



(12) 发明专利申请

(10) 申请公布号 CN 103944954 A

(43) 申请公布日 2014. 07. 23

(21) 申请号 201410029889. 6

(22) 申请日 2014. 01. 22

(30) 优先权数据

13/747, 545 2013. 01. 23 US

(71) 申请人 A10 网络股份有限公司

地址 美国加利福尼亚州

(72) 发明人 L·韩 Z·曹

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 张东梅

(51) Int. Cl.

H04L 29/08 (2006. 01)

H04L 29/06 (2006. 01)

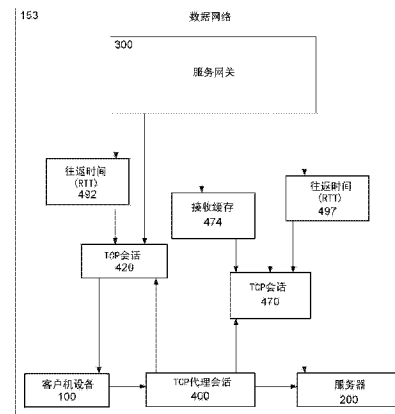
权利要求书2页 说明书6页 附图2页

(54) 发明名称

基于延迟的确认减少对 TCP 代理会话的缓存使用

(57) 摘要

通过服务网关减少客户机和服务器之间的 TCP 代理会话的缓存使用包括：确定服务器侧 TCP 会话的第一往返时间 (RTT)，并确定客户机侧 TCP 会话的第二 RTT；将第一 RTT 与第二 RTT 进行比较；确定第二 RTT 是否超出第一 RTT 一阈值；若是，则基于第二 RTT 计算要求的 RTT；根据计算出的要求 RTT 设定定时器，其中服务器侧 TCP 会话的 TCP 确认被延迟直到定时器到期为止。要求的 RTT 可被计算为第二 RTT 的百分比或第二 RTT 减去一预定值。服务网关在将 TCP 确认数据分组发送至服务器之前等待直到定时器到期为止。



1. 一种减少在客户机和服务器之间的传输控制协议 (TCP) 代理会话的缓存使用的方法,所述方法包括:

(a) 确定服务网关和服务器之间的 TCP 代理会话的服务器侧 TCP 会话的第一往返时间 (RTT),并确定服务网关和客户机之间的 TCP 代理会话的客户机侧 TCP 会话的第二 RTT;

(b) 通过服务网关将第一 RTT 与第二 RTT 进行比较;

(c) 确定第二 RTT 是否超出第一 RTT;

(d) 响应于确定第二 RTT 超出第一 RTT,基于第二 RTT 通过服务网关计算要求的 RTT;以及

(e) 根据计算出的要求 RTT 通过服务网关设定定时器,其中服务器侧 TCP 会话的 TCP 确认被延迟直到定时器到期为止。

2. 如权利要求 1 所述的方法,其特征在于,所述确定 (c) 和计算 (d) 包括:

c1) 确定第二 RTT 是否超出第一 RTT 一预定阈值;以及

d1) 响应于确定第二 RTT 超出第一 RTT 所述预定阈值,基于第二 RTT 通过服务网关计算要求的 RTT。

3. 如权利要求 1 所述的方法,其特征在于,所述计算 (d) 包括:

d1) 通过服务网关将要求的 RTT 计算为第二 RTT 的百分比。

4. 如权利要求 1 所述的方法,其特征在于,计算 (d) 包括:

(d1) 通过网关服务器将要求的 RTT 计算为第二 RTT 减去一预定值。

5. 如权利要求 1 所述的方法,其特征在于,所述设定 (e) 包括:

(e1) 通过服务网关在服务器侧 TCP 会话上从服务器接收数据分组;

(e2) 通过服务网关确定将 TCP 确认发送至服务器的需要;

(e3) 通过服务网关将定时器设定至要求的 RTT;以及

(e4) 响应于定时器的到期,通过服务网关将 TCP 确认发送至服务器。

6. 一种具有嵌入在内的计算机可读程序代码以减少客户机与服务器之间的传输控制协议 (TCP) 代理会话的缓存使用的非临时计算机可读介质,所述计算机可读程序代码被配置成:

(a) 确定服务网关和服务器之间的 TCP 代理会话的服务器侧 TCP 会话的第一往返时间 (RTT),并确定服务网关和客户机之间的 TCP 代理会话的客户机侧 TCP 会话的第二 RTT;

(b) 将第一 RTT 与第二 RTT 进行比较;

(c) 确定第二 RTT 是否超出第一 RTT;

