



(12) 发明专利申请

(10) 申请公布号 CN 104144088 A

(43) 申请公布日 2014. 11. 12

(21) 申请号 201410355604. 8

(22) 申请日 2014. 07. 24

(71) 申请人 重庆邮电大学

地址 400065 重庆市南岸区黄桷垭崇文路 2 号

(72) 发明人 唐红 刘静娴 赵国锋 邓娅茹

王大瑞 戴俊 王新恒 唐文广

(74) 专利代理机构 重庆市恒信知识产权代理有限公司 50102

代理人 刘小红

(51) Int. Cl.

H04L 12/26 (2006. 01)

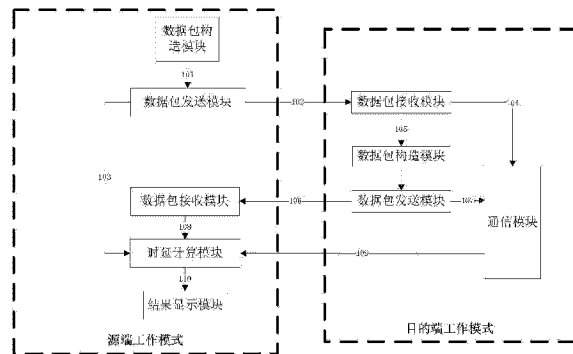
权利要求书1页 说明书5页 附图4页

(54) 发明名称

一种提高时延测量精度的网络时延测量方法

(57) 摘要

本发明公开了一种网络时延测量方法。该方法包括：时延测量过程中自动减去目的端设备的处理时间，使得时延测量值能更精确地反应当下网络运行的状况；通过 UDP 无连接的传输层协议发送测试数据包，解决有些情况下 ICMP 数据包被屏蔽的问题；针对 UDP 传输不可靠的问题利用 TCP 可靠传输来传递时间戳的值；在数据包到达网卡的时刻计时，将收发包时间戳记录位置从应用层移动到物理层，以便更加接近数据包离开和到达的真实时刻。能够去除目的端设备的处理时延，使得收发包时间戳更加接近真实收发包时刻，所测时延值可以更好的反应当下网络的运行状况。



1. 一种提高时延测量精度的网络时延测量方法,其特征在于,包括以下步骤:

101、当进行网络时延测量时,源端按照 UDP 协议构造测量数据包 I,然后将构造好的测量数据包 I 发送给目的端,并记录测量数据包 I 离开源端网卡的时间作为发送时间戳 T1,将该发送时间戳 T1 存放到源端 MAC 层的时延计算模块中;

102、当目的端接收到源端发送来的测量数据包 I 后,目的端将测量数据包 I 到达目的端网卡的时间作为接收时间戳 T2,并将 T2 存放在目的端 MAC 层的通信模块中;

103、目的端对测量数据包 I 进行响应,按照 UDP 协议构造新的测量数据包 II,并将新的测量数据包 II 发送给源端,将测量数据包 II 离开目的端网卡的时间记为发送时间戳 T3,将该时间戳 T3 存放在目的端 MAC 层的通信模块中;

104、当源端接收到新的测量数据包 II 后,将新的测量数据包 II 到达源端网卡的时间为接收时间戳 T4,并将 T4 传递到源端 MAC 层的时延计算模块;

105、同时目的端构造 TCP 数据包,将 T2 和 T3 作为数据传送到源端,当源端接收到 TCP 数据包后,得到时间戳 T2 和 T3,源端 MAC 层的时延计算模块结合本地存储的 T1 和 T4,计算往返时延值  $RTT = (T2-T1)+(T4-T3) = (T4-T1)-(T3-T2)$ ,并显示网络时延测量结果值。

2. 根据权利要求 1 所述的提高时延测量精度的网络时延测量方法,其特征在于:步骤 101 中所述的测量数据包 I 为 UDP 数据包。

## 一种提高时延测量精度的网络时延测量方法

### 技术领域

[0001] 本发明涉及网络时延测量技术领域,具体是一种通过减去目的端设备的处理时延来提高时延测量精度的网络时延测量方法。

### 背景技术

[0002] 随着时延是评价网络性能的基本指标,同时也是测量时延抖动、网络带宽等性能指标的基础。时延测量在网络性能监测、网络行为分析、网络应用设计等领域有着广泛的应用。随着网络的发展,特别是光纤的诞生应用,网络带宽已经达到 10Gbps,桌面也有很多采用百兆以太网或者 ADSL 等宽带接入,时延变得很小,常用工具的测量精度已经不能满足测量分析的需求。随着网络速率的提升,对高精度的时延测量方法进行研究非常有必要。

