



(12) 发明专利

(10) 授权公告号 CN 111316254 B

(45) 授权公告日 2023.05.05

(21) 申请号 201880072247.0

(22) 申请日 2018.11.20

(65) 同一申请的已公布的文献号
申请公布号 CN 111316254 A

(43) 申请公布日 2020.06.19

(30) 优先权数据
62/589,163 2017.11.21 US

(85) PCT国际申请进入国家阶段日
2020.05.08

(86) PCT国际申请的申请数据
PCT/US2018/062133 2018.11.20

(87) PCT国际申请的公布数据
W02019/104098 EN 2019.05.31

(73) 专利权人 塔拉里网络股份有限公司
地址 美国加利福尼亚

(72) 发明人 J·奥斯瓦尔德 R·J·哈奈德

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

专利代理师 冯薇

(51) Int.Cl.
G06F 15/16 (2006.01)
H04L 12/28 (2006.01)
H04L 12/46 (2006.01)
H04L 9/32 (2006.01)
H04L 9/40 (2022.01)

(56) 对比文件
US 2016182319 A1, 2016.06.23
US 2007250552 A1, 2007.10.25
CN 104662994 A, 2015.05.27
US 2004117438 A1, 2004.06.17
US 2002095470 A1, 2002.07.18
US 2007047901 A1, 2007.03.01
曹风华. 一种基于授权机制的分布式文件系统小文件访问优化策略.《计算机系统应用》.2013, (第07期), 第178+185页-第188页.

审查员 包文韬

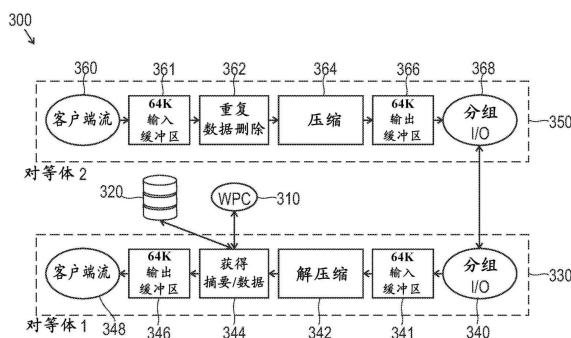
权利要求书2页 说明书12页 附图8页

(54) 发明名称

自适应专用网络 (APN) 中的智能写保护高速缓存 (IWPC)

(57) 摘要

在处理和固态硬盘 (SSD) 上的盘高速缓存之间引入写保护高速缓存 (WPC)。该高速缓存提供了保持区域, 并且仅有已被看到多于一次的数据才被提交到盘高速缓存。WPC通过避免对驱动器进行不必要的写入来为磁盘上的数据提供保护层。通过仅将广域网 (WAN) 处理看到多于一次的数据或块提交到盘来完成这种保护。



1. 一种用于减少数据存储的广域网WAN优化方法,该方法包括:
第一次接收数据和数据的摘要;
响应于确定这是第一次接收到数据和摘要,将摘要存储在写保护高速缓存中;
第二次接收数据和数据的摘要;
响应于确定这是第二次接收到数据和摘要,将数据和数据的摘要存储在主盘上并且从写保护高速缓存中移除摘要;
接收关于另外的传送的摘要;以及
基于接收到的摘要从主盘访问数据。
2. 如权利要求1所述的方法,其中将数据和数据的摘要存储在主盘上将仅在WAN出口处理期间发生。
3. 如权利要求1所述的方法,其中所述第一次接收数据和数据的摘要发生在WAN出口处理中。
4. 如权利要求1所述的方法,其中所述摘要包括安全散列算法摘要。
5. 如权利要求1所述的方法,其中数据和摘要一起包括块,其中典型的块在1.5kB至22kB之间,平均尺寸为8kB。
6. 如权利要求3所述的方法,还包括:
通过WAN入口处理接收到达局域网的数据;以及
将数据分组成块。
7. 如权利要求6所述的方法,还包括:
通过WAN入口处理来计算针对每个块的摘要。
8. 如权利要求7所述的方法,还包括:
向出口处理发送针对第一块的第一摘要。
9. 如权利要求8所述的方法,还包括:
如果出口处理具有存储在主盘中的第一块,那么由出口处理以确认响应ACK进行响应;
以及
如果出口处理不具有存储在主盘中的第一块,那么由出口处理以非确认响应NACK进行响应。
10. 如权利要求9所述的方法,还包括:
如果响应是NACK,那么将数据的第一块发送到出口处理。
11. 一种用于减少数据存储的广域网WAN优化装置,包括:
用于第一次接收数据和数据的摘要的部件;
用于响应于确定这是第一次接收到数据和摘要而将摘要存储在写保护高速缓存中的部件;
用于第二次接收数据和数据的摘要的部件;
用于响应于确定这是第二次接收到数据和摘要而将数据和数据的摘要存储在主盘上并且从写保护高速缓存中移除摘要的部件;
用于接收关于另外的传送的摘要的部件;以及
用于基于接收到的摘要从主盘访问数据的部件。
12. 如权利要求11所述的装置,其中所述用于第一次接收数据和数据的摘要的部件包

括WAN出口处理器。

13. 如权利要求12所述的装置,其中将数据和数据的摘要存储在主盘上将仅在WAN出口处理期间发生。

14. 如权利要求11所述的装置,其中所述摘要包括安全散列算法摘要。

15. 如权利要求11所述的装置,其中数据和摘要一起包括块,其中典型的块在1.5kB至22kB之间,平均尺寸为8kB。

16. 如权利要求12所述的装置,还包括:

WAN入口处理器接收到达局域网的数据,WAN入口处理器将数据分组成块。

17. 如权利要求16所述的装置,还包括:

WAN入口处理器计算针对每个块的摘要。

18. 如权利要求17所述的装置,还包括:

WAN入口处理器向出口处理器发送针对第一块的第一摘要。

19. 如权利要求18所述的装置,还包括:

如果出口处理器具有存储在主盘中的第一块,那么出口处理器以确认响应ACK进行响应;以及

如果出口处理不具有存储在主盘中的第一块,那么出口处理器以非确认响应NACK进行响应。

20. 如权利要求19所述的装置,还包括:

如果响应是NACK,那么用于发送数据的第一块的部件。

21. 一种其上存储有可执行指令的非瞬态计算机可读介质,所述可执行指令在由处理器执行时,使得所述处理器执行如权利要求1-10中任一项所述的方法。

22. 一种系统,包括:

处理器;以及

存储器,耦合到所述处理器并且所述存储器上存储有可执行指令,所述可执行指令在由所述处理器执行时,使得所述处理器执行如权利要求1-10中任一项所述的方法。

自适应专用网络 (APN) 中的智能写保护高速缓存 (IWPC)

[0001] 本申请要求于2018年11月20日提交的序列号为16/196,358的美国专利申请的优先权,并要求于2017年11月21日提交的序列号为62/589,163的美国临时专利申请的权益,这两个申请都通过引用整体并入本文。

[0002] 相关申请的交叉引用

[0003] 本申请还涉及于2014年1月3日提交的序列号为14/146,786的美国专利申请,其是2012年8月23日提交的序列号为13/592,460的美国专利申请的分案,13/592,460被发布为编号为8,644,164的美国专利,该专利是于2012年1月19日提交的序列号为13/353,693的美国专利申请的继续申请,13/353,693被发布为编号为8,274,891的美国专利,该专利要求于2009年6月11日提交的编号为12/482,766的美国专利申请的权益和优先权,12/482,766被发布为题为“Flow-Based Adaptive Private Network with Multiple WAN-Paths”的编号为8,125,907的美国专利,所有这些都要求于2008年6月12日提交的题为“Flow-based Adaptive Private Network with Multiple WAN-Paths”的编号为61/060,846的美国临时专利申请的权益;于2014年5月30日提交的序列号为14/291,776的美国专利申请,其是于2013年3月26日提交的序列号为13/850,411的美国专利申请的继续申请,13/850,411被发布为编号为8,775,547的美国专利,并且是于2011年8月12日提交的题为“Adaptive Private Network Asynchronous Distributed Shared Memory Services”的序列号为13/208,825的美国专利申请的继续申请,13/208,825被发布为编号为8,452,846的美国专利,所有这些都要求于2010年8月12日提交的题为“Adaptive Private Network Asynchronous Distributed Shared Memory Services”的序列号为61/372,904的美国临时专利申请的权益;于2012年12月19日提交的题为“An Adaptive Private Network with Geographically Redundant Network Control Nodes”的序列号为13/719,433的美国专利申请;于2013年9月6日提交的题为“An Adaptive Private Network with Path Maximum Transmission Unit (MTU) Discovery Process”的编号为14/019,723的美国专利申请;于2014年9月9日提交的题为“Adaptive Private Network with Dynamic Conduit Process”的序列号为14/481,335的美国专利申请;于2015年12月17日提交的题为“Methods and Apparatus for Providing Adaptive Private Network Centralized Management System Discovery Processes”的序列号为14/972,270的美国专利申请;于2015年12月17日提交的题为“Methods and Apparatus for Providing Adaptive Private Network Centralized Management System Timestamp Correlation Processes”的序列号为14/972,353的美国专利申请;于2015年12月17日提交的题为“Methods and Apparatus for Providing Adaptive Private Network Database Schema Migration and Management Processes”的序列号为14/972,514的美国专利申请;于2015年12月17日提交的题为“Methods and Apparatus for Providing Adaptive Private Network Centralized Management System Data Visualization Processes”的序列号为14/973,193的美国专利申请;于2015年12月17日提交的题为“Methods and Apparatus for Providing Adaptive Private Network Centralized Management System Time Correlated Playback of Network