(d) 响应于确定第二 RTT 超出第一 RTT,基于第二 RTT 计算要求的 RTT;以及

(e) 根据计算出的要求 RTT 设定定时器,其中服务器侧 TCP 会话的 TCP 确认被延迟直到定时器到期为止。

7. 如权利要求 6 所述的介质,其特征在于,被配置成确定 (c) 和计算 (d) 的计算机可读程序代码被进一步配置成:

c1) 确定第二 RTT 是否超出第一 RTT 一预定阈值;以及

d1) 响应于确定第二 RTT 超出第一 RTT 所述预定阈值,基于第二 RTT 通过服务网关计算要求的 RTT。

8. 如权利要求 6 所述的介质,其特征在于,被配置成计算 (d) 的计算机可读程序代码被

进一步配置成：

d1) 将要求的 RTT 计算为第二 RTT 的百分比。

9. 如权利要求 6 所述的介质,其特征在于,被配置成计算 (d) 的计算机可读程序代码被进一步配置成：

(d1) 将要求的 RTT 计算为第二 RTT 减去一预定值。

10. 如权利要求 6 所述的介质,其特征在于,被配置成设定 (e) 的计算机可读程序代码被进一步配置成：

(e1) 在服务器侧 TCP 会话上从服务器接收数据分组；

(e2) 确定将 TCP 确认发送至服务器的需要；

(e3) 将定时器设定至要求的 RTT；以及

(e4) 响应于定时器的到期,将 TCP 确认发送至服务器。

11. 一种用于减少客户机和服务器之间的传输控制协议 (TCP) 代理会话的缓存使用的系统,包括：

安全网关,其中在安全网关和服务器之间建立 TCP 代理会话的服务器侧 TCP 会话,在安全网关和客户机之间建立 TCP 代理会话的客户机侧 TCP 会话,所述安全网关：

(a) 确定服务网关和服务器之间的 TCP 代理会话的服务器侧 TCP 会话的第一往返时间 (RTT),并确定服务网关和客户机之间的 TCP 代理会话的客户机侧 TCP 会话的第二 RTT；

(b) 将第一 RTT 与第二 RTT 进行比较；

(c) 确定第二 RTT 是否超出第一 RTT；

(d) 响应于确定第二 RTT 超出第一 RTT,基于第二 RTT 计算要求的 RTT；以及

(e) 根据计算出的要求 RTT 设定定时器,其中服务器侧 TCP 会话的 TCP 确认被延迟直到定时器到期为止。

12. 如权利要求 11 所述的系统,其特征在于,所述确定 (c) 和计算 (d) 包括：

c1) 确定第二 RTT 是否超出第一 RTT 一预定阈值；以及

d1) 响应于确定第二 RTT 超出第一 RTT 所述预定阈值,基于第二 RTT 计算要求的 RTT。

13. 如权利要求 11 所述的系统,其特征在于,所述计算 (d) 包括：

d1) 将要求的 RTT 计算为第二 RTT 的百分比。

14. 如权利要求 11 所述的系统,其特征在于,所述计算 (d) 包括：

(d1) 将要求的 RTT 计算为第二 RTT 减去一预定值。

15. 如权利要求 11 所述的系统,其特征在于,所述设定 (e) 进一步包括：

(e1) 在服务器侧 TCP 会话上从服务器接收数据分组；

(e2) 确定将 TCP 确认发送至服务器的需要；

(e3) 将定时器设定至要求的 RTT；以及

(e4) 响应于定时器的到期,将 TCP 确认发送至服务器。

基于延迟的确认减少对 TCP 代理会话的缓存使用

发明领域

[0001] 本发明总地涉及数据通信,更具体地涉及服务网关。

背景技术

[0002] 诸如防火墙和服务器负载平衡器之类的许多服务网关一段时间提供传输控制协议 (TCP) 代理功能。TCP 代理的典型服务应用包括网络分析、安全和由于非对称客户机和服务器状态的通信量适应。TCP 代理服务器一般分配一定量的存储器缓存以应付客户机设备和服务器之间的 TCP 代理会话的数据分组缓存。存储器缓存被用来应付客户机侧会话和服务器侧会话的数据分组缓存。客户机侧会话发送和接收缓存及服务器侧会话发送和接收缓存间的存储器空间分配通常不将性能考虑在内。TCP 代理服务器从服务器侧会话接收数据分组,根据需要的服务应用处理数据分组,并将结果数据分组发送至客户机侧会话。在理想的场景下,这些步骤是在来自服务器侧会话的下一数据分组被传递至 TCP 代理服务器之前完成的。然而,在许多部署的情况下,客户机设备通过移动宽带接入或剩余因特网接入来接入服务,这种接入具有由于长距广域网引起的较长传输时间以及基于订户接入服务的较慢传输带宽。然而,一个或多个 TCP 代理服务器位于同一数据中心内并享有短传输时间和高容量带宽。在该部署场景中,当 TCP 代理服务器从服务器侧会话接收数据分组时,所接收的数据分组被放置在服务器侧会话接收缓存中,并等待由服务应用处理的时机,进而等待客户机侧会话以释放客户机侧会话发送缓存,由于之前发送的数据分组的缓慢传输,该客户机侧会话发送缓存被填充以早前处理的待决数据分组,进而等待它们传输的时机。

