(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105940644 B
(45)授权公告日 2019.11.12

(21)申请号 201480074612.3
(22)申请日 2014.12.02
(65)同一申请的已公布的文献号
申请公布号 CN 105940644 A
(43)申请公布日 2016.09.14
(30)优先权数据
61/910,851 2013.12.02 US
(85)PCT国际申请进入国家阶段日
2016.07.29
(86)PCT国际申请的申请数据
PCT/US2014/068217 2014.12.02
(87)PCT国际申请的公布数据
WO2015/084878 EN 2015.06.11
(73)专利权人 阿卡麦科技公司
地址 美国马萨诸塞州
(72)发明人 布兰登·O·威廉姆斯 马丁·K·洛纳 凯文·哈蒙 杰弗里·鲍尔
(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
代理人 周靖 郑霞
(51)Int.Cl.
H04L 12/723(2006.01)
H04L 12/901(2006.01)
H04L 12/28(2006.01)
H04L 29/06(2006.01)
(56)对比文件
US 2011289311 A1,2011.11.24,
CN 101232519 A,2008.07.30,
CN 101834793 A,2010.09.15,
CN 102143487 A,2011.08.03,
审查员 周萍

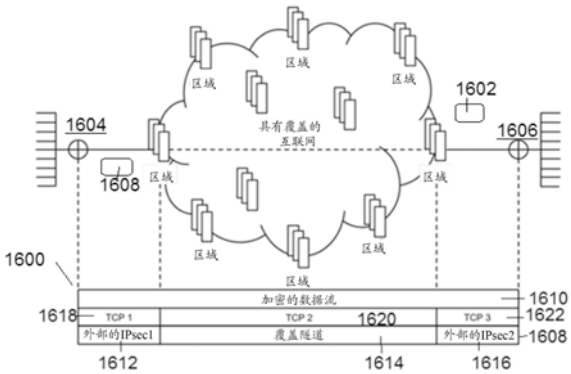
权利要求书2页 说明书11页 附图7页

(54)发明名称

在保持端对端数据安全的同时具有分发优化的虚拟专用网络(VPN)即服务

(57)摘要

一种优选地在覆盖网络内实现的覆盖IP路由机构的上下文内促进专有网络(VPN)即服务的机构。网络即服务的客户操作被期望安全地以及专门地利用覆盖IP(OIP)路由机构连接到彼此的端点。覆盖提供在端点处定位的覆盖网络设备之间端对端的包的分发。在这种分发期间,设备被配置成使得每一个包的数据部分具有与包的TCP/IP部分的加密上下文不同的加密上下文。通过建立和维持这些不同的加密上下文,覆盖网络可解密以及访问TCP/IP流。这使覆盖网络提供商能够应用一个或多个TCP优化。同时,独立的加密上下文确保在传输期间在任何点处每一个包的数据部分在明文中绝不可用。



CN 105940644 B

1.一种将第一端点位置耦合到第二端点位置的系统,包括:

路由覆盖网络,所述路由覆盖网络包括至少部分地跨互联网位置分布的一组机器,所述路由覆盖网络包括与所述第一端点位置相关联的入口点以及与所述第二端点位置相关联的出口点;

第一设备,其连接到所述入口点;

第二设备,其连接到所述出口点;

所述第一设备和所述第二设备中的每一个包括:

计算机存储器,其保存能够被硬件处理器执行的计算机程序指令,且所述计算机程序指令能够操作来:

配置和维持第一加密上下文和第二加密上下文,所述第一加密上下文在所述第一设备和所述第二设备之间端对端地延伸,所述第二加密上下文具有一组传输协议段,所述一组传输协议段包括在所述第一设备和所述入口点之间的第一外部传输协议段、在所述出口点和所述第二设备之间的第二外部传输协议段以及在所述第一外部传输协议段和所述第二外部传输协议段之间的中间协议段;

所述第一加密上下文保护在所述第一设备和所述第二设备之间的端对端数据流;

所述第二加密上下文保护包括所述数据流的每个包的传输和网络协议层报头;

在没有暴露被所述第一加密上下文保护的所述数据流的情况下选择性地解密被所述第二加密上下文保护的所述传输和网络协议层报头;以及

在解密之后,将路由覆盖网络优化应用于所述数据流。

2.如权利要求1所述的系统,其中,在每一个设备中的所述计算机存储器存储密钥,所述密钥适合于结合所述第二加密上下文来实现解密以实现通过所述路由覆盖网络对所述传输和网络协议层报头的访问。

3.如权利要求2所述的系统,其中,传输协议是传输控制协议TCP。

4.如权利要求3所述的系统,其中,所述计算机程序指令还能够操作来提供TCP优化。

5.如权利要求2所述的系统,其中,所述密钥不使能对所述数据流的访问,当所述数据流穿过在所述第一设备和所述第二设备之间的所述路由覆盖网络时,所述数据流保持加密。

6.如权利要求3所述的系统,其中,在每一个TCP段上实现了逐跳认证和加密。

7.如权利要求1所述的系统,其中,所述路由覆盖网络从实现多协议标签交换(MPLS)的广域网(WAN)卸载所述数据流的分发。

8.如权利要求1所述的系统,其中,所述入口点包括内容分发网络(CDN)的边缘区域。

9.如权利要求1所述的系统,其中,所述路由覆盖网络支持来自多个租户的多个数据流。

10.一种将第一端点位置耦合到路由覆盖网络的设备,所述路由覆盖网络包括至少部分地跨互联网位置分布的一组机器,所述路由覆盖网络包括与所述第一端点位置相关联的入口点以及与第二端点位置相关联的出口点,所述设备包括:

一个或多个硬件处理器;

计算机存储器,其保存能够被所述一个或多个硬件处理器执行的计算机程序指令,以及所述计算机程序指令相对于与所述出口点相关联的第二设备能够操作来:

配置和维持第一加密上下文和第二加密上下文,所述第一加密上下文在所述设备和所述第二设备之间端对端地延伸,所述第二加密上下文具有一组传输协议段,所述一组传输协议段包括在所述设备和所述入口点之间的第一外部传输协议段、在所述出口点和所述第二设备之间的第二外部传输协议段以及在所述第一外部传输协议段和所述第二外部传输协议段之间的中间传输协议段;

所述第一加密上下文保护在所述设备和所述第二设备之间的端对端数据流;

所述第二加密上下文保护包括所述数据流的每个包的传输和网络协议层报头;

在没有暴露被所述第一加密上下文保护的所述数据流的情况下选择性地解密被所述第二加密上下文保护的所述传输和网络协议层报头;以及

在解密之后,将路由覆盖网络优化应用于所述数据流。

11.如权利要求10所述的设备,其中,所述计算机存储器存储密钥,所述密钥适合于结合所述第二加密上下文来实现解密以实现通过所述路由覆盖网络对所述传输和网络协议层报头的访问。

12.如权利要求10所述的设备,其中,传输协议是传输控制协议TCP。

13.如权利要求12所述的设备,其中,所述计算机程序指令还能够操作来提供TCP优化。

14.如权利要求11所述的设备,其中,所述密钥不使能对所述数据流的访问,当所述数据流穿过在所述设备和所述第二设备之间的所述路由覆盖网络时,所述数据流保持加密。

15.如权利要求12所述的设备,其中,在每一个TCP段上实现了逐跳认证和加密。

在保持端对端数据安全的同时具有分发优化的虚拟专用网络(VPN)即服务

[0001] 背景

技术领域

[0002] 本申请通常涉及在公开路由的互联网上覆盖网络路由。

[0003] 相关技术的简要描述

[0004] 在现有的技术中分布式计算机系统是公知的。一种这样的分布式计算机系统是由服务提供商操作和管理的“内容分发网络”(CDN)或“覆盖网络”。服务提供商通常为了使用服务提供商的共享的基础设施的第三方(客户)的利益提供内容分发服务。这种类型的分布式系统指通过网络或多个网络连接的自主的计算机的集合,连同软件、系统、被设计用来促进各种服务的协议和技术,比如内容分发、网络应用加速或外包的起源站点基础设施的其他支持。CDN服务提供商通常通过数字财产(比如网站)提供服务配送,其在客户门户中被提供然后被部署到网络。数字财产通常被绑定到一个或多个边缘配置,允许服务提供商解释流量并且给客户开账单。

[0005] 广域网(WAN)是通常覆盖广阔的地理区域的电信网,例如链接于大都市的、地区的、国家的或国际的界线,其通常使用租用的电信线路。企业和政府实体利用WAN在员工、客户、买主和来自各种地理位置的供应商之间转播数据。例如,WAN通常被用于将局域网(LAN)和其他类型的网络连接在一起,使得在一个位置的用户和计算机可与在其他位置的用户和计算机通信。许多WAN是为一个特定的组织建立的,并且是私有的。其他类型的WAN包括通过互联网服务商建立的那些,以及这些可被用于提供从组织的LAN到互联网的连接。当WAN使用租用的线路被建立时,在租用线路的每一个末端定位的路由器将每一侧上的LAN连接到彼此。

[0006] 使用租用的线路的一个常见的WAN方式实现多协议标签交换(MPLS)。MPLS是用于加速网络流量的基于标准的技术。在MPLS中,用于给定的包序列的特定的路径(通过标签标识的)被建立,因此避免路由器查找转发该包的下一个地址。MPLS和各种类型的网络协议(比如,IP、ATM和帧中继)一起工作。虽然在MPLS上的分发是有效的且安全的,但是主要由于租用的线路的花费,其也是昂贵的。作为备选方案,WAN也可使用成本更低的包交换方法来建立,比如可以充分利用互联网的包交换网的那些方法。

[0007] 实际上,企业现在期望有效地利用互联网链接作为优化的广域网(WAN),在互联网上将分支、数据中心、远程工作人员和移动用户连接到应用。而且通过云计算和移动性的驱动,企业需要可以将最优的和可预测的云体验分发给用户的网络服务,优选地是具有安全性的和内置的优化的低成本的、易于使用的和全局的网络。

[0008] 简述

[0009] 本公开提供了各种机制来促进虚拟专用网络(VPN)即服务(或更普遍地,“网络即服务”),优选地在覆盖(或“内容分发”)网络(“CDN”)内实现的覆盖IP(OIP)路由机制的上下文内。联网“即服务”的概念使企业能够有效地利用互联网链接作为优化的广域网(WAN),在