Traffic”的序列号为14/973,343的美国专利申请;于2017年1月18日提交并且被公开为序列号为2017-0207996A1的题为“Methods and Apparatus for Configuring a Standby WAN Link in an Adaptive Private Network”的美国专利申请的序列号为15/409,001的美国专利申请;于2017年1月18日提交并且被公开为序列号为2017-0207997A1的题为“Methods And Apparatus For Accessing Selectable Application Processing Of Data Packets In An Adaptive Private Network”的美国专利申请的序列号为15/409,006的美国专利申请;于2017年1月18日提交并且被公开为序列号为2017-0207963A1题为“Methods and Apparatus for Accessing Dynamic Routing Information from Networks Coupled to a Wide Area Network(WAN) to Determine Optimized End-To-End Routing Paths”的美国专利申请的序列号为15/409,016的美国专利申请;以及于2017年1月18日提交并且被公开为序列号为2017-0207976A1题为“Adaptive Private Network (APN)Bandwidth Enhancements”的美国专利申请的序列号为15/409,019的美国专利申请,所有这些都与本申请具有相同的受让方,并且由此通过引用整体并入本文。

技术领域

[0004] 本发明一般而言涉及改进的广域网(WAN)优化和智能高速缓存技术。更具体而言,本发明涉及在自适应专用网络(APN)的上下文中使用盘和RAM资源来管理用于重复数据删除、重复数据的消除或减少的网络数据的高速缓存。

背景技术

[0005] WAN优化(WANop)是被设计为为了多于一个用户频繁查看的大块业务文件(bulk traffic file)和其它信息而提高跨网络的带宽效率的联网改进。这种类型的数据的常见示例是软件更新和共享文档,这将有益于传输一次并供多个用户使用。WANop设备可以将穿越(traverse)网络的数据进行高速缓存,并识别何时复制该数据,以便可以跨WAN发送对高速缓存的数据的引用来代替重复数据。WANop还通过向尚未压缩数据的网络协议添加压缩来改进网络。使用WANop的益处是现有的网络应用可以利用更少的WAN带宽,而无需对应用本身进行改变。

[0006] 在进行重复数据删除时,进行广域网(WAN)优化(WANop)的产品必须处理关于设备上可用的RAM和盘存储空间(storage)有限的难题。当看到网络数据时,网络节点设备将无法知道是否将再次看到该数据。在一种先前的方法中,WANop将把它在入口或出口上接收到的任何新数据直接写到盘。在本发明的几个方面当中,本发明认识到这种方法可能导致显著的驱动器磨损并且可以用非重复数据代替频繁使用的数据暖块(warm chunk)。本发明还认识到,将所有数据进行高速缓存需要大量的RAM和盘存储空间来处理高带宽WAN链路,并且这大大提高了WANop网络设备的价格。本发明还认识到,理想地,网络设备仅希望将未来会再次出现的数据进行高速缓存。本发明还认识到,最常见的数据传送来自存储在数据中心的中心的信息,该信息被给定客户端站点处的多个用户重复访问。

[0007] 而且,固态驱动器(SSD)技术虽然速度惊人,但其使用寿命却比硬盘驱动器(HDD)短得多。例如,在仅写入14TB数据之后,128MB驱动器就可能发生故障。如果驱动器在设备保修期内发生故障,那么会对提供保修的公司产生财务影响。

发明内容

[0008] 在本发明的几个方面当中,本发明认识到以较低的成本在网络中提供改进的广域网(WAN)优化是有利的。代替总是在本文所述的上下文中保存数据,发送数据和数据的摘要。在第一次看到数据后,摘要被存储在WPC中。在第二次看到数据并传送了相同的摘要后,摘要和数据被存储在主盘上,并且摘要从WPC中被移除。此后,在进行第三次及后续的任何传送时,仅摘要被发送,因为数据已经存储在主盘上。为此,本发明的实施例应用减少数据存储的方法。该方法包括:第一次接收数据和数据的摘要;响应于确定这是第一次接收到数据和摘要,将摘要存储在写保护高速缓存中;第二次接收数据和数据的摘要;将数据和数据的摘要存储在主盘上,并从写保护高速缓存中移除摘要;接收关于另外的传送的摘要;以及基于接收到的摘要从主盘访问数据。

[0009] 通过下面的具体实施方式、附图和权利要求,对本发明以及本发明的其它特征和优点的更全面理解将是显而易见的。

附图说明

[0010] 结合附图,根据以下描述和所附权利要求,本发明的示例性实施例将变得更加显而易见。应该理解的是,这些附图仅描绘了示例性实施例,因此,不应被认为限制本发明的范围,将通过使用所附文档中包含的附图以附加的具体性和细节对本发明的示例性实施例进行描述。

[0011] 图1A图示了根据本发明实施例的端到端网络,该网络配置有在自适应专用网络(APN)的控制下的在耦合到广域网(WAN)的局域网(LAN)上的客户端节点,示出了服务路径并采用如结合图3-图10进一步描述的改进的WANop处理;

[0012] 图1B图示了根据本发明实施例的自适应专用网络(APN)管道(conduit),其提供包括在客户端站点A和客户端站点B之间的集成防火墙的双端服务并且采用如结合图3-图10进一步描述的改进的WANop处理;

[0013] 图2图示了根据本发明实施例的结合如结合图3-图10所描述的改进的WANop处理采用的APN,该APN具有APN网络控制节点(NCN)和耦合到十六个APN客户端站点的十六个APN管道;

[0014] 图3示出了根据本发明的采用写保护高速缓存(WPC)的说明性WANop处理的实施例的框图;

[0015] 图4图示了根据本发明的入口处理的方面;

[0016] 图5示出了根据本发明的用于利用WANop处理的各种平台的建模的设备数据的表;

[0017] 图6图示了根据本发明的先前未看到的数据的数据流;

[0018] 图7图示了根据本发明的第二次看到的数据的数据流;

[0019] 图8图示了根据本发明的第三次及后续时间看到的数据的数据流;

[0020] 图9A和图9B(统称为图9)示出了根据本发明的出口对等体处理的示例;以及

[0021] 图10示出了本发明的虚拟设备实施例的数据的表。

具体实施方式

[0022] 图1A图示了根据本发明实施例的示例性自适应专用网络(APN) 100,该APN 100具

有在该APN的控制下的分别在耦合到广域网(WAN) 102的局域网LAN1 101和LAN2 103上的本地站点 124_{A1} 、 124_{A2} 、 \dots 、 124_{AN} 、本地站点 125_{B1} 、 125_{B2} 、 \dots 、 125_{BN} ，示出了服务路径。APN 100包括一个或多个广域网(WAN) (诸如WAN 102)、APN设备(APNA) 104-106、WAN路由器 110_1 - 110_3 和网络应用服务以及APNA之间的APN管道。如下面更详细地描述的。APN 100由充当网络控制节点(NCN)的单个APNA配置，该网络控制节点为APN提供单个控制点。但是，首先，定义本文中使用的多个术语以它们在本发明的上下文中使用时的含义。

[0023] APN路径是跨WAN位于不同地理站点的两个WAN链路之间建立的逻辑连接，其中一个WAN链路将业务发送到另一个WAN链路。由于路径是单向实体(一个发送方和一个接收方)，因此彼此连接的两个WAN链路之间有两条路径。每个WAN链路将一条路径视为其传输路径，并将另一条路径视为接收路径。在正常情况下，APN路径用于发送用户业务。

[0024] APN管道是两个APN节点(也称为客户端站点)之间的虚拟连接，并且通过聚合一条或多条APN路径及其被分配的WAN链路资源而形成。管道服务是一条或多条路径的逻辑组合。利用APN的全部价值，管道服务通常用于企业站点到站点的内联网通信。利用管道服务，取决于配置，跨多个WAN链路管理业务以创建端到端隧道(tunnel)。管道将虚拟网络覆盖在底层网络之上。

[0025] APN设备(APNA)是包含APN客户端站点功能的设备，该设备包括管控其参与APN的软件模块。高可用性(HA)站点包含两个APNA，一个处于活动状态，一个处于操作待机模式并且可以在需要时代替另一个APNA变为活动状态。

[0026] WAN链路表示到广域网(WAN)的物理接入点，诸如数字订户线(DSL)连接或线缆调制解调器。WAN链路的显著特点是带宽，其表示可用于传输和接收的数据容量的量。WAN链路可以在APN管道以及内联网和互联网网络服务之间共享。在本实施例中，APN设备不直接附接到WAN链路。APN设备通过逻辑连接(诸如图1A的WAN路由器 110_1 - 110_3)与WAN链路进行通信。

[0027] 专用WAN链路提供到非公共WAN目的地的物理接入点。这种专用WAN链路的示例包括与ATM虚拟电路的异步传送模式(ATM)链路、与帧中继电路的帧中继链路、多协议标签交换(MPLS)隧道、虚拟专用网(VPN)隧道或租用的点对点线路。在具有专用WAN链路的网络上进行到网络另一端的目的地专用列表的连接。公共WAN链路表示到互联网的物理接入点。可以假设任何公共WAN链路都可以建立到任何其它公共WAN链路的连接。

[0028] 本地WAN链路(LWL)是APN客户端站点到WAN的接入点。站点A的LWL耦合到用于站点B的对应远程WAN链路(RWL)。对于站点A和站点B之间的管道，站点A的本地WAN链路是站点B的远程WAN链路。

[0029] 互联网服务被用于企业站点与公共互联网站点之间的业务。互联网被视为单独的安全区。互联网上的业务被认为比APN中的管道业务更不可信，因为互联网业务没有像管道中那样被封装和加密。而且，互联网业务一般来自不受拥有可信网络的企业控制的实体。

[0030] 内联网服务被用于尚未被定义用于跨管道传输的企业内联网业务的任何部分。就像互联网业务一样，内联网业务保持未封装，并且APN通过在拥塞期间相对于其它服务类型限制内联网业务的速率来管理网络中的带宽。应注意的是，在某些条件下，并且如果被配置用于在管道上进行内联网后备(fallback)，那么代替地可以将通常经由管道行进的业务视为内联网业务，以便维持网络的可靠性。由于管道业务是站点到站点的，因此客户一般在没