[0003] 网络性能测量中,往返时延常常被用作反映网络状态的一个重要指标。往返时延可以看做由三部分组成:源端到目的端的传输时延,目的端设备的处理时延,目的端到源端的传输时延。在这三部分中,目的端设备的处理时延并非是数据包在网络中传输的时间,并且这段时间也会随着目的端硬件设备的不同而不同,这就导致测量得到的时延值也不同。事实上,这段时延只是目的端设备的处理时间,并不是数据包在网络中经过的时间,不能反映当下网络的运行状况,所以将目的端设备的处理时延略去有利于时延更准确的反应网络状况。然而最常用的时延测量工具 PING 为了方便起见并没有将这一段时延去掉,这就导致测量得到的时延值不准确。此外,PING 采用 ICMP 协议,当时延测量的目的主机或者是路径中其它节点过滤 ICMP 数据包后,PING 将不能进行测量。针对这一问题可以采用 UDP 数据包来测量往返时延。数据包在发送和接收的过程中存在缓存,缓存的时间与主机设备以及主机当下的负载状况有关。网卡会产生中断来通知驱动程序对数据包进行处理,对中断的响应所需的时间也是随着主机设备以及主机当下的负载情况有关。由于数据包缓存延迟以及终端响应时间延迟不确定等因素的影响,记录数据包收发时间戳的精度不够高。此外,想要在时延计算过程中减去目的端设备处理时延的影响就要实现源端和目的端之间时间戳数据的通信,即将目的端记录的两个时间戳传送到源端。UDP 协议是面向无连接的,采用 UDP 报文来传输时间戳数据信息不是最佳选择。如果直接将 T2 和 T3 作为数据用测量包传输到源端,那么该情况下目的端数据包发送时间戳 T3 必须是在数据包构造之前就获取,这就会导致获取的时间戳距离真实的发包时刻相差更远。上述问题在时延测量的过程中均会引起误差,影响时延对网络状态的真实反应。

### 发明内容

[0004] 针对以上现有技术中的不足,本发明的目的在于提供一种提高时延测量精度的网络时延测量方法。本发明的技术方案如下:一种提高时延测量精度的网络时延测量方法,其包括以下步骤:

[0005] 101、当进行网络时延测量时,源端按照 UDP 协议构造测量数据包 I,然后将构造好的测量数据包 I 发送给目的端,并记录测量数据包 I 离开源端网卡的时间作为发送时间

戳 T1, 将该发送时间戳 T1 存放到源端 MAC 层的时延计算模块中;

[0006] 102, 当目的端接收到源端发送来的测量数据包 I 后, 目的端记录测量数据包 I 到达目的端网卡的时间作为接收时间戳 T2, 并将 T2 存放在目的端 MAC 层的通信模块中;

[0007] 103、目的端对测量数据包 I 进行响应, 按照 UDP 协议构造新的测量数据包 II, 并将新的测量数据包 II 发送给源端, 将测量数据包 II 离开目的端网卡的时间记为发送时间戳 T3, 将该时间戳 T3 存放在目的端 MAC 层的通信模块中;

[0008] 104、当源端接收到新的测量数据包 II 后, 记录新的测量数据包 II 到达源端网卡的时间为接收时间戳 T4, 并将 T4 传递到源端 MAC 层的时延计算模块;

[0009] 105, 同时目的端构造 TCP 数据包, 将 T2 和 T3 作为数据传送到源端; , 当源端接收到 TCP 数据包后, 得到时间戳 T2 和 T3, 源端 MAC 层的时延计算模块结合本地存储的 T1 和 T4, 计算往返时延值  $RTT = (T2-T1)+(T4-T3) = (T4-T1)-(T3-T2)$ , 并显示网络时延测量结果值。

