任意一个堆栈划分值之间的任意关系 (例如, 如图 7 及图 7B 所示, 其中, 按照递归的堆栈方式所使用的本发明的 MPLS 传输网络方案 (例如, 经由通过 MPLS 的 MPLS 的 MPLS (MPLS-over-MPLS-over-MPLS), 等))。

[0133] 在本发明的一个实施方式中, 将全局保留值 (即, 在保留的标签集合 0-15 的值中的一个) 指派给全部 MPLS 网络所使用的堆栈划分值。但是, 使用全局保留值对于实现本发明并不是必要的。任意值均可用于堆栈划分值, 只要将该值指派给了 LSP 的 MPLS 网络能够确保在承载该堆栈划分值的 LSP 的上下文内可以理解该堆栈划分值即可。在该 MPLS 网络中, 堆栈划分值可以是唯一的, 也可以不是唯一的。

[0134] 对于本领域普通技术人员而言, 对这里所述的本发明各个特征的各种修改及功能上的等同物是显而易见的, 并且除非说明书明确地排除, 否则在说明书中旨在隐含地包括

这种修改的及功能上等同的特征,同样,对于实现本发明而言是必要的但出于清楚和简要的目的而省略的任意特征,对于本领域普通技术人员是显而易见的。

[0135] 在本发明的一些实施方式中,术语“技术”可以指段层技术,除非该术语的上下文表示相反含义。例如,在本发明的一个实施方式中,“支持在边缘节点之间通信的通信技术”可以指“段层”技术。

[0136] 这里,术语“多协议标签交换的”或“多协议标签交换”及首字母缩写“MPLS”用于表示标准协议,并且隐含地表示具有类似的报头结构限制的任意变型协议,例如,传输 MPLS(TMPLS, TMPLS 由国际电信联盟 (ITU) 所支持,更多关于 TMPLS 的信息可以从网站 <http://www.transport-mpls.com> 得到) 等。因此,只要在上述说明书中合适,应当认为术语“MPLS”包含了与诸如 T-MPLS 等 MPLS 协议相关联的变化及衍生协议。

[0137] 术语“业务单元”、“分组”、“信元”及“帧”都旨在彼此用作同义概念,并且为了本发明的目的而将它们视为等同物,虽然以上引用的 RFC 3031 及 RFC 3032 使用这里为了一致而使用的术语“MPLS 分组”。

[0138] 对于本领域普通技术人员而言显而易见的是,这里使用“层 2”及“层 3”可以表示 OSI 层 2(数据链路)及层 3(网络)技术、和/或表示无连接的分组交换及面向连接的分组交换(或面向连接的电路交换)通信协议。

[0139] 这里重复摘要文本,以构成说明书的一部分:

[0140] 一种对数据分组的报头堆栈进行划分的方法,所述报头堆栈符合标签交换协议并包括多个报头,这些报头中的每一个包括可对其指派标签值的标签字段,以使得标签交换节点能够在包括多个网络的通信系统中的标签交换协议的网络路由域中对所述分组进行路由,该方法包括以下步骤:在所述通信系统中的路由域的边缘节点处接收所述分组;以及在所述边缘节点处理所述标签报头堆栈,以增加具有标如下签字段的新的堆栈条目,该标签字段包括一个用于对位于一个堆栈域与至少一个其它堆栈域之间的所述报头堆栈条目进行定界的值,所述一个堆栈域可以由属于所述通信系统的所述标签交换协议的第一操作域的所述节点来设置,所述至少一个其它堆栈域可以由属于所述通信系统的另一路由域的节点来设置,由此可以实现 MPLS 传输网络方案。