[0003] 在典型情况下,一旦成功地从服务器接收到适当量的 TCP 数据,TCP 代理服务器根据 TCP 协议发送 TCP 确认。当服务器接收到之前发送的 TCP 数据的 TCP 确认时,服务器将附加的 TCP 数据分组发送至 TCP 代理服务器。TCP 代理服务器应当进一步增加服务器侧会话接收缓存的存储器空间以存储附加的 TCP 数据分组,同时等待之前的 TCP 数据被处理并被发送至客户机。这种级联效应使 TCP 代理服务器消耗大量的存储器空间,用于容纳服务器侧会话的接收的 TCP 数据分组所需的服务器侧会话接收缓存。所使用的缓存空间越大,可供 TCP 代理服务器使用以应付附加 TCP 代理会话的存储器资源越少;尽管 TCP 代理服务器可能具有其它丰裕的资源以应付附加负载。

发明内容

[0004] 根据本发明的一个实施例,一种减少客户机和服务器之间的传输控制协议 (TCP) 代理会话的缓存使用的方法,包括:(a) 确定服务网关和服务器之间的 TCP 代理会话的服务器侧 TCP 会话的第一往返时间 (RTT),并确定服务网关和客户机之间的 TCP 代理会话的客户机侧 TCP 会话的第二 RTT;(b) 通过服务网关将第一 RTT 与第二 RTT 进行比较;(c) 确定第二 RTT 是否超出第一 RTT;(d) 响应于确定第二 RTT 超出第一 RTT,基于第二 RTT 通过服务网关计算要求的 RTT;以及 (e) 根据计算出的要求 RTT 通过服务网关设定定时器,其中服务器侧 TCP 会话的 TCP 确认被延迟直到定时器到期为止。

[0005] 在本发明的一个方面,确定 (c) 和计算 (d) 包括 :c1) 确定第二 RTT 是否超出第一 RTT 一预定阈值 ;以及 d1) 响应于确定第二 RTT 超出第一 RTT 该预定阈值,基于第二 RTT 通过服务网关计算要求的 RTT。

[0006] 在本发明的一个方面,计算 (d) 包括 :d1) 通过服务网关将要求的 RTT 计算为第二 RTT 的百分比。

[0007] 在本发明的一个方面,计算 (d) 包括 :(d1) 通过网关服务器将要求的 RTT 计算为第二 RTT 减去一预定值。

[0008] 在本发明的一个方面,设定 (e) 包括 :(e1) 通过服务网关在服务器侧 TCP 会话上从服务器接收数据分组 ;(e2) 通过服务网关确定将 TCP 确认发送至服务器的需要 ;(e3) 通过服务网关将定时器设定至要求的 RTT ;以及 (e4) 响应于定时器的到期,通过服务网关将 TCP 确认发送至服务器。

[0009] 对应于前面总结的方法的系统 and 计算机可读介质也在这里被描述和要求保护。

附图说明

[0010] 图 1 示出根据本发明一实施例为客户机设备和服务器之间的 TCP 代理会话提供服务的服务网关。

[0011] 图 1a 示出根据本发明一实施例的服务网关的组件。

[0012] 图 2 示出根据本发明一实施例的延迟发送 TCP ACK 分组的过程。

具体实施方式

[0013] 本发明能采用完全硬件实施例、完全软件实施例或包含硬件和软件部件两者的实施例的形式。在一优选实施例中,本发明以软件实现,该软件包括但不限于固件、驻留软件、微代码等。

[0014] 此外,本发明能采用可从计算机可用或计算机可读介质访问的计算机程序产品的形式,该计算机程序产品提供由计算机或任何指令执行系统使用或与之联系地使用的程序代码。出于说明的目的,计算机可用或计算机可读介质可以是包含、存储、通信、传播和传输程序的任何装置,以由指令执行系统、装置或设备使用或与之联系地使用。