[0010] 进一步的, 步骤 101 中所述的测量数据包 I 为 UDP 数据包。

[0011] 本发明的优点及有益效果如下:

[0012] 本发明网络时延测量方法通过去除目的端设备处理时延可以更精确的获得时延测量值, 更好地通过时延反应当下网络运行状况, 避免因服务端设备性能差异而带来的误差; 通过 TCP 可靠传输解决目的端测量时延值传送到源端的问题; 通过将时间戳记录位置由应用程序转移到网卡位置来获得更准确的时间戳。该方案所测的的网络时延值将更加精确。

## 附图说明

[0013] 图 1 所示为本发明优选实施例系统流程示意图;

[0014] 图 2 数据包接收过程中可以添加时间戳位置的示意图;

[0015] 图 3 本发明采用的时延计算方法图;

[0016] 图 4 数据包发送和接收过程;

[0017] 图 5 时间戳传递原理图;

[0018] 图 6 是传输时间戳数据的 TCP 报文格式;

[0019] 图 7 是 UDP 测量数据报格式。

## 具体实施方式

[0020] 下面结合附图给出一个非限定性的实施例对本发明作进一步的阐述。

[0021] 下面结合附图给出一个非限定的实施例对本发明作进一步的阐述。但是应该理解, 这些描述只是示例的, 而非非要限制本发明的范围。此外, 在以下说明中, 省略了对公知结构和技术的描述, 以避免不必要地混淆本发明的概念。为了达到目的, 本发明的技术方案如下:

[0022] 1) 为了保证时延值是网络状态的真实反映, 减去目的端设备的处理时延。

[0023] 2) 为了解决 ICMP 协议被目的端主机或者路径中的中间节点过滤掉的情况, 采用 UDP 报文作为测量数据包。

[0024] 3) 采用面向连接的 TCP 协议, 将目的端测量得到的时间戳数据通过 TCP 可靠传输

传递到源端。

[0025] 4) 时间戳记录的真实性影响时延值的准确性,将时间戳的记录位置从应用层转移到物理层网卡的位置。

[0026] 下面对本发明进行进一步说明:

[0027] 1) 为了保证时延值是网络状态的真实反映,减去目的端设备的处理时延。从源端发送 UDP 数据包到测试的目的端,目的端接收到该测试数据包后向源端发送 UDP 数据包。分别按照顺序在源端和目的端记录四个时间戳:源端测量数据包发送时记录时间戳 T1;目的端接收测量数据包时记录时间戳 T2;目的端发送测量数据包时记录时间戳 T3;源端接收到测量数据包时记录时间戳 T4。往返时延值  $RTT = (T2-T1)+(T4-T3) = (T4-T1)-(T3-T2)$ 。

[0028] 2) 为了解决 ICMP 被目的端主机或者路径中的中间节点过滤掉的情况,采用 UDP 报文作为测量数据包。UDP 和 ICMP 较为类似,都是面向无连接的,区别是 UDP 需要指定一个端口。

[0029] 3) 想要实现源端和目的端之间的测量数据的传输,直接将测量得到的数据放在测试数据包 UDP 中传递不是最佳方案,因为 UDP 协议是面向无连接的。此外,如果直接将 T2 和 T3 作为数据用测量包传输到源端,那么该情况下目的端数据包发送时间戳 T3 必须是在数据包构造之前就获取,这就会导致获取的时间戳距离真实的发包时刻相差更远。针对这两个问题的解决方案是采用 TCP 来传送 T2 和 T3,在保证数据可靠传输的同时也避免了因需要传送 T2 和 T3 而带来的发包时间戳的误差。在源端记录时间戳 T1 和 T4,在目的端记录时间戳 T2 和 T3,将 T2 和 T3 作为数据利用 TCP 协议从目的端传送到源端,以供计算往返时延。

[0030] 4) 时间戳记录的真实性影响时延值的准确性。下面以接收数据包为例来进行说明。以太网上的物理帧被网卡芯片获取之后,网卡将数据传递到主机内存,同时通过中断通知驱动程序处理接收到的数据包。接着主机响应中断,将数据包拷贝到上层协议栈中进行处理,处理后的数据包到达应用层。在不同位置插入时间戳所受缓存和中断响应时间带来的延迟的影响是不同的,这会导致时间戳与数据包到达的真实时间偏差也不同。为了数据包的接收时间戳更加接近测试数据包接收的真实时间,将时间戳的计时位置从应用层转移到物理层网卡的位置。在数据包到达网卡时刻记录数据包接收时间。