现有技术

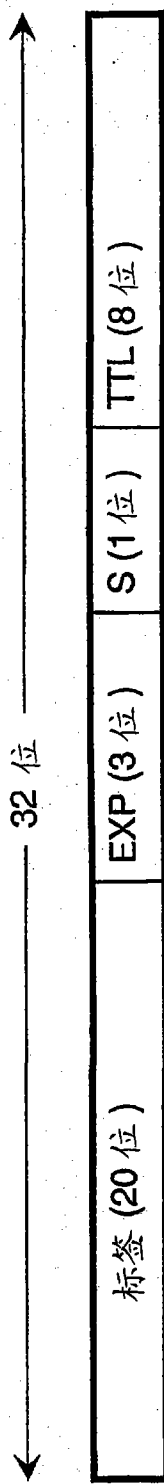


图 1A

MPLS 报文格式

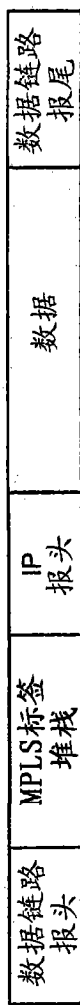


图 1B



图 1C

现有技术

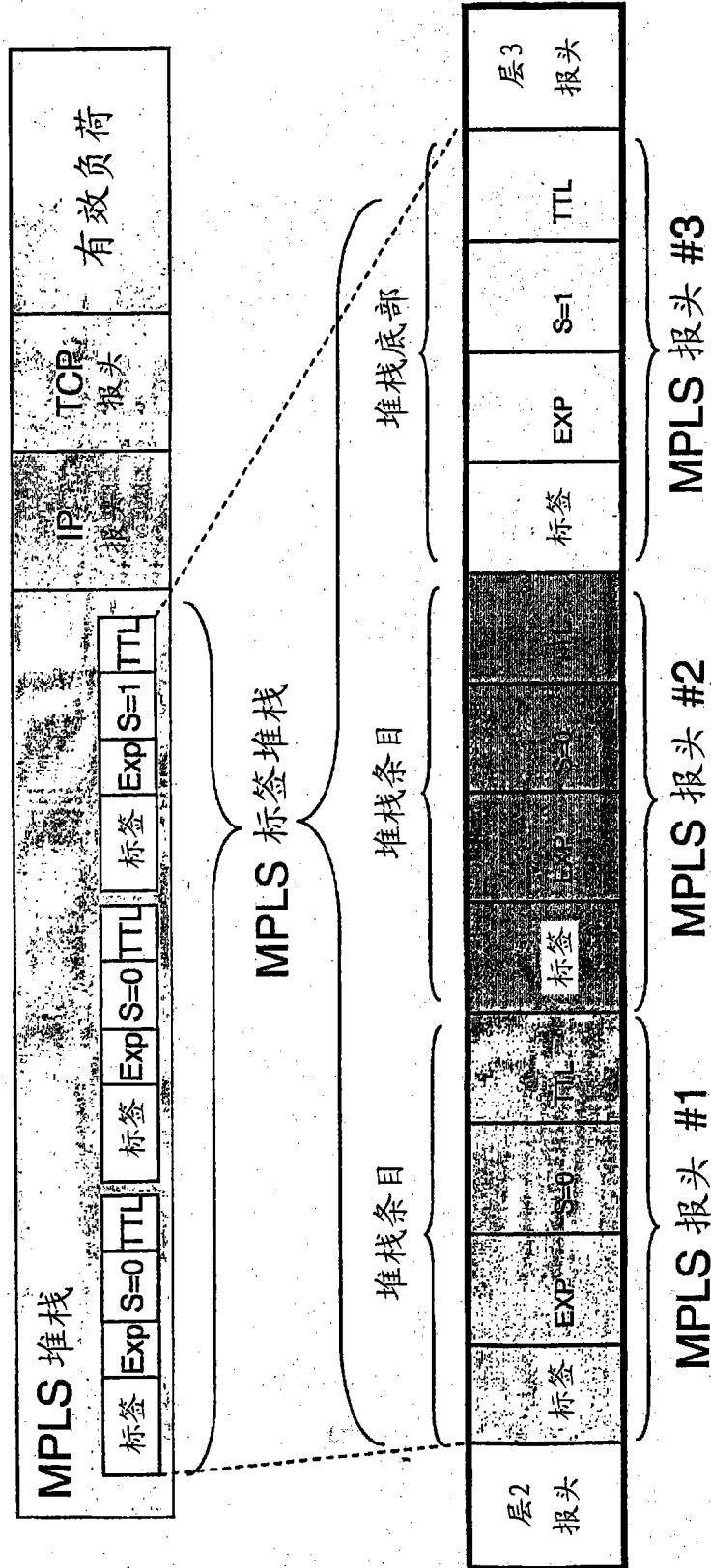


图 2

现有技术

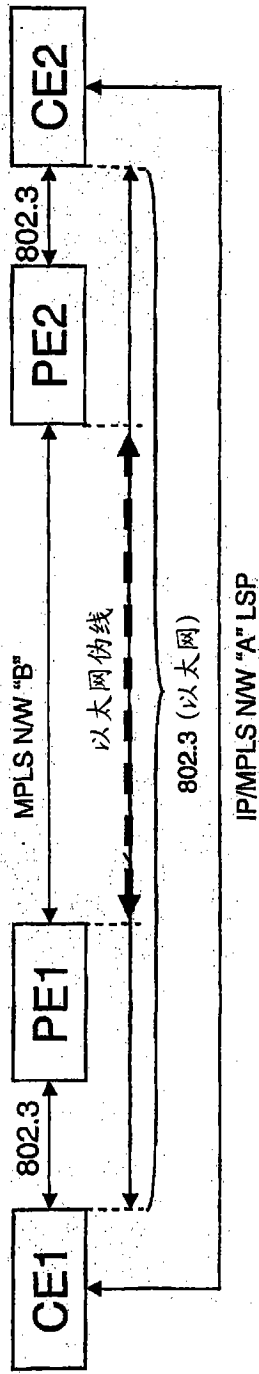


图 3A

