(19)中华人民共和国国家知识产权局

(12)发明专利说明书

(10)申请公布号 CN 1625169 A

(43)申请公布日 2005.06.08

(21)申请号 CN200310117163.X

(22)申请日 2003.12.02

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 李斌 李丰林 李德丰 董伟嗣

(74)专利代理机构 北京德琦知识产权代理有限公司

代理人 宋志强

(51)Int.CI

H04L29/02

H04L12/56

权利要求说明书 说明书 幅图

(54)发明名称

一种在虚拟专用网的骨干网中保证业务质量的方法

(57)摘要

本发明公开了一种在虚拟专用网的骨干网中保证业务质量的方法，该方法对于需要保证 QoS 的 VRF－VRF 间连接，设置所需要的带宽资源；在 LSP 路径的每一跳上生成<外层标签、VPN 标签集合>，并从 LSP 预留的带宽资源中分配其对应的

VRF－VRF 间连接专用的带宽资源；数据报文转发时，入口 PE 在该数据报文的头部增加<外层标签、VPN 标签>信息，在 LSP 路径的每一跳上当该数据报文的头部与一个<外层标签、VPN 标签集合>的外层标签和 VPN 标签集合中的一个 VPN 标签匹配时，使用专用的带宽资源进行转发。本发明可以保证每个 VPN 的资源独立，达到与 ATM/FR VPN 相当的 QoS，LSP 数量少，计算工作量小，维护的状态少，简便可行。

法律状态

法律状态公告日	法律状态信息	法律状态

权 利 要 求 说 明 书

1、一种在虚拟专用网的骨干网中保证业务质量的方法，其特征在于，该方法包括以下步骤：

1)在 PE-PE 间建立 LSP，为该 LSP 预留带宽资源，并建立 VRF-VRF 间连接；

2)在入口 PE 上，对于每个 VRF-VRF 间连接，提取该连接所有路由的 VPN 标签，组成一个对应的 VPN 标签集合；

3)对于需要保证 QoS 的 VRF-VRF 间连接，在其入口 PE 上设置其所需要的带宽资源；

4)入口 PE 为该 VRF-VRF 间连接向出口 PE 发起资源请求；

5)在 LSP 路径的每一跳上生成<外层标签、VPN 标签集合>，并为该<外层标签、VPN 标签集合>从 LSP 预留的带宽资源中分配其对应的 VRF-VRF 间连接专用的带宽资源；

6)数据报文转发时，入口 PE 在该数据报文的头部增加<外层标签、VPN 标签>信息，在 LSP 路径的每一跳上当该数据报文的头部与一个 <外层标签、VPN 标签集合>的外层标签和 VPN 标签集合中的一个 VPN 标签匹配时，使用步骤 5)中为该<外层标签、VPN 标签集合>分配的专用的带宽资源进行转发。

2、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：为骨干网中每个 PE 上的每个 VRF 设置一个对应的 RD；步骤 2)所述组成 VPN 标签集合的过程包括以下步骤：

21)在入口 VRF 上，根据路由的下一跳信息区分其来自哪一个出口 PE；

22)如果 VPN 在出口 PE 上有一个以上的 VRF，则根据路由的 RD 信息判断其来自出口 PE 上的那个 VRF；

23)对于匹配了上述一个<下一跳、RD>的多个 VPN 路由，提取该多个路由中携带的 VPN 标签，组成一个 VPN 标签集合。

3、如权利要求 1 所述的方法，其特征在于，步骤 3)所述设置 VRF-VRF 间连接所需要的带宽资源的方法为：在 CE-CE 间运行 RSVP，入口 CE 向入口 PE 发送包括带宽资源需求的 RSVP PATH 消息；

入口 PE 识别该消息后，先根据报文的入接口确定入口 VRF；然后根据出口 CE 查找路由表，查找到其对应的 RD 和下一跳，及出口 PE，确定出口 VRF；将 CE 向 PE 发送的资源请求合并到 VRF-VRF 间的资源请求。

4、如权利要求 1 所述的方法，其特征在于，步骤 3)所述设置 VRF-VRF 间连接所需要的带宽资源的方法为：通过命令行配置，直接在入口 PE 上配置，入口 PE 将配置的所需要的带宽资源信息加入到 VRF-VRF 间的资源请求中。