[0031] 图 1 是本发明时延测量原理,包括如下步骤:

[0032] 步骤 101,源端构造测量数据包;

[0033] 步骤 102,源端发送测量数据包,记录数据包离开网卡的时间作为发送时间戳 T1;

[0034] 步骤 103,将该时间戳 T1 存放在时延计算模块中;

[0035] 步骤 104,目的端接受测量数据包,记录数据包到达网卡的时间作为接收时间戳 T2,并将 T2 存放在通信模块中;

[0036] 步骤 105,目的端对测量数据包进行响应,构造新的测量数据包;

[0037] 步骤 106,目的端发送测量数据包,记录数据包离开网卡的时间作为发送时间戳 T3;

[0038] 步骤 107,将该时间戳 T3 存放在通信模块中;

[0039] 步骤 108,源端接收测量数据包,记录数据包到达网卡的时间作为接收时间戳 T4,并将 T4 传递到时延计算模块;

[0040] 步骤 109,目的端构造 TCP 数据包,将 T2 和 T3 作为数据传送到源端;

[0041] 步骤 110,源端接收 TCP 数据包,得到时间戳 T2 和 T3,结合本地存储的 T1 和 T4 计算往返时延值并显示结果。

[0042] 图 2 是数据包接收过程中可以添加时间戳位置的示意图。

[0043] 影响数据包收发处理时间的因素主要为几下几点:中断服务开销;数据拷贝和内存分配操作;在内核协议栈上多余的协议处理时间。每个数据包的响应时间是在内核里的传输时间和用户空间程序响应时间之和。内核里的传输时间包括硬终端延时、中断服务开销、中断和内核的上下文切换时间、协议栈处理的时间以及内核和用户空间的上下文切换时间。数据包在用户空间的处理时间不仅取决于程序本身的效率,还与 LINUX 内核行为有关。如图 2 所示,加入时间戳的位置可以有多种选择,记录的时间戳分别为用户空间、内核空间、硬件时间戳。最常用的时延测量工具 PING 采用的是用户空间时间戳,如图 2 中 t1、t2 位置。内核空间网卡驱动是通过软件可以控制的最接近数据包收发的真实时刻,见图 2 中 t3、t4 位置。硬件时间戳须修改硬件,虽然实现较为复杂,但是最接近数据包收发的真是时刻。本方案将时间戳记录位置移动到 MAC 层,即图 2 中 t5、t6 位置,以便更接近数据包到达网卡的真实时间。

[0044] 图 3 是本发明实施例的数据包发送和接收过程示意图,包括如下步骤:

[0045] 步骤 301, Linux 网卡驱动程序为 IP 数据包添加 14 字节的 MAC 包头,构成 MAC 数据包,主机的 CPU 首先检查网卡是否有足够大的发送缓冲区,用于存放待发送的数据包;

[0046] 步骤 302,主机 CPU 将内存缓冲区中的 MAC 数据包拷贝到网卡的发送缓冲区内;

[0047] 步骤 303,网卡芯片将其发送缓冲区内 MAC 数据包封装成物理帧,并添加头部信息和 CRC 校验码,然后将 MAC 数据包传到以太网上并记录此时的时间戳。等到 MAC 数据包传到以太网后,网卡设备产生一个发送中断,随后进入中断处理程序,清除缓冲区里的数据包;

[0048] 步骤 304,来自以太网的帧会被网卡芯片获取网卡芯片检测物理帧的 CRC(Cyclical Redundancy Check),从而保证物理帧的完整。检验完成后,物理帧的头部会被去掉,从而得到 MAC 包。网卡芯片对比 MAC 包内的目的 MAC 地址是否与本网卡 MAC 相同,不同则丢弃,若相同则记录时间戳,然后将 MAC 帧拷贝到网卡上接收缓冲区中;