互联网上将分支、数据中心、远程工作人员和移动用户连接到应用。

[0010] 在本方法中，假设网络即服务的客户操作端点(例如，局域网(LAN))，端点被期望安全地且私有地使用覆盖IP(OIP)路由机制连接彼此。覆盖层(overlay)提供在端点处定位的覆盖网络设备之间端对端的包的分发。然而，在这种分发期间，设备被配置成使得每一个包的数据部分具有与包的TCP/IP部分的加密上下文不同的加密上下文。通过建立和保持这些不同的加密上下文，覆盖网络可解密以及访问TCP/IP流。这使覆盖网络提供商能够将一个或多个TCP优化应用到TCP流。同时，然而，用于数据流的独立的加密上下文保证在覆盖层中进行传输期间每一个包的数据部分在任何点是绝不可以明文可用的。

[0011] 本方法是非常有利的，因为它允许端点之间的覆盖网络客户数据流量利用和很好地使用全部TCP和由覆盖网络路由机制提供的路由优化，同时避免了(全部或部分)传统的传输机制(如MPLS)及它们相关联的高成本。通过建立和施行关于数据的独立的加密上下文(一方面)以及包头(另一方面)，当其穿过公共互联网(在其上，覆盖层被建立且运行)时，客户数据被保护免于非加密的访问。覆盖层也可实现额外的认证以及加密保护以防止发现(例如，从遍历公共互联网的包)关于客户内部网络的细节，以及其他方式以防止未被授权访问覆盖网络而获得对客户的内部网络的访问或获得被优化的互联网传输和路由。

[0012] 前述已经概述了公开的主题的一些较相关的特征。这些特征应被解释为仅仅是说明性的。许多其它有益的结果可通过以不同的方式施加所公开的主题或通过修改如将要描述的主题来获得。

[0013] 附图简述

[0014] 为更完整的理解主题及其优点，现在参考结合附图进行以下描述，其中：

[0015] 图1示出了被配置为内容分发网络(CDN)的已知的分布式计算机系统的框图；

[0016] 图2是有代表性的CDN边缘机器的配置；

[0017] 图3是已知的覆盖解决方案，借由此覆盖层被定位在可公开路由的互联网的顶部；

[0018] 图4示出了在如在图1中示出的内容分发网络内实现的多路覆盖IP(OIP)路由机制；

[0019] 图5示出了具有典型地通过广域网多协议标签交换(MPLS)连接到一个或多个分支位置的企业数据中心的用于企业的典型的用例方案。

[0020] 图6示出了使用MPLS的端对端的路径；

[0021] 图7对应于图5，但是还包括在每一个端点的覆盖网络设备的内含物；

[0022] 图8对应于图7，但是示出了图1-图4的覆盖网络如何结合在每一个端点处的覆盖网络设备一起使用，其中，覆盖网络提供MPLS路由方法的可选方案；

[0023] 图9示出了在一个端点处的覆盖网络设备如何建立到附近的边缘区域或覆盖网络的区域的安全的逻辑VPN隧道；

[0024] 图10示出了在安装网络设备以及在第一侧(LAN-A)上建立到覆盖网络(表示为平台)的VPN隧道之后加入客户网络(ABC)的与覆盖网络客户相关联的第一办公室A；

[0025] 图11示出了在其侧(LAN-B)上安装网络设备之后加入客户网络ABC的与覆盖网络客户相关联的第二办公室B；

[0026] 图12示出了在建立到它们的附近边缘(多个)区域的安全VPN隧道之后，多个覆盖网络客户在它们的各种位置如何使用具有网络设备的覆盖网络平台；

[0027] 图13示出了在到覆盖网络区域或多个区域的安全VPN隧道的建立之后，多个封装的和加密的流如何被提供在网络设备和覆盖网络平台之间流动；

[0028] 图14示出了端对端多站点VPN，其中，MPLS和覆盖网络路由可在端点间被实现；

[0029] 图15示出了覆盖隧道的方法，其中，外隧道被拆分且被分割，以及内隧道是安全的端对端的；以及

[0030] 图16示出了本主题公开的技术，其中，覆盖网络建立和保持一方面用于包数据并且另一方面用于TCP/IP报头的不同的加密上下文；这个拆分加密上下文的方法促进对于TCP流的一个或多个TCP优化(通过覆盖网络)的应用，当流量被路由以及穿过覆盖网络被分发时同时确保数据流不可以明文可用。

[0031] 详细描述

[0032] 在已知的系统中，如在图1中所示的，分布式计算机系统100被配置为内容分发网络(CDN)，以及被假定具有围绕互联网分布的一组机器102a-n。通常，大部分的机器是位于互联网的边缘附近的服务器，即，处于或邻近终端用户接入网络。网络操作命令中心(NOCC)104管理在系统中的各种机器的操作。第三方站点(如，网站106)将分发的内容(例如，HTML、嵌入的页面对象、流媒体、软件下载，等等)卸载到分布式计算机系统100，并且，具体的，到"边缘"服务器。通常，内容提供商通过将给定的内容提供商域或子域混叠(例如，通过DNS CNAME)到由服务提供商授权的域名服务管理的域来卸载他们的内容分发。期望内容的终端用户是针对分布式计算机系统的以获得更可靠和更有效的内容。尽管没有详细示出，分布式计算机系统也可包括其他基础设施，如分布式数据收集系统108，其从边缘服务器收集使用数据和其他数据，穿过区域或一组区域聚集该数据以及将该数据传递到其他后端系统110、112、114和116以促进监控、日志、警报、开账单、管理以及其他操作的和管理的功能。分布式网络代理118监控网络和服务器负载，并且将网络、流量和负载数据提供给DNS查询处理机构115，这对于正由CDN管理的内容域是权威的。分布式数据传输机构120可被用于将控制信息(例如，促进负载平衡的、管理内容的元数据，等等)分布到边缘服务器。

[0033] 如在图2中示出的，在内容分发网络中给定的机器200包括运行支持一个或多个应用206a-n的操作系统核(如Linux或变型)204的商用硬件(例如，Intel Pentium处理器)202。为了促进内容分发服务，例如，给定的机器通常运行一组应用，如HTTP代理207(有时被称为“全局主机”或“ghost”进程)、名称服务器208、本地监控进程210、分布式数据收集进程212，等等。对于流媒体，机器可包括一个或多个媒体服务器，比如正如由所支持的媒体格式需要的Windows Media Server(WMS)或Flash服务器，或其可利用组成流的分块内容片段的基于HTTP的分发。

[0034] CDN边缘服务器被配置为，优选地基于特定域、特定客户，优选地利用使用配置系统被分布到边缘服务器的配置文件，提供一个或多个扩展的内容分发特征。给定的配置文件优选地是基于XML的以及包括促进一个或多个先进的内容处理特征的一组内容处理规则和指令。配置文件可经由数据传输机制被分发到CDN边缘服务器。美国专利第7,111,057号示出了用于分发和管理边缘服务器内容控制信息的有用的基础设施，并且这个和其他边缘服务器控制信息可由CDN服务提供商本身或(经由外联网或类似物)操作源服务器的内容提供商客户来提供。

[0035] CDN可包括存储子系统，比如在美国专利第7,472,178号中描述的那样，其中的公

开内容通过引用并入本文。

[0036] CDN可操作服务缓存层次结构以提供客户内容的中间缓存;一个这样的缓存层次结构子系统在美国专利第7,376,716号中被描述,其中的公开内容通过引用并入本文。

[0037] CDN可以以在公布号为20040093419美国专利中描述的方式在客户端浏览器、边缘服务器和客户源服务器之间提供安全的内容分发。如在其中所描述的,安全内容分发一方面在客户端和边缘服务器进程之间以及另一方面在边缘服务器进程和源服务器进程之间执行基于SSL的链接。这使SSL保护的网页和/或其组件能够经由边缘服务器被分发。

[0038] 在通常的操作中,内容提供商识别期望通过CDN服务的内容提供商域或子域。CDN服务提供商将内容提供商域与边缘网络主机名称关联(例如,经由规范名,或CNAME),以及CDN提供商然后将该边缘网络主机名称提供给内容提供商。当对于内容提供商域或子域的DNS查询在内容提供商域名服务器被接收时,那些服务器通过返回边缘网络主机名称响应。边缘网络主机名称指向CDN,而且该边缘网络主机名称然后通过CDN名称服务进行解析。为此,CDN名称服务返回一个或多个IP地址。请求客户端浏览器然后做出对与IP地址相关联的边缘服务器的内容请求(例如,经由HTTP或HTTPS)。请求包括主机报头,其包括原始的内容提供商域或子域。具有主机报头的请求一经接收,边缘服务器就检验其配置文件以确定所请求的内容域或子域实际上正被CDN处理。如果这样,边缘服务器根据在配置文件中指定的那样应用关于该域或子域应用的其内容处理规则和指令。这些内容处理规则和指令可位于基于XML的“元数据”配置文件内。

[0039] 通过进一步的背景,CDN客户可订购“在防火墙后面”的管理服务产品以加速在客户的企业防火墙后面托管的内联网网络应用,以及加速在防火墙后面的它们的用户到在内联网云中托管的应用之间桥接的网络应用。为了实现这两个用例,CDN软件可执行在一个或多个客户数据中心中托管的虚拟机上以及在远程“分支机构”中托管的虚拟机上。在客户数据中心中执行的CDN软件通常提供服务配置、服务管理、服务报告、远程管理访问、客户SSL证书管理以及关于配置的网络应用的其他功能。在分支办公室中执行的软件为位于那里的用户提供最后一英里网络加速。CDN本身通常提供在CDN数据中心中托管的CDN硬件以在客户防火墙后面运行的节点和服务提供商的其他基础设施(例如,网络和操作设施)之间提供网关。这种类型的受管解决方案给企业提供机会来利用关于他们的公司内联网的CDN技术。

[0040] 作为覆盖层,CDN资源(比如上述的)还可被用于促进企业数据中心(其可以是私有受管的)和第三方软件即服务(SaaS)的提供商之间的广域网(WAN)加速服务。以下提供关于这种类型的解决方案的额外的细节。