有管道的情况下具有递送这种站点到站点业务的途径。这种未封装的服务(称为内联网服务)无法享受管道带来的益处。如果无法建立管道隧道,那么使用该管道的路由将被忽略,并且这意味着将本应使用管道的业务重定向为使用未封装的站点到站点传输方法。

[0031] 流(flow)由n元组(n-tuple)定义,该元组由<IP源地址,IP目的地地址,IP协议号,传输控制协议(TCP)/用户数据报协议(UDP)源端口(如果IP协议是TCP或UDP)、TCP/UDP目的地端口(如果IP协议是TCP或UDP)>组成。取决于上下文,可以将其它项添加到元组,包括:区分的服务代码端口(DSCP)标签、路由域和服务标识符等。而且,流是单向的。例如,如果节点A和B正在通信,那么存在表示从A到B的业务的流以及表示从B到A的业务的流。

[0032] APN服务是对通过APN传输的分组执行的处理步骤的集合。如图1A中所示,通过APN 100和APN设备106移动的数据业务取决于发送站和接收站的位置而可以要求不同类型的服务。APN服务实例是例如在APN设备106内部的APN设备存储器107中保持的APN服务的特定配置的上下文实例。APN服务实例的存储器包含但不限于特定于上下文的配置数据、统计数据和跟踪状态数据。例如,APN客户端站点可以具有连接到远程APN客户端站点的多个APN管道。对于每个APN管道,对于APN管道服务类型都存在单独的APN服务实例。

[0033] 与路径112相关联的APN管道服务管理通过APN 100从APN设备105通过路由器110₁、通过WAN 102、通过另一个路由器110₃到APN设备(APNA)104传输的网络业务分组。用于路径112的APN管道服务在APN设备104和105两者上进行操作。APN管道服务利用由APN管道服务提供的全部益处在于具有APNA 105的第一地理位置和具有APNA 104的不同地理位置之间发送和接收数据,以进行WAN资源分配和网络自适应。通过也被其它APN服务利用的WAN链路,与路径114相关联的APN内联网服务被用于管理具有APN设备105的第一地理位置与不具有APN设备的企业非APN站点120内的不同地理位置之间的数据的发送和接收。

[0034] 在另一个实施例中,APN内联网服务(诸如与路径112相关联的APN内联网服务)可以被用于向具有APN设备的第一地理位置发送数据和从具有APN设备的第一地理位置接收数据,但是管理员选择性地将APN配置为对于特定类型或类别的业务不使用APN管道服务112。与路径116相关联的APN互联网服务被用于通过也被其它APN服务利用的WAN链路在具有APNA 105的第一地理位置与企业网络外部的不同地理位置之间发送和接收数据。例如,使用APN互联网服务的业务可以与访问公共互联网web服务器122的网络用户相关联。APN直通服务118被用于在具有APNA 105的第一地理位置与相同第一地理位置内的本地站点124_{A1}之间发送和接收数据。在另一个实施例中,APN直通服务可以被用于使用与任何其它APN服务相关联的任何WAN链路在具有APN设备105的第一地理位置与不具有APN设备且不穿越WAN的企业网络内的不同地理位置之间发送和接收数据。

[0035] 在另外的实施例中,路径126具有第一本地站点124_{A2},该第一本地站点124_{A2}连接到LAN1 101到APNA 105到WAN路由器110₁,通过WAN 102连接到WAN路由器110₃到APNA 104到LAN2 103到第二本地站点125_{B2}。LAN1 101和LAN2 103是具有多个路由器和路由路径的示例性网络,这些路由器和路由路径被管理并且可以改变以改进网络性能。

[0036] 管道由多条路径组成。在与管道相关联的2条WAN链路之间形成路径。通过收集质量度量(诸如分组丢失和延时),监视APN中每个管道中的每条路径的通信质量。这种监视是通过控制消息来完成的,并且在每条路径上进行,而不管该路径是否用于传输用户业务。因而,除非其无法操作,否则没有任何路径是完全空闲的(free)。由于管道中的所有路径都被

测量,而不管是否有用户业务通过该路径,因此该管道维持被APN用来选择最佳网络路径以传输用户数据的最新的每路径度量。

[0037] 静态管道是在配置文件中配置并在APNA启动时创建的管道。在不改变配置文件的情况下,静态管道不被移除。

[0038] 动态管道是在需要时在APN客户端之间创建并且在不再需要时可以被移除的管道。动态管道解决了跨网络实时发生的静态配置的网络中的改变。动态管道实时地优化网络性能,以适应网络中节点之间不断改变的通信模式。动态管道还可以被用于分流(offload)遇到拥塞的中间节点的业务。

[0039] 根据本发明的自适应专用网络(APN)软件产品在虚拟机内作为集中式管理系统运行,以创建APN配置并监视系统资源、分析系统资源并在操作中管理已配置的APN,如本文进一步解决的。APN系统还允许网络控制节点(NCN)对物理网络进行集中式虚拟单点控制,其中NCN提供系统范围的定时同步。集中式单点控制不限于节点网络内的中心位置,可以在网络内的任何点处、并且可以在网络边界之外考虑的点处耦合。

[0040] 图1B图示了根据本发明实施例的支持双端服务150的自适应专用网络(APN)管道,其包括在APN客户端站点A 152和APN客户端站点B 154之间的集成防火墙192、194、196和198。每个APN客户端站点也被视为APN中的节点并且包含软件模块的集合,这些软件模块管控该节点在APN中的参与。APN客户端站点A 152和APN客户端站点B 154的软件模块分别包括控制平面模块156和158、WAN入口处理器模块160和162、WAN出口处理器模块164和166以及节点管理和接口软件程序模块168和170。如图1B中所示,WAN入口处理器模块160和162包括管道处理级(stage) 172和174,而WAN出口处理器模块164和166包括重复的管道处理级176和178。有利地,防火墙192和194各自分别被集成到WAN入口处理器管道处理级172和174的输入级。防火墙196和198各自分别被集成到WAN出口处理器管道处理级176和178的输入级。集成防火墙在下面进一步详细描述。在每个APN客户端站点处还提供内联网服务、互联网服务和直通服务。每种APN服务类型(包括管道、内联网、互联网和直通服务类型)针对分别与传送到WAN和从WAN传送的每种类型的数据业务实现了处理。

[0041] 如图1B中所示,由粗虚线箭头路径180和182识别的APN管道业务作为穿越该APN的业务流经两个APN客户端站点152和154。APN客户端站点B 154的WAN入口处理模块162在经由WAN 184向APN客户端站点A 152传输业务180之前执行WAN入口管道服务处理174。APN客户端站点A 152的WAN出口处理器模块164在向位于LAN 186上的一个或多个节点传输业务180之前执行WAN出口管道服务处理176。一个APN客户端站点的WAN入口管道处理174与对等APN客户端站点的WAN出口管道服务处理176的绑定构成APN管道188,在其中跨多个WAN资源主动地监视和管理业务。在转发用户数据的同时,被称为t2_apps 190和191的控制程序在与APN中其它APNA通信的每个APNA上运行。

[0042] APN能够使用不同的非对称WAN链路,这些非对称WAN链路链路在带宽、延时、抖动、分组丢失和拥塞方面的行为经常随时间频繁变化。例如,APN可以使用通过公共网络在上游以512kbps向WAN传输数据并且以6Mbps从WAN传输数据的非对称DSLWAN链路与在上游和下游以1544kbps传输数据的专用对称租用电路T1WAN链路以及在上游以312kbps向WAN传输数据并且以3Mbps从WAN到对等体传输数据的线缆宽带连接相结合,该对等体具有这些速率的足够聚合带宽用于以理论上2368kbps的传输速率和10544kbps或10.544Mbps的接收速率进

行单个传输控制协议 (TCP) 文件传送会话。在实践中,在良好的网络行为下,实际速率将接近这些速率的90%。如果连接的行为发生了改变,例如到DSL链路的路径损耗巨大,那么APN将使用其高频性能反馈机制通过使用替代资源或试图从损耗中恢复来使网络适应以避免或缓解问题。

[0043] 在路径选择中,评估管道路径并选择最佳可用路径。当前处于路径质量良好状态的任何路径都有资格被优先选择。如果多条路径处于路径质量良好状态,那么针对每条路径评估并比较所估计的端到端时间,并选择具有最低端到端时间的路径。如果没有路径处于路径质量良好状态,那么选择具有最高带宽的路径质量不良状态的路径。“单向时间”(one way time,OWT)是指分组从源到接收器穿越网络所花费的时间量。在本发明的上下文中,通过从来自WAN入口模块160的发送时间戳中减去来自WAN出口模块166的接收时间戳来测量单向时间(图1B)。例如,于2009年6月11日提交的题为“Flow-Based Adaptive Private Network with Multiple WAN-Paths”的编号为8,125,907并且通过引用整体并入本文的美国专利在第6列第1行-第19列第27行提供了自适应专用网络(APN)中进行定时和网络控制的当前优选方法的示例性细节。

[0044] APN路径处理服务负责提供跨网络将用户数据和控制信息从一个APN节点传送到另一个APN节点的手段。特别地,用户数据和控制信息可以从一个APN节点的WAN入口处理器模块160跨WAN传输并且在WAN出口处理器模块166处被接收,例如图1B中所示。

