



(12)发明专利申请

(10)申请公布号 CN 111314199 A

(43)申请公布日 2020.06.19

(21)申请号 202010104713.8

(22)申请日 2014.12.24

(62)分案原申请数据

201480082628.9 2014.12.24

(71)申请人 柏思科技有限公司

地址 中国香港九龙长沙湾青山道481号香港纱厂工业大厦6期5楼A5室

(72)发明人 宋浩伟 吴锦超 陈浩明

(74)专利代理机构 深圳新创友知识产权代理有限公司 44223

代理人 江耀纯

(51)Int.Cl.

H04L 12/46(2006.01)

H04L 1/16(2006.01)

H04L 5/00(2006.01)

权利要求书2页 说明书15页 附图15页

(54)发明名称

用于管理误差校正模式的方法和系统

(57)摘要

本发明揭示用于通过包括多个隧道的聚合连接将数据包从第一通信路由器传输到第二通信路由器的方法和系统。所述第一通信路由器从第一主机接收属于数据会话的第一数据包。所述第一通信路由器随后将所述第一数据包传输到所述第二通信路由器，并且将所述第一数据包存储于本地存储媒体中。所述第一通信路由器确定是否激活误差校正模式。如果激活误差校正模式，则所述第一通信路由器重新传输所述第一数据包，并且还传输对应于所述第一数据包的误差校正包。



1. 一种用于在第一通信路由器处执行管理误差校正模式的方法,其包括以下步骤:

(a) 传输第一数据包到第二通信路由器;

(b) 将所述第一数据包存储在本地存储媒体中;

(c) 当从所述第二通信路由器接收到延迟查询消息时,激活所述误差校正模式;其中所述延迟查询消息用于表示所述第二通信路由器尚未成功接收到所述第一数据包。

(d) 重新传输所述第一数据包到所述第二通信路由器;

(e) 传输对应于所述第一数据包的误差校正包;其中所述延迟查询消息包括所述第一数据包的全局序列号;

(f) 当从所述第二通信路由器接收恢复正常消息时,停用所述误差校正模式;其中所述恢复正常消息用于表示所述第一通信路由器不再需要处于所述误差校正模式中。

2. 根据权利要求1所述的方法,其中所述第一数据包是通过局域网接口从第一主机接收,其中所述第一数据包通过所述第二通信路由器去往可到达的第二主机。

3. 根据权利要求1所述的方法,其中所述误差校正包是多个包的组合,其中所述多个包是包括所述第一数据包。

4. 根据权利要求1所述的方法,其中所述误差校正包是前向纠错包、自动重复请求包或奇偶校验包。

5. 根据权利要求1所述的方法,其中当多个数据包未成功从所述第二通信路由器接收时,所述延迟查询消息包括所述多个数据包的全局序列号。

6. 根据权利要求1所述的方法,其中所述延迟查询消息是IP包,其中所述IP包包括标头、其它信息字段及净负荷;其中所述第一数据包的所述全局序列号是在所述净负荷中;其中所述其它信息字段包括所述第一数据包的会话信息及所述IP包是所述延迟查询消息的信息。

7. 根据权利要求1所述的方法,其中步骤(d)及步骤(e)在等待时间周期后执行。

8. 根据权利要求1所述的方法,其中所述第一数据包是通过从所述第一通信路由器及所述第二通信路由器建立的聚合连接传输;其中所述聚合连接包括多个端到端连接。

9. 根据权利要求8所述的方法,其中所述多个端到端连接包括第一端到端连接、第二端到端连接及第三端到端连接;

其中所述第一数据包在步骤(a)是通过所述第一端到端连接传输;

其中所述第一数据包在步骤(d)是通过所述第二端到端连接重新传输;以及

其中所述误差校正包在步骤(e)是通过所述第三端到端连接传输。

10. 根据权利要求1所述的方法,其中所述本地存储媒体是高速缓冲存储器;其中所述高速缓冲存储器能存储至少20秒数据传输。

11. 一种用于在第一通信路由器处执行管理误差校正模式的系统,其中所述第一通信路由器包括:

多个网络接口;

至少一个处理单元;

至少一个本地存储媒体,其存储可由所述至少一个处理单元执行以将数据包从所述第一通信路由器传输到第二通信路由器的程序指令,其包括以下步骤:

(a) 传输第一数据包到所述第二通信路由器;

(b) 将所述第一数据包存储在本地存储媒体中；

(c) 当从所述第二通信路由器接收到延迟查询消息时，激活所述误差校正模式；其中所述延迟查询消息用于表示所述第二通信路由器尚未成功接收到所述第一数据包。

(d) 重新传输所述第一数据包到所述第二通信路由器；

(e) 传输对应于所述第一数据包的误差校正包；其中所述延迟查询消息包括所述第一数据包的全局序列号；

(f) 当从所述第二通信路由器接收恢复正常消息时，停用所述误差校正模式；其中所述恢复正常消息用于表示所述第一通信路由器不再需要处于所述误差校正模式中。

12. 根据权利要求11所述的系统，其中所述第一数据包是通过局域网接口从第一主机接收，其中所述第一数据包通过所述第二通信路由器去往可到达的第二主机。

13. 根据权利要求11所述的系统，其中所述误差校正包是多个包的组合，其中所述多个包是包括所述第一数据包。

14. 根据权利要求11所述的系统，其中所述误差校正包是前向纠错包、自动重复请求包或奇偶校验包。

15. 根据权利要求11所述的系统，其中当多个数据包未成功从第二通信路由器接收时，所述延迟查询消息包括所述多个数据包的全局序列号。

16. 根据权利要求11所述的系统，其中所述延迟查询消息是IP包，其中所述IP包包括标头、其它信息字段及净负荷；其中所述第一数据包的所述全局序列号是在所述净负荷中；其中所述其它信息字段包括所述第一数据包的会话信息及所述IP包是所述延迟查询消息的信息。

17. 根据权利要求11所述的系统，其中步骤(d)及步骤(e)在等待时间周期后执行。

18. 根据权利要求11所述的系统，其中所述第一数据包是通过从所述第一通信路由器及所述第二通信路由器建立的聚合连接传输；其中所述聚合连接包括多个端到端连接。

19. 根据权利要求18所述的系统，其中所述多个端到端连接包括第一端到端连接、第二端到端连接及第三端到端连接；

其中所述第一数据包在步骤(a)是通过所述第一端到端连接传输；

其中所述第一数据包在步骤(d)是通过所述第二端到端连接重新传输；以及

其中所述误差校正包在步骤(e)是通过所述第三端到端连接传输。

20. 根据权利要求11所述的系统，其中所述本地存储媒体是高速缓冲存储器；其中所述高速缓冲存储器能存储至少20秒数据传输。

用于管理误差校正模式的方法和系统

技术领域

[0001] 本发明大体上涉及计算机网络的领域。更具体地说,本发明揭示用于通过包括多个隧道的聚合连接(aggregated connection)将数据包从第一通信路由器传输到第二通信路由器的方法和系统。

背景技术

[0002] 多广域网(WAN)站点到站点VPN路由器是支持集中多个互连(例如,用于访问一个或多个远程专用网络的WAN连接)的带宽的路由器。在一些实施方案中,每个TCP/IP会话仅路由到一个WAN。在此配置中,单个TCP文件传送会话仅能够利用每个终端上的一个WAN连接的带宽。例如,在基于会话的站点到站点虚拟专用网络(VPN)连接中,VPN流量被路由到两个站点(例如,站点A和B)之间的多个WAN连接。

[0003] 在一个实施方案中,在WAN连接之间先形成 $M \times N$ 个隧道,其中M和N分别是站点A和站点B的WAN网络连接的数目。随后经由不同的隧道路由应用程序TCP/IP会话。然而,应注意,虽然基于会话的站点到站点VPN能够针对不同的会话利用不同的隧道,但是此类型的连接中的单个下载会话仅能够利用一个隧道。

[0004] 在无线通信中,包传输的质量可能不可预测且丢包率可能经常改变。这会降低总体包传输的质量。即使每个隧道的带宽限制较高,丢包率也可能不会提高。需要一种利用多个隧道来增加成功传输数据的可能性的解决方案,所述解决方案可以通过不止一次传输数据包以及还通过传输可以用于检索原始数据包的包来实现。

发明内容

[0005] 本发明揭示用于通过包括多个隧道的聚合连接将数据包从第一通信路由器传输到第二通信路由器的方法和系统。第一通信路由器通过局域网(LAN)接口从第一主机接收属于数据会话的第一数据包。第一数据包去往第二主机,其中通过第二通信路由器可到达所述第二主机。第一通信路由器随后将第一数据包传输到第二通信路由器,并且将所述第一数据包存储于本地存储媒体中。第一通信路由器确定是否激活误差校正模式。如果激活误差校正模式,则第一通信路由器重新传输第一数据包,并且还传输对应于所述第一数据包的误差校正包。

[0006] 根据本发明的一个实施例,当第一通信路由器从第二通信路由器接收到第一延迟查询消息时,在所述第一通信路由器处激活误差校正模式。第一延迟查询消息表示第二通信路由器未成功接收到第一数据包。第一延迟查询消息可以包括第一数据包的第一全局序列号。当第二通信路由器未成功接收到多个数据包时,延迟查询消息包括多个数据包的多个全局序列号。

[0007] 根据本发明的一个实施例,当第一通信路由器从第二通信路由器接收恢复正常消息时,第一通信路由器停用误差校正模式。恢复正常消息用于表示第一通信路由器不再需要处于误差校正模式。在一个实施例中,可以由第一通信路由器的管理员激活误差校正模

式。

[0008] 根据本发明的一个实施例,第一数据包通过聚合连接的第一隧道传输,第一数据包通过聚合连接的第二隧道重新传输,并且误差校正包通过聚合连接的第三隧道传输。

[0009] 根据本发明的一个实施例,当激活误差校正模式时,仅当在等待时间周期内未从第二主机接收到对应于第一数据包的确认时才重新传输第一数据包且传输误差校正包。

[0010] 根据本发明的一个实施例,等待时间周期优选小于第一时间周期减去第三时间周期,并且还优选小于第二时间周期减去第三时间周期。第一时间周期是第一主机传输第一数据包与第一主机决定重新传输第一数据包之间的时间。第二时间周期是第一主机传输第一数据包与第二主机将请求发送到第一主机以重新传输第一数据包之间的时间。第三时间周期是第一数据包从第一通信路由器传输到第二通信路由器所需的时间。

[0011] 根据本发明的一个实施例,在第一主机重新传输第一数据包之前,重新传输第一数据包且传输误差校正包。

附图说明

[0012] 图1A说明根据本发明的实施例的用于优化多个可变带宽连接的吞吐量的整个系统;

[0013] 图1B说明根据本发明的各种实施例的网络环境;

[0014] 图1C说明根据实施例调适的系统100,所述系统被配置成优化结合的多个可变带宽连接的吞吐量。

[0015] 图2A说明根据本发明的实施例的描绘用于增加结合连接的吞吐量的方法的流程图;

[0016] 图2B说明根据本发明的实施例的描绘用于增加结合连接的吞吐量的方法的流程图;

[0017] 图3是说明根据本发明的实施例的可以封装在所传输IP包中的信息的类型的示例实施例;

[0018] 图4A是说明根据本发明的实施例的可以封装在反馈包中的信息的类型的示例实施例;

[0019] 图4B是说明图4A的反馈包的字段的可能值的图表;

[0020] 图5描绘被调适成实施本发明的处理系统的框图;

[0021] 图6说明根据本发明的各种实施例的网络环境;

[0022] 图7是说明根据本发明的一个实施例的如何在主机之间传送数据的时序图;

[0023] 图8是说明根据本发明的一个实施例的如何在主机之间传送数据的时序图;

[0024] 图9是说明根据本发明的一个实施例的在第一通信路由器处执行的过程的流程图;

[0025] 图10是说明根据本发明的一个实施例的在第二通信路由器处执行的过程的流程图;

[0026] 图11说明根据本发明的一个实施例的延迟查询消息的结构。

具体实施方式

[0027] 以下说明仅提供优选的示例性实施例且并不意图限制本发明的范围、适用性或配置。实际上,优选的示例性实施例的以下说明将为本领域技术人员提供实施本发明的优选示例性实施例的有利描述。应理解,在不脱离如在所附权利要求书中阐述的本发明的精神和范围的情况下可以对元件的功能和布置进行各种改变。

[0028] 在以下描述中给出具体细节以提供对实施例的深入理解。然而,所属领域的技术人员应理解,所述实施例可以在没有这些具体细节的情况下实践。例如,可以以框图示出电路以免以不必要的细节混淆实施例。在其它情况下,可以在没有不必要的细节的情况下示出熟知的电路、过程、算法、结构以及技术以便避免混淆实施例。