[0041] 具体的,图3示出了已知的“覆盖”网络解决方案,由此“覆盖层”被定位在可公开路由的互联网顶部之上。这个结构有时被称为“路由覆盖”或“路由覆盖网络”。路由覆盖网络可利用现有的内容分发网络(CDN)基础设施,比如,在上面的图1-图2中示出的基础设施,以及如由商业服务提供商(如Akamai Technologies,Inc.of Cambridge,Massachusetts(马萨诸塞州坎布里奇市的阿卡迈技术有限公司))提供的基础设施。这种类型的覆盖网络对于使用互联网协议(IP)作为传输协议通过围绕下行链路或寻找具有最小延迟的路径路由的任何应用提供重要的性能增强。众所周知,互联网协议(IP)通过交换成群的称为包(包括报头和正文的字节的短序列)的信息工作。报头描述了包的目的地,互联网路由器使用该目的地传送包直到其到达它的最终的目的地。正文包含应用数据。通常,IP包通过传输控制协议

(TCP)行进,传输控制协议(TCP)提供可靠的按次序的字节的流的分发。TCP重新布置无序的包,最小化网络拥塞以及重新传输丢弃的包。

[0042] 在覆盖层中的许多机器是位于互联网边缘附近的服务器,即,在或邻近最终用户访问的网络。如在上面已经描述的(例如,图1),第三方网站和应用提供商将分发的内容和应用卸载到网络,作为受管的服务操作。覆盖网络包括用于数据收集、监控、日志、警报、开账单、管理和其他操作的和管理的功能的分布式基础设施。如已经描述的以及如在图2中示出的,通常的CDN机器包括运行支持一个或多个应用的操作系统核(比如,Linux™或变型)的商用硬件(例如,Intel® Pentium®处理器)。为促进内容分发服务,例如,给定的机器通常运行一组应用,如HTTP网络代理、名称服务器、本地监控进程以及一个或多个数据收集进程。网络代理包括或具有与其相关的边缘服务器管理器进程以促进与内容分发网络相关联的一个或多个功能。

[0043] 已知的OIP路由机构包括有代表性的组件集合,如在图3中示出的:

[0044] • 边缘服务器302－通常,如在以下描述的,运行OIP边缘服务软件进程(oidp)的CDN边缘服务器。如将要被描述的,这个软件负责接收、封装以及转发IP包。

[0045] • 边缘区域300－通常,被配置为用于覆盖机构的CDN边缘区域。

[0046] • 中间服务器306－通常,从边缘区域300或其他中间服务器接收封装的包并且将它们转发到其他中间服务器上或网关区域的服务器。

[0047] • 中间区域304－中间服务器的区域。

[0048] • 网关服务器310－通常,已被配置为从覆盖层接收封装的包,而且将源网络地址转换(NAT)应用到源包且将它们转发到目标服务器上的边缘服务器。

[0049] • 网关区域308－通常,包括网关服务器而且常常在客户驻地上部署的一种类型的边缘区域。

[0050] • 目标服务器312－其流量通过覆盖层被隧穿的机器。

[0051] • 目标地址－目标服务器的IP地址;当和CDN虚拟IP地址比较时,这个地址有时被称作直接地址。

[0052] • 槽－覆盖层的单个“实例”;优选地,槽是对应于单个的目标地址的编号的索引。

[0053] • 虚拟IP地址－通常,对应于槽的CDN地址;优选地,每槽每边缘区域有一个IP地址。它有时被称作VIP。

[0054] • 路径314－在边缘区域和网关区域之间的一组有序的CDN区域。

[0055] • 路径段－路径的单个跳跃。

[0056] • 隧道318－从边缘服务器到网关服务器的一组一个或多个路径。

[0057] • 会话320－从客户端到目标服务器的单个端对端连接;优选地,会话由五个元组(IP有效载荷协议、源地址、目的地地址、源端口、目的地端口)定义。源是客户端并且目的地是目标。

[0058] 在覆盖网络的一个已知的应用场景中,一个或多个客户端期望将包发送到单一的IP地址。这在图4中被示出且现在进行描述。在步骤1,客户端400做出对解析主机名称(通常,主机名称与可网络访问的应用相关联)的DNS请求。这个主机名称是正被授权的DNS402管理的域的别名(例如,通过CNAME);通常,授权的DNS由CDN服务提供商管理。优选地,这个主机名称对应于单一的网关区域(和目标地址)404。如上所述,这也被称作槽。在步骤2,DNS

查询返回关于主机名称的单个的IP地址。这个地址识别表现最好的可用的边缘区域406，以及，优选地，该区域是主机名称专用的。如上所述，地址被称作虚拟IP地址。在步骤3，客户端400开始将IP包发送到虚拟IP地址。这些包由在边缘区域406中的服务器接收。边缘区域406基于在IP包报头中的目的地地址知道向其发送包的网关区域404。该包然后被封装。在步骤4，基于由CDN映射系统优选提供的路由，在边缘区域406中的边缘服务器沿着多个路径发送出封装的包的多个副本。用来执行这个多个路径包传输操作的一个技术，在授予给Akamai Technologies,Inc(阿卡迈技术有限公司)的美国专利第6,665,726号和第6,751,673号中被描述。如在步骤5示出的，几个中间服务器接收封装包，并且，再一次优选地基于从CDN映射系统提供的路由，将它们转发(直接地，或通过其他中间区域(未示出))到网关区域404。在步骤6，包由在网关区域404中的服务器接收，副本被删除。目的地NAT将虚拟IP转换到目标地址，并且源网络地址端口转换在包被发送之前被应用到包，使得返回流量也将在覆盖网络上被发送。优选地，信息被存储，使得返回流量被发送到客户端包起源的边缘区域406。在步骤7，网关区域404从目标地址接收IP包，以及de-NAT包。该包然后被封装。在步骤8，包的多个副本沿着多个路径进行发送。在步骤9，中间服务器发送包返回到关于该会话的源边缘区域。在步骤10，包由边缘服务器接收，且副本被删除。包起源于虚拟IP地址，且然后被发送返回到边缘区域。

[0059] 在覆盖网络中使用的以及如通常被描述的各种连接经由SSL或其他传输层安全(TLS)技术是安全的。

[0060] 虚拟专用网络(VPN)即服务

[0061] 一般地，本公开提供了各种机制来促进在覆盖IP(OIP)路由机构的上下文(如在图3中所示)内的虚拟专用网络(VPN)即服务(或更普遍地，“网络即服务”)，而且包括覆盖层(或内容分发)网络(“CDN”)的一部分，如在图1中所示。然而，这个实现不旨在进行限制，因为解决方案可在任何无线或有线网络、公共或专用、虚拟或其他网络中被支持。联网“即服务”的概念使企业能够有效地利用互联网链接作为优化的广域网(WAN)，通过互联网将分支、数据中心、远程工作人员和移动用户连接到应用。通过云计算和移动性的驱动，企业需要可以将最优的和可预测的云体验分发给用户的网络服务，优选地是具有内置的安全性和优化的低成本的、易于使用的，并且是全局的。

[0062] 图5示出了关于企业的典型的用例场景。企业具有通常在广域网多协议标签交换(MPLS)504上的一个或多个分支位置502所连接到的企业数据中心500。如上所述，MPLS是用于加速网络流量的基于标准的技术。在MPLS中，用于给定的包序列的特定的路径(通过标签来识别)被建立，由此避免路由器查找转发包的下一个地址。如在图5中所见，企业也可希望利用第三方软件即服务(SaaS)的解决方案506，如Microsoft Office365、Google App(谷歌应用)、salesforce.com，以及许多其他解决方案。

[0063] 图6示出使用MPLS的端对端路径。在这个例子场景中，企业分支办公室600与LAN-A相关联，并且企业数据中心602与LAN-B相关联。

[0064] 图7对应于图5但包括在每个端点处的覆盖网络设备700的内含物。设备700可以是具有一个或多个处理器、数据存储、存储器、联网支持和软件(操作系统、实用程序、软件应用，等等)的机架安装的硬件设备。如在图2中所示，边缘机器可以被实现为这类设备。此外，在设备中执行的软件可提供各种类型的端对端处理(例如，加密、去重，等等)以促进流过广

域网的流量。此外，以及依赖于配置，这个软件也提供额外的功能以使企业利用覆盖层，覆盖层包括例如WAN服务配置、服务管理、服务报告、远程管理访问、客户SSL证书管理，以及用于配置的网络应用的其他功能。设备可以是物理的或虚拟的(例如，被支持作为在虚拟环境内的虚拟机)。

[0065] 图8对应于图7，但是示出了图3－图4的覆盖网络800如何结合在每一个端点处的覆盖网络设备802被使用的。以这种方式，覆盖网络提供了现有技术的MPLS路由方法(路径804)的可选方案。

[0066] 图9示出了在一个端点处的覆盖网络设备900如何建立到附近的边缘区域或覆盖网络906的区域904的安全的逻辑VPN隧道902。在这种情况下，网络设备900通常提供基于IPsec的认证和基于流的加密，以及安全的逻辑的VPN隧道902可合并一个或多个加密流。优选地，边缘904使用通信(例如保活信号)以保持穿过企业防火墙的隧道。

[0067] 图10示出了在安装网络设备1002以及在第一侧(LAN-A)上建立到覆盖网络(表示为平台1006)的VPN隧道1004之后加入客户网络(ABC)的与覆盖网络客户相关联的第一办公室A1000。例如，使用基于云的门户UI，这种布置可被建立，一旦设备被安装，通过基于云的门户UI，客户的管理员手动地编辑/添加LAN子网以与设备相关联。

[0068] 图11示出了在其一侧(LAN-B)上安装网络设备1102以及建立VPN隧道1104之后加入客户网络ABC(图10的)的与覆盖网络客户相关联的第二办公室B1100。

[0069] 图12示出了在建立到它们的附近边缘(多个)区域的安全VPN隧道之后，多个覆盖网络客户如何使用在其各种位置处具有网络设备的覆盖网络平台1200。在这个实施例中，每个客户的不同的加密机密级被保持。

[0070] 图13示出了在到覆盖网络区域或多个区域的安全VPN隧道1306的建立之后，多个封装的和加密的流1300(这里，基于UDP)如何在网络设备1302和覆盖网络平台1304之间被提供。