[0045] 路径状态表示由WAN出口APN节点的路径状态监视处理接收的反馈确定的网络路径的最当前状况。在接收到分组时,跟踪分组的序列号,以查看在WAN入口APN节点与WAN出口APN节点之间的传输中是否丢失了任何分组。使用方法来触发倾向于在接收到的分组的短时段内对丢失具有更大的容忍度而在较长时段内对丢失的容忍度很小的路径状态过渡。这种方法的独特之处在于能够在几乎同时且持续地在数个持续时间内跟踪路径的分组丢失阈值,同时仍维持低处理器开销。这一方面通过以下来获得:利用高分辨率和高度同步的APN时间戳对跨WAN发送的管道业务进行通用路径标记以启用传输延时的高度预测性估计和延时的统计变化。并行地,控制平面模块的路径状态监视服务被用于检测分组丢失和业务跨APN使用的最佳路径。结果是能够检测偶然的短期网络丢失与长期持久问题之间的差异。

[0046] 在当前的优选实施例中,客户端站点处的APN节点的软件模块在相同物理APN设备中存储并操作;但是,在替代实施例中,模块也可以存在于单独的物理APN设备中。结合本文公开的实施例描述的方法可以直接实施在由处理器和存储器复合体(complex)执行的一个或多个软件模块中,诸如具有一个或多个中央处理单元设备的自适应专用网络(APN)设备(APNA)、机架式安装处理设备、个人计算机,服务器等中使用的。例如,处理器和存储器复合体可以被配置为在存储在计算机可读非瞬态存储介质上的软件模块程序的控制下执行访问数据和对数据进行操作的指令,该软件模块程序直接与处理器和存储器复合体本地相关联(诸如可以通过指令高速缓存获得)或者可通过I/O设备访问。软件模块可以驻留在计算机可读非瞬态存储介质中,该计算机可读非瞬态存储介质可以包括随机存取存储器(RAM)、闪存、动态随机存取存储器(DRAM)、同步动态随机存取存储器(SDRAM)、只读存储器(ROM)、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM)、硬盘、可移动盘、CD-ROM、数字视频盘(DVD)、其它类型的可移动盘、或任何

其它合适的非瞬态存储介质。非瞬态存储介质还可以耦合到处理器和存储器复合体,使得硬件处理器可以通过内联网或互联网从该存储介质读取信息或向该存储介质写入信息。

[0047] 自适应专用网络节点(APN客户端站点)包含支持参与自适应专用网络的软件模块。APN节点可以存在于一个位置处的一个或多个APN设备中。APN节点包含软件模块的集合,这些软件模块由位于APN节点中的处理器和存储器复合体执行,这些软件模块管控APN节点在APN内的参与,诸如图1B中的控制平面模块156和158、WAN入口处理器模块160和162以及WAN出口处理器模块164和164。控制平面模块156和158负责与网络中其它APN节点并行地控制和参与APN节点的控制。

[0048] WAN入口处理器模块160可以适当地被实施为负责处理网络业务以从局域网(LAN)传输到WAN的软件和硬件组件。WAN出口处理器模块164可以适当地被实施为在硬件组件上操作的软件,诸如负责处理用于从WAN传输到LAN的网络业务的处理器和存储器复合体。WAN入口和WAN出口处理器模块将在下面进一步详细讨论。APN客户端站点的控制平面模块156可以适当地被实施为在硬件组件上运行的软件,诸如利用APN客户端站点的WAN入口处理器模块160和WAN出口处理器模块164作为用于跨WAN向APN节点传输和接收APN节点到APN节点控制数据的手段的处理器和存储器复合体。

[0049] 用于APN的软件分组使用作为变更管理软件的一部分的被称为Tapplication协议(TAP)分组的控制分组通过WAN进行分发,或者通过管理接口(诸如使用接口168和170将软件下载到APN客户端站点)进行分发。TAP是用于通过WAN发送的消息的协议,以允许不同设备上的t2_app外部的处理彼此通信。TAP可以被视为点对点或类似以太网的设备,该设备不是从物理介质接收分组,而是从用户程序接收分组,并且该设备不是经由物理介质发送分组,而是将分组写入用户程序。在软件更新之后,然后重启APN客户端站点152和154上的APN服务,从而使APN软件节点配置同步。

[0050] 图2图示了根据本发明实施例的APN 200,APN 200具有耦合到管道部分220的APN网络控制节点(NCN)202和分别耦合到十六个APN客户端站点204-219的十六个APN管道部分221-236。如图2中所示,在当前的优选实施例中,APN 200被集中配置。网络管理员通过由NCN 202处理的APN配置文件来配置整个APN 200。然后,NCN 202将配置设置分发到APN 200中的所有客户端站点。这种配置APN 200的方法旨在通过向网络提供单个配置点来为管理员带来益处。它还通过严格的版本检查几乎同时确保网络中所有APN客户端站点的配置一致性和兼容性。通过对APN资源及其初始分配进行映射,中央配置还为网络提供了附加配置带宽优化。此外,在将配置加载到生产网络上之前,集中式配置可以向管理员提供有关配置行为的信息和警告,这些信息和警告可以不是根据该配置而显而易见或预期的。

[0051] 站点204-219和主要NCN站点202中的每一个都包含提供APN功能的APN设备。APN 200的配置一般而言提供站点A(诸如站点205)与站点B(诸如站点208)之间的连接性,其中从站点A的角度看,连接性是站点A→LWL→“WAN”→RWL→站点B。从站点B的角度看,连接性是站点B→LWL→“WAN”→RWL→站点A。WAN 201表示分配的WAN链路资源和APN选择的路径。在图2中,站点A和站点B之间的管道是通过使用管道部分222和225形成的,并且是对应的站点A和站点B之间的虚拟连接。管道包括路径的集合并且涵盖从站点A处的本地WAN链路(LWL)→“WAN”→站点B处的RWL的路径。

[0052] 在一个当前的优选实施例中,如图2中所示,在NCN与例如十六个APN客户端站点之

间存在APN管道。将认识到的是,虽然出于说明的目的示出了十六个APN站点,但是可以适当采用更多或更少数量的潜在APN客户端站点。每个APN管道可以具有管理员为与特定APN相关联的每个地理位置的特定需求定制的独特配置参数。

[0053] 对于APN路径状态的定义,下面提供对路径处理服务的描述。当前处于路径质量好状态的任何路径都有资格被优先选择。如果多条路径处于路径质量好状态,那么针对每条路径评估并比较所估计的端到端时间,并选择具有最低端到端时间的路径。如果没有路径处于路径质量好状态,那么选择具有最高带宽的处于路径质量不良状态的路径。

[0054] 示例性APN 200的十六个客户端站点204-219一般彼此远离,并且可以包括地理上不同的客户端站点。如果设备物理上位于不同的位置(诸如不同的建筑物、城市、州、时区或国家),那么将站点定义为远端(remote)。例如,主要NCN 202可以位于第一国家中的公司总部位置,其中客户端站点204-209和客户端站点217-219也位于第一国家中。其它客户端站点210-216可以位于第二国家中。

[0055] 如本文所使用的,APN设备是包含根据软件模块的APN节点功能的设备,其中软件模块诸如控制平面模块156和158、WAN入口处理器模块160和162、以及WAN出口处理器模块164和166,如上面参考图1B更详细地描述的。十六个客户端站点204-219分别通过管道部分221-236耦合,并且管道部分可以连接在一起,以在客户端站点处的两个连接的APN设备之间提供可配置的虚拟连接。应注意的是,虽然图示了十六个客户端站点204-219,但是APN可以支持所需数量的客户端站点。

[0056] 图2的网络控制点(NCP)202是APN 200的管理点。在一个实施例中,NCP 202驻留在APN节点内。APN控制节点是指还作为APN的网络控制点执行的APN节点。在另一个实施例中,NCP驻留在与APN节点分开的设备中,并且管理和控制APN内的APN节点。NCP为APN提供管理和控制,包括但不限于将配置对象分发到APN客户端节点以及与APN时间同步。

[0057] 动态管道是在需要在APN客户端之间创建的管道,并且可以基于配置的阈值在不再需要时被移除。例如,客户端站点205可以配置有两个本地WAN链路,一个来自第一网络提供商,一个来自第二网络提供商。多个管道可以连接到站点205,站点205可以被配置为使用本地WAN链路中的一个或两个。在连接到站点205的所有管道都使用两个本地WAN链路的示例性场景中,当任一本地WAN链路的使用量超过配置的阈值时,可以触发动态管道的创建。

[0058] APN业务流是针对特定APN流记录识别出的网络会话业务的管理员指定。APN业务流要求是管理员配置的管控与APN业务流相关的APN的预期行为的要求。例如,APN业务流要求可以包括持久路径流要求、重复流要求和可靠流要求。

[0059] APN流记录保持在APN设备的存储器中。APN流记录跟踪定义的APN业务流,从而确保遵循APN业务流的预先配置的要求。APN流记录包含APN业务流要求和APN业务流的状态两者。特定APN流记录的要求是从APN业务流匹配的路由和服务规则得出的。APN流记录的状态包括但不限于APN服务类型、APN服务实例、与选择的最后一条APN路径有关的信息、当前APN流序列号、接收到最后一个分组的时间、传输最后一个分组的时间、处理的分组的数量和字节的数量的计数、用于序列重新排序的未决分组的集合、用于分段的未决分组的集合,以及先前处理的分组的历史记录的集合。

[0060] 关于WAN入口管道处理的处理级的更多详细信息参见于2012年2月8日发布的题为

“Flow-Based Adaptive Private Network with Multiple WAN-Paths”的编号为8,125,907的美国专利,以获取更多细节,并且该专利通过引用整体并入本文。

[0061] 本发明涉及改进的WAN优化(WANop)处理的各方面,WANop是改进WAN网络如何处理数据业务的途径。为此,如本文进一步解决的,采用两个主要工具,即重复数据删除和压缩。重复数据删除通过认识到相同的位模式常常多次穿越WAN链路而工作。例如,当从事同一项目时,许多员工可以从数据中心的服务器下载相同的电子邮件和文件。本方法允许WAN附近的网络设备将该数据进行高速缓存并使用小代码表示实际数据。这种方法减少了必须穿越WAN网络的数据量。