[0015] 介质可以是电子、磁、光、电磁、红外或半导体系统 (或装置或设备) 或传播介质。计算机可读介质的例子包括半导体或固态存储器、磁带、可移动计算机磁盘、随机存取存储器 (RAM)、只读存储器 (ROM)、硬磁盘和光盘。光盘的当前例子包括紧致盘 - 只读存储器 (CD-ROM)、紧致盘 - 读 / 写 (CD-R/W) 和 DVD。

[0016] 适于存储和 / 或执行程序代码的数据处理系统将包括通过系统总线直接或间接耦合至存储器元件的至少一个处理器。存储器元件可包括在实际执行程序代码期间利用的本地存储器、大容量存储以及高速缓存存储器,这些存储器提供至少一些程序代码的临时存储以减少在执行过程中必须从大容量存储中检索代码的次数。

[0017] 输入 / 输出或 I/O 设备 (包括但不限于键盘、显示器、定点设备等) 可直接或通过中间 I/O 控制器耦合至系统。

[0018] 网络适配器也可耦合至系统以允许数据处理系统通过中间私用网或公用网耦合至其它数据处理系统或远程打印机或存储设备。调制解调器、线缆调制解调器和以太网卡

只是当前可用类型的网络适配器中的一些。

[0019] 附图中的流程图和框图示出根据本发明各实施例的系统、方法和计算机程序产品的可能实现的架构、功能和操作。在这方面,流程图或框图中的每个方框可代表模块、区段或代码的一部分,其包括一个或多个可执行指令以实现规定的本地功能。也应当注意,在一些替代实现中,框内记录的功能可按附图中记录的顺序之外的顺序发生。例如,连续示出的两个框实际上可基本同时地执行,或者某些框有时可按相反顺序执行,这依赖于所涉及的功能。还要注意框图和 / 或流程图的每个框以及框图和 / 或流程图中的这些框的组合可通过实现特定功能或动作的基于专用硬件的系统或专用硬件和计算机指令的结合来实现。

[0020] 本文中使用的术语仅是为了描述具体实施例而不旨在对发明构成限制。本文中使用的单数形式“一”、“一个”和“该”旨在也包括复数形式,除非上下文中清楚指明了相反情形。要进一步理解,术语“包括”和 / 或“包含”,当在本说明书中使用,规定了所声称的特征、整体、步骤、操作、要素和 / 或组件的存在,但不排除一个或多个其它特征、整体、步骤、操作、要素、组件和 / 或它们的组的存在或附加。

[0021] 本发明的实施例,如下文所述,调整了服务侧会话传输时间以减少缓存使用,这进而增加了 TCP 代理服务器的 TCP 代理会话的能力。根据本发明的实施例,TCP 代理服务器延迟服务器发送附加的 TCP 数据,其中该延迟允许 TCP 代理服务器在服务器侧会话接收缓存中处理和发送拟被处理和发送至客户机的当前 TCP 数据。当服务器在延迟后发送附加 TCP 数据时,TCP 代理服务器在服务器侧会话接收缓存中将具有足够的空间以接收附加的 TCP 数据。该延迟延长了该服务器和 TCP 代理服务器之间的服务器侧会话的传输时间。

[0022] 图 1 示出根据本发明的实施例经由数据网络 153 为客户机设备 100 和服务器设备 200 之间的 TCP 代理会话 400 提供服务的服务网关 300。

[0023] 在一个实施例中,数据网络 153 包括网际协议 (IP) 网络、团体数据网络、区域团体数据网络、因特网服务提供者网络、住户数据网络、诸如以太网的有线网络、诸如 WiFi 网络的无线网络或蜂窝网络。在一个实施例中,数据网络 153 位于数据中心内,或连接至一网络或应用网络云。

[0024] 客户机设备 100 一般是具有网络接入能力的计算设备。在一个实施例中,客户机设备 100 是工作站、台式个人计算机或膝上个人计算机、个人数据助理 (PDA)、平板计算设备、智能电话或蜂窝电话、机顶盒、因特网媒体浏览器、因特网媒体播放器、智能传感器、智能医疗设备、网顶盒、联网电视机、联网 DVR、联网蓝光播放器、联网手持游戏设备或媒体中心。

[0025] 在一个实施例中,客户机设备 100 是住户宽带网关、商业因特网网关、商业 Web 代理服务器、网络消费者房屋设备 (CPE) 或因特网接入网关。

[0026] 在一个实施例中,客户机设备 100 包括宽带远程接入服务器 (BRAS)、数字订户线接入多路复用器 (DSLAM)、线缆调制解调器终端系统 (CMTS) 或服务提供者接入网关。

[0027] 在一个实施例中,客户机设备 100 包括诸如网关 GPRS 支持节点 (GGSN)、家庭代理 (HA) 或 PDN 网关 (PGW) 之类的移动宽带接入网关。