[0071] 图14示出了端对端多站点VPN，其中，在覆盖网络1402上的MPLS1400和覆盖网络路由两者都在端点(即(例如)LAN-A 1404和LAN-B 1406)之间可被实现。在这里，覆盖网络提供加密的网状覆盖网络，其具有从设备到边缘的独立的认证、具体的客户和隔离的专有IP路由、具有数据包复制的多路径分发(使用在图3中所示的方法)以及如在图16的上下文中将要被描述的，专有的TCP/IP流的TCP终端。如在图14中示出的，覆盖网络客户可保持其专用的MPLS容量的链接1400，但其也可利用覆盖网络1402来提供低成本的可选路径。在一个实施例中，并且在安装网络设备之后，客户使用MPLS和覆盖网络链接两者。在另一个实施例中，并且根据给定的条件或事件，一些流量从MPLS链接迁移到覆盖层。客户的一定比例的流量可能被拆分，并使用覆盖层进行分发。一种类型的客户流量可以使用MPLS链路，而另一种类型的客户流量可以使用覆盖网络。因为一旦网络设备被定位及被配置，多个客户同时使用加密的网状覆盖网络，如在图12中所示的覆盖层优选地是多租户的。

[0072] 在图14中示出的方法将MPLS的可选方案和其他类似的专用容量的解决方案提供给覆盖网络客户。覆盖网络优化在关于全部流量类型网络设备之间的传输和路由，同时仍允许这些设备提供数据流和应用层优化(例如，加密、去重，等等)。

[0073] 图15示出了覆盖网络设备可如何被控制以配置端对端的VPN隧道来展示不同的安全性上下文。如在以前描述的，假定覆盖网络客户期望经由基于互联网的覆盖层1504连接

LAN 1500和LAN 1502。为此，设备1506被定位在每一个端点处或邻近每一个端点。在这个实施例中，设备则是可操作来创建和保持拆分的外部的安全性上下文1508，以及是安全的端对端的内部的安全性上下文1510。优选地，内部的安全性上下文1510被用于客户数据。拆分的外部的安全性上下文1508可包括第一段1512(是外部IPsec段)、中间安全性上下文段1514和第二外部段1516(其也是外部IPsec段)。因此，并可以看出，覆盖隧道被分割和分段，以及内部隧道是安全的端对端。外部隧道封装内部隧道。

[0074] 在图15中描绘的“嵌套的”方法中，单一的分段的VPN隧道端对端地分发包，但使用不同的安全性上下文。在这种方法中，每一个外部侧到覆盖网络段在段边界提供实现身份认证的能力。因为身份认证可在每一段的基础上实现，覆盖网络提供商可确保第三方不能不合适地使用覆盖网络，而同时使这种使用对于合法的覆盖网络客户开账单。该方法也阻止未被授权访问覆盖网络以用于获取对客户的内部网络的访问或获取优化的互联网传输及路由。

[0075] 在一个实现中，在网络设备(在每一个端)和覆盖网络之间传输的包使用IPsec封装安全协议(ESP)用于认证来限制对覆盖网络的访问。这些包优选地使用有自定义的ESP框架以保护客户联网的细节，例如，通过只加密网络和传输层协议报头。在这个示例实施例中，IPsec安全关联(SAs)被使用基于证书的相互认证来协商。在网络设备和覆盖网络之间交换的包上用于认证和报头加密的会话层端点是设备和覆盖网络区域。客户联网细节在覆盖层的入口区域被重新加密(使用共享的对称的每个客户密钥)以在运输到覆盖层出口区域途中时保护它们。覆盖网络提供商可提供合适的密钥管理基础设施用于管理及保护这种客户密钥。在覆盖网络区域之间交换的包上用于网络细节加密的会话层端点是覆盖层入口和出口区域。通过覆盖层中间区域进行数据加密/解密不是所要求的。

[0076] 具有传输协议层优化和端对端数据安全性的VPN即服务

[0077] 借助以上述背景，现在描述本公开优选的技术。

[0078] 说服客户把他们的流量转移到如上所述的覆盖网络，主要要求之一是将在公共互联网上相对的相当程度的隐私提供为他们所期望从他们现有的专用容量链接(如MPLS)提供的。本公开的技术提供了这个保证。具体地，如示出的及描述的(例如，见图15)，经由覆盖网络在互联网上流动的流量优选地使用认证、加密和网络管理以提供连接客户网络设备的虚拟专用网络(VPN)。客户的数据保持完全安全，但是流量的流动可利用由覆盖网络本身提供的好处。

[0079] 具体地，密钥的好处(从覆盖网络可获得的)是提供一个或多个TCP优化，如缓解包丢失、TCP缓冲区管理，以及其他TCP优化。为了使覆盖网络客户能够利用从覆盖网络提供商可获得的全范围的TCP优化，在图15中示出的技术被扩展为现在所描述的。具体地，图16示出了本公开主题的技术，其中，覆盖网络在一对设备、分发端对端的包(穿过路由覆盖层)的单一的分段的隧道之间建立及维持，但是，其中，对于每个数据包的特定部分实施拆分安全性上下文。在本公开中，“安全性上下文”通常指由一个或多个加密密钥实现的或与一个或多个加密密钥相关联的以及是相关的加密算法的“加密上下文”。加密上下文的特定的性质不是本公开的限制；而是，这里的概念是有穿过单一的分段的隧道1600实施的两(2)个有区别的和不同的加密上下文。第一加密上下文由在网络设备1604和1606(端对端隧道1600穿过网络设备1604和1606建立)之间共享的第一加密密钥(或“第一密钥”)1602定义。第二加

密上下文由在设备(如设备1604)和一些边缘机器(如设备关联的边缘服务器)之间共享的第二加密密钥(或“第二密钥”)1608定义。第一密钥不同于第二密钥,并且不跨加密上下文共享密钥。由于第一密钥和第二密钥以这种不同的方式使用,可见,第一加密上下文不同于第二加密上下文并区别于第二加密上下文。如上所述,第二加密上下文被拆分及分段。

[0080] 由于密钥不共享且不同,设备到设备的加密上下文不同于设备到边缘的加密上下文。以这种方式,穿过覆盖层的TCP连接可被拆分及分段以创建不同的TCP段1618(TCP1)、1620(TCP2)和1622(TCP3),而加密的数据流1610从不终止。然后,作为结果,TCP/IP报头流(其在不同的TCP段出现)可以,且优选地,根据这个优选的实施例被终止以使覆盖网络能够将一个或多个TCP优化应用到TCP流。

[0081] 为此,以及根据本公开,拆分加密上下文的方法被应用,一方面,相对于数据流,另一方面,相对于网络和传输协议层(例如,TCP/IP)报头。因此,第一设备和第二设备配置为建立和维持上述的第一加密上下文和第二加密上下文,其中,第一加密上下文在设备间端对端地延伸,以及,其中,第二加密上下文包括一组传输协议段,该组传输协议段包括在第一设备1604和入口点之间的第一外部传输协议段1618、在出口点和第二设备1606之间的第二外部传输协议段1622、在第一外部段和第二外部段之间的中间传输协议段1620。有可能是多个中间传输协议段。第一加密上下文保护第一设备和第二设备之间的端对端数据流,以及,第二加密上下文保护传输和网络协议层报头。具体地,在包的分发期间,包的数据部分具有第一加密上下文,以及包的TCP/IP部分具有第二加密上下文。因此,包括第二加密密钥的加密密钥可被用于在各种段边界启用TCP/IP流的解密。由于TCP/IP流(但不是相关的客户数据流)终止,覆盖网络可应用其各种分发优化。这些分发优化潜在地包括用于TCP流的TCP优化,以及穿过覆盖层的路由优化两者。因此,由于流量穿过覆盖层进行分发,那些分发的优化可被应用(基于分段式),同时仍然确保流的数据部分在明文中绝不可获得。

[0082] 然后如所见,这个方法扩展在图15的方法,使传输协议(例如TCP)和路由优化以能够被应用。

[0083] 加密被应用在的特定加密上下文内的特定方式可变化。这些已知的技术包括预建立对称密钥、基于证书的相互的认证和许多其他技术。优选的技术是为每一个网络设备优选地使用具有基于证书的相互认证的TLS来提供用于设备到设备的TCP连接的数据加密。如前所述,在这个水平经由拆分的加密上下文应用加密允许覆盖网络的现有的TCP层优化技术的应用而不需要在边缘上的数据解密。优选地,在网络设备和覆盖网络边缘之间的流量流动使用具有X.509证书的基于证书的相互认证。此外,优选地,每个客户设备白名单也被分布到覆盖网络以用于设备认证。在一个示例实施例中,认证和对称密钥生成/交换是利用IKEv2(RFC 5996)实施的,其中包在隧道模式中在设备和覆盖网络之间使用IPsec封装安全净荷(RFC4303)进行交换的。如上所述,这些ESP包为整个包提供数据完整性,但由于数据已知是加密的,不需要为了传输到覆盖网络而被重新加密。

[0084] 优选地,在隧穿的包中的IP和传输协议报头在公共互联网上传输之前被加密。仅加密网络和传输层报头提高了包处理的性能特性。

[0085] 如在图16中示出的,除了用于数据流优化的TCP连接,优选地,网络设备也支持对于面向数据报的流(如,UDP、ICMP、非终止的TCP,等等)的优化的路由。如果它们通过基于TCP的SSL流被发送,在这个类别中的大部分的流将受到负面影响,因此,优选地使用基于数

据报的加密方法。可用于这个目的的两个例子选项包括数据报TLS(DTLS),以及合并的互联网密钥交换(IKE)和IP安全性(IPSec)。DTLS提供相互认证、数据完整性和数据安全功能。IKE和IPSec一起起到DTLS相同的作用。IKE提供相互认证和安全的对称密钥交换。经由IKE交换的密钥可通过IPsec来使用来以提供每个包的认证、数据完整性和数据安全性。

[0086] IP Sec的负载平衡

[0087] 在单一的网络设备后可有任何数量的客户机,这意味着每一个设备可与大量的不同的连接相关联。此外,单一的覆盖网络区域可作为任何数量的不同的设备的隧道入口点。由于这些原因,理想的是能够通过在该区域内多个机器间分布进入的包来平衡在区域内的IPsec处理负载。以下部分提供关于这个优化的额外的细节。