[0062] 在一个实施例中,本发明在WANop处理与SSD上的盘高速缓存之间引入写保护高速缓存(WPC)。写保护高速缓存提供了保持(holding)区域,使得仅有已被看到多于一次的数据才被提交到盘高速缓存。

[0063] 写保护高速缓存(WPC)为盘上的数据提供了保护层。它当前描述的使用试图避免不必要地写入驱动器。它仅通过将看到多于一次的数据(块(chunk))提交到盘来实现这个目的。这个特征解决了至少两个单独的问题:

[0064] • SSD盘磨损:SSD技术虽然速度惊人,但其使用寿命却比HDD短得多。例如,在仅写入14TB之后,设备上的128MB驱动器就可能发生故障。如果驱动器在设备保修期内发生故障,那么对驱动器供应商产生财务影响。

[0065] • 盘高速缓存的改进利用:通过仅将先前已记录在WPC中的数据写入盘,避免了将非重复数据(诸如加密数据)写入盘。

[0066] 此外,不利用永远不会再看到的数据来替换(displace)频繁使用的良好的暖数据是有利的。

[0067] WPC是已接收到的安全散列算法(SHA)1摘要(digest)的高速缓存。它们存储在设备存储器的保留部分中。4GB的WPC包含大约7000万个摘要,表示超过500GB的数据。摘要存储在WPC中,直到检测到重复为止。然后,将块、数据和摘要添加到盘高速缓存中,并将条目从WPC中移除。

[0068] 本文使用的术语包括:DWPD:每天的驱动器写入;WPC:写保护高速缓存;以及WANop:用于提高跨WAN的数据传送效率的技术的集合。

[0069] 如上所述,一种当前的WANop方法将块写入WAN入口和WAN出口设备两者上的盘。这种方法确保在一个方向发送的数据可以立即用于相反方向的重复数据删除。但是,本发明认识到,在许多情况下,由于数据趋于单向,因此这两种方式的可用性是不必要的。

[0070] 相反,通过不像图3的WANop处理300中所示的那样在WAN入口上写入盘,WANop处理300仅在WAN出口上学习。因此,如果在一个方向上将块学习或写入盘,那么该信息不与相反方向共享。在大多数情况下,这种方法是非常好的权衡。如果相同的数据在两个方向上流动,那么WANop处理300将快速学习它。但是,如果像预期的那样WAN出口和WAN入口所共有的数据很少,那么可以识别出相当大的处理节省。

[0071] 更详细地转到图3,对等体1(peer 1)330可以适当地是WANop出口处理器,该WANop出口处理器处理来自WAN的网络业务(诸如云分组I/O连接340)以从云存储空间下载数据。例如,数据被下载到LAN,诸如连接到客户端设备(诸如膝上型计算机)的LAN上的客户端流348。在云连接340和客户端流348之间,数据通过64k输入缓冲器341到达解压缩处理342,并

且从那里获得摘要数据344,以检索摘要和/或数据,如下面从WPC 310和盘320进一步解决的。然后,数据通过64k输出缓冲区346前进到客户端流348。

[0072] 对等体2(peer 2)350可以适当地是WAN入口处理器,该WAN入口处理器处理网络业务以从连接到客户端设备(诸如膝上型计算机)的局域网(LAN)进行传输,客户端从该客户端设备发送客户端流360用于通过WAN云分组输入-输出(I/O)连接368将数据传输到云存储空间。在这两者之间,数据通过64k输入缓冲区到达重复数据删除处理362、压缩处理364和64k输出缓冲区366。

[0073] 在WAN出口上,WANop的WPC 310存储块摘要,并将新到达的摘要与先前看到的摘要进行比较以识别重复。当如本文进一步解决的那样检测到重复时,摘要和块被写入盘高速缓存320。当前描述的方法的缺点在于,必须从WAN入口对等体发送两次数据,而不是如本文中进一步详细解决的那样在先前方法中发送一次数据。

[0074] 为了进一步限制盘写入,图4中所示的WAN入口处理的处理400将不向盘写入数据。在步骤402,数据到达LAN上并且在步骤404中被分组为块。典型的块在1.5kB至22kB之间,平均约为8kB。在步骤405中,为块计算安全散列算法(SHA)1摘要,并将摘要转发给对等体。将认识到的是,可以适当地采用其它散列算法。在步骤407中,如果对等体在其高速缓存中具有块/摘要,那么以ACK进行响应,或者如果对等体不具有块/摘要,那么以NACK进行响应。在步骤408中,如果响应是NACK,那么将块数据发送到对等体。如所描述的,对于单向流,在入口上将不存在对高速缓存(WPC或SSD)的写入。在用户界面(UI)WAN入口统计信息中,ACK被示为命中(hit),并且NACK被报告为未命中(miss)。

[0075] 转向用于出口处理的设备建模,用于存储摘要的数据结构为64字节,并且平均数据块为8kB。通过仅存储摘要,WANop先前看到的数据的存储器变得大约是WPC的128倍大。例如,一个兆字节的WPC包含16k摘要,表示128兆字节的块数据。

[0076] 设备建模示例如下:

[0077] 设备支持100Mbps的持续WANop吞吐量,

[0078] 100Mbps对应于12.5MBps,对应于每秒1600(8kB)个块;以及

[0079] WPC填充速率:WPC的102.4kBps,对应于352MB/hr(352MB/小时)WPC。

[0080] 利用100Mbps的数据速率,3.5GB WPC提供10小时的独特块存储器,表示439GB的数据。图5的表500中提供了用于其它平台的类似数据。

[0081] 转向数据流,图6示出了用于处理对WAN入口对等体接收的先前未看到的块的请求的处理600。换句话说,该块不在WPC或盘中。首先,在步骤604中,将摘要发送到WAN出口对等体。因为这是新块并且摘要不在其盘高速缓存中,因此在步骤606中,WAN出口对等体以NACK进行答复。在接收到NACK时,WAN入口对等体在步骤608中发送块的数据。在确定摘要不在WPC中后,在步骤610中,WAN出口对等体将摘要添加到WPC并向LAN转发数据。

[0082] 图7示出了用于处理第二次看到的块的处理700。如图7中所示,在步骤702中,当WAN入口对等体接收到相同的块时,摘要被再次发送到WAN出口对等体。在步骤704中确定摘要不在盘上,并且在步骤706中,WAN出口对等体以NACK进行答复。在步骤708中,入口对等体将块的数据发送到WAN出口对等体。由于在步骤610中添加了块,因此在WPC中找到针对块的摘要。在步骤710中,将数据添加到盘并从WPC中移除摘要。在步骤712中,向LAN转发数据。

[0083] 图8示出了用于处理第三次及后续时间看到的块的处理800。如图8中所示,当第三

次或后续时间看到块时,在步骤804中,将包括数据和摘要的块发送到出口对等体。在步骤806中,在步骤804中发送的摘要与已经存储在盘上的块对应并且ACK被发送到入口对等体。在步骤808中,出口对等体向LAN转发数据。

[0084] 图9A和图9B(统称为图9)分别图示了处理流程900和950。在图9A的处理900的步骤902中,WAN出口对等体从入口对等体接收块摘要。在步骤904中,出口对等体执行查找操作以判定接收到的摘要是否存储在当前优选实施例中作为半导体存储设备(SSD)的盘上。在步骤906中,如果步骤904中的判定为“否”,那么处理900通过向入口对等体发送NACK响应来继续。如果在步骤906中的判定为“是”,那么在步骤908中将块数据发送到LAN,并且在步骤910中将ACK响应发送到入口对等体。

[0085] 在图9B的处理950的步骤952中,出口对等体从入口对等体接收对具有包括摘要的块数据的NACK答复。出口对等体在步骤954中继续,以在盘高速缓存中查找摘要,以查看摘要是否存在。如果在步骤956中判定摘要在盘高速缓存上,那么在步骤958中,出口对等体继续判定块数据是否存储在WPC中。如果是,那么在步骤960中,将块添加到盘存储空间中。在步骤962中,从WPC中删除摘要,并且在步骤966中,将块数据发送到LAN。在步骤968中,将ACK响应发送到入口对等体。如果在步骤958中判定为“否”,摘要不在WPC中,那么在步骤964中将摘要添加到WPC。

[0086] 返回到步骤956,如果是,那么摘要位于盘存储空间上,然后在步骤966中,将块数据发送到LAN。在步骤968中,ACK响应被发送到入口对等体。

[0087] 转到WPC摘要管理,当前优选的WPC中的每个条目都是64字节(B)并包含块摘要。最初,所有条目都在空闲(free)列表上。在将摘要添加到WPC时,条目从空闲列表移至使用中(in use)列表。当WPC命中时,该条目从使用中列表中被移除,然后被放回到空闲列表上。当要求新条目并且空闲列表为空时,使用中列表上最旧的条目将被释放并用于新条目。可替代地,可以释放一定尺寸的最旧的条目。例如,最旧的条目或小于特定尺寸的条目将在更大尺寸的类型旧条目之前被删除。

[0088] 将认识到的是,本发明可以在虚拟设备中实施。在这种情况下,虚拟设备的操作方式与常规设备相同,但是虚拟设备必须基于底层虚拟存储器(VM)配置来调整其可伸缩性和资源要求。在这种上下文中,WANop存储器要求尺寸是基于总存储器的。

[0089] 虽然已经在当前优选实施例的各个方面的上下文中公开了本发明,但是将认识到的是,本发明可以适当地应用于与所附权利要求一致的其它环境。通过示例的方式,虽然本申请讨论了直到第二次看到块才将摘要和数据、块存储到盘高速缓存,但是将认识到的是,可以仅在第三次或后续时间看到相同的块后才发生存储操作,并且该次数可以是适合于特定用户的体验和上下文的用户可选择参数。另外,该次数可以取决于块尺寸而变化,例如,最大尺寸的块在第二次被看到后被存储并且最小尺寸的块在第三次被看到后被存储。还可以考虑其它因素,诸如一天中的时间、剩余的可用盘存储空间的量等。