[0029] 同样,应注意,实施例可以描述为过程,过程描绘为流程图、流程图表、数据流图、结构图或方块图。尽管流程图可以将操作描述为连续过程,但许多操作可以并行或同时执行。另外,可以重新布置操作的顺序。当操作完成时,过程终止,但是过程可以具有不包含在图中的另外步骤。过程可以对应于方法、函数、程序、子例程、子程序等。当过程对应于函数时,其终止对应于函数返回到调用函数或主函数。

[0030] 实施例或其各部分可以以程序指令来实施,所述程序指令可在处理单元上操作以用于执行如本文中所描述的功能和操作。构成各个实施例的程序指令可以存储于存储媒体中。

[0031] 构成各个实施例的程序指令可以存储于存储媒体中。此外,如本文所揭示,术语“存储媒体”可以表示用于存储数据的一个或多个装置,包含只读存储器(ROM)、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、随机存取存储器(RAM)、磁性RAM、磁芯存储器、软盘、软磁盘、硬盘、磁带、CD-ROM、快闪存储器装置、存储卡和/或用于存储信息的其它机器可读媒体。术语“机器可读媒体”包含但不限于,便携式或固定存储装置、光学存储媒体、磁性媒体、存储器芯片或盒式磁盘、无线信道以及能够存储、容纳或载送指令和/或数据的不同其它媒体。机器可读媒体可以通过虚拟化来实现,且可以是虚拟机器可读媒体,包含在基于云的实例中的虚拟机器可读媒体。

[0032] 如本文中所使用的术语“计算机可读媒体、主存储器或辅助存储装置”是指参与将指令提供到处理单元以用于执行的任何媒体。计算机可读媒体仅是机器可读媒体的一个实例,所述机器可读媒体可以载送指令以用于实施本文中所描述的方法和/或技术中的任一个。此媒体可以采用许多形式,包含但不限于非易失性媒体、易失性媒体和传输媒体。非易失性媒体包含(例如)光盘或磁盘。易失性媒体包含动态存储器。传输媒体包含同轴电缆、铜线以及光纤。传输媒体还可以采用声波或光波的形式,例如在无线电波和红外线数据通信期间产生的声波或光波。

[0033] 易失性存储装置可以用于在通过处理单元执行指令期间存储临时变量或其它中间信息。非易失性存储装置或静态存储装置可以用于存储用于处理器的静态信息和指令,以及各种系统配置参数。

[0034] 存储媒体可以包含多个软件模块,所述软件模块可以实施为通过处理单元使用任何合适的计算机指令类型来执行的软件代码。软件代码可以作为一系列指令或命令、或作为程序存储在存储媒体中。

[0035] 各种形式的计算机可读媒体可以涉及将一个或多个指令的一个或多个序列载送

到处理器以便执行。例如,指令可以首先承载在远程计算机的磁盘上。可替代地,远程计算机可以将所述指令加载到其动态存储器中,且将指令发送到运行一个或多个指令的一个或多个序列的系统。

[0036] 处理单元可以是微处理器、微控制器、数字信号处理器 (DSP)、那些装置的任何组合、或被配置成处理信息的任何其它电路。

[0037] 处理单元执行程序指令或代码段以用于实施本发明的实施例。此外,实施方案可以由硬件、软件、固件、中间件、微码、硬件描述语言或其任何组合来实施。当在软件、固件、中间件或微码中实施时,用于执行必要任务的程序指令可以存储于计算机可读存储媒体中。处理单元可以通过虚拟化来实现,且可以是虚拟处理单元,包含在基于云的实例中的虚拟处理单元。

[0038] 本发明的实施例涉及使用计算机系统来实施本文所描述的技术。在实施例中,本发明的处理单元可以驻留在计算机平台等机器上。根据本发明的一个实施例,本文中所描述的技术由计算机系统执行,以响应于处理单元执行易失性存储器中所包含的一个或多个指令的一个或多个序列。此类指令可以从另一计算机可读媒体读取到易失性存储器中。执行易失性存储器中所包含的指令的序列使得处理单元执行本文中所描述的过程步骤。在替代实施例中,硬接线电路可以用于取代或结合软件指令以实施本发明。因此,本发明的实施例不限于硬件电路以及软件的任何具体组合。

[0039] 程序指令等代码段可以表示步骤、函数、子程序、程序、例程、子例程、模块、软件包、类或者指令、数据结构或程序语句的任何组合。代码段可以通过传递和/或接收信息、数据、自变量、参数或存储器内容而耦合到另一代码段或硬件电路。可以经由包含存储器共享、消息传递、令牌传递、网络传输等任何合适的手段传递、转发或传输信息、自变量、参数、数据等。

[0040] 可替代地,硬接线电路可以用于取代或结合软件指令以实施符合本发明的原理的过程。因此,符合本发明的原理的实施方案不限于硬件电路和软件的任何特定组合。

[0041] 可以由节点提供的网络接口是以太网接口、帧中继接口、光纤接口、电缆接口、DSL 接口、令牌环接口、串行总线接口、通用串行总线 (USB) 接口、火线接口、外围组件互连 (PCI) 接口等。

[0042] 网络接口可以通过独立的电子组件实施或者可以与其它电子组件整合。取决于配置,网络接口可以不具有网络连接或具有至少一个网络连接。网络接口可以是以太网接口、帧中继接口、光纤接口、电缆接口、数字订户线 (DSL) 接口、令牌环接口、串行总线接口、通用串行总线 (USB) 接口、火线接口、外围组件互连 (PCI) 接口、蜂窝网络接口等。

[0043] 网络接口可以连接到有线或无线接入网络。接入网络可以载送一个或多个网络协议数据。有线接入网络可以使用以太网、光纤、电缆、DSL、帧中继、令牌环、串行总线、USB、火线、PCI或可以传递信息的任何材料来实施。无线接入网络可以使用红外线、高速包接入 (HSPA)、HSPA+、长期演进 (LTE)、WiMax、通用包无线电服务 (GPRS)、全球移动通信系统 (GSM)、GSM演进增强数据速率 (EDGE)、码分多址 (CDMA)、WiFi、CDMA2000、宽带CDMA (WCDMA)、时分CDMA (TD-SCDMA)、蓝牙、WiBro、演进数据优化 (EV-DO);数字增强型无绳通信 (DECT);数字AMPS (IS-136/TDMA);集成数字增强型 (iDEN) 或任何其它无线技术来实施。例如,网络接口可以用作局域网 (LAN) 接口或广域网 (WAN) 接口。

[0044] 实施例或其各部分可以以计算机数据信号来实施,所述计算机数据信号可以采用用于经由传输媒体进行通信的任何合适形式,使得所述计算机数据信号是可读的以用于通过功能装置(例如,处理单元)来执行从而执行本文中所描述的操作。计算机数据信号可以包含能够经由传输媒体传播的任何二进制数字电子信号,所述传输媒体例如电子网络信道、光纤、空气、电磁媒体、射频(RF)链路等,且因此数据信号可以采用电信号、光信号、射频或其它无线通信信号等形式。在某些实施例中,代码段可以经由计算机网络下载,所述计算机网络例如因特网、企业内部网、LAN、城域网(MAN)、广域网(WAN)、公共交换电话网络(PSTN)、卫星通信系统、电缆传输系统和/或其类似者。

[0045] 图1A说明根据实施例调适的系统100,所述系统被配置成通过调节在数据传送会话期间的隧道带宽加权模式来优化结合的多个可变带宽连接的吞吐量。系统100包含多个站点102和104,每个站点各自包括至少一个网络节点。网络节点可以称为通信路由器。然而,本发明的范围不限于通信路由器,使得可以在网关、路由器、服务器或任何其它类型的网络节点处执行本发明。为简单起见,图1A说明站点102和104分别包括通信路由器106和108。通信路由器106和108可以实施为支持集中多个因特网连接的带宽的多WAN路由器。通信路由器106和108通过网络110连接。网络110可以包括LAN、MAN、WAN、无线网络、PSTN、因特网、企业内部网、外联网等。

[0046] 站点102和路由器106可以包括M个连接112,且站点104和路由器108可以包括N个连接114。连接112和114是用于在站点102与104之间的网络110内传送信息的数据连接。在所说明的实施例中,M等于3且N等于2;然而,这些值可以根据所需的路由器和配置改变。连接112和114可以具有类似的或不同的带宽容量。此外,连接112和114可以包括不同类型的WAN连接,例如WiFi、电缆、DSF、T1、3G、4G、卫星连接等。还应注意,站点102和站点104可以被认为是发送器或接收器两者,并且关于任一站点的功能性的论述可以在另一站点上实施。换句话说,系统100可以实施为对称网络。

[0047] 图1B说明根据本发明的一个实施例的网络环境。在通信路由器106与通信路由器108之间建立隧道103A、103B和103C。可以结合隧道103A、103B和103C以形成聚合连接。

[0048] 根据一个实施例,通信路由器106和108可以具有多个网络接口。通信路由器106经由其多个网络接口中的一个或多个与通信路由器108的一个或多个网络接口建立隧道103A、103B和103C。

[0049] 通信装置106和108可以充当网关、路由器、交换机、接入点、集线器、网桥等。

[0050] 图1C说明根据实施例调适的系统100,所述系统被配置成优化结合的多个可变带宽连接的吞吐量。系统100类似于系统101,除 $M \times N$ 个虚拟隧道116之外。当例如通过实施结合的站点到站点VPN连接来在站点102与104之间建立结合的连接时,可以创建 $M \times N$ 个隧道116。隧道116对应于站点102的网络连接和站点104的网络连接的独特排列。可以在通信路由器106与108之间形成聚合连接。隧道116可以是虚拟隧道。

[0051] 多个建立的隧道116可以集中、组合或结合在一起以形成一个聚合连接。所属领域的技术人员将理解,存在用于集中、组合或结合多个建立的隧道以形成一个集中式隧道的多种方法。使用聚合连接的会话或应用将所述聚合连接视为一个隧道。聚合连接可以是端到端连接、虚拟专用网络连接或面向无连接连接。例如,聚合连接可以是TCP连接或UDP连接。在另一实例中,聚合连接是多个隧道的集中,并且每个隧道连接在通信路由器106与通

信路由器108之间。在另一实例中,聚合连接可以是VPN隧道,包括多个建立的隧道,并且每个建立的隧道连接在通信路由器106与通信路由器108之间。

[0052] 图2A示出系统100的操作的高阶流程图,其描绘了用于增加结合的连接的吞吐量的方法200。应了解,图2中提供的具体功能、功能的顺序等意图是根据本发明的概念的示例性操作。因此,本文中的概念可以用与所说明的实施例不同的方式来实施。

[0053] 在所说明的实施例的步骤201处,当例如通过实施结合的站点到站点VPN连接而在路由器102与104之间建立结合的连接时,可以创建 $M \times N$ 个虚拟隧道116,如在图1C中所说明。虚拟隧道116对应于站点102的网络连接和站点104的网络连接的独特排列。

[0054] 在所说明的实施例的步骤202处,为隧道确定和/或分配默认权重。为了确定默认权重,实施例在站点102与104之间交换连接112和114的上行和下行带宽数据。使用此带宽数据,可以根据以下来计算默认权重:假设站点102的连接1至 m 的下行带宽是 d_1, d_2, \dots, d_m ,并且站点104的连接1至 n 的上行带宽是 u_1, u_2, \dots, u_n ;站点102的连接 x 与站点104的连接 y 之间的隧道的默认权重可以定义为 $DW(x, y)$,其中 $DW(x, y) = d_x \cdot u_y$ 。使用以上方法来计算默认权重,如果连接112-1到112-3是具有10M/6M、8M/4M和6M/6M的相应上行带宽/下行带宽的多WAN路由器的WAN连接,且连接114-1到114-2是具有7M/5M和9M/3M的相应上行带宽/下行带宽的多WAN路由器的WAN连接,那么每个隧道的相应默认权重将是如下:

[0055] [表0001]

[0056] 表0001

对于站点102	对于站点104
$DW(1, 1) = 6 \cdot 7 = 42$	$DW(1, 1) = 5 \cdot 10 = 50$
$DW(1, 2) = 6 \cdot 9 = 54$	$DW(1, 2) = 5 \cdot 8 = 40$
$DW(2, 1) = 4 \cdot 7 = 28$	$DW(1, 3) = 5 \cdot 6 = 30$
$DW(2, 2) = 4 \cdot 9 = 36$	$DW(2, 1) = 3 \cdot 10 = 30$
$DW(3, 1) = 6 \cdot 7 = 42$	$DW(2, 2) = 3 \cdot 8 = 24$
$DW(3, 2) = 6 \cdot 9 = 54$	$DW(2, 3) = 3 \cdot 6 = 18$