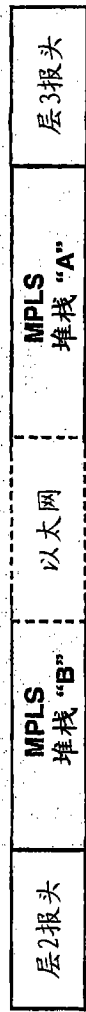


图 3B

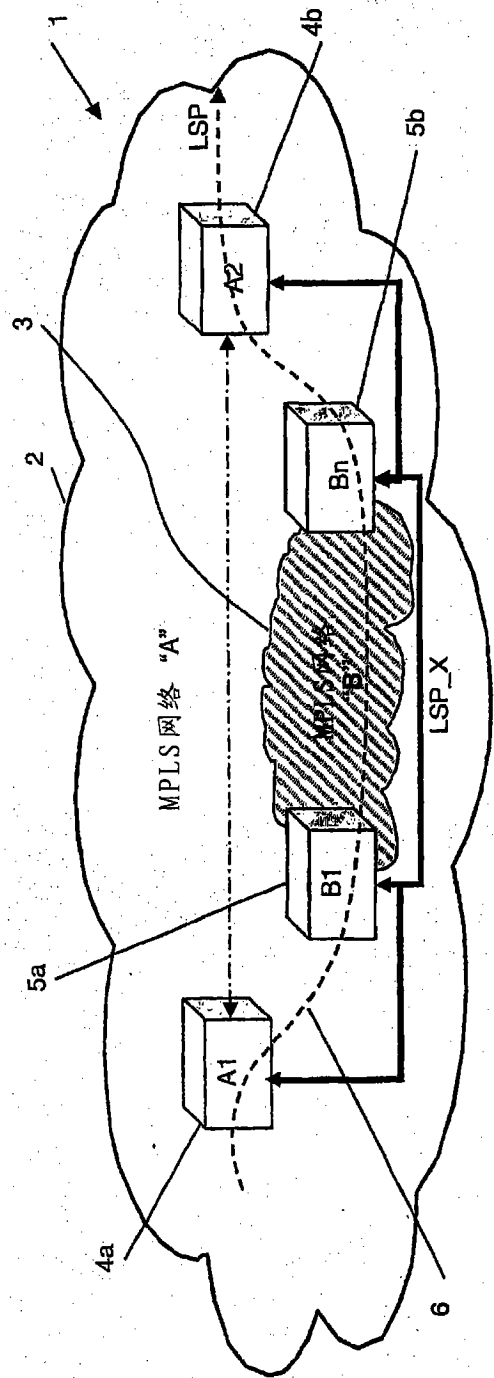


图 4A

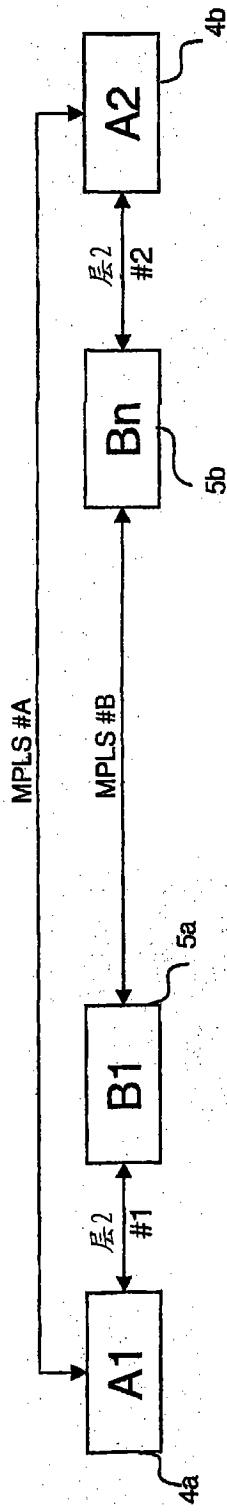


图 4B