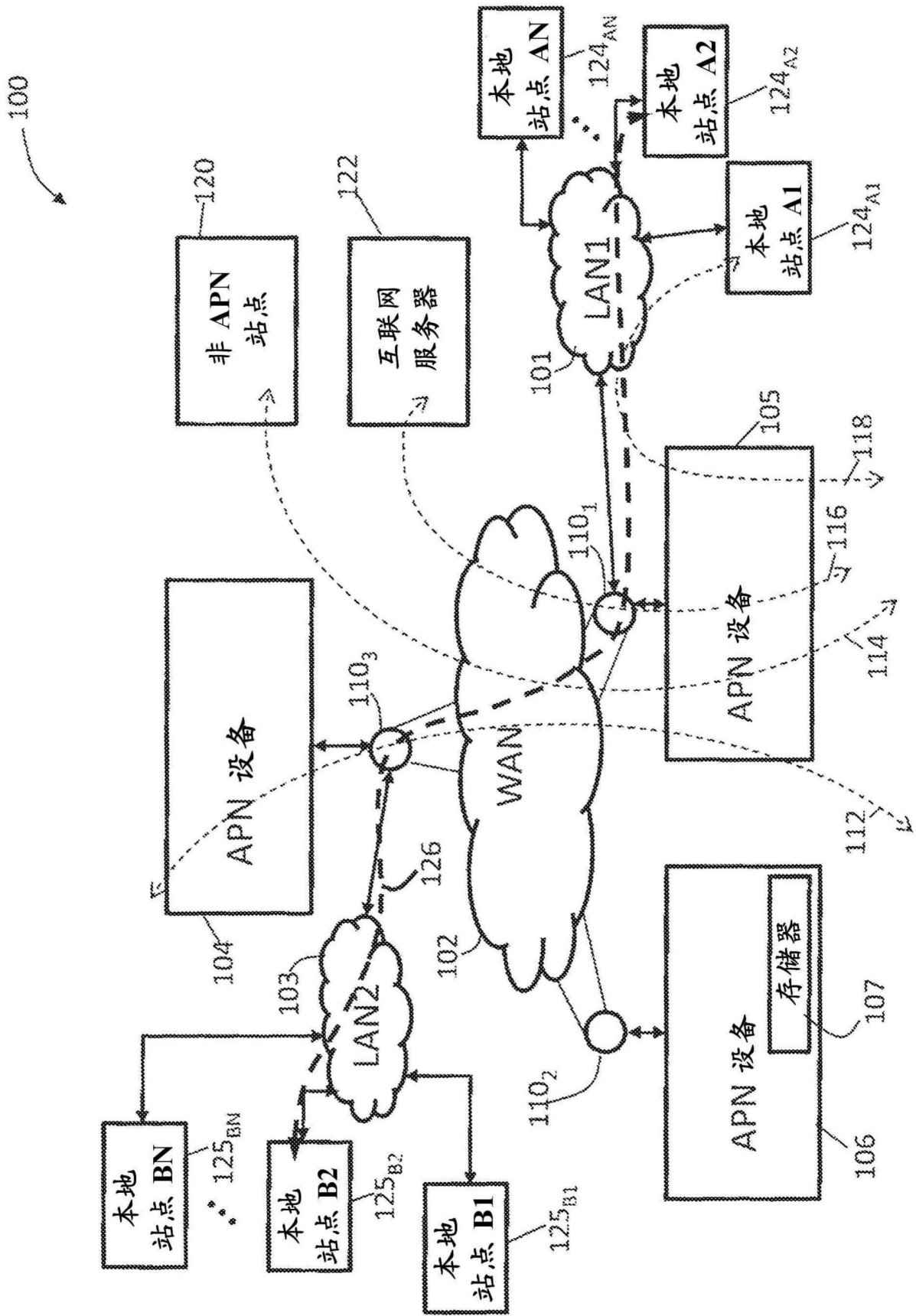


图1A

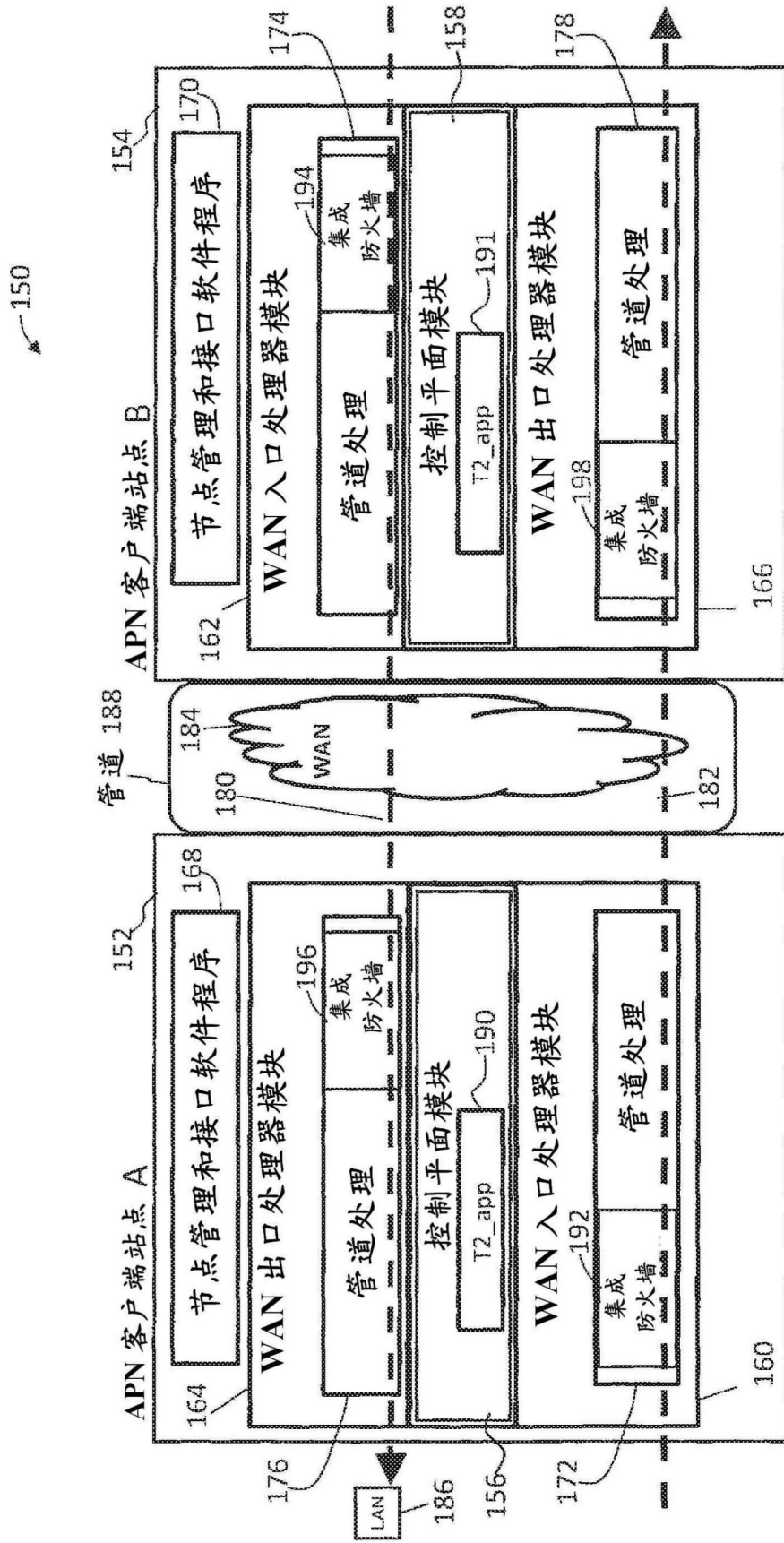


图1B

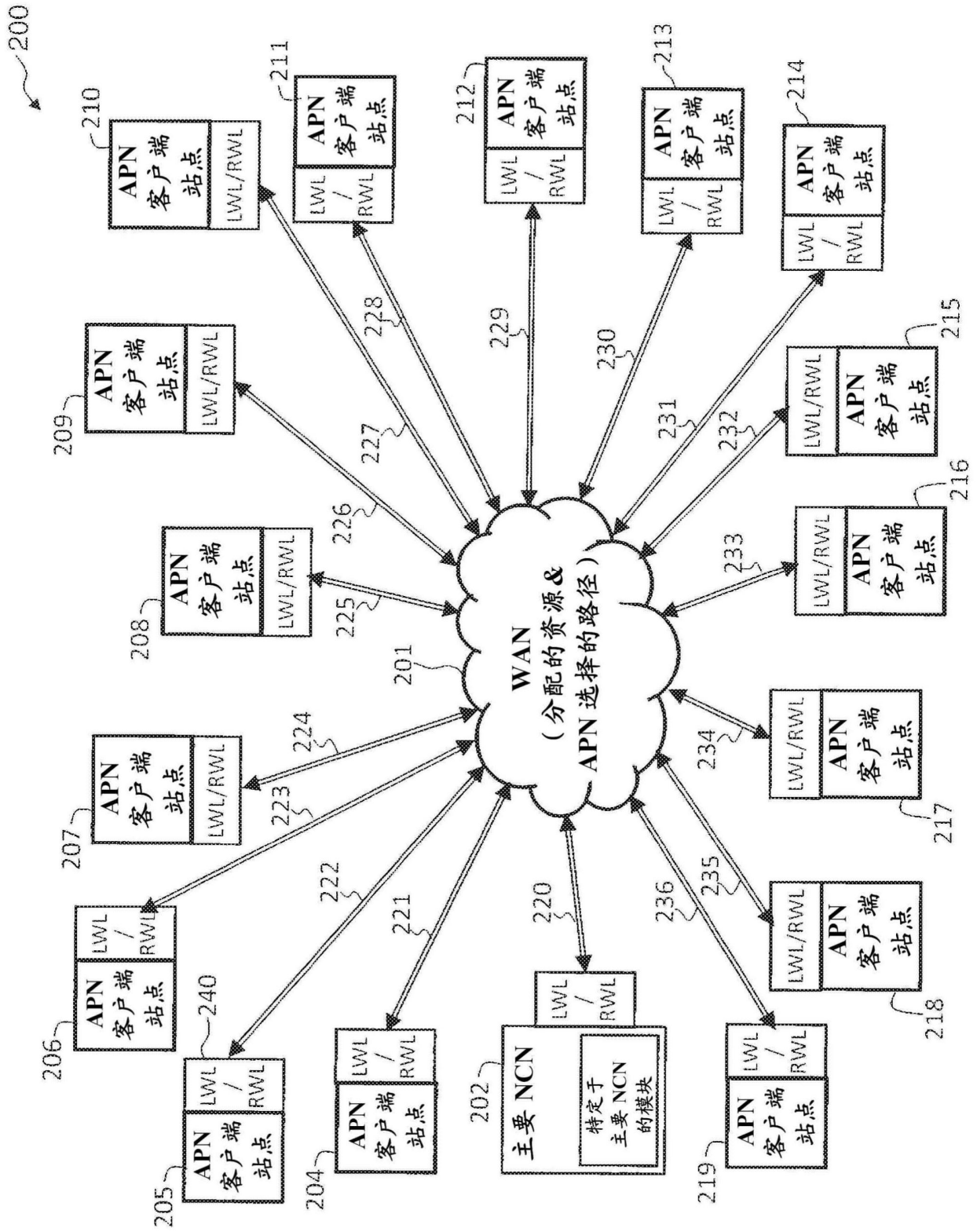


图2

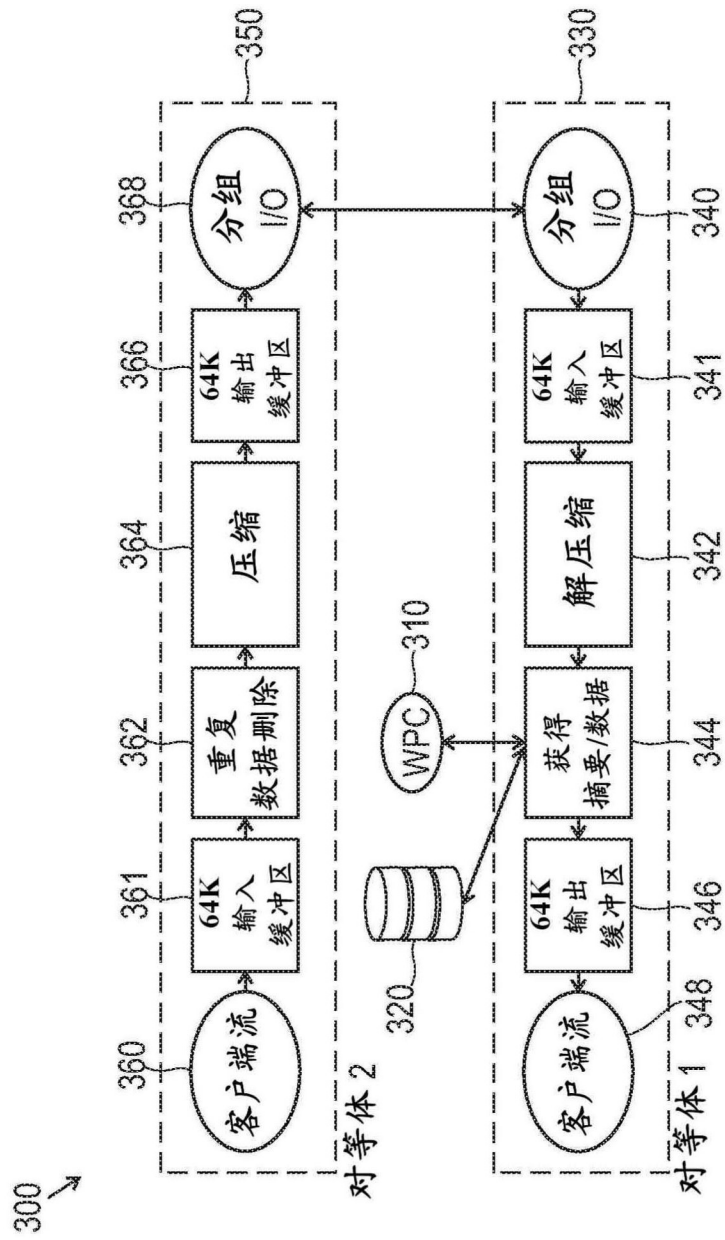


图3

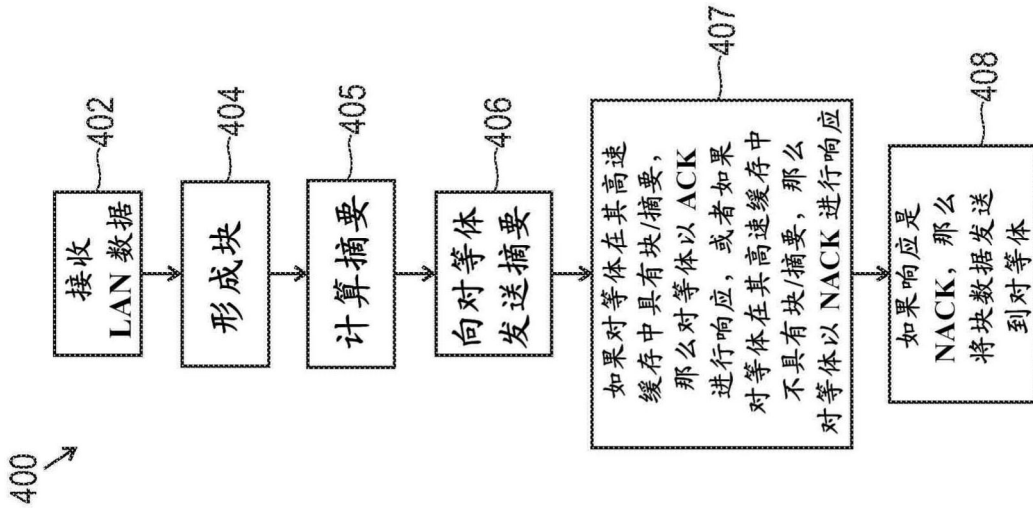


图4

500

出口 Mbps	出口 MB/s	8KB CK/秒	WPC MB/小时	一天中的工作小时数	WPC MB/天
25	3.1	400	88	10	879
100	12.5	1,600	352	10	3,516
200	25.0	3,200	703	12	8,438
300	37.5	4,800	1,055	12	12,656
300	37.5	4,800	1,055	15	15,820