[0088] 为了限制与区域中转发相关联的负载,理想的是在它们在区域中被转发前避免负载平衡器对验证和解密包的需要。同时,当包被分布在多个机器之间时,期望的是,与具体的数据流相关联的包通过相同的机器来处理,以最小化与包重新排序相关联的问题。

[0089] 对于在区中权量轻的转发,系统使用自定义的IPsec协议安全性参数索引(SPI)字段,允许分派的机器嵌入在SPI中的主机标识符。除了在SPI中的主机标识符,IPsec包也将一个字节流ID散列增加到ESP有效载荷格式。流ID具有多个用途,其中一个用途是帮助做出低成本的负载平衡决策。

[0090] 优选地,每一个安全关联(SA)与一对安全性参数索引(SPI)相关联,一个用于“隧道”的每一个端。每一个SA主机使用其自己的SPI作为在其本地数据库内的唯一的密钥,这允许其查找SA并且取回用于认证和加密的密钥,以及用于重演保护的相关的序列号。被发送的每个IPsec包开始于SPI,预期的接收器将用于在其自己的SA数据库中查找SA。SPI是32位数字,其对于选择它的SA端点主机必须是唯一的标识符,这允许个体端点(即,网络设备或覆盖网络区域)支持大量激活的SA。

[0091] 当SA经由IKE进行协商时,SPI选择优选地被限制在可用25位表示的数字的范围。每个所选的SPI(数字的主机ID部分)的高7位将因此是0。当在负载平衡区域的隧道端点将包发送到它的同层,其在SPI的高7位中编码其自己的主机ID,导致机器专门的SPI出现在包中。通过跟踪在接收的包中编码的主机ID,非负载平衡端点能够智能地选择主机ID以在其发送的包中进行编码,使得负载平衡器能够识别应当接收每一个包而无需隧穿的包报头的认证或加密的机器。当非负载平衡端点还不具有关联流的合适的主机ID时,它使用特定的主机ID127(在高7位中全是1)来表示该情形。

[0092] 除了修改SPI字段,系统优选地将一个字节的流ID散列添加到ESP有效载荷。该值是由该包使用的地址和端口散列一起产生。这个流ID散列不能用于可靠地识别连接,但其可被接收器用来在负载平衡区域内做出一致的负载平衡的决策。当更具体的主机ID在SPI中尚未被提供时,流ID散列可被负载平衡器用来将流分配给机器。

[0093] IPsec使用64位的序列号用于重演保护。每当包被发送,序列号递增一。接收器应该跟踪它已经使用滑动窗口接收的序列号,丢弃在窗外或在窗口中但是先前接收的包。虽然当计算包的完整性校验值时高32位被包括,但只有序列号的低阶32位出现在IPsec包中。换句话说,高阶32位是必须由发送器和接收器两者已知而没有曾经被明确地发送的值。

[0094] 重演保护的这个方法对于负载平衡的隧道端点是成问题的,其中,完整的区域将作为单一的逻辑的IPsec端点,共享安全关联。出于这个原因,优选地,在该区域中的每一个

个体计算机和非负载平衡端点维持对于特定主机的SPI值中的每一个序列号是唯一的。全局SPI共享序列号，但是这个SPI仅用于传输到负载平衡区域，而不是来自它。对于可能需要验证使用全局SPI以正常运行的包的所有机器，全局SPI的接收的序列号周期性地在区域内同步。只要它比32位低阶序列号进行包装所需要的更频繁地同步，该地区所有的机器会知道高阶32位是什么，允许它们计算必要的完整性校验值。

[0095] 每一个上述进程在计算机软件中优选地实现为在作为专用机器的一个或多个处理器中可执行的一组程序指令。

[0096] 在本文中所提供的主题上有代表性的机器可以是运行Linux或Linux变种的操作系统和一个或多个应用来执行所描述的功能的基于Intel Pentium的计算机。上述的一个或多个进程被实现为计算机程序，即，作为用于执行描述的功能的一组计算机指令。

[0097] 虽然以上描述了由本发明的某些实施例执行的操作的特定顺序，但应理解的是，这样的顺序是示例性的，因为替代的实施例可以以不同的顺序执行操作，组合某些操作，重叠某些操作，或类似操作。在本说明书中对给定的实施例的参考指示了所描述的实施例可包括特定特征、结构或特性，但是每个实施例可以不必包括该特定特征、结构或特性。

[0098] 而所公开的主题已经在方法或进程的上下文中进行了描述，该主题还涉及用于执行在本文中的操作的设备。此类设备可以是被特别构造用于期望的目的特定的机器，或其可以包括通过在计算机中存储的计算机程序选择性激活或重新配置的计算机。这种计算机程序可以存储在计算机可读存储介质中，比如，但不限于，包括光盘、CD-ROM和磁光盘、只读存储器(ROM)、随机存取存储器(RAM)、磁卡或光卡，或任何类型的适于存储电子指令的介质且每一个耦合到计算机系统总线的任何类型的盘。本发明的给定的实现是以给定的编程语言编写的结合兼容的DNS名称服务器(例如，BIND)在运行操作系统(如Linux)的标准的Intel硬件平台上运行的软件。功能可以被构建到名称服务器代码中，或者它可以作为辅助代码被执行。本文中实现技术的机器包括处理器、保存由处理器执行的以执行上述方法的指令的计算机存储器。

[0099] 虽然该系统的给定的组件被分开描述，但是普通技术人员将理解的是，功能中的一些在给定的指令、程序序列、代码部分，以及类似物中可被组合或被共享。

[0100] 虽然该系统的给定的组件被分开描述，但是普通技术人员将理解的是，功能中的一些在给定的指令、程序序列、代码部分，以及类似物中可被组合或被共享。通过将钩子提供到另一个应用程序中，通过促进使用机构作为插件，通过链接到机构，等等，本文中描述的任何应用或功能可以被实现为本地代码。

[0101] 本文中的技术通常提供对技术或技术领域的上述改进，以及包括分布式联网、基于互联网的覆盖层、基于WAN的联网(使用MPLS或其他)、互联网链接的安全利用等等(全部如上所述)的具体技术改进。

[0102] 已经描述了我们的发明，我们现在所要求保护的在下面被陈述。

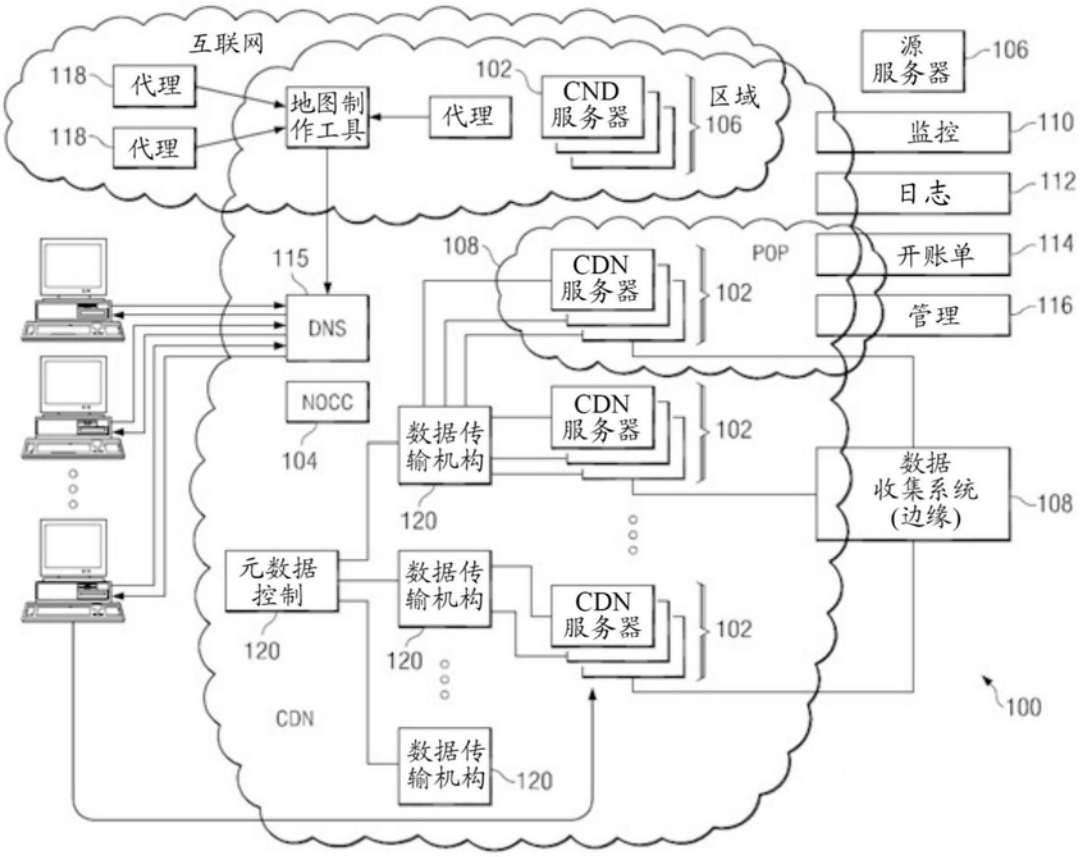

互联网
118
代理
118
代理
地图制作工具
代理
102
CND服务器
区域
106
源服务器
106
监控
110
日志
112
开账单
114
管理
116
115
DNS
108
CDN服务器
POP
102
NOCC
104
数据传输机构
120
CDN服务器
102
数据收集系统(边缘)
108
元数据控制
120
数据传输机构
120
CDN服务器
102
CDN
数据传输机构
120
100