[0058] 应注意,能构想计算默认权重的其它方式,并且上文仅是本发明的实施例的实施方案的实例。应注意,可以使用许多不同的加权模式来定义隧道的初始带宽。例如,可能希望仅使用接收站点的下行容量和发送站点的上行容量对一个方向上的隧道加权。用来表征在建立结合的连接时的隧道容量的任何加权方案可以用于本发明的目的。

[0059] 当根据实施例将包从站点102路由到站点104时,以根据有效权重 $EW(x, y)$ 的比率将包分布到隧道。初始地,实施例的有效加权设置成等于默认权重, $EW(x, y) = DW(x, y)$,并且如果隧道116的带宽从初始设置保持不变,则所述有效权重对于包分布是最佳的。然而,如果用户在具有存在丢包情况的一个或多个隧道的TCP会话中正经由结合的网络连接下载文件,则所述会话的整体吞吐量将急剧下降。这部分是因为丢包将继续引起TCP重新传输,并且TCP的流量控制将保持较低吞吐量,即使无丢包情况的隧道未被全部占用。

[0060] 提高吞吐量的一种有效方式可以是避免此类丢包。为了这样做,在所说明的实施例的步骤203处,本发明的实施例对隧道正经历丢包率的提高或降低的时间进行辨别。在步骤204处,实施例进一步用于修改正经历或已经历丢包率的改变的隧道的有效权重。可以连续监控或基于特定时间周期监控丢包率信息。一旦确定隧道正经历不可接受的丢包率(步

骤204-1),在步骤204-2处,所说明的实施例就减小隧道的有效权重。在一些实施例中,不可接受的可能意味着丢包率是非零数量,而其它实施例可以确定不可接受的比率是除预定义阈值以外的任何比率。实施例通过逐步方式、连续方式、一次与丢包率的增加成比例减小的方式等来实施这些减小。当通过逐渐方式完成减小时,实施例可以继续监控隧道,以便优化实施的减小的量。

[0061] 可以通过经由每个隧道从路由器106或路由器108发送心跳包来建立或监控隧道116。在一些实施例中,当接收终端在一段时间未能从隧道接收到心跳包时,所述接收终端会将所述隧道看作下行并且所述隧道将不会用于路由流量。如果心跳包再次开始被接收到,则可以重新建立所述隧道并连同其它隧道一起加权所述隧道。由此,在所有包在隧道中被丢弃并且所述隧道的有效权重减小至零的情况下,实施例可以利用心跳包来监控并重新建立连接。

[0062] 此外,当隧道恢复其相应带宽的全部或部分时,例如,确定丢包率降低(步骤204-3)时,所说明的实施例用于增加此类隧道的有效权重(步骤204-4),以便充分或更充分地利用所述带宽。一些实施例使用预定步长增加隧道的有效权重,直到恢复精确的有效权重。其它实施例与最新测量的带宽成比例地增加有效权重,最新测量的带宽可以对应于最新测量的丢包率。此外,实施例可以基于预定线性或指数比例增加隧道的有效权重。

[0063] 在调节隧道的有效权重之后,或在确定无需调节之后,在所说明的实施例的步骤205处更新系统的加权模式。此更新可以包括存储任何处理后的信息、在进一步处理中使用此类信息、使得系统采取行动等。例如,关于步骤205执行的处理可以用来在一段时间内对加权模式取平均值以便降低与高度瞬态异常相关联的误差。此外,可以在系统100上使用更新后的信息以修改数据传送会话的包分布,如关于图2B所述。系统100可以在整个数据传送会话中连续地或周期性地继续实施步骤203到205。

[0064] 图2B说明其中在实施加权方法200之后至少部分基于隧道的修改后权重分布包的实施例。具体来说,所说明的实施例的步骤206用来根据通过方法200的操作确定的加权方案在隧道上分布包。在一些实施例中,此分布将在整个数据传送会话中改变,并且因此图2B的步骤被示为重复的。每当在步骤205处更新系统时,一些实施例改变包分布。此外,响应于某些丢失率变化阈值等,步骤205可以引起周期性地实施变化。应了解,通过方法200的操作确定加权以及在步骤206处将所确定的加权应用于包分布可以具有不同周期性。例如,方法200可用以使用相对较短的迭代循环来提供加权模式信息的更新,而包的分布使用较长的迭代循环基于此类加权模式信息而变化。

[0065] 为了监控各个隧道116的带宽,本发明的一些实施例用各种信息封装每个所传输的IP包。图3说明示出可以封装在所传输的IP包中的信息300的类型的示例实施例。版本字段302可以含有关于所利用的协议版本的信息且协议类型字段303可以含有净负荷包的协议类型。一般来说,此字段的值将对应于包的以太网协议类型。然而,另外的值可以在其它文档中定义。隧道ID字段304可以是32位字段并且可以含有用以识别IP包的当前隧道的标识符。高级加密标准(AES)初始化向量字段306可以是32位字段并且可以含有AES加密的初始化向量。全局序列号字段308可以是32位字段并且可以含有序列号,所述序列号用来在用于各个会话的包中的每个包已经从其相应隧道出现时将所述包重新排序为适当的顺序。每隧道序列号字段310可以是32位字段,所述字段可以表示分配给路由到具体隧道的每个包

的序列号。AES加密后的净负荷字段312可以用来传送IP包的净负荷。AES加密可以适用于净负荷的较高安全性,以便防止第三方攻击。

[0066] 上文所论述的每个隧道序列号可以用来监控隧道中的丢包。在一个实施例中,接收终端上的路由器通过监控接收到的包的每个隧道序列号来计算每f秒每个隧道的丢包率 $DR(x, y)$ 。 $DR(x, y)$ 可以表征为在周期f内错过的序列号除以序列号增加量。周期f的长度可以改变,并且在一个实施例中,f等于5秒。

[0067] 还可以用其它方法监控丢包,例如:发送器可以周期性地通知接收终端其已发送的包的数量,发送器每恒定时间周期将心跳包发送到接收终端以及接收终端可以通过监控心跳包的丢失率、通过从物理接口/装置/层等获取丢失率图来估计总体丢失率。

[0068] 接收终端可以将具体隧道的丢失率、有效权重或其它带宽指示返回到发送路由器。当发送器接收到关于丢包的信息时,一些实施例将隧道的有效权重 $EW(x, y)$ 减小 $EW(x, y) \cdot DR(x, y)$ 。可以使用其它度量来修改隧道的有效权重。在一些实施例中,发送器可以接收反馈,并且有效权重可以减小大于或小于丢包率的数目。可以根据通信系统的具体需要来配置此类变化。以上示例表示尝试将隧道的有效权重减小至防止进一步丢包同时将隧道的可用带宽量的权重的度量最大化。发现此平衡的任何度量可以是优选的。

[0069] 图4A说明可以封装在反馈包中的信息400的类型的示例实施例,将所述反馈包发送到传输路由器,以便报告丢包率或在接收终端路由器处接收到的其它带宽相关数据。类型字段402可以包含关于数据1字段404和数据2字段406中将包含的数据类型的数据。数据1字段404和数据2字段406可以含有可以用于帮助路由器确定关于隧道数目、隧道带宽、隧道中的丢包次数等的隧道信息的任何信息。在图4B的图表中示出数据字段404和406中的类型字段402的可能值的示例。

[0070] 封装在例如图3和图4中所示的所传输的IP包中的信息还可以用于包缓存和重新排序。因为每个隧道的时延可能不同,所以当同一TCP会话的两个连续包通过结合的VPN隧道发送到VPN对等装置时,所述两个连续包不按顺序到达,因为所述两个连续包经由两个不同隧道路由。如果TCP会话从VPN接收无序包,则TCP会话将由于TCP重新传输而减慢。因此,在较慢包到达之前或在已经过到期时间之前,接收终端应缓存太早出现的包。通过此缓存,在到期时间之前出现的迟延包将按顺序转发到目标装置。此缓存有助于优化端对端吞吐量。

[0071] 应注意,时常在VPN连接的上下文中论述本文中所描述的实施例。展现这些论述,以便示出结合的连接的示例实施例。在本文的权利要求中描述的发明性概念不限于此连接。实际上,可以通过本发明的实施例来实施其中可以获得和交换足够数据,以便动态地监控多个通信路径的带宽的任何连接,所述带宽用于数据传送会话。

[0072] 如上文所论述,每个包可以分配有两个不同的序列号,即全局序列号(global sequence number,GSN)和每个隧道序列号(per tunnel sequence number,PTSN)。这些数字可以用于帮助包缓存和重新排序操作。在将包传递到上层之后,接收终端可以更新下一预期的每个隧道序列号(next expected per tunnel sequence number,NE-PTSN)和下一预期的全局序列号(next expected global sequence number,NE-GSN)。

[0073] 下文将描述在接收和解密包之后包可以如何进行缓存或转发到目标装置的一个方法。

- [0074] 1. 如果包的GSN等于零,则立即将包转发到目标装置。
- [0075] 2. 检查包的PTSN是否等于NE-PTSN。如果不是,则按顺序将具有较小GSN的所有包出列(转发到目标装置)。保持包未处理。
- [0076] 3. 更新NE-PTSN(即,将NE-PTSN设定成PTSN+1)。
- [0077] 4. 如果GSN小于NE-GSN,则转发到目标装置。
- [0078] 5. 如果包的GSN等于NE-GSN,则更新NE-GSN(即,将NEGSN设定成GSN+1)且转发到目标装置。重复更新NE-GSN,并且如果缓存头的GSN等于新的NE-GSN,则将缓存头从缓存器出列。
- [0079] 6. 否则(GSN大于NE-GSN),按GSN的顺序将包入列。
- [0080] 7. 如果包处于比固定时间量长的队列中,则将NE-GSN设定成包的GSN+1且按顺序将包以及具有较小GSN的所有包出列。
- [0081] 因此,在图2和图3中论述的封装包信息可以包含通过响应于监控丢包率,而帮助优化隧道带宽以及通过帮助有效重新排序数据传送会话中所接收包两者,来优化数据传输系统(例如,100)的总体吞吐量的信息。
- [0082] 图5说明根据某些实施例的可以用于实施系统、装置和方法的示例性基于处理器的系统500。基于处理器的系统500可以表示通信路由器106和108的架构。中央处理单元(CPU)501耦合到系统总线502。CPU 501可以是任何通用CPU,或可以是被设计成实施以上教示的专用CPU。只要CPU 501(和系统500的其它组件)支持如本文所描述的发明性操作,则本发明不受CPU 501(或示例性系统500的其它组件)的架构限制。CPU 501可以执行本文所描述的各种逻辑指令。例如,CPU 501可以根据上文结合图2描述的示例性操作流执行机器级指令。当执行表示图2中所说明的操作步骤的指令时,CPU 501变为具体来说被配置成根据本文所描述的教示的各种实施例操作的专用计算平台的专用处理器。
- [0083] 系统500还包含随机存取存储器(RAM)503,其可以是SRAM、DRAM、SDRAM等等。RAM 503可以是存储可由CPU 501执行的程序指令的辅助存储装置。系统500包含只读存储器(ROM)504,其可以是PROM、EPROM、EEPROM等等。如在本领域中众所周知,RAM 503和ROM 504保存用户和系统数据以及程序。
- [0084] 系统500还包含输入/输出(I/O)适配器505、通信适配器511、用户接口适配器508和显示适配器509。在某些实施例中,I/O适配器505、用户接口适配器508和/或通信适配器511可以使用户能够与系统500交互,以便输入信息。
- [0085] I/O适配器505将例如硬盘驱动器、压缩光盘(CD)驱动器、软盘驱动器、磁带驱动器等的存储装置506连接到系统500。除了RAM 503之外,存储装置用于执行以上实施例中所述的操作的相关存储器需求。通信适配器511被调适成将系统500耦合到网络512,这可以使信息能够经由此网络512(例如,因特网或其它广域网、局域网、公共或私有交换电话网络、无线网络、上述项的任何组合)输入系统500和/或从系统500输出。通信适配器511可以被视为网络接口,并且系统500可以包括多个通信适配器511。用户接口适配器508将用户输入装置,例如键盘513、指向装置507和麦克风514和/或输出装置,例如扬声器515耦合到系统500。显示适配器509由CPU 501驱动以控制在显示装置510上的显示。显示适配器509传输用于变换或操纵供显示装置510使用的各种数目个像素的状态的指令,以在视觉上向用户呈现所需信息。此类指令包含用于从开到关改变状态、设定特定的色彩、强度、持续时间或类