[0028] 在一个实施例中,客户机设备 100 包括服务器负载平衡器、应用传递控制器、通信量管理器、防火墙、VPN 服务器、远程接入服务器或企业或数据中心接入网关。

[0029] 在一个实施例中,客户机设备 100 是类似于服务网关 300 的设备。

[0030] 客户机设备 100 经由服务网关 300 向服务器 200 发起 TCP 会话 400。

[0031] 服务器 200 是一般耦合至处理器和存储计算机可读程序代码的计算机可读介质的计算设备。具有处理器和计算机可读程序代码的服务器 200 实现 Web 服务器、文件服务器、视频服务器、数据库服务器、应用服务器、语音系统、会议服务器、媒体网关、媒体中心、应用服务器或网络服务器的功能,这些服务器使用 TCP 会话 400 向客户机设备 100 提供基于 TCP 的服务或应用服务。

[0032] 在一个实施例中,服务器 200 是类似于服务器网关 300 的设备。

[0033] 在一个实施例中,TCP 会话 400 包括 HTTP 会话、FTP 文件传输会话、基于 TCP 的视频流会话、基于 TCP 的音乐流会话、文件下载会话、组会议会话、数据库接入会话、远程终端接入会话、远程通信网会话、电子商务交易、远程程序调用或基于 TCP 的网络通信会话。

[0034] 图 1a 所示的服务网关 300 可操作地耦合至处理器 310、存储器模块 320、网络接口模块 330 以及计算机可读介质 340。计算机可读介质 340 存储计算机可读程序代码,该计算机可读程序代码当由使用存储器模块 320 的处理器 310 执行时,实现如本文所述的本发明的各实施例。在一些实施例中,服务网关 300 被实现为服务器负载均衡器、应用传递控制器、服务传递平台、通信量管理器、安全网关、防火墙系统的组件、虚拟个人网络 (VPN) 的组件、视频服务器的负载均衡器、将负载分配至一个或多个服务器的网关、Web 或 HTTP 服务器、网络地址转换 (NAT) 网关或 TCP 代理服务器。

[0035] 在一个实施例中,计算机可读介质 340 包括服务应用 350 的指令,并且处理器 310 执行服务应用 350。

[0036] 在一个实施例中,服务应用 350 实现 VPN 防火墙、网关安全应用、HTTP 代理、基于 TCP 的音频或视频流会话代理、Web 会话代理、内容过滤、服务器负载均衡、防火墙或网络应用会话代理的功能。

[0037] 回到图 1,在向客户机设备 100 和服务器 200 之间的 TCP 代理会话 400 提供服务的一个实施例中,服务网关 300 建立与客户机设备 100 的客户机侧 TCP 会话 420 以及与服务器 200 的服务器侧 TCP 会话 470。

[0038] 在一个实施例中,服务网关 300 为服务器侧 TCP 会话 470 分配一接收缓存 474。在一个实施例中,接收缓存 474 驻留在存储器模块 320 中。

[0039] 在一个实施例中,服务网关 300 使用 TCP 会话 470 的往返时间 (RTT) 497 监测 TCP 会话 470 的性能。服务网关 300 测量或估计 TCP 会话 470 的 RTT 497。在一个示例实施例中,服务网关 300 基于服务网关 300 向服务器 200 发送 TCP 会话 470 的数据分组的时间与服务网关 300 接收到对发送数据分组的确认的时间之间的时长来测量 RTT 497。在一个实施例中,服务网关 300 在 TCP 会话 470 期间周期地或偶尔地测量 RTT 497。在一个实施例中,服务网关 300 基于与服务器 200 的一个或多个之前的服务器侧 TCP 会话估计 RTT 497。在一个实施例中,服务网关 300 将 RTT 497 估计为 10 毫秒、100 毫秒、3 毫秒、22 毫秒或 3 秒。

[0040] 在一个实施例中,服务网关 300 从接收缓存 474 检索数据,在一个实施例中通过服务应用 350 处理该数据,并通过 TCP 会话 420 将经处理的数据发送至客户机设备 100。在一个实施例中,服务网关 300 每当 TCP 会话 420 准备传输时处理来自接收缓存 474 的数据。TCP 会话 420 的缓慢传输造成服务网关 300 延迟处理来自接收缓存 474 的数据。在一个实施例中,服务网关 300 使用 TCP 会话 420 的往返时间 (RTT) 492 监测 TCP 会话 420 的性能。