[0049] 步骤 305, MAC 帧接受完后,网卡会产生接收中断,通知主机有数据包被接收, CPU 进入接收中断处理;

[0050] 步骤 306, CPU 会在内存中分配块缓冲区,来存放数据包;

[0051] 步骤 307,网卡驱动通过中断将网卡缓冲区中的数据包拷贝到系统分配的内存中的缓冲区中;

[0052] 步骤 308,协议帧会去掉缓冲区传来的 MAC 包头,得到数据包。

[0053] 图 4 是本发明时延计算原理图图。将目的端对数据包进行处理的时间去掉,使得测量所得时延能更好的反应当下网络的运行状况,同时避免目的端设备差异带来的影响。计算公式为:往返时延值  $RTT = (T2-T1)+(T4-T3) = (T4-T1)-(T3-T2)$ 。具体步骤如下:

[0054] 步骤 401,源端向目的端发送测量数据包;

[0055] 步骤 402,目的端接收到数据包后对源端进行响应,向源端发送测量数据包。

[0056] 图 5 是本发明时间戳数据传输原理图。首先,源端建立 TCP 三次握手协议与目的

端建立连接；其次，发送UDP测量数据包；目的端接受后记录T2，并返回相应UDP包记录T3；同时TCP将记录结果T2和T3传递到源端。每一次测量过程都是上述过程的重复，所有测量包发送结束后，TCP四次挥手结束连接。

[0057] 图6是传输时间戳数据的TCP报文格式，TCP数据被封装在一个IP数据包中，在构造TCP数据包时将T2和T3填充到TCP数据部分。

[0058] 图7是UDP测量数据报格式，用户数据报协议在IP数据报服务至上增加了端口功能和差错功能。UDP报文中有两个部分：数据部分和首部，首部部分由四个字段组成：源端口号、目的端口号、长度、检验和。

[0059] 以上这些实施例应理解为仅用于说明本发明而不用于限制本发明的保护范围。在阅读了本发明的记载的内容之后，技术人员可以对本发明作各种改动或修改，这些等效变化和修饰同样落入本发明方法权利要求所限定的范围。