似的指令。每一此种指令组成呈现指令,用以控制如何在显示装置510上显示以及在显示装置510上显示什么内容。

[0086] 图6说明基于图1A中所说明的网络环境的网络环境。通信路由器106的LAN接口中的一个连接到主机103,并且通信路由器108的LAN接口中的一个连接到主机105。主机103和105分别处于站点102和104中。可以通过网络110在通信路由器106与108之间建立多个隧道。通信路由器106可以通过通信路由器106的一个或多个WAN接口以及通信路由器108的一个或多个WAN接口与通信路由器108建立隧道。仅出于说明的目的,通信路由器108通过两个连接来连接到因特网110,而通信路由器106通过两个连接来连接到网络110。网络110可以是互连网络,例如因特网。

[0087] 图7是根据本发明的一个实施例的说明如何在主机105与主机103之间传送数据的时序图。在流701中,将数据-1从主机103传输到通信路由器106。数据-1源于主机103且预期传输到主机105。数据-1可以包括在一个或多个数据包、一个或多个帧或任何其它数据单元中。数据-1包括在一个数据包还是多个数据包中取决于数据-1的大小。如果数据-1的大小使得数据-1可以封装在一个数据包中,则数据-1包括在一个数据包中。如果数据-1太大而无法封装在一个数据包中,则数据-1可以分段成多个部分并且多个部分分别封装在多个数据包中。通信路由器106随后通过互连网络110在流702中将数据-1发送到通信路由器108。通信路由器108将数据-1转发到主机105。当主机105接收数据-1时,所述主机产生去往主机103的确认包,即Ack-1。Ack-1在流704处在通信路由器108处接收且在流705处转发到通信路由器106。通信路由器106随后在流706中将Ack-1发送到主机103。因此,主机103目前知晓数据-1成功地到达主机105。主机103随后通过通信路由器106和108将数据-2传输到主机105。在流707中,将数据-2从主机103传输到通信路由器106。通信路由器106在流708中将数据-2转发到通信路由器108。通信路由器108并不按时接收数据-2,因为当传输到通信路由器108时数据-2可以被舍弃或可以延迟。通信路由器108随后在流709中产生和发送延迟查询消息,以通知通信路由器106尚未接收到数据-2并且检查为何存在延迟。当通信路由器106接收延迟查询消息时,所述通信路由器确定数据-2已丢失且因此激活误差校正模式。在误差校正模式下,通信路由器106分别在流710和711中发送数据-2-1和数据-2-EC。数据-2-1可以与数据-2相同,而数据-2-EC包括对应于数据-2的误差校正数据。使用数据-2-1和/或数据-2-EC,通信路由器108可以在流712中产生数据-2且将数据-2发送到主机105。当主机105接收数据-2时,所述主机产生去往主机103的确认包,即Ack-2。Ack-2在流713处在通信路由器108处接收且在流714处转发到通信路由器106。通信路由器106随后在流715中将Ack-2发送到主机103。主机103通过通信路由器106和108将数据-3发送到主机105。在流716中将数据-3传输到通信路由器106。由于在通信路由器106中激活误差校正模式,因此通信路由器106分别在流717、718和722中将数据-3和数据-3-1两者以及数据-3-EC发送到通信路由器108。通信路由器108在流719中成功地接收数据-3-1且将数据-3转发到主机105。主机105随后产生且发送去往主机103的确认包,即Ack-3。Ack-3在流720处在通信路由器108处接收且在流721处转发到通信路由器106。通信路由器106随后在流723中将Ack-3发送到主机103。在通信路由器108接收数据-3-1和数据-3-EC两者时,所述通信路由器确定数据不再丢失,且因此在流724中将恢复正常消息发送到通信路由器106。当通信路由器106接收恢复正常消息时,所述通信路由器可以停用误差校正模式。当停用误差校正模式时,通信路由

器106不再发送误差校正包,例如,数据-2-EC和数据-3-EC。

[0088] 在例如TCP的一些通信协议中,当数据传送会话结束时,结束消息可以从发送器发送到接收器,以便指示数据传送会话已结束,即,将不再发送属于数据传送会话的数据包。通信路由器108可以确定例如数据-2的数据延迟或丢失,因为在接收数据-1之后,所述通信路由器可能未接收到任何结束消息。因此,所述通信路由器在流709中发送延迟查询消息,以便请求通信路由器106重新发送数据-2。

[0089] 当在通信路由器106处从主机103接收数据时,通信路由器106可以将数据保存在例如高速缓冲存储器的本地存储媒体中,使得在必要时通信路由器可以发送数据的副本,而不需要主机103重新发送数据。在当传输到通信路由器108时数据丢失的情况下,通信路由器106可以快速重新发送数据。不需要等待主机103重新发送数据或通知主机103所述丢失。本领域技术人员将理解,在一些通信协议中,当主机103了解数据丢失时,主机103可以减慢数据传输。由于减少等待时间,因此通信路由器108能够从通信路由器106接收数据的重新传输,而不是向主机103发出重新传输请求。

[0090] 在流721中,通信路由器106发出数据-3-EC。出于说明的目的,巧合的是,在通信路由器106处接收到Ack-3之前发送数据-3-EC。在一个变型例中,如果早期已在通信路由器106处接收到Ack-3,则通信路由器106将不具有传输数据-3-EC。

[0091] 数据-2-1的内容可以与数据-2的内容相同,因为通信路由器106通过发送数据-2-1来重新发送数据-2。数据-2-EC封装在误差校正包中。在一个变型例中,封装数据-2-1的数据包与封装数据-2的数据包相同,因为数据-2-1的内容与数据-2的内容相同,数据包的源地址相同且数据包的目标地址相同。

[0092] 误差校正包可以是由通信路由器106传输以校正或检测从通信路由器106传输到通信路由器108的数据包中的任何误差的包。误差校正包包含前向纠错(forward error correction,FEC)包或自动重复请求(Automatic repeat request,ARQ)包、奇偶校验包等。在一个变型例中,误差校正包还可以是两个或多于两个包的组合。例如,如果第一包包括数据-2且第二包包括数据-2-EC,则第三包可以包括基于应用于第一和第二包的内容的异或(XOR)操作的数据内容。使用数据-2-EC,通信路由器108可能检索数据-2且将数据-2发送到主机105。

[0093] 在一个变型例中,数据-2可以分段成两个部分。数据-2的第一部分封装在第一数据包中,并且数据-2的第二部分封装在第二数据包中。出于说明的目的,当通信路由器106在流708中将数据-2发送到通信路由器108时,通信路由器108接收第一数据包,但舍弃或丢失第二数据包。通信路由器108随后在流709中发送延迟查询消息。数据-2-EC可以包括误差校正包,所述误差校正包的数据内容是基于应用于第一数据包和第二数据包的内容的XOR操作,即,应用于数据-2的第一部分和数据-2的第二部分的XOR操作。当通信路由器106发送数据-2-EC时,通信路由器108可以通过使用数据-2和数据-2-EC的第一部分来检索数据-2的第二部分。本领域技术人员应了解,可以如何使用数据-2和数据-2-EC的第一部分来检索数据-2的第二部分,因为数据-2-EC是基于数据-2的第一和第二部分。

[0094] 在一个变型例中,当通信路由器106接收延迟查询消息且激活误差校正模式时,通信路由器106可以仅发送误差校正包,例如数据-2-EC,且不发送数据-2-1。在另一变型例中,当通信路由器106接收延迟查询消息时,通信路由器106可以仅发送数据-2-1且不发送

数据-2-EC。仅发送数据-2-1和数据-2-EC中的任一个的益处在于,传输更少数目的包,且因此消耗较低带宽。这可以有助于减少流量拥挤。然而,在一些情形中,如果未接收到数据-2-1,则数据-2-EC单独无法帮助检索数据-2。例如,当数据-2-EC仅含有校验和、或奇偶校验包时,所述数据-2-EC可以仅用于误差检测。在另一实例中,当通过在数据-2的两个包上应用XOR操作来产生数据-2-EC时,如果已接收到数据-2的两个包中的至少一个,则可以仅从数据-2-EC检索到数据-2。因此,发送数据-2-1可能是有帮助的,使得可以使用数据-2-1和数据-2-EC检索数据-2。在一个变型例中,当在通信路由器108处接收到的数据-2-1中可能存在一些误差时,数据-2-EC可以用于校正误差。

[0095] 图11说明根据本发明的一个实施例的延迟查询消息的结构。当舍弃或丢失任何数据包时,即,存在丢失的数据包时,可以发送延迟查询消息1100。延迟查询消息1100可以是包括IP标头1101、其它信息字段1104和净负荷1105的IP包。IP标头1101包括源地址字段1102和目标地址字段1103。例如,参考图7,在流709中发送延迟查询消息,源地址是通信路由器108的IP地址并且目标地址是通信路由器106的IP地址。其它信息字段1104可以包含例如IP包的本质等的不同信息,即,IP包是延迟查询消息的信息。这向通信路由器106表明IP包不是数据包,而仅仅是管理消息。通信路由器106可以通过不同方式处理数据包和管理消息。例如,通信路由器106可以将来自通信路由器108接收的数据包存储在高速缓冲存储器中,但是通信路由器106可以不将例如延迟查询消息等的管理消息存储在高速缓冲存储器中。其它信息字段1102可以进一步包含会话信息,以便指示丢失包属于哪一会话。净负荷1105可以包含丢失包的GSN。例如,在流709中,由通信路由器108发送到通信路由器106的延迟查询消息1100的净负荷1105可以包含对应于数据-2的一个或多个数据包的GSN,因为数据-2尚未成功到达通信路由器108。当通信路由器106从通信路由器108接收延迟查询消息1100,并且净负荷1105包含对应于数据-2的一个或多个数据包的GSN时,通信路由器106可以确定通信路由器108尚未接收到数据-2。因此,通信路由器106随后将数据-2-1和数据-2-EC发送到通信路由器108。本发明的范围不限于延迟查询消息是IP包。延迟查询消息可以是短消息服务(SMS)消息、多媒体消息服务(MMS)消息或可以由通信路由器108发送到通信路由器106以提供任何丢失数据包的信息的任何其它类型的消息。

[0096] 图8是根据本发明的一个实施例的说明如何在主机105与主机103之间传送数据的时序图。主机103通过通信路由器106和108将数据-1发送到主机105。在流801中,将数据-1从主机103传输到通信路由器106。通信路由器106随后在流802中将数据-1传输到通信路由器108,并且通信路由器108在流803中将数据-1转发到主机105。通信路由器106最初处于误差校正模式中,并且因此在等待时间周期之后,所述通信路由器分别在流806和807中发送数据-1-1和数据-1-EC。传输数据-1-1和数据-1-EC的益处在于,在丢失数据-1的情况下,主机103将不必重新传输数据-1。数据-1-1和数据-1-EC可以有助于通信路由器108产生数据-1,并且在于流802中丢失数据-1的情况下将数据-1发送到主机105。当主机105接收数据-1时,所述主机产生去往主机103的确认包,即,Ack-1。Ack-1在流804中传输到通信路由器108,并且在流805中由通信路由器108传输到通信路由器106。当通信路由器108从通信路由器106接收数据-1和数据-1-1两者时,所述通信路由器108在流808中将恢复正常消息发送到通信路由器106。在接收恢复正常消息之后,通信路由器106停用误差校正模式,因为不再丢失包。主机103传输去往主机105的数据-2。当在流810中在通信路由器106处接收数据-2

时,通信路由器106在流811中将数据-2转发到通信路由器108。通信路由器108随后在流812中将数据-2转发到主机105。在已停用误差校正模式时,通信路由器106不发送数据-2的任何误差校正包。当主机105接收数据-2时,所述主机产生去往主机103的确认包,即,Ack-2。Ack-2在流813中被发送到通信路由器108,并且在流814中被转发到通信路由器106。通信路由器106随后在流815中将Ack-2发送到主机103。