服务网关 300 测量或估计 TCP 会话 420 的 RTT492。在一个示例性实施例中,服务网关 300 基于服务网关 300 将 TCP 会话 420 的数据分组发送至客户设备 100 的时间和服务网关 300 接收对发送的数据分组的确认的时间之间的时长来测量 RTT492。在一个实施例中,服务网关 300 在 TCP 会话 420 过程中周期地或偶尔地测量 RTT492。在一个实施例中,服务网关 300 基于与客户机设备 100 的一个或多个在前客户机侧 TCP 会话估计 RTT492。在一个实施例中,服务网关 300 将 RTT492 估计为 10 毫秒、100 毫秒、3 毫秒、22 毫秒或 3 秒。

[0041] 在一个实施例中,服务网关 300 将 RTT497 与 RTT492 进行比较。在一个实施例中,当服务网关 300 确定 RTT492 超出 RTT497 某一阈值时,服务网关 300 如下文所述地施加处理,以调整 RTT497 以收窄 RTT492 和 RTT497 之间的间隙。在一个实施例中,当 RTT492 比 RTT497 高过至少 2 倍、5 倍或 10 倍时或当 RTT492 至少大过 RTT497 一预定量(例如 20 毫秒、50 毫秒或 200 毫秒)时,RTT492 被确定为超出 RTT497 该阈值。

[0042] 在一个实施例中,服务网关 300 确定 RTT492 未超出 RTT497 该阈值,并且服务网关 300 不调整 RTT497。

[0043] 在一个实施例中,服务网关 300 规则地或偶尔地测量 RTT492 和 RTT497,并将 RTT492 与 RTT497 进行比较。

[0044] 图 2 示出根据本发明一实施例用于调整服务器侧 TCP 会话 470 的 RTT497 的过程。在一个实施例中,服务网关 300 从服务器 200 接收 TCP 会话 470 上的数据分组 480。服务网关 300 将数据分组 480 存储到接收缓存 474 中。在一个实施例中,服务网关 300 从接收缓存 474 确定按照 TCP 协议发送 TCP 确认的需要。不是立即发送 TCP ACK 数据分组 479,而是服务网关 300 使用定时器 487 安排在稍后时间发送 TCP ACK 数据分组 479。服务网关 300 将定时器 487 设定至要求的 RTT498。当定时器 487 到期时,服务网关 300 发送 TCP ACK 数据分组 479。在一个实施例中,服务网关 300 包括时钟(未示出),该时钟允许服务网关 300 确定定时器 487 是否到期。

[0045] 在一个实施例中,服务网关 300 基于 RTT492 计算要求的 RTT498。在一个实施例中,计算要求的 RTT498 使其落在 RTT492 的实际范围内。例如,要求的 RTT498 被计算为 RTT492 的预定百分比,例如 RTT492 的 30%、40%、60%或 75%。在一个实施例中,要求的 RTT498 被计算为 RTT492 减去一预定值,例如 10 毫秒、5 毫秒或 25 毫秒。要求的 RTT498 提供发送 TCP 会话 470 的 TCP 确认的定时延迟,并由此增加 TCP 会话 470 的往返时间。当在发送 TCP ACK 数据分组 479 之后服务网关 300 如图 1 所示地测量 RTT497 时,RTT497 被预期具有与要求的 RTT498 类似的值。

[0046] 在一个实施例中,服务网关 300 执行一过程,即测量 RTT497、RTT492,将 RTT492 与 RTT497 进行比较,并当服务网关 300 确定 RTT492 显著大于 RTT497 时执行图 2 中的处理步骤,从而减小接收缓存 474 的存储器容量,这进而增加了服务网关 300 处理额外 TCP 代理会话的能力。

[0047] 在一实施例中,由用户凭经验确定预定 RTT492 的百分比或预定值,即针对不同客户机和服务器的不同 TCP 代理会话使用多种百分比和值。典型地,RTT492 和 RTT497 之间的差越小,则所需的接收缓存 474 的存储器容量越小。在一个实施例中,用户配置要求的 RTT498 以减少 RTT497 和 RTT492 之间的差。在一个实施例中,预定的百分比在 30%和 50%之间,并由用户对服务网关 300 配置。用户能配置对较小的接收缓存 474 容量配置较高值的

预定百分比或要求的 RTT498, 并对较大的接收缓存 474 容量配置较小值的预定百分比或要求的 RTT498。用户可考虑一预定百分比或预定值以在接收缓存 474 容量和要求的 RTT498 之间取得平衡。

[0048] 尽管已根据所示实施例描述了本发明, 然而本领域内技术人员将容易理解, 对这些实施例可能存在变型并且这些变型可落在本发明的精神和范围内。因此, 可由本领域内技术人员作出许多修正而不脱离所附权利要求书的精神和范围。