图1

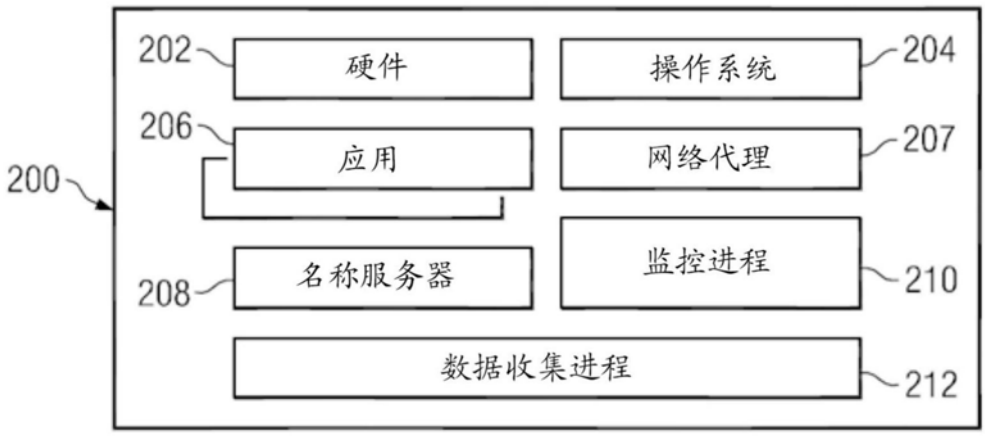

202
硬件
操作系统
204
206
应用
网络代理
207
200
名称服务器
208
监控进程
210
数据收集进程
212

图2

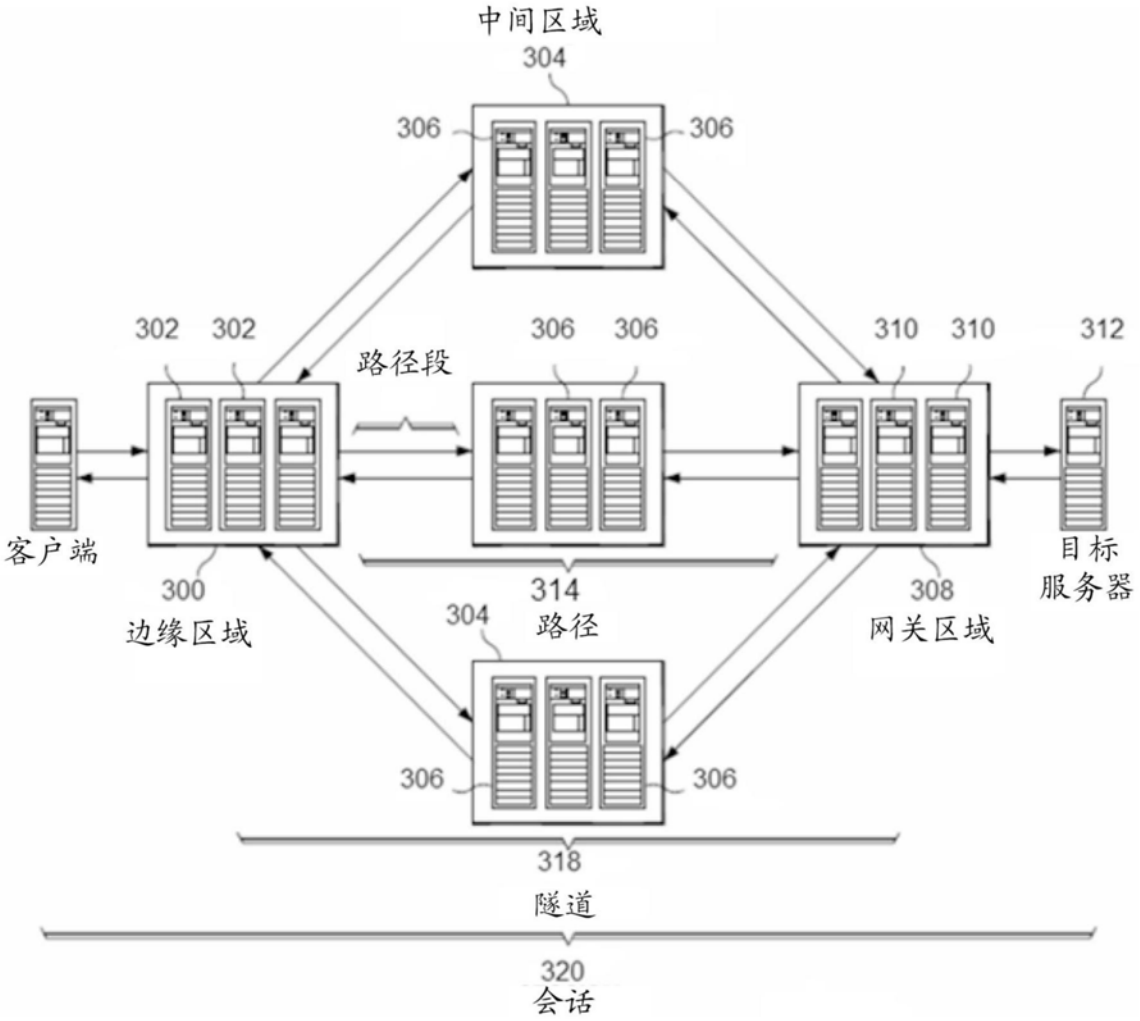

中间区域
304
306
306
302
302
路径段
306
306
310
310
312
客户端
300
边缘区域
314
304
路径
308
网关区域
目标
服务器
306
306
318
隧道
320
会话

图3

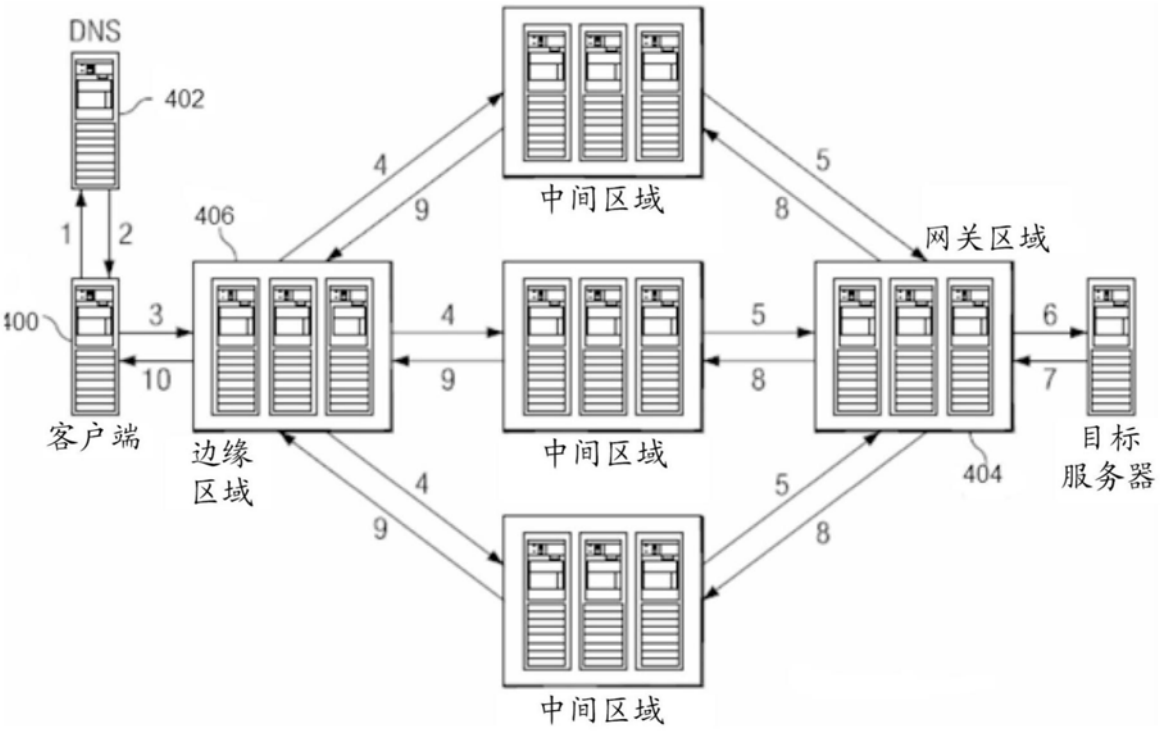

DNS
402
406
100
客户端
边缘区域
中间区域
中间区域
中间区域
网关区域
目标服务器
404
1
2
3
4
5
6
7
8
9
10

图4

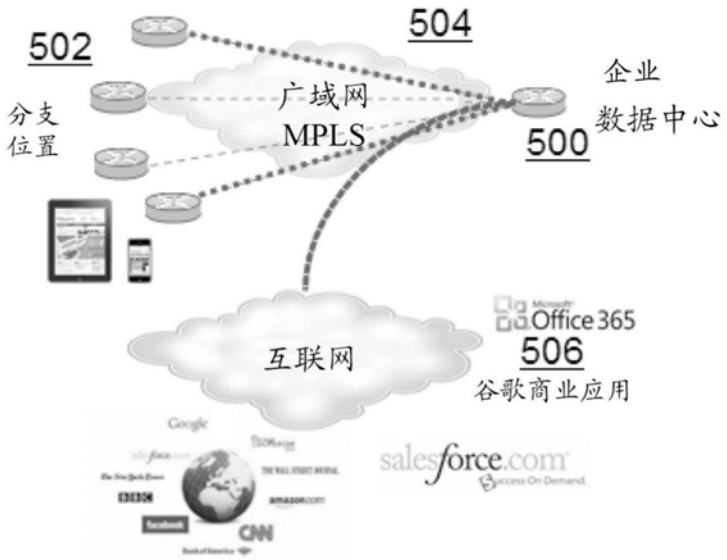

502
504
分支位置
广域网
MPLS
企业数据中心
500
互联网
Office 365
506
谷歌商业应用
Google
salesforce.com
BBC
CNN
amazon.com