[0097] 可以将通信路由器106发送数据-1与发送数据-1-1之间的等待时间周期调节或调适到不同网络需求。发射数据-1-1的目的是减小主机103重新传输数据-1的可能性。出于以下原因中的一个或两个,主机103可以重新传输数据-1:(i)当主机103确定数据-1丢失时,因为未接收到确认;以及(ii)当主机105将请求发送到主机103以重新发送数据-1时。因此,在一个变型例中,等待时间周期优选小于(i)主机103发送数据-1与主机103决定重新传输数据-1之间的第一时间周期以及(ii)主机103传输数据-1与主机105决定将请求发送到主机103以重新发送数据-1之间的第二时间周期。可以在例如TCP协议的传输协议中定义第一时间周期和第二时间周期。由于特定延迟引起数据从通信路由器106行进到通信路由器108,因此对于设定等待时间周期,还可以考虑数据包从通信路由器106传输到通信路由器108所需的第三时间周期。第三时间周期可以根据网络条件变化。等待时间周期随后可以小于第一时间周期减去第三时间周期,并且还可以小于第二时间周期减去第三时间周期。减小主机103重新传输数据-1的可能性的益处在于,主机103不会由于丢包而减慢数据传输。在另一变型例中,等待时间周期大于第四时间周期,所述第四时间周期一般来说是Ack-1从通信路由器108传输到通信路由器106所需的时间。这可以使通信路由器108在所需时间内等待接收Ack-1,并且如果在第四时间周期内未接收Ack-1,则传输数据-1-1和数据-1-EC。等待时间周期大于第四时间周期的益处在于可以消耗更少带宽,因为如果接收到Ack-1,则不传输数据-1-1和数据-1-EC。然而,第四时间周期可以根据网络条件变化。等待时间周期大于第四时间周期的缺点在于,当已经过第四时间周期时,主机103可以重新传输数据-1,并且还可以减慢数据传输。因此,优选考虑如上文所论述的第一、第二和第三时间周期,并且忽略第四时间周期。这确保成功地传输数据,并且即使带宽消耗可能较高,也不减慢数据传输。在另一变型例中,可由通信路由器106和108的管理员调节等待时间周期。

[0098] 图7和图8中的时序图之间的差别在于,在图7中,当在通信路由器106处接收延迟查询消息时,在通信路由器106处激活误差校正模式。或者,在图8中,误差校正模式可以默认被激活,或可以由通信路由器106的用户或管理员激活。如果误差校正模式默认被激活,则会存在由误差校正包和数据包的重新发送引起的流量拥挤。这会减慢数据传输并且显著消耗高带宽。因此,仅当接收延迟查询消息时激活误差校正模式,以便保存带宽且减少流量拥挤可能非常有益。然而,默认激活误差校正模式会使重新发送数据包和误差校正包的过程更快,因为通信路由器106不需要等待通信路由器108发送延迟查询消息。当通信路由器106已知晓隧道的弃包率或丢包率高时,这可能非常有益。当弃包率或丢包率高时,通信路由器106重新发送数据包并发送误差校正包。

[0099] 在示例中,当主机103使用TCP通过通信路由器106和108将数据传输到主机105,并且多个包被丢弃或丢失时,主机103可以减小传输速率以便减小丢包和弃包。即使丢包的增加是临时的,这也可以使总体传输慢许多。此外,当主机103确定包已被丢弃或丢失时,所述主机将重新发送包。与当主机103重新发送包时相比,当通信路由器106重新发送包时,总体

传输可以更快。为了避免主机103减小传输速率以及重新发送包,通信路由器106发送同一包的多于一个的副本,例如数据-2和数据-2-1,并且还可以发送误差校正包,例如,数据-2-EC。当通信路由器106发送同一包的多于一个的副本时,冗余较高且更可能将在主机105处接收包。因此,确认包可以从主机105发送并且及时在主机103处接收,并且主机103将不需要减小传输速率或重新发送包。

[0100] 在优选实施例中,如果可能,通过聚合连接的不同隧道传输同一包的副本。例如,参考图7,如果通过隧道103A在流708中传输数据-2,则通过隧道103B在流710中传输数据-2-1且通过隧道103C在流711中传输数据-2-EC。这可以允许基本上在同一时间传输数据-2、数据-2-1和数据-2-EC。此外,如果使用由不同服务提供商提供的网络建立隧道103A、103B和103C,则数据-2、数据-2-1和数据-2-EC中的至少一个成功地到达通信路由器108的机率较高。这是因为隧道的网络质量可能不同。如果隧道103A的网络质量不令人满意或隧道103A被破坏或隧道103A正经受大量丢包,则数据-2不可以成功地到达通信路由器108。如果通过隧道103A发送数据-2,则数据-2-1可能经受与数据-2刚经受的网络问题相同的网络问题。因此,优选地通过可以经受更佳网络性能的不同隧道(例如,隧道103B)发送数据-2-1,因为隧道103B使用的路径、路由或连接可以不同于隧道103A使用的路径、路由或连接。不同隧道的使用更可能增加数据-2-1成功地到达通信路由器108的可能性。

[0101] 图9是说明根据本发明的一个实施例的在通信路由器106处执行的过程的流程图。在步骤901中,通信路由器106从主机103接收数据,其中数据去往主机105。在步骤902中,通信路由器106确定是否激活误差校正模式。如果激活误差校正模式,则在步骤903中通信路由器106发送原始数据和原始数据的误差校正数据两者。原始数据与最初从主机103接收到的数据相同。或者,如果未激活误差校正模式,则在步骤904中通信路由器仅发送原始数据。在步骤905中,过程结束。

[0102] 图10是说明根据本发明的一个实施例的在通信路由器108处执行的过程的流程图。在步骤1001中,通信路由器108从通信路由器106接收数据。在步骤1002中,通信路由器108随后确定是否及时接收所有数据。当在通信路由器106中激活误差校正模式时,所有数据包括主机103发送的原始数据以及源于通信路由器106的误差校正数据。如果及时接收所有数据,则在步骤1003中通信路由器108将恢复正常消息发送到通信路由器106。通信路由器108可以将原始数据转发到主机105。或者,如果未及时接收所有数据,则在步骤1004中通信路由器108将请求发送到通信路由器106,其中所述请求是重新发送数据。在步骤1005中,过程结束。

[0103] 在一个变型例中,如果误差校正模式已停用,则省略1003。通信路由器108确定是否激活误差校正模式。

[0104] 优选的是,在主机103确定尚未成功地传输数据之前,在步骤1004中通信路由器108发送用于重新发送数据的请求。当通信路由器106接收用于重新发送数据的请求时,所述通信路由器将数据的副本重新发送到通信路由器108。通信路由器106还可以将误差校正数据发送到通信路由器108。

[0105] 通信路由器106和108优选地具有大型高速缓冲存储器。根据本发明,通信路由器106必须基于主机103发送的且保存在通信路由器106的高速缓冲存储器中的包来创建和发送误差校正包以及包的副本。因此,通信路由器106可能需要将多个包保存在其高速缓冲存

存储器中。另外,多个管理消息,例如,恢复正常消息和延迟查询消息可能需要在通信路由器106与108之间交换。这些管理消息还可能需要保存在通信路由器106和108的高速缓冲存储器中。

[0106] 应了解,通信路由器108能够执行与通信路由器106相同的过程,且反之亦然。在以上描述中,为便于阅读,通信路由器106已描述为数据发送器且通信路由器108已描述为数据接收器。应注意,本发明的范围不限于仅主机103将数据发送到主机105,使得还可以发生从主机105到主机103的数据传输。数据传输还可以是双向的,使得主机103和主机105两者可以将数据发送到彼此。

[0107] 由于通信路由器需要存储数据用于重新发送,因此存储媒体用于存储数据。所需的存储量应能够将所传输数据保持几秒。优选地存储约20秒数据传输。例如,对于能够每秒发送100兆位的接入连接,存储量应约为20秒乘以每秒100兆字节且产生约250兆字节的存储量。本领域的技术人员将理解,存储量越大,则可以重新传输越多数据。