图5

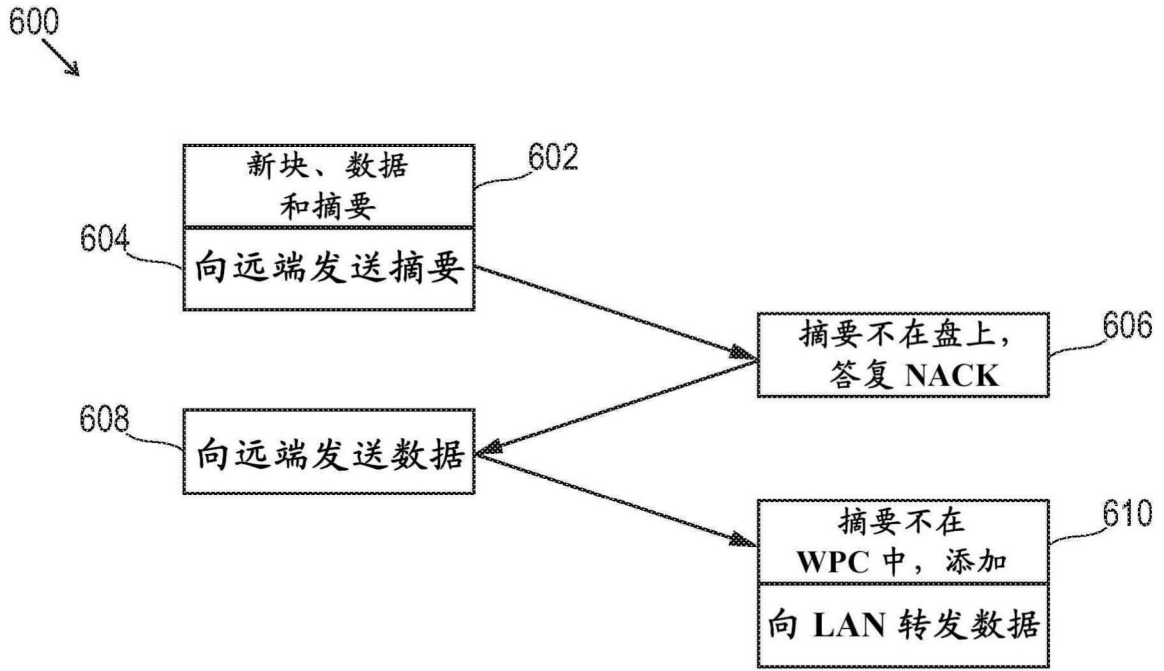


图6

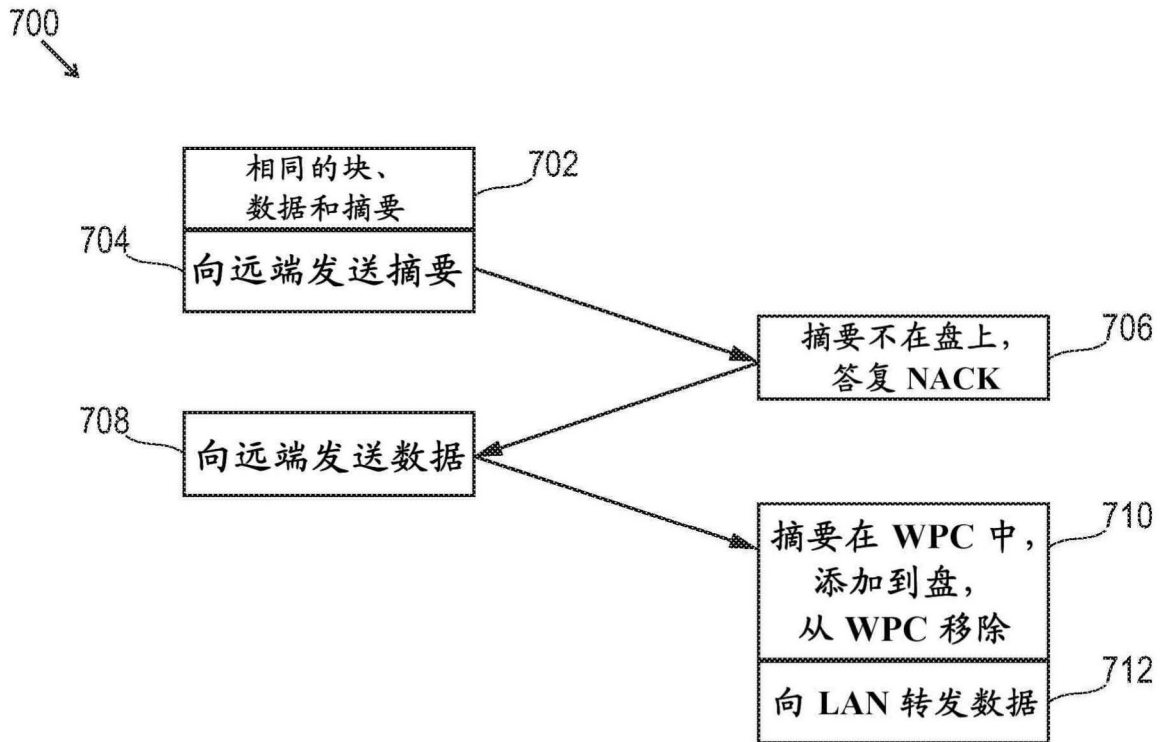


图7

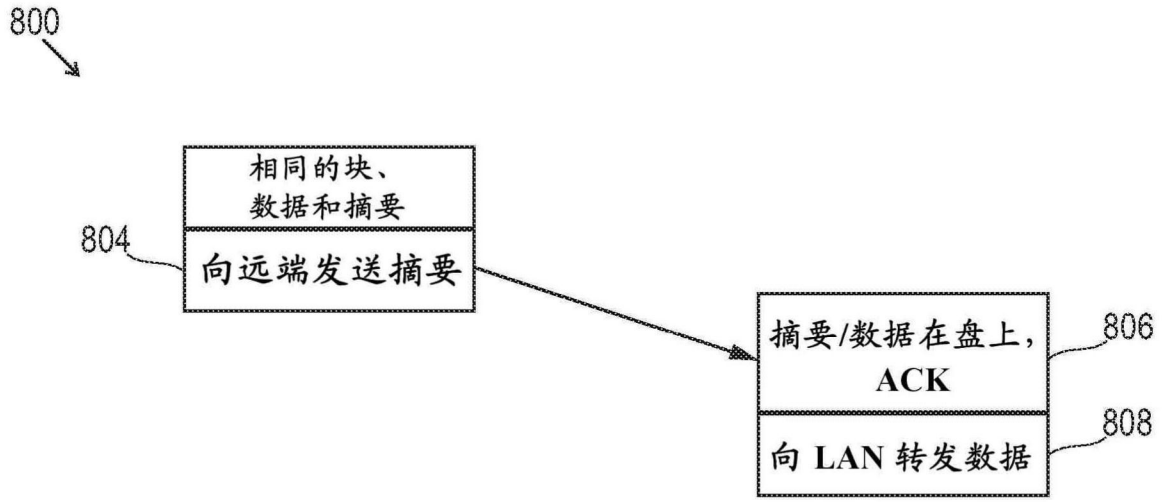


图8

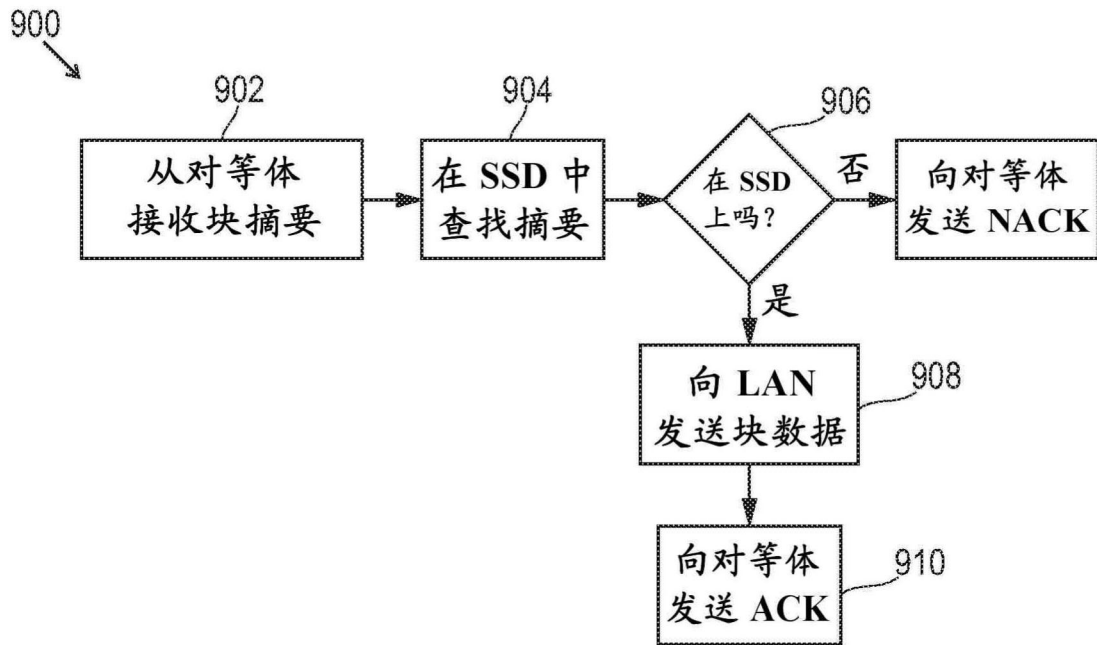


图9A

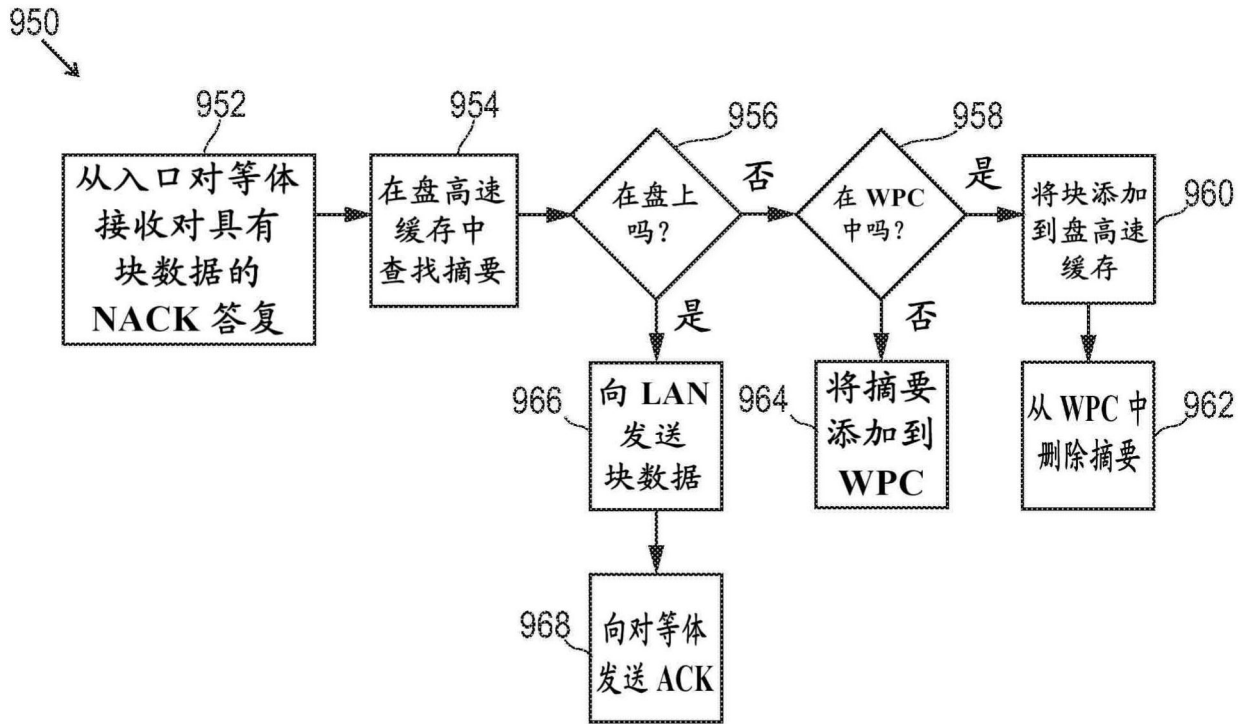


图9B

1000

存储器 GB	最大流 (TCP & WANOP)	WPC 尺寸 (MB)
0-4	500	128
5-9	1,500	2,048
10-15	5,000	4,096
16-31	10,000	6,144
>31	16,000	8,192

图10