5、如权利要求 1 所述的方法，其特征在于，步骤 3)所述设置 VRF-VRF 间连接所需要的带宽资源的方法为：网管配置，通过网管系统下发配置给入口 PE，入口 PE

将配置的所需要的带宽资源信息加入到 VRF-VRF 间的资源请求中。

6、如权利要求 1 所述的方法，其特征在于，步骤 4)所述入口 PE 为 VRF-VRF 间连接向出口 PE 发起资源请求，是采用 RSVP-TE 或 CR-LDP 协议发起的资源请求。

7、如权利要求 1 所述的方法，其特征在于，步骤 5)所述在 LSP 路径的每一跳上生成<外层标签、VPN 标签集合>的方法为：在入口 PE 上查找到 VRF-VRF 间连接对应的 VPN 标签集合，随资源请求发送给 LSP 路径上的每一跳；在 LSP 的每一跳上根据 LSP 确定外层标签，将外层标签和 VPN 标签集合组成<外层标签、VPN 标签集合>。

8、如权利要求 1 所述的方法，其特征在于，步骤 5)所述为<外层标签、VPN 标签集合>对从 LSP 预留的带宽资源中分配其对应的 VRF-VRF 间连接的带宽资源的方法为：

先为<外层标签、VPN 标签集合>生成一个对应的 QoS 队列，然后用外层标签和 VPN 标签集合中的各个 VPN 标签分别组成<外层标签、VPN 标签>对，将每个<外层标签、VPN 标签>对同 QoS 队列关联；为这个 QoS 队列从原有 LSP 所预留的带宽资源中分配 VRF-VRF 间连接专用的带宽资源。

9、如权利要求 8 所述的方法，其特征在于，步骤 6)所述的使用步骤 5)中为该<

外层标签、VPN 标签集合>对分配的专用的带宽资源进行转发的方法为：使用 LSP 上每一跳上预先分配的 VRF-VRF 间连接的专用带宽资源，对 QoS 队列中的数据报文进行调度和转发。

10、如权利要求 1 所述的方法，其特征在于，所述的步骤 6)中入口 PE 进行数据报文转发前进一步包括：入口 PE 对 VRF-VRF 间连接流量进行流量监管 CAR。

11、如权利要求 10 所述的方法，其特征在于，所述的流量监管的方法为：在入口 PE 的出接口上对匹配了<外层标签、VPN 标签集合>的数据报文的流量进行测量和整型。

12、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：当由于 LSP 带宽资源不足时，为该 LSP 分配足够的带宽资源，并从其中分配 VRF-VRF 间连接的带宽资源。

13、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：当 VRF-VRF 间连接资源请求失败时，放弃为该 VRF-VRF 间连接预留带宽资源，用原来的 LSP 进行数据报文转发。

14、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：数据报文在发端 CE 内部根据数据报文的优先级进行 DS 标记；

步骤 1)所述建立 LSP 的方法为：在 PE-PE 间为不同优先级的数据报文建立专门的 LSP；

步骤 3)所述设置 VRF-VRF 间连接所需要的带宽资源的方法为：根据数据报文的优先级设置 VRF-VRF 间连接所需要的带宽资源；

所述步骤 5)进一步包括：根据数据报文的优先级选择 LSP，在入口 PE 上，根据数据报文的 DS 标记进入选定的 LSP。

15、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：当骨干网采用 ASBR 方式的跨域 AS 方案时，上游 ASBR 将包含 VPN 标签集合的资源请求传递到下游 ASBR，在下游 ASBR 上，根据该 VPN 标签集合同本 AS 内的外层 LSP 和 VPN 标签的对应关系，进行转换，然后继续向出口 PE 发起资源请求。

16、如权利要求 15 所述的方法，其特征在于，上游的 ASBR 同下游 ASBR 运行另外一个 RSVP 实例；

上游 ASBR 在终结域内的 RSVP-TE/CR-LDP 后，将对<VPN 标签集合>的资源请求传递到下游 ASBR；

通过 RSVP-TE/CR-LDP 继续向出口 PE 发起资源请求。

17、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：当骨干网采用 Multi-hop EBGP 方式的跨域 AS 方案时，在进行资源预留时，识别 VRF-VRF 间连接通过三层标签来完成，所述的 VPN 标签集合为三层标签中的最底的一层。