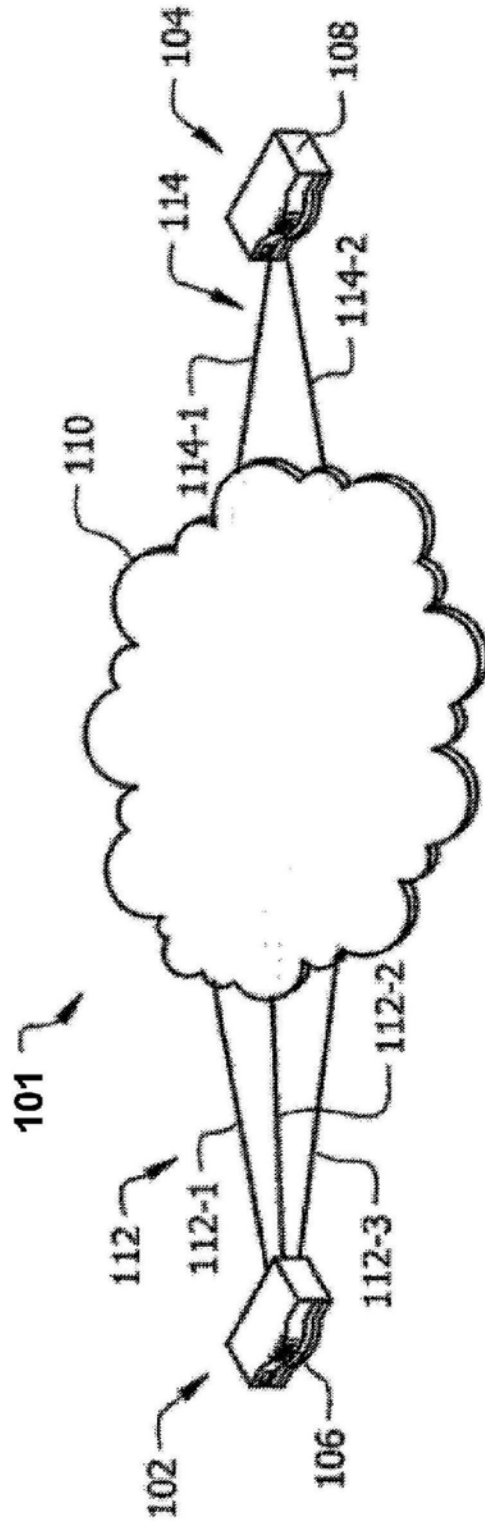


图1A

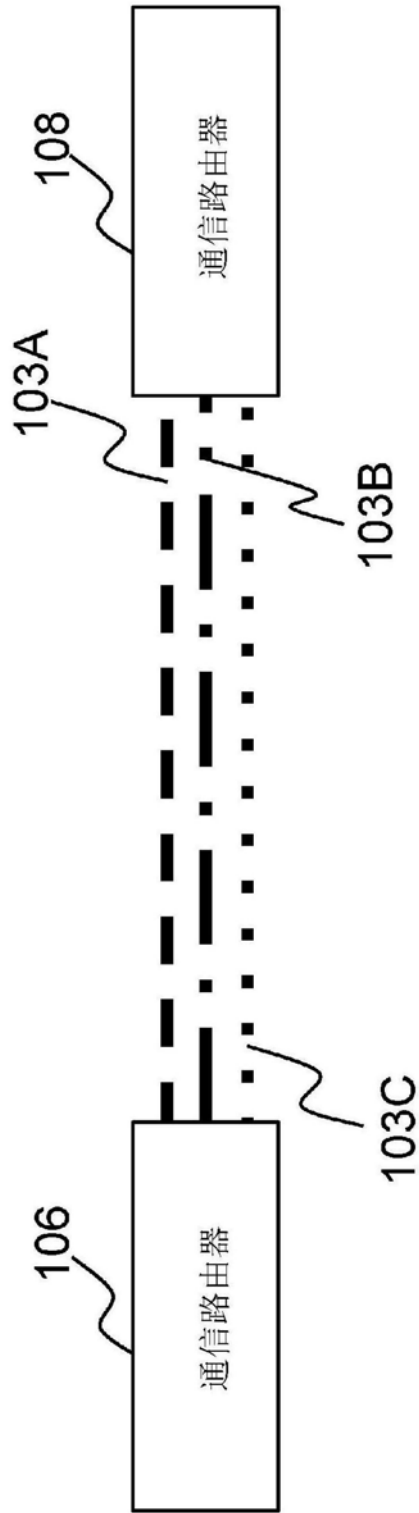


图1B

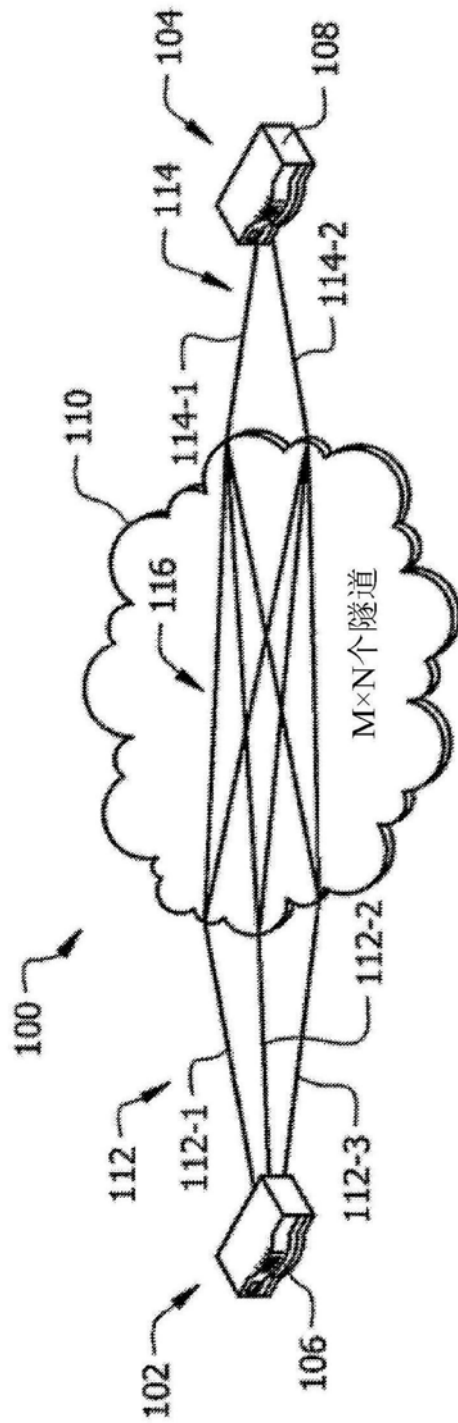


图1C

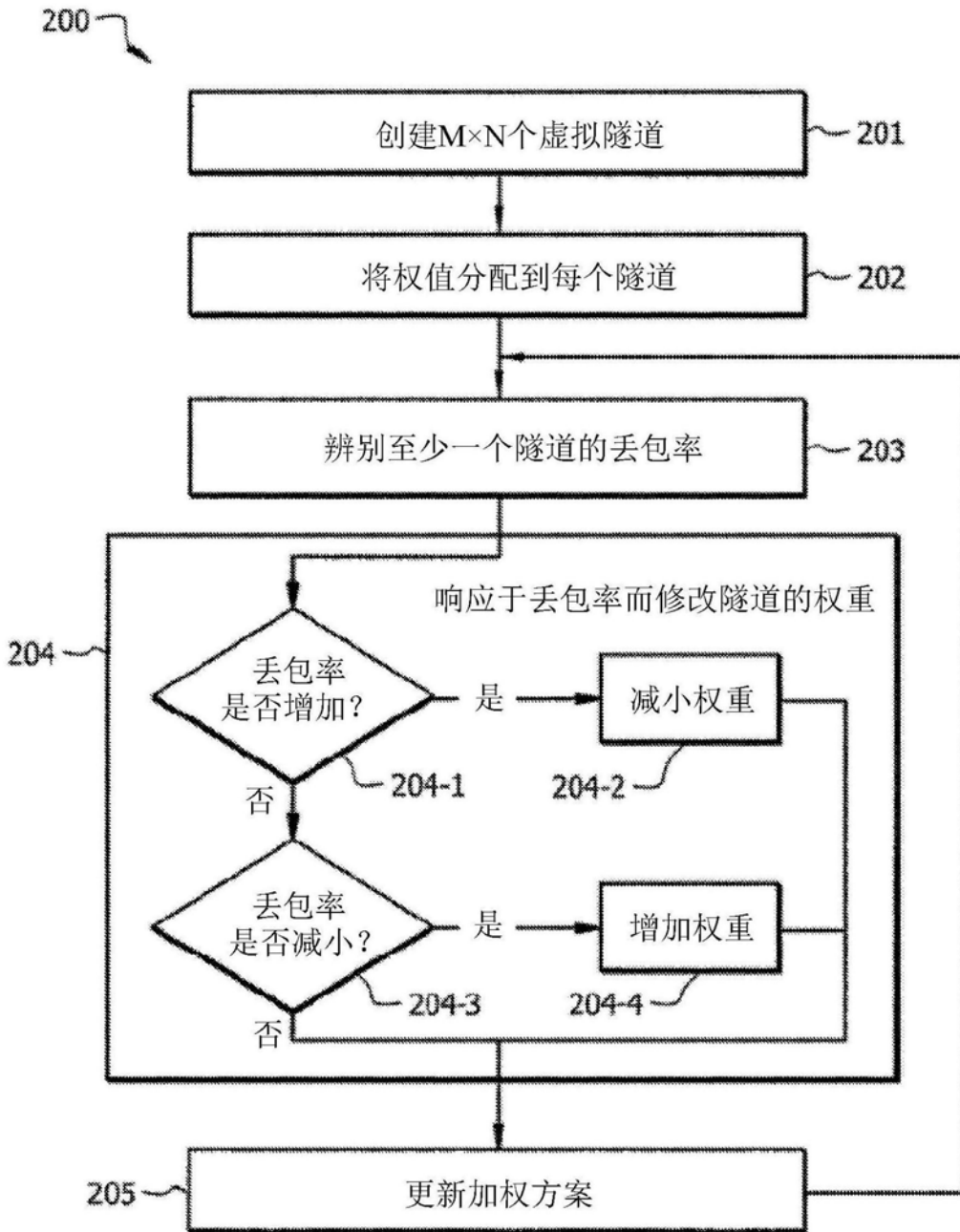


图2A

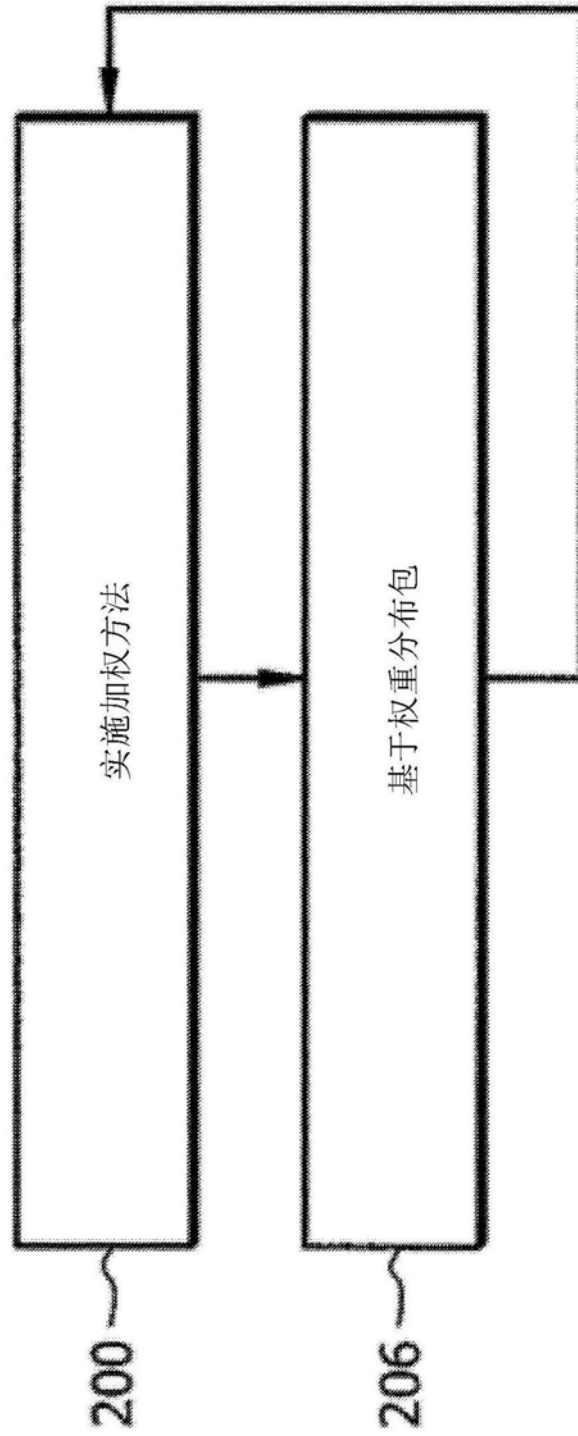


图2B

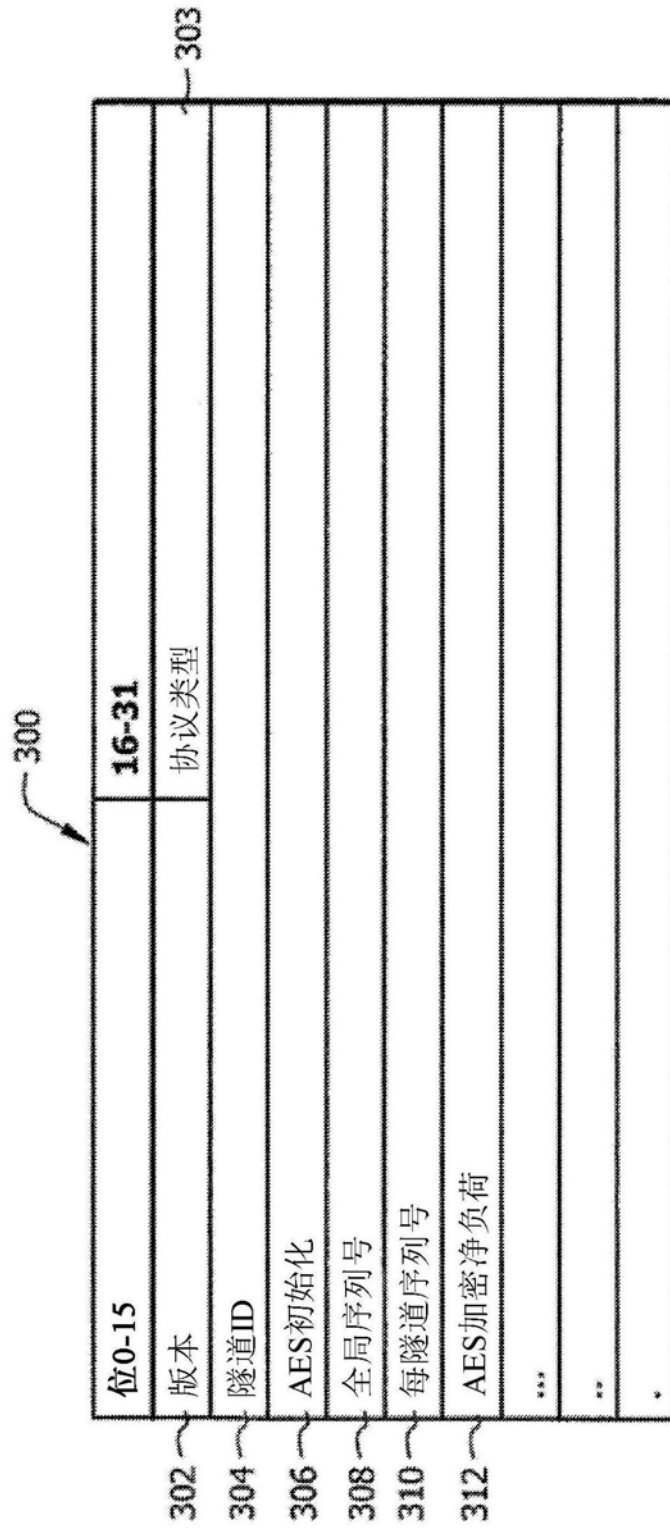


图3

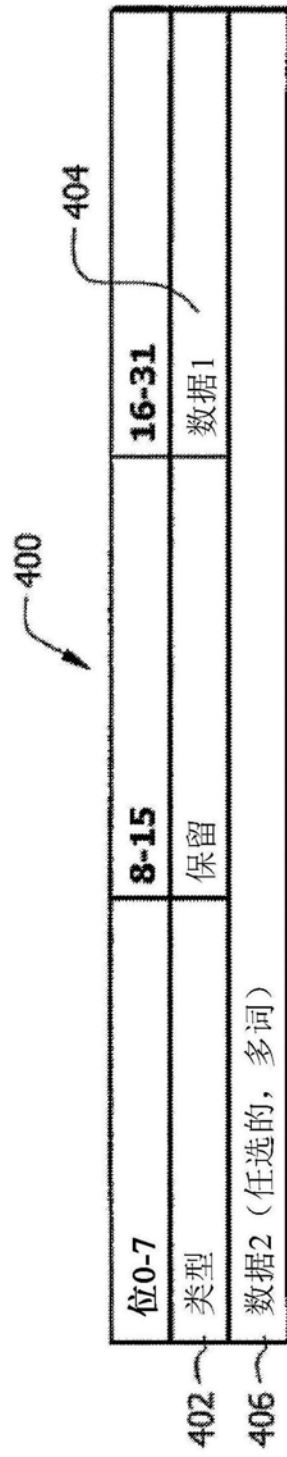


图4A

类型	数据1	数据2
BWOPT_VERSION (0x01)	此包的版本。目前为1。	零
BWOPT_MAX_WAN (0x02)	发送器上的WAN连接的最大数目	零
BWOPT_WAN_MAXBW (0x03)	数据字段中的32位词的数目	以每秒千字节为单位的WAN连接的最大下行带宽。每个WAN32位词。
BWOPT_DROP_RATE (0x04)	数据字段中的32位词的数目。此数目必须为2。	第一32位词是所丢弃包的数目。第二32位词是接收到的全部包。