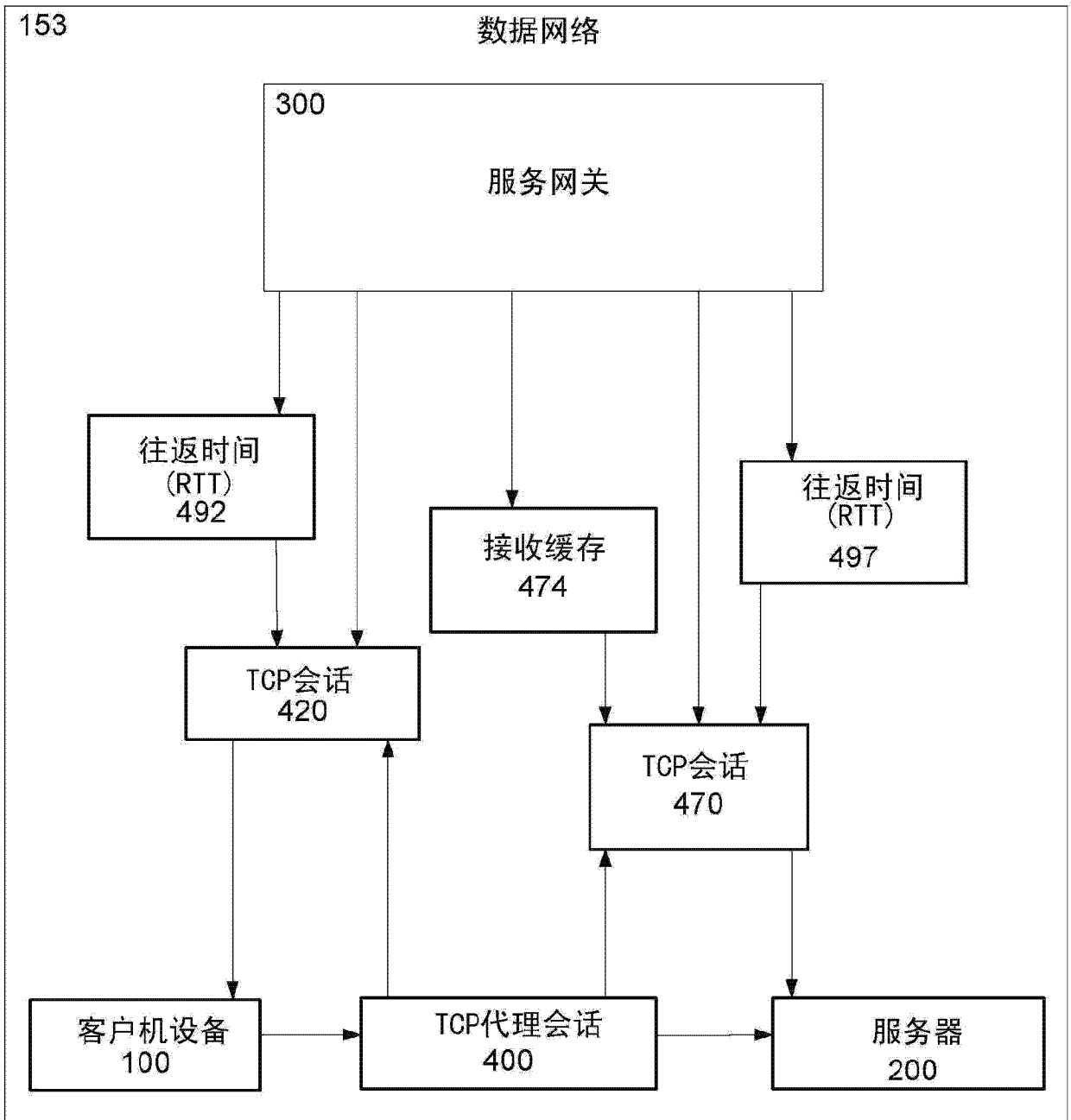


图 1

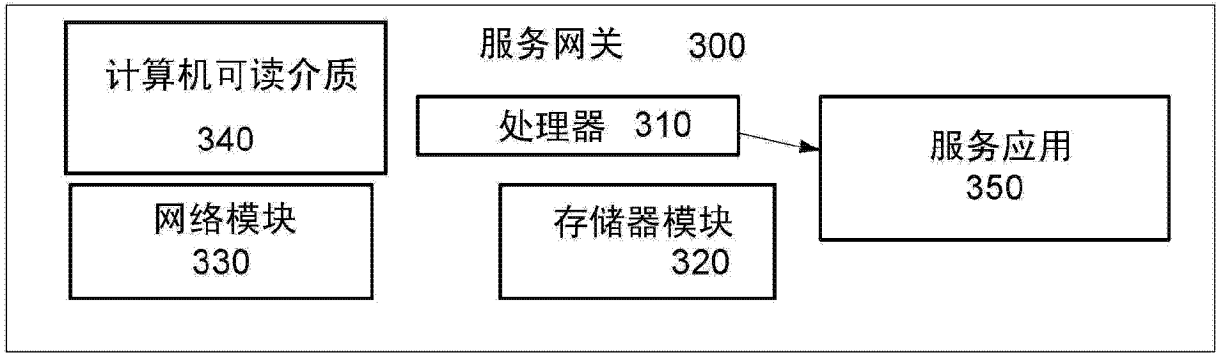