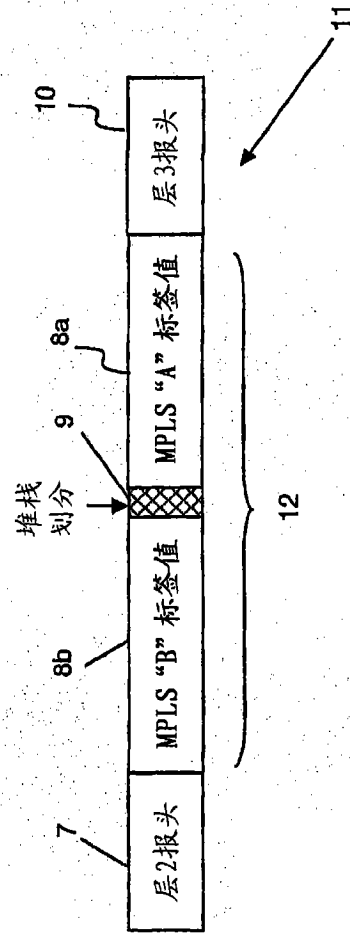


图 4C

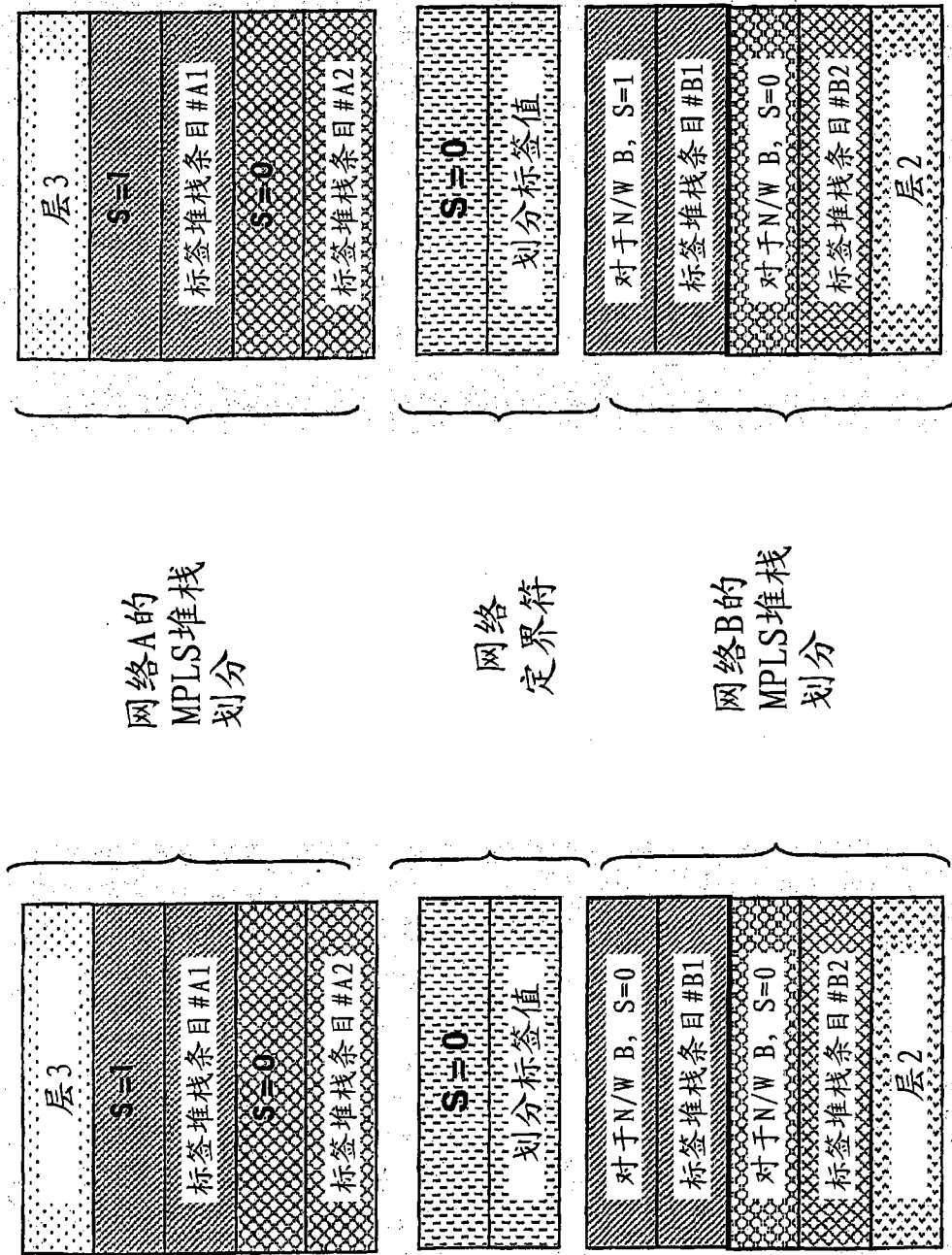


图5A

图5B

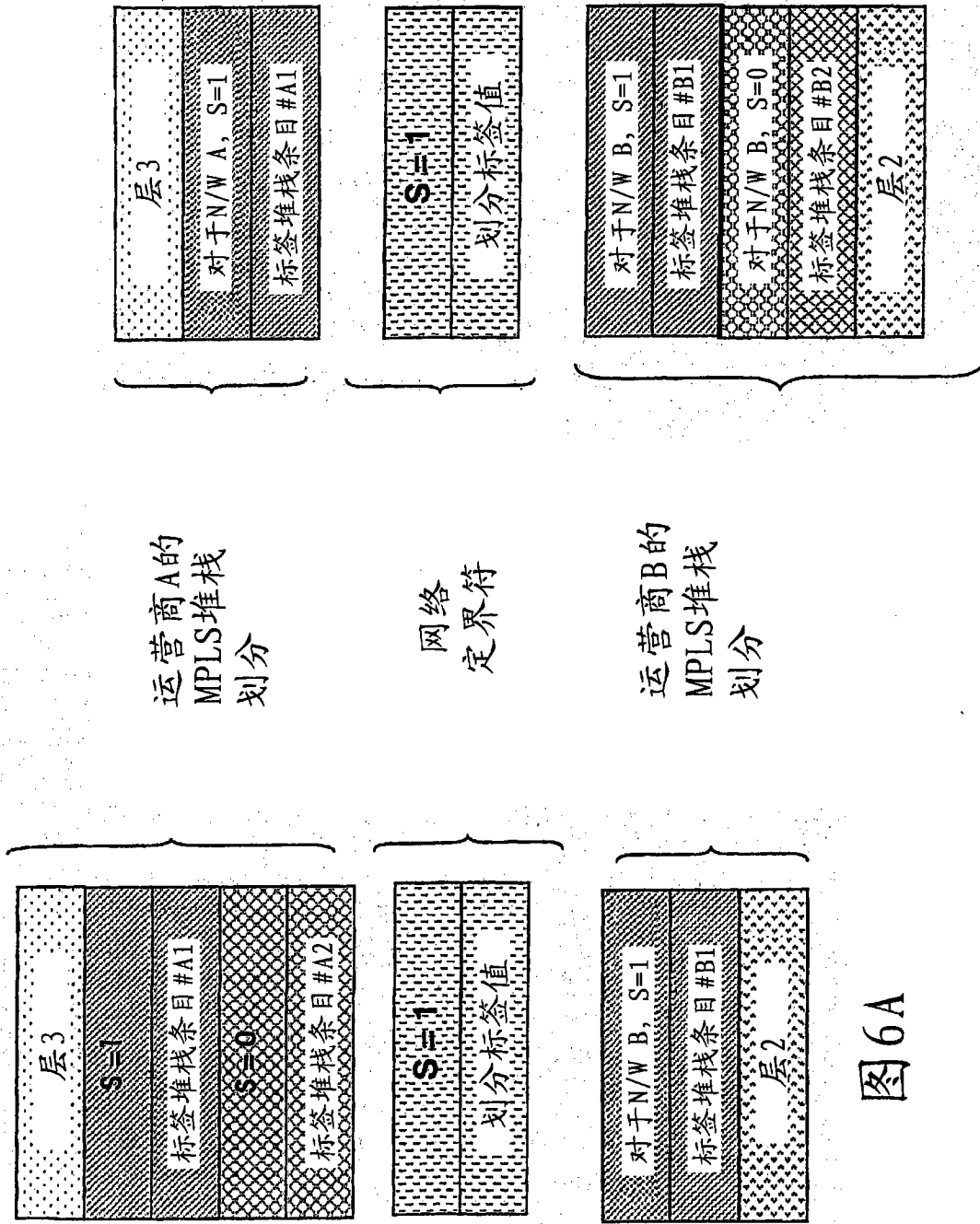