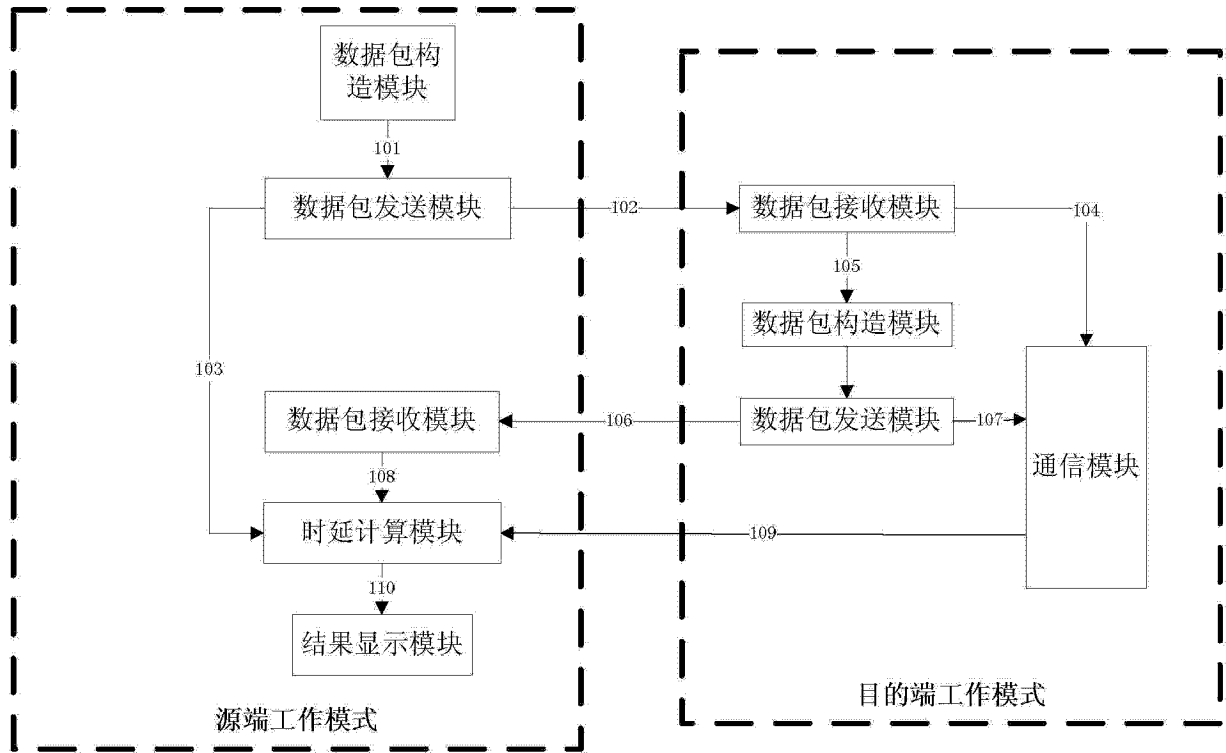


图 1

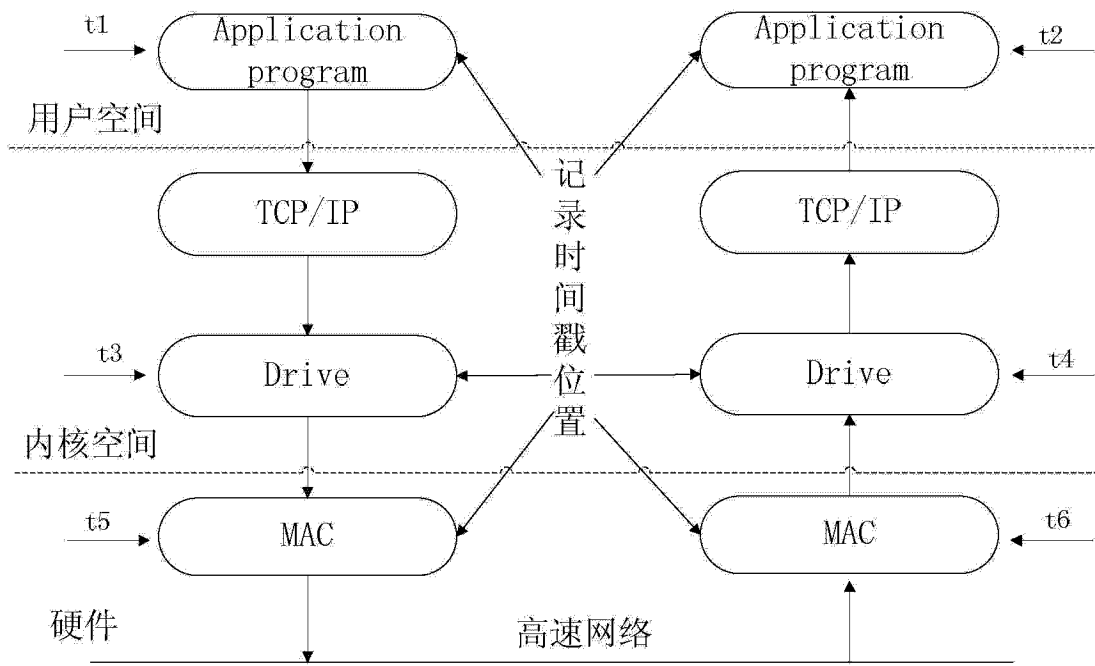


图 2