图4B

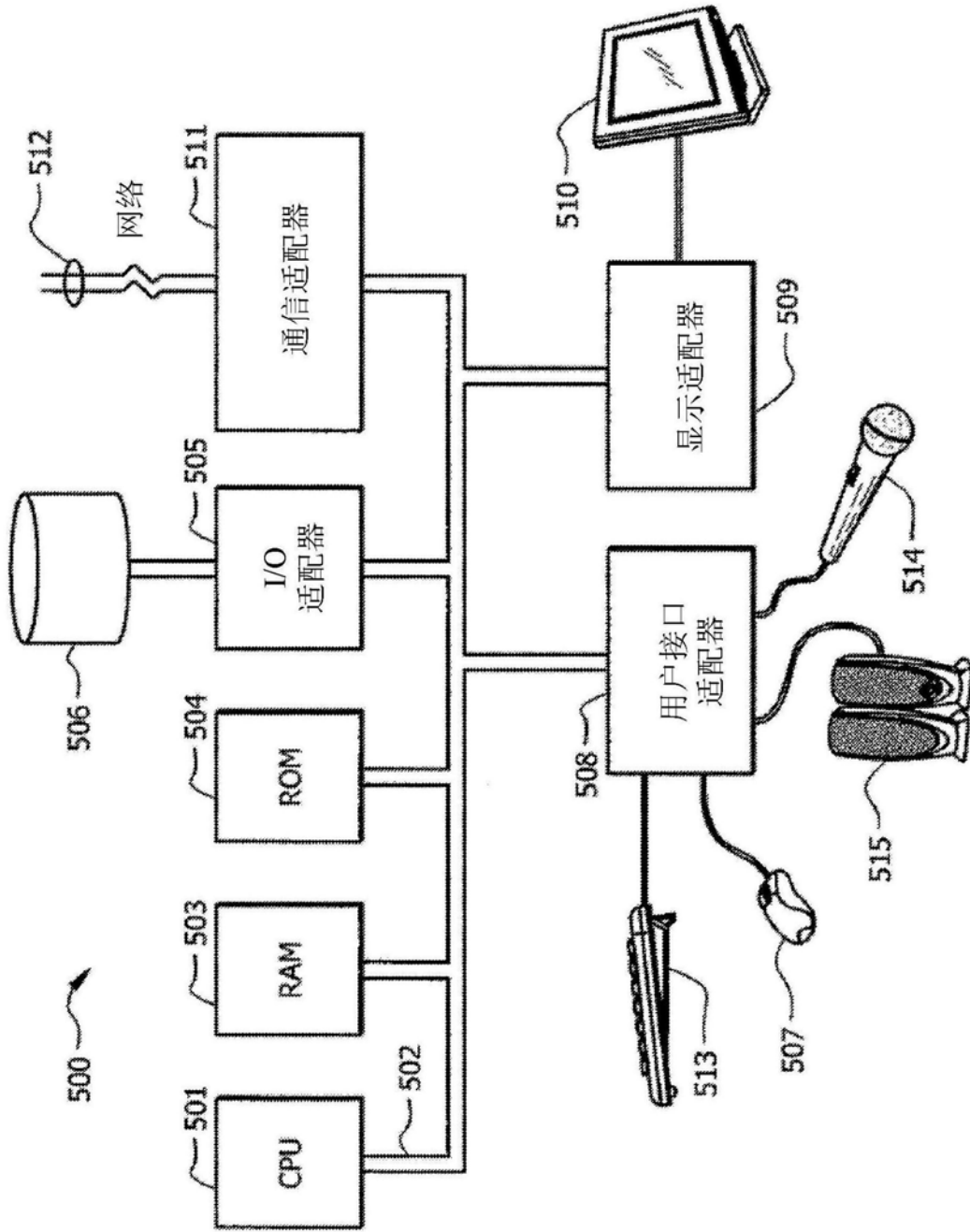


图5

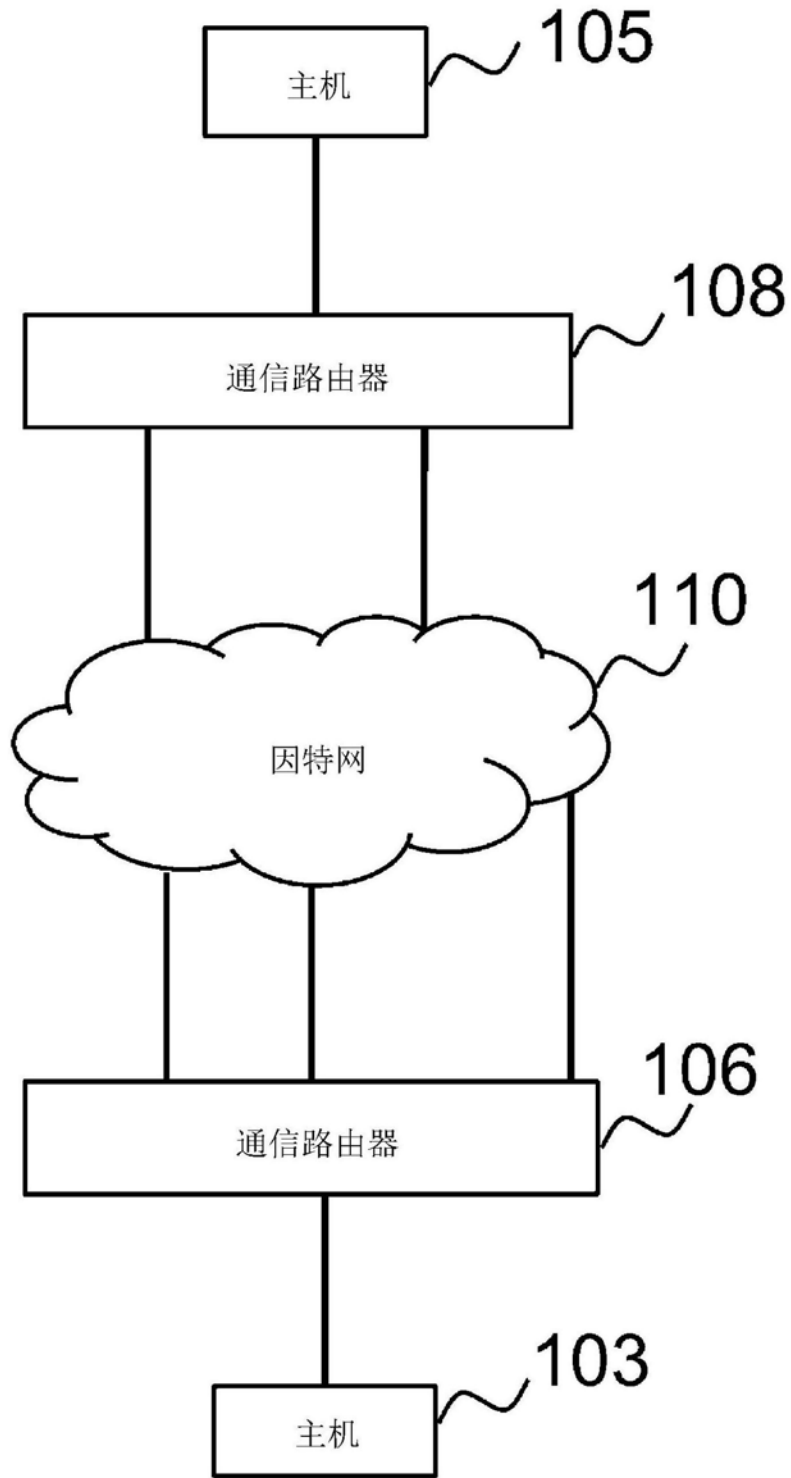


图6

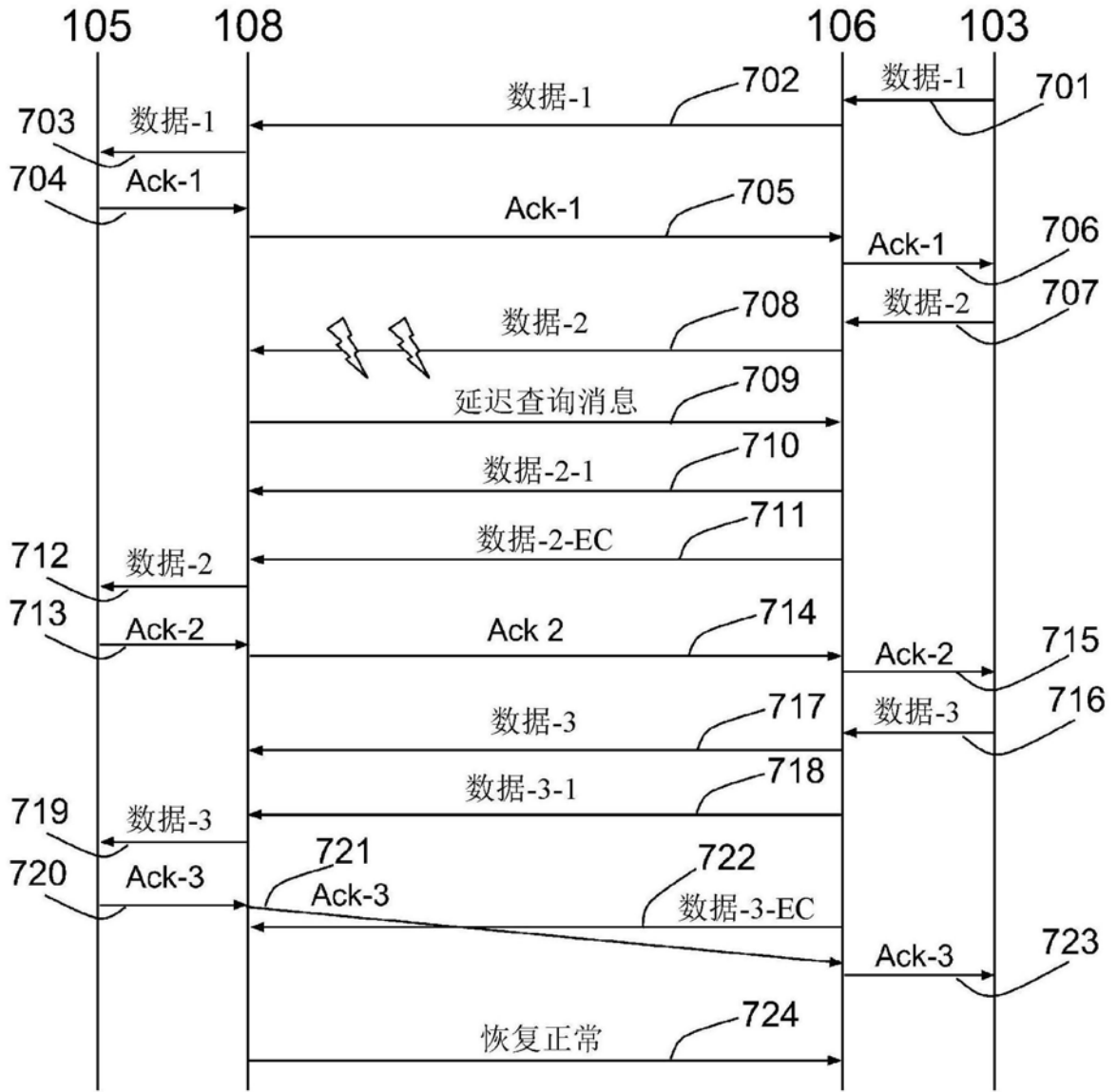


图7