说　明　书

<u>技术领域</u>

本发明涉及一种保证网络业务质量的方法，特别涉及一种在多协议标签交换虚拟专用网(BGP/MPLS VPN)的骨干网中保证业务质量(QoS)的方法。

<u>背景技术</u>

BGP/MPLS VPN 使用边界网关协议(BGP)把 VPN 路由信息分布到提供商的骨干网中，并使用 MPLS 把 VPN 流量从一个客户站点(site)转发到另一个客户站点上。

BGP/MPLS VPN 的网络构造由服务提供者(SP)来完成，参见图 1，图 1 为 BGP/MPLS VPN 结构示意图。该网络由 SP 的骨干网与客户的 VPN 构成，所谓 VPN 就是对客户 site 集合的划分，一个 VPN 就对应一个由若干 site 组成的集合。每个客户 site 包含了一个客户边缘(CE)设备，每个 CE 与骨干网中的一个提供商边缘(PE)路由器相连。骨干网中，PE 与提供商路由器(P)相连。例如：图 1 中，VPN1 的 CE1 和 VPN2 的 CE2 与骨干网中的 PE1 相连；VPN2 的 CE3 和 CE4 与骨干网中的 PE2 相连；VPN1 的 CE5 与骨干网中的 PE3 相连。

其中，客户边缘(CE)设备允许客户通过连接一台或多台提供商边缘(PE)路由器的一条数据链路接入服务提供商骨干网络。CE 设备可以是一台主机或一台第二层交换机，但典型的 CE 设备是一台 IP 路由器，它与其直接连接的 PE 路由器建立连接关系。在建立连接后，CE 路由器把站点的本地 VPN 路由广播到 PE 路由器，并从 PE 路由器上学习远程 VPN 路由。

PE 路由器使用静态路由、路由信息协议版本 2(RIPv2)、开放最短路径优先协议(OSPF)或外部边界网关协议(EBGP)与 CE 路由器交换路由信息。尽管 PE 路由器维护着 VPN 路由信息，但它只需为其直接相连的那些 VPN 维护 VPN 路由。这种设

模型的扩充能力，因为 PE 路由器不需维护服务提供商的所有 VPN 路由。

每台 PE 路由器为其直接相连的每个站点维护一个逻辑上分离的路由表，每个站点在 PE 路由器上对应一个虚拟路由转发实例(VRF)，每个 VRF 有独立的路由表。每个客户连接：如帧中继 PVC、ATM PVC 和 VLAN，映射到某个 VRF 上。因此，PE 路由器上的一个端口与 VRF 相关，PE 路由器上的多个端口可以与一个 VRF 相关。PE 路由器能够维护多个转发表，支持按 VPN 分隔路由信息。一个 VRF 实际上综合了和它所对应 site 的 VPN 成员关系和路由规则。分离的路由表除了能防止数据泄漏出 VPN 之外，同时能防止 VPN 之外的数据进入。采用路由标识符(Route Distinguisher，RD)来标示每个 VRF，实际上是每个 site 一个。RD 在骨干网中保持唯一性，是网络中 site 的标识符。

在从 CE 路由器上学习本地 VPN 路由后，PE 路由器通过使用内部边界网关协议(IBGP)来向其它路由器传播 VPN 组成信息和路由。PE 路由器可以保持到路由反射器的 IBGP 会话，作为全网状 IBGP 会话的替代方案。部署多个路由反射器增强了 BGP/MPLS VPN 模型的扩充能力，因为它不需任何单个网元维护所有 VPN 路由。

最后，使用 MPLS 在提供商骨干网中转发 VPN 数据流量时，入口 PE 路由器作为入口标签交换路由器(LSR)使用，出口 PE 路由器作为出口 LSR 使用。

提供商路由器(P)是哪些没有连接 CE 设备的骨干网中的任何路由器。例如：图 1 中，P1-P4。它们在 PE 路由器之间转发 VPN 数据流量时，提供商路由器(P)作为 MPLS 转接 LSR 使用。