图 1a

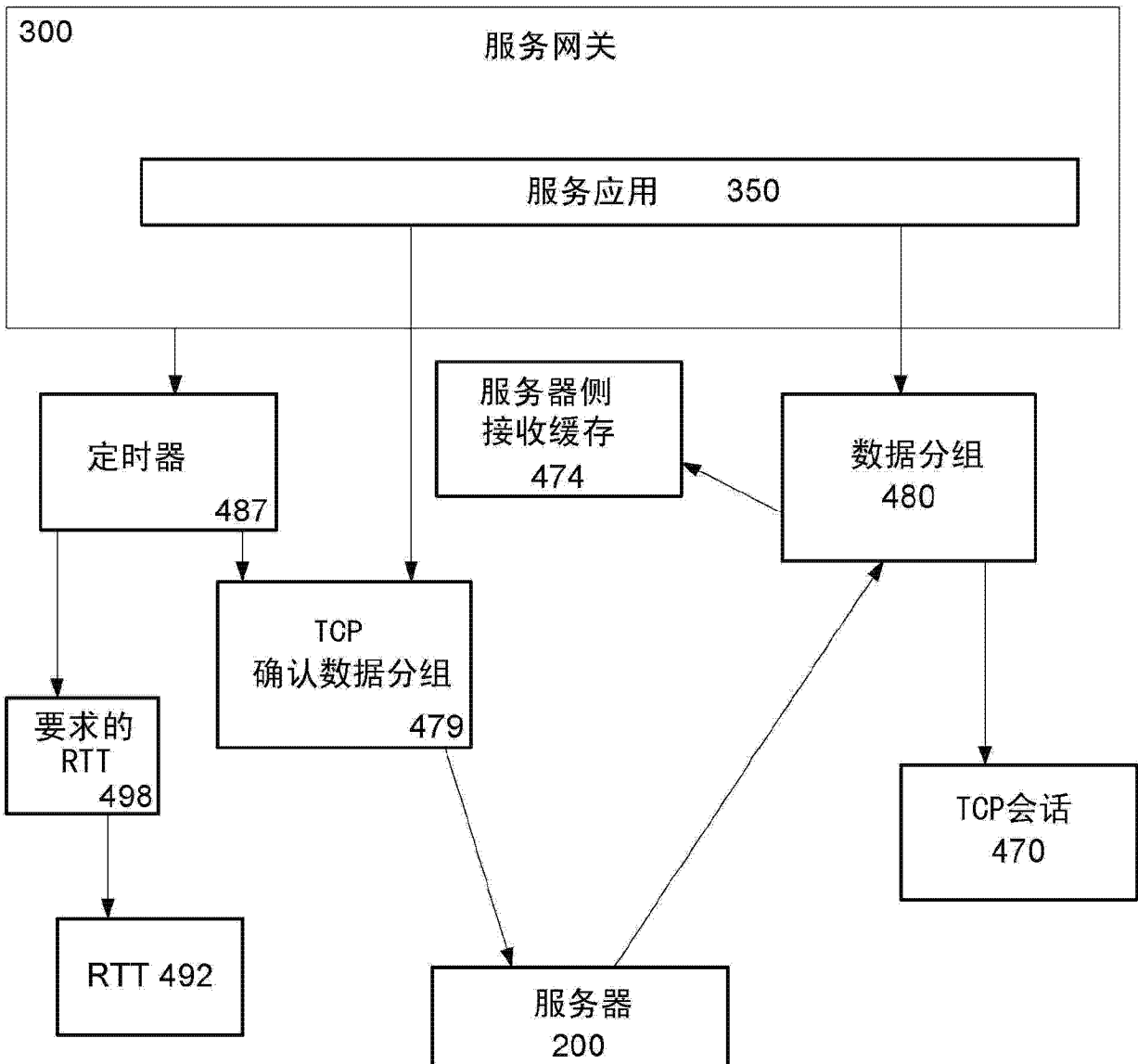


图 2