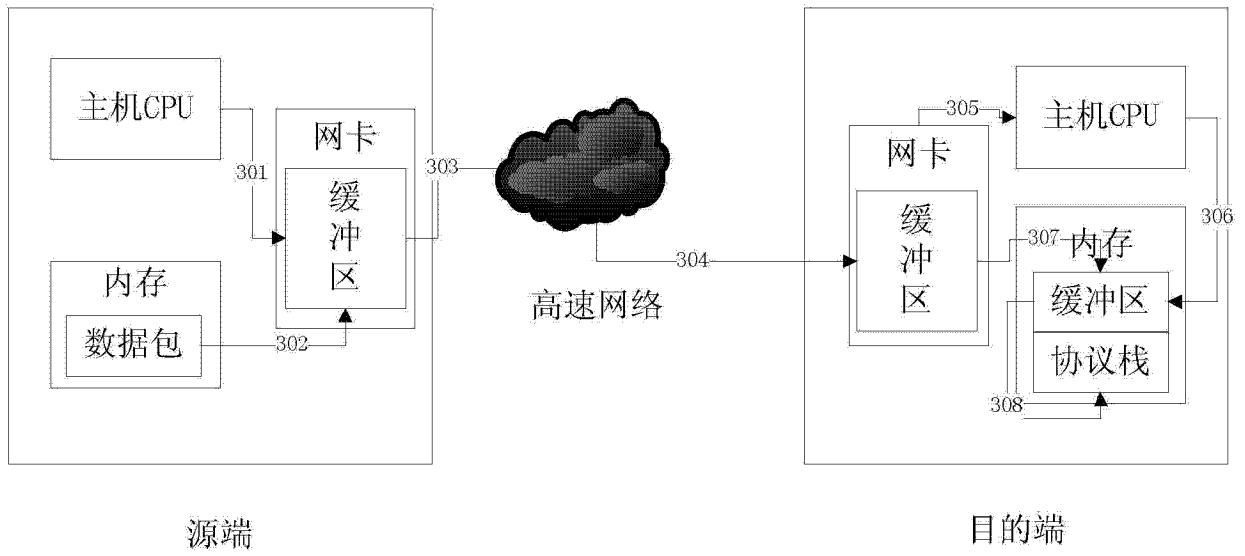


图 3

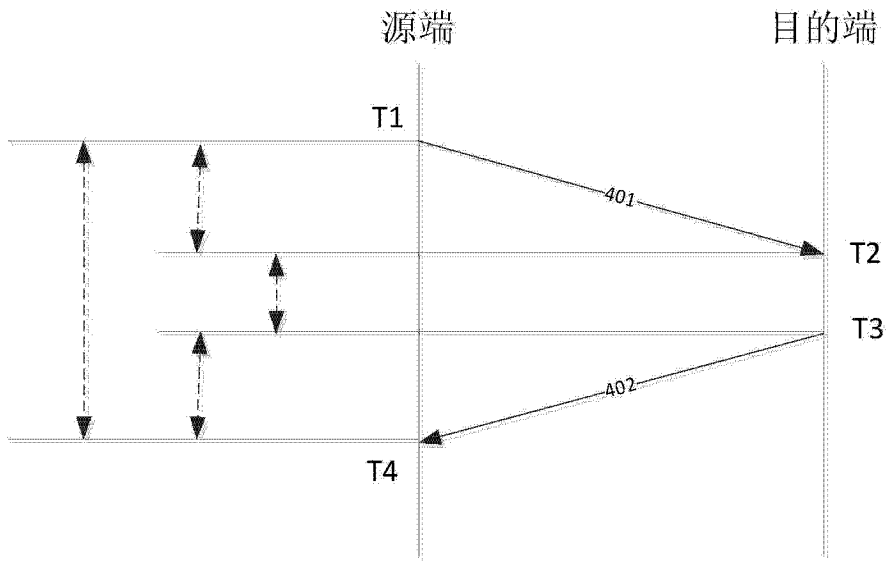


图 4