MPLS/BGP VPN 提供了灵活的地址管理。由于采用了单独的路由表，允许每个 VPN 使用单独的地址空间中，称为 VPN-IPv4 地址空间，RD 加上 IPv4 地址就构成了 VPN-IPv4 地址。

BGP/MPLS VPN 中，属于同一的 VPN 的两个 site 之间转发报文使用两层标签来解决，在入口 PE 上为报文打上两层标签，第一层(外层)标签在骨干网内部进行交换，代表了从 PE 到对端 PE 的一条隧道，VPN 报文打上这层标签，就可以沿着 LSP 到达对端 PE，这时候就需要使用第二层(内层)标签，这层标签指示了报文应该到达哪个 site，或者更具体一些，到达哪一个 CE，这样，根据内层标签，就可以找到转发的接口。

由于是在采用两层标记堆栈的 MPLS 骨干网中转发流量，因此提供商路由器只需维护到提供商 PE 路由器的路由，而不需维护每个客户站点(site)专用的 VPN 路由信息。

VPN 的成员关系是通过路由所携带的路由目标(Route Target)属性来获得的。PE 路由器中每个 site 的路由表可以有一或多个 Route Target 属性。它表示了该路由可以被哪些 site 所接收，接收哪些 site 传送来的路由。一个具有这种属性的路由必须发送给所有在 Route Target 中指明的 site 所连接的 PE 路由器，PE 路由器接收到包含此属性的路由后，若此属性指明的 site 与自己关联，则加入到相应的路由表中。对一个 PE 路由器，有一个 Route Target 属性的集合用于附加到从某个 site 接收的路由上，称为 Export Targets，另一个 Route Target 属性的集合用于决定哪些路由可以引入到此 site 的路由表中，称为 Import Targets。它们是不同的集合。这两个集合的组合可以构造任何拓扑类型的 VPN。

目前，可以采用共用的流量工程隧道保证 VPN 带宽。

流量工程允许将流从内部网关协议(IGP)发现的最短路径上移开到其它物理路径上，因为最短路径可能是最拥塞的路径，而其它路径反而比较空闲。这可以最大限度的提高网络资源的利用率，平衡网络内部的流量分布，使得网络内部没有明显的过载节点和空闲节点。从而使得网络被更高效的管理和提供更可预测的服务。

扩展现有的基于链路状态的 IGP(如 OSPF、IS-IS)，可以在链路状态通告中附加链路属性，这些属性包括最大链接带宽、最大保留链接带宽、当前保留带宽和链路类别(link coloring)等，它们通过标准的基于链路状态的 IGP 的散播算法分布到域内的各个 LSR。

链路属性和拓扑信息保存在流量工程数据库(TED)中，这个数据库专门用于计算显式路径，进而在物理网络中部署 LSP。这个数据库同链路状态数据库是分离的。

当通过 IGP 在整个网络中散播了链路属性以及拓扑信息后，每个入口 LSR 就可以根据流量工程数据库(TED)中的信息来计算从它起始的 LSP 路径，这些路径可以代表严格或松散的显式路由。显式路由是指 LSP 必须经过的一些预先定义的路由器的序列。如果这个序列包括路径上所有的路由器，称为严格的显式路由，如果只包括部分路由器，称为松散的显式路由。

计算显式路由有两种方式，在线和离线算法。在线算法在入口 LSR 上实时的计算显式路由，其算法是约束最短路径优先(Constrained Shortest PathFirst，CSP)算法，它在特定的限制条件下计算从入口到出口的最短路径。其输入有 TED 中的链路状态拓扑信息、TED 中的表示了网络资源的链路属性信息以及用户配置的管理性属性，其输出是符合上述约束的最短路径中的各个 LSR 组成的显式路由。在线计算的结果是同顺序相关的，首先计算的 LSP 一般会获得较多的资源，当计算顺序改变后，计算出的 LSP 的集合会相应的改变，因此需要一种全局的计算方式。离线计算工具收集整个网络的拓扑信息和网络资源，综合考虑每个 LSP 的约束，从而计算出全局最优的结果，包括每个 LSP 的显式路由。

显式路由最后由 RSVP-TE/CR-LDP 来部署。它能够建立符合严格和松散显式路由的 LSP。