图6B

图6A

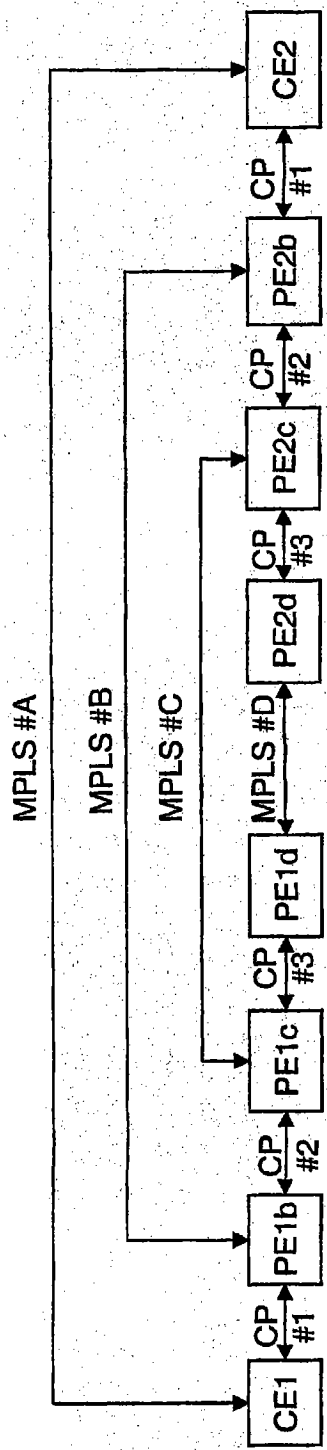


图 7A

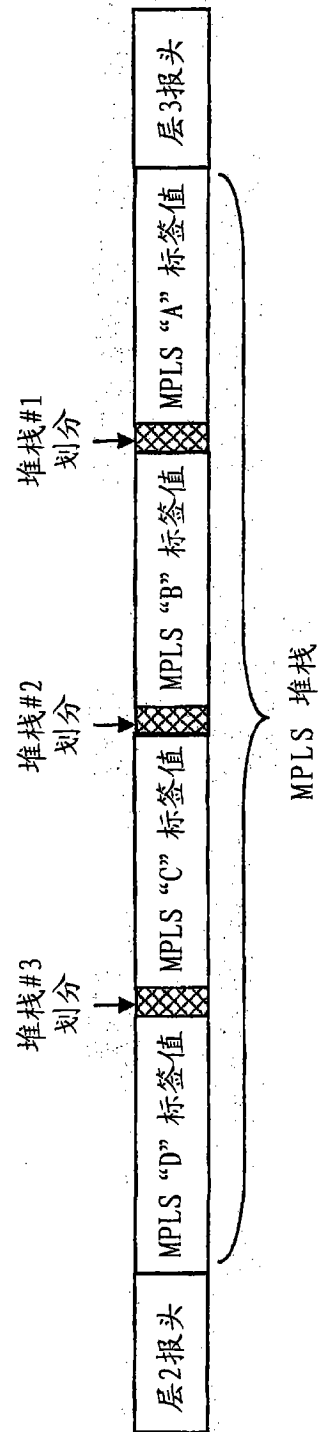


图 7B