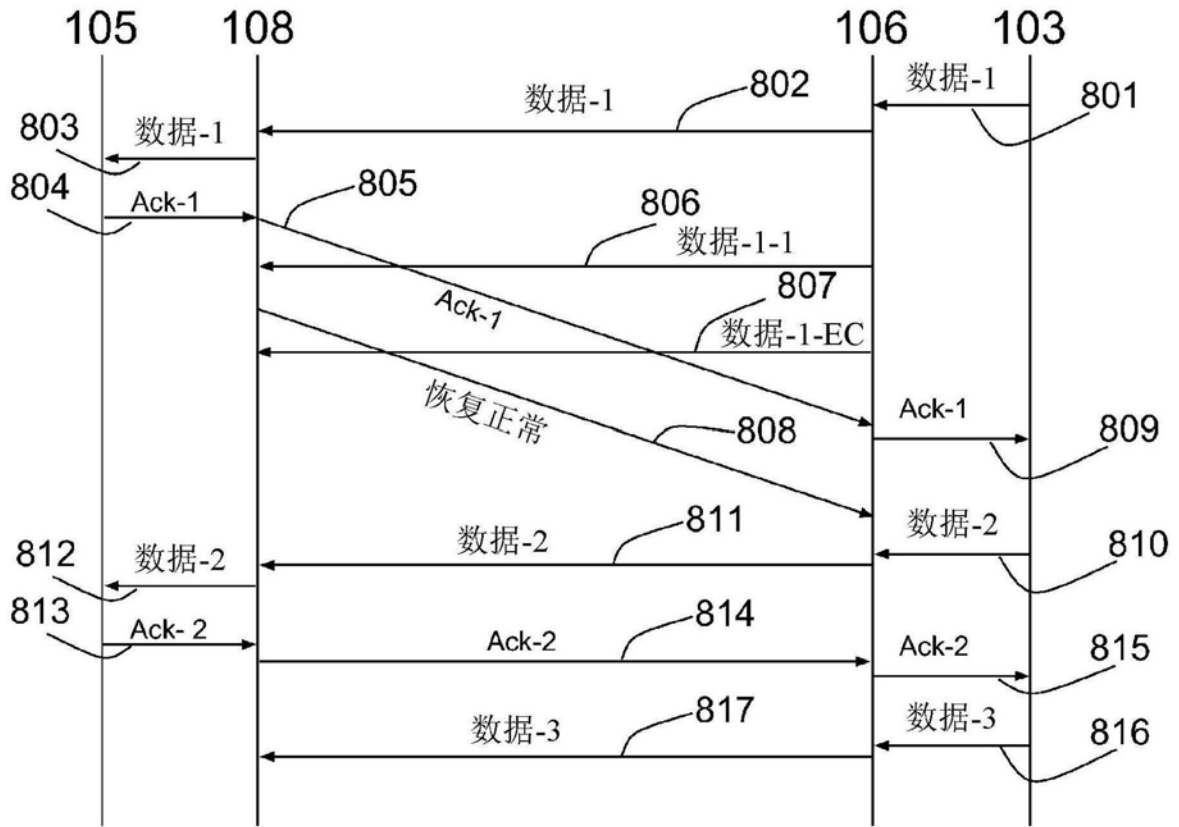


图8

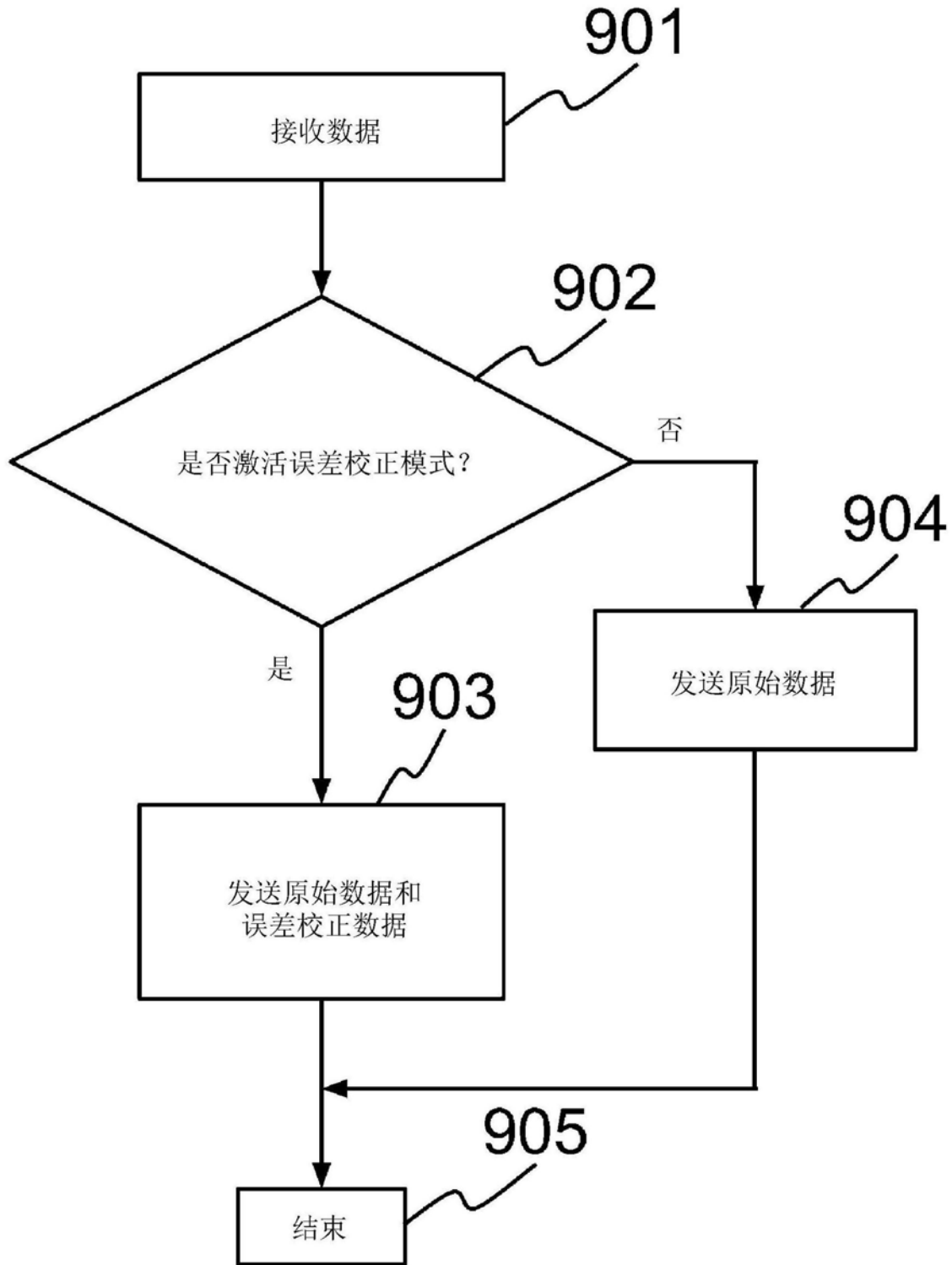


图9

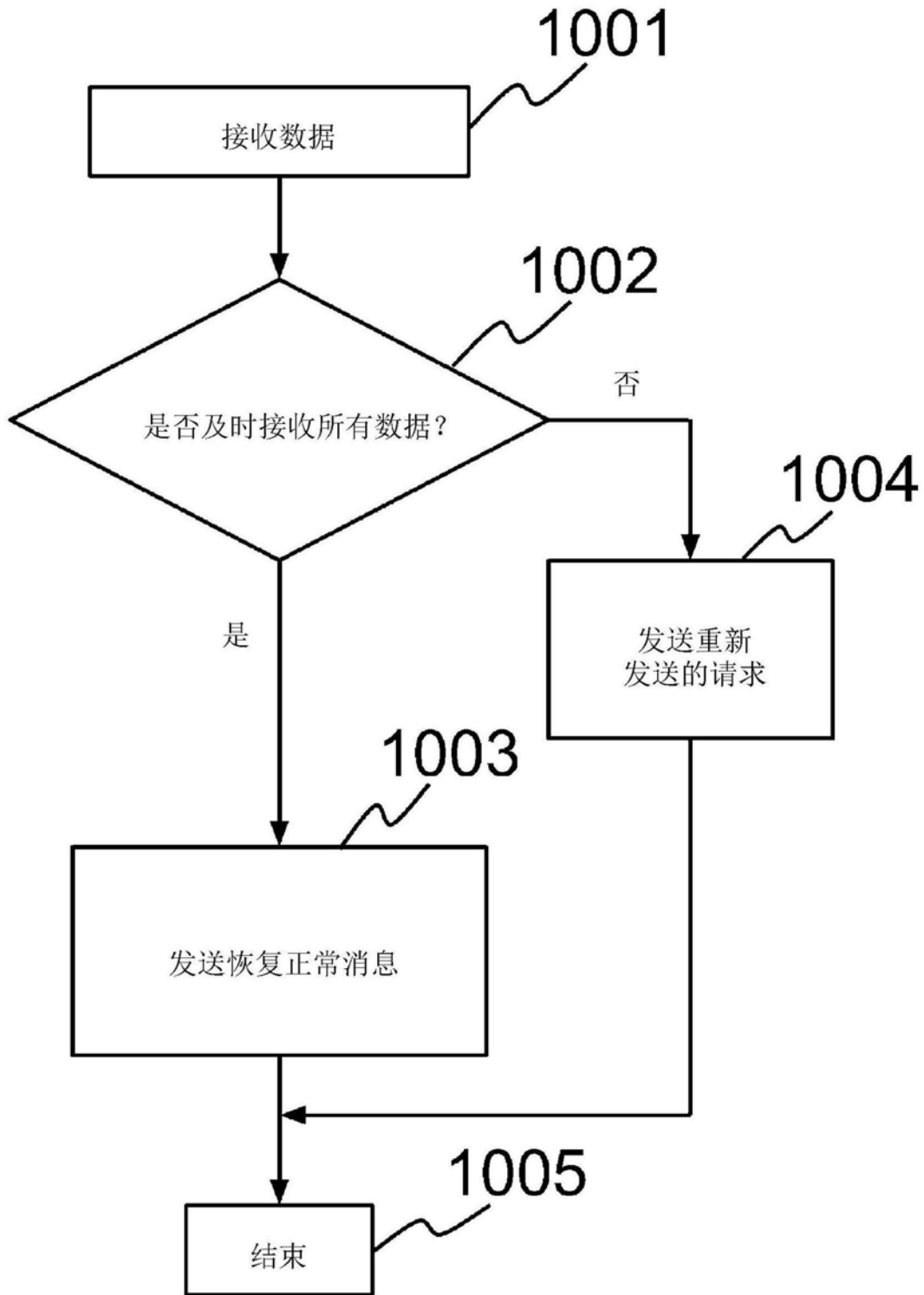


图10

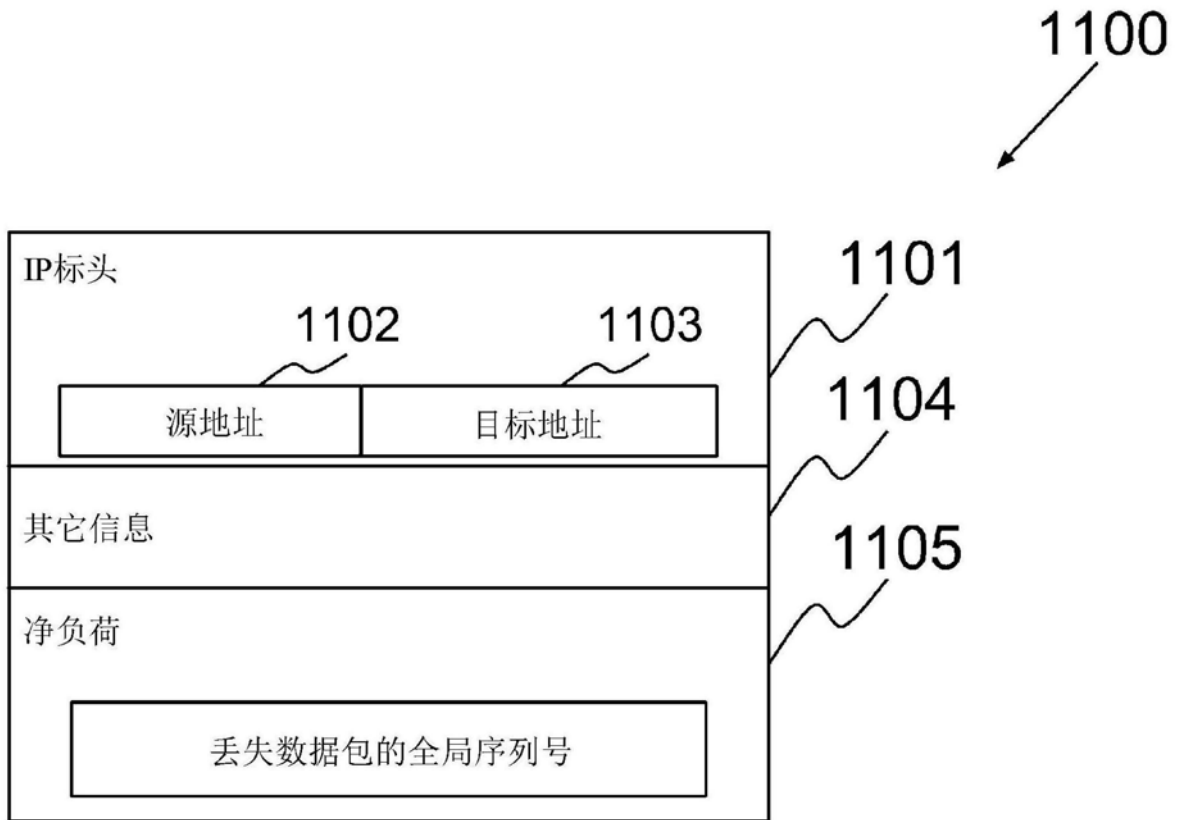


图11