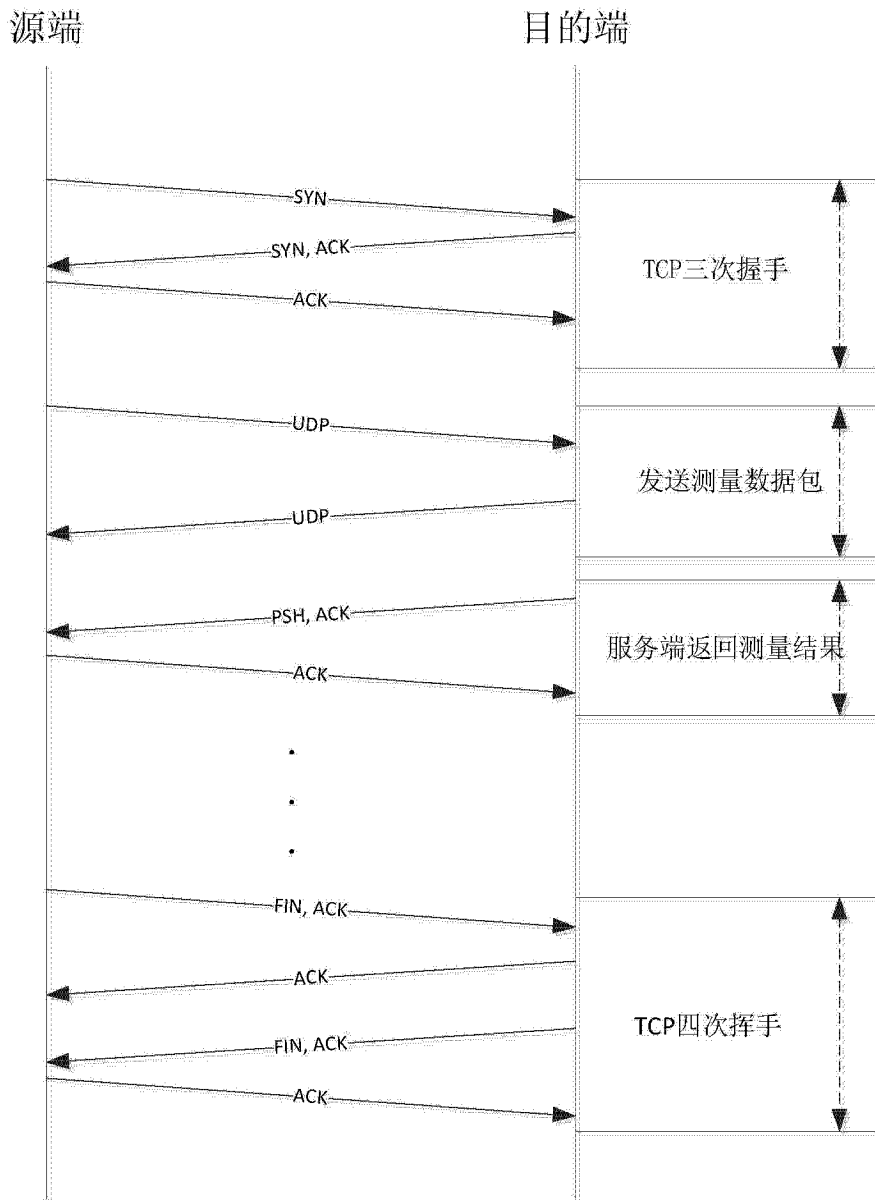


图 5



图 6

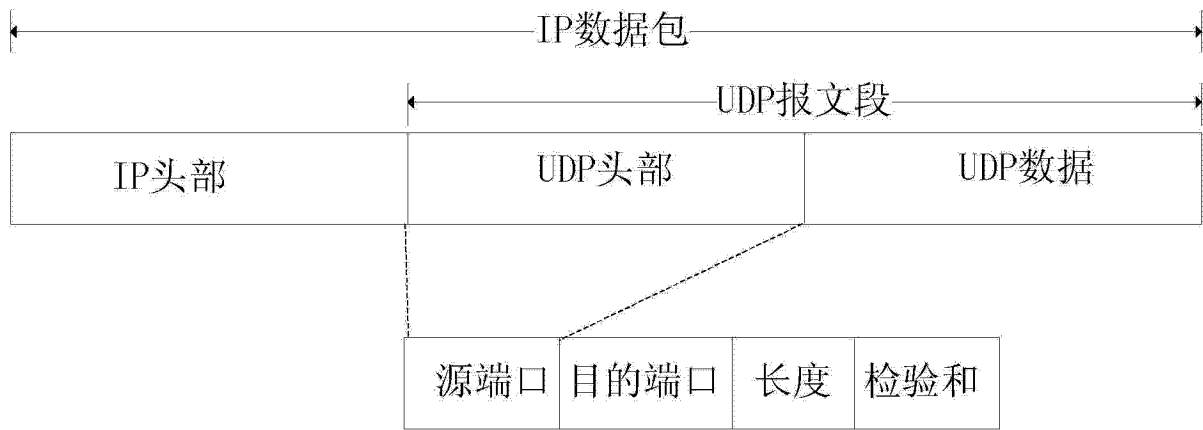


图 7