图5

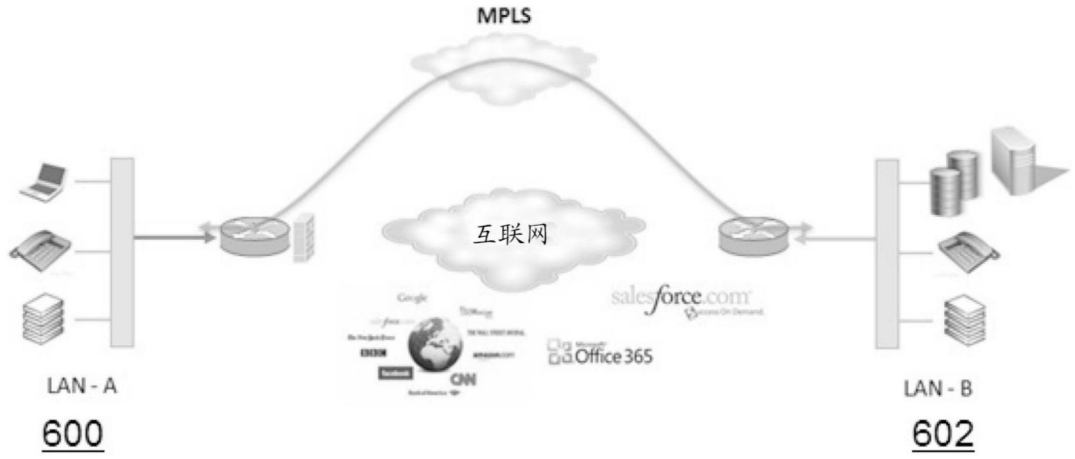

MPLS
互联网
LAN - A
600
LAN - B
602
Office 365

图6

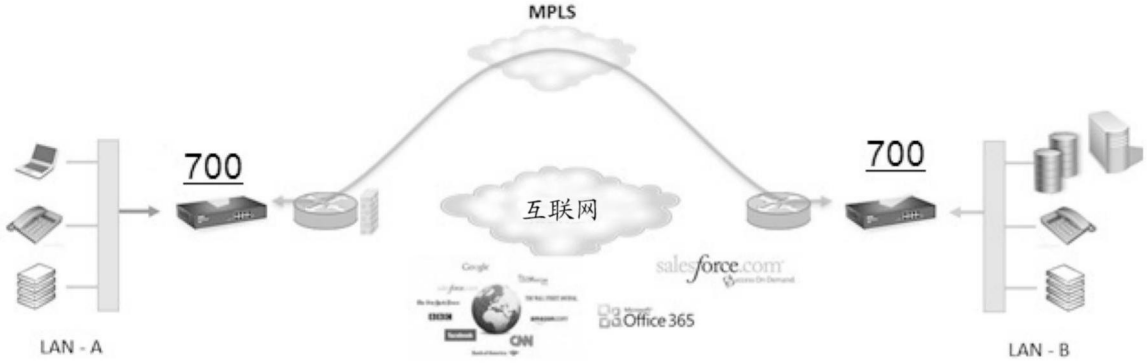

MPLS
700
700
互联网
LAN - A
LAN - B
Office 365

图7

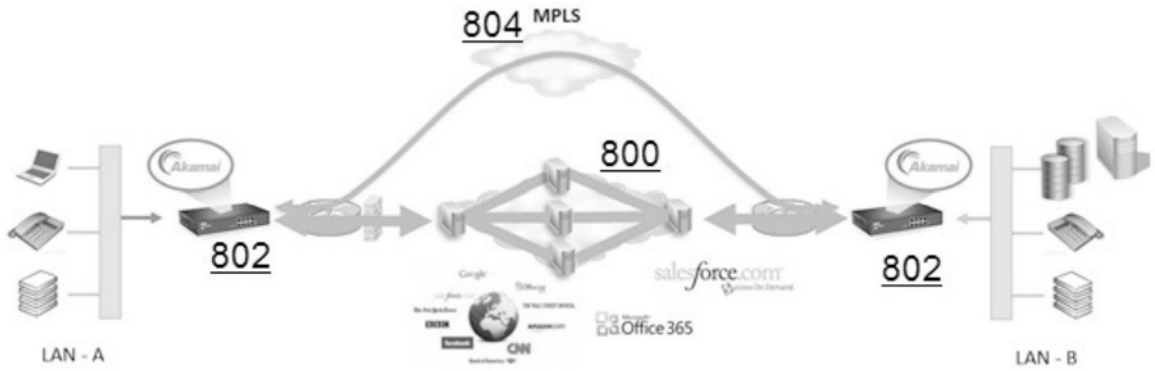

804
MPLS
800
802
802
LAN - A
LAN - B
Office 365

图8

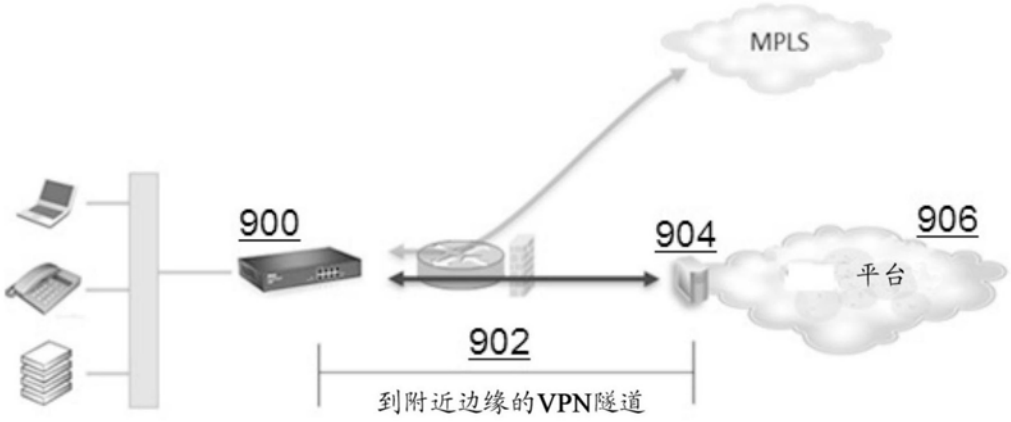

MPLS
900
904
906
平台
902
到附近边缘的VPN隧道

图9

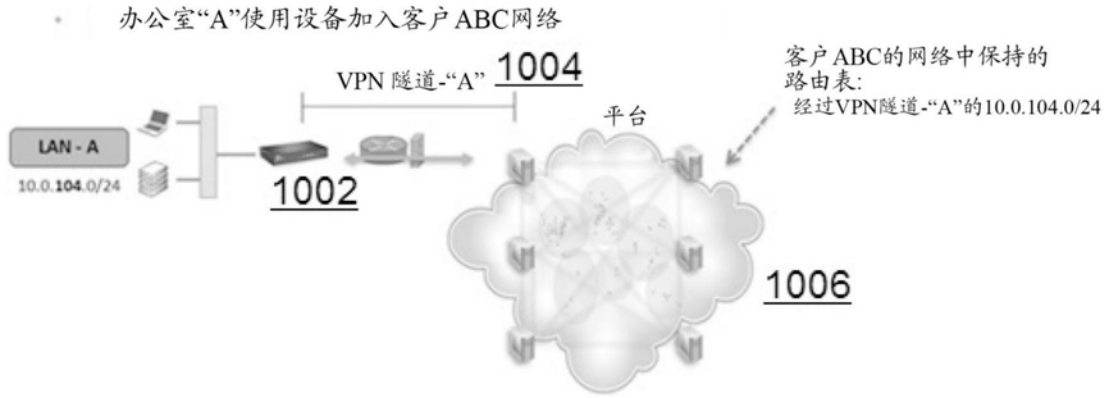

办公室“A”使用设备加入客户ABC网络
VPN 隧道-“A” 1004
客户ABC的网络中保持的
路由表:
经过VPN隧道-“A”的10.0.104.0/24
平台
LAN - A
10.0.104.0/24
1002
1006

图10

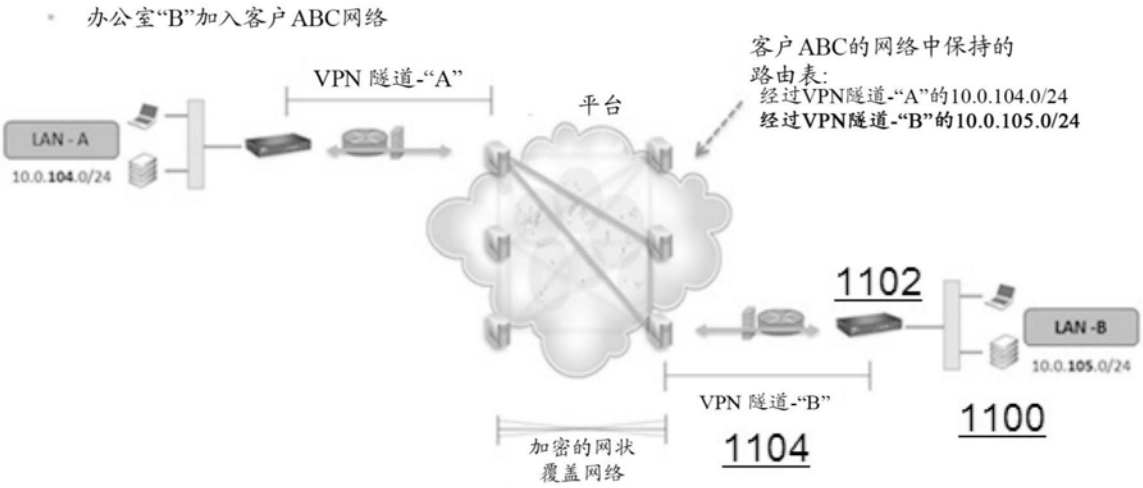

办公室“B”加入客户ABC网络
客户ABC的网络中保持的
路由表:
经过VPN隧道-“A”的10.0.104.0/24
经过VPN隧道-“B”的10.0.105.0/24
VPN 隧道-“A”
平台
LAN - A
10.0.104.0/24
1102
LAN -B
10.0.105.0/24
VPN 隧道-“B”
1100
加密的网状
覆盖网络
1104

图11

安全VPN网。在互联网上每个客户不同的加密机密

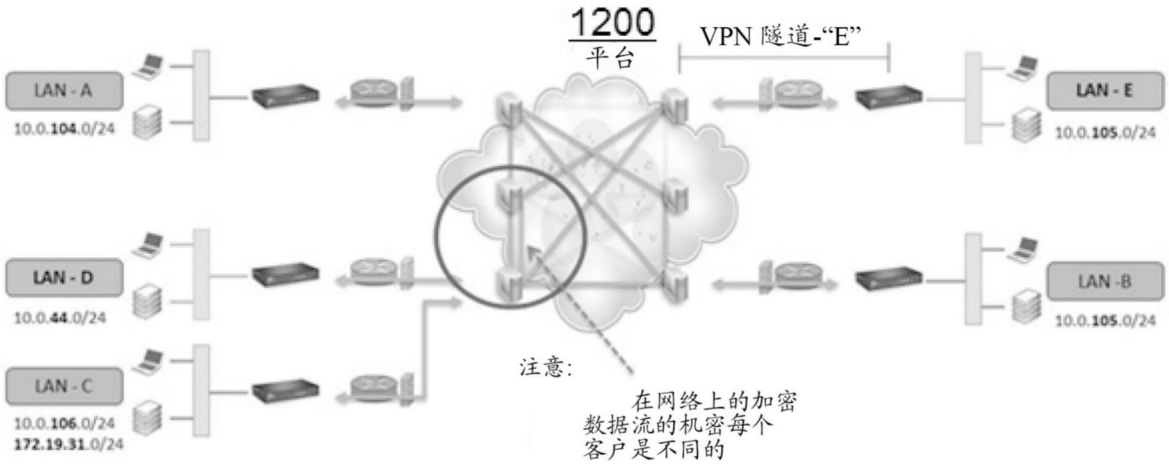


图12

多个封装的和加密的UDP 流

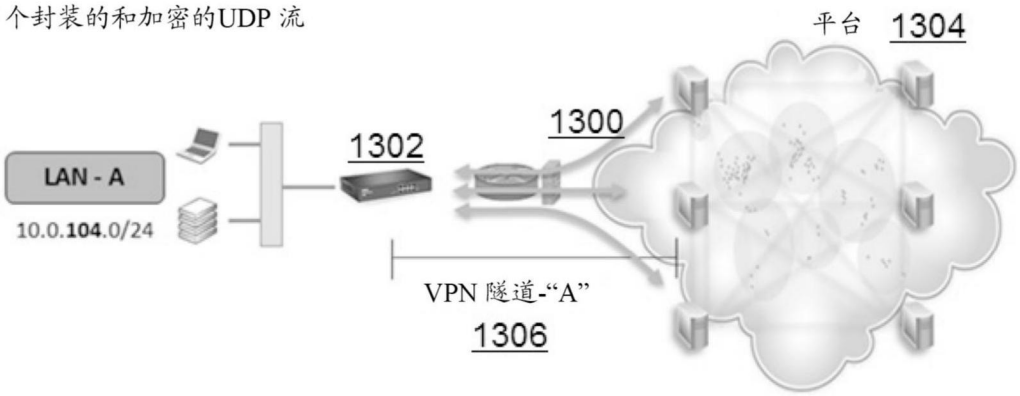


图13

多站点VPN扩展的特征

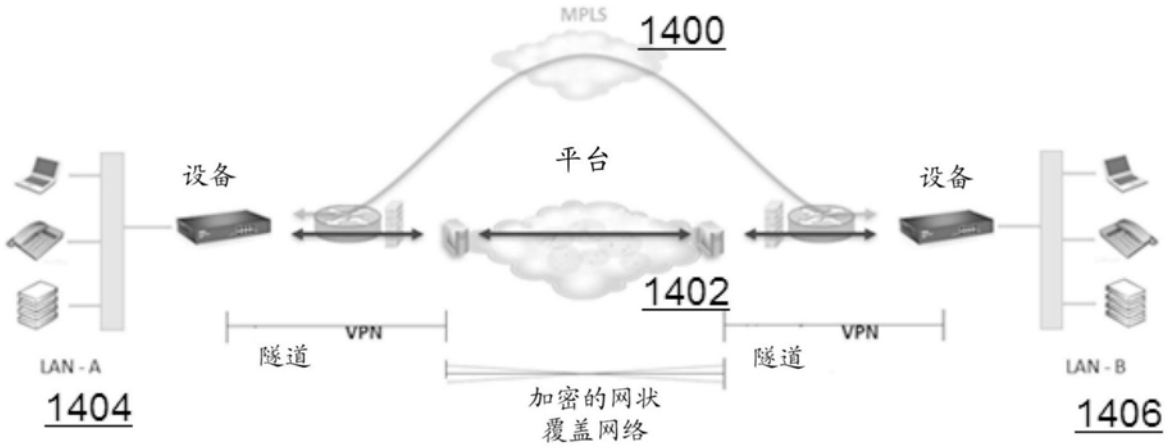


图14

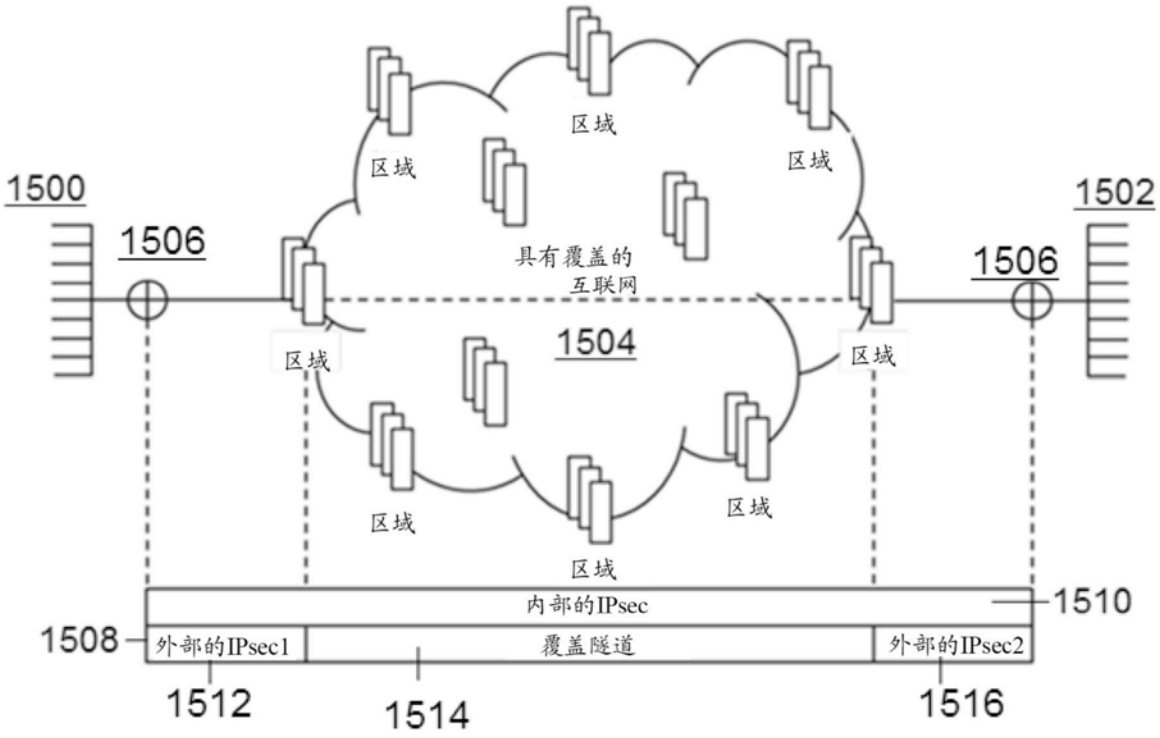

1500
1502
1506
1506
具有覆盖的
互联网
1504
区域
区域
区域
区域
区域
区域
区域
区域
内部的IPsec
1510
1508
外部的IPsec1
覆盖隧道
外部的IPsec2
1512
1514
1516

图15

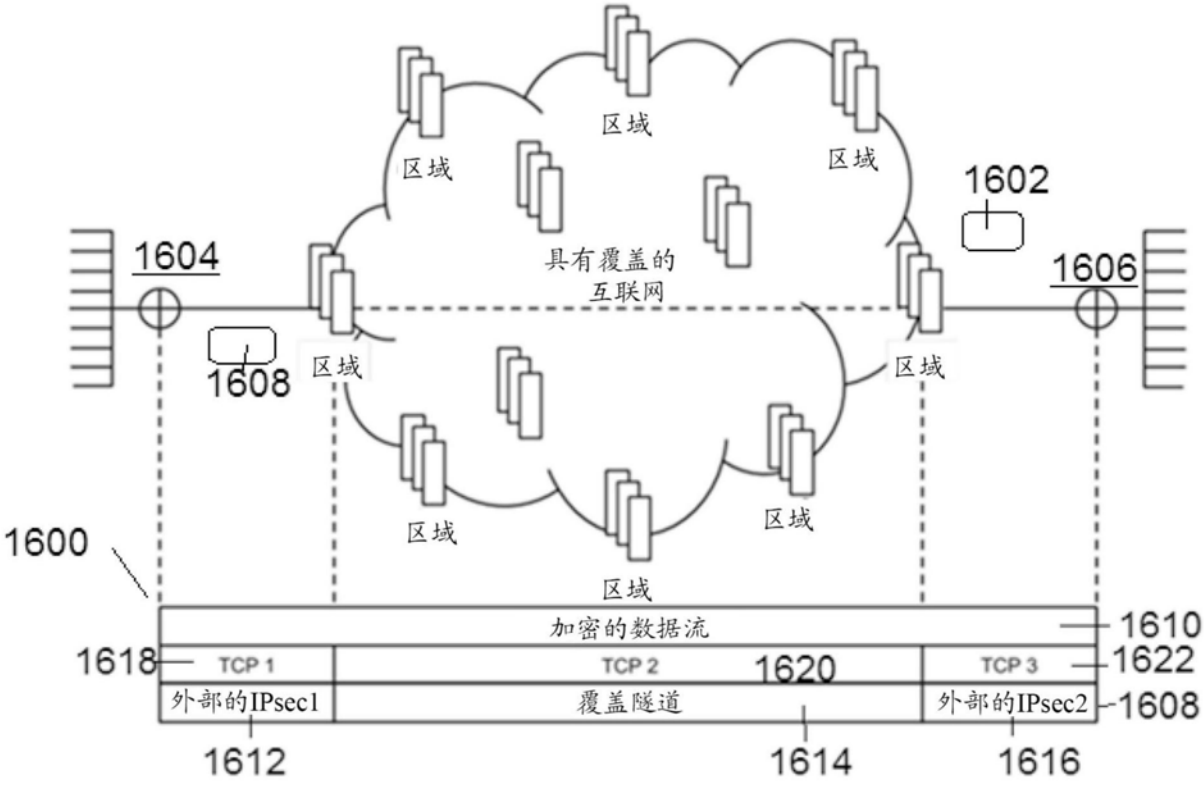

1602
1604
1606
具有覆盖的
互联网
1608
区域
区域
区域
区域
区域
区域
区域
1600
加密的数据流
1610
1618
TCP 1
TCP 2
1620
TCP 3
1622
外部的IPsec1
覆盖隧道
外部的IPsec2
1608
1612
1614
1616

图16