

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

G06F 15/16

G06F 17/30



[12] 发明专利申请公开说明书

[21] 申请号 02811288.1

[43] 公开日 2004 年 7 月 14 日

[11] 公开号 CN 1513142A

[22] 申请日 2002.6.4 [21] 申请号 02811288.1

[30] 优先权

[32] 2001.6.4 [33] US [31] 60/295,721

[32] 2001.6.4 [33] US [31] 60/295,672

[32] 2001.6.4 [33] US [31] 60/295,676

[32] 2001.6.4 [33] US [31] 60/295,720

[32] 2001.6.4 [33] US [31] 60/295,671

[86] 国际申请 PCT/US2002/017932 2002.6.4

[87] 国际公布 WO2002/099677 英 2002.12.12

[85] 进入国家阶段日期 2003.12.4

[71] 申请人 NCT 集团公司

地址 美国康涅狄格

[72] 发明人 迈克尔·J·SR·帕雷拉

约翰·拉什

[74] 专利代理机构 中原信达知识产权代理有限公司

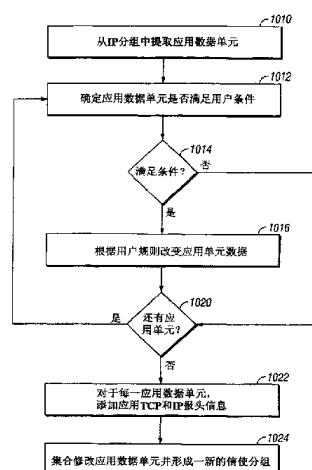
代理人 陆弋钟强

权利要求书 2 页 说明书 20 页 附图 9 页

[54] 发明名称 用于使用元素解析来修改数据流的系统及方法

[57] 摘要

提供了用于提高网络中信息传输效率并用于修改从服务器到用户的数据流中的数据的一种系统及方法。来自用户请求的万维网服务器的应用数据，例如 HTML、XML、SGML、scripts 或其他软件代码，可以由位于用户的 PC 和万维网服务器之间的中间服务器解析为元素。该中间服务器可以基于一组用户定义的规则修改、删除、增加、查询、过滤或替换一个或多个元素(1016)并将修改的应用数据转发给用户(1024)。



1. 一种用于改变第一计算机系统经由通信网络发送至第二计算
机系统的应用数据的方法，其中所述第二计算机具有用于显示所述应
用数据的浏览器，所述方法包括：

由所述第一计算机将所述应用数据解析为元素；

如果所述多个元素中的一个元素满足预定的用户条件，则根据预
定动作改变所述元素，其中所述改变包括替换、修改和添加；以及
将所述经改变的元素发送至所述浏览器。

10

2. 如权利要求 1 的方法，其中改变进一步包括删除和过滤。

3. 如权利要求 1 的方法，其中所述预定的用户条件和预定动作
是基于内容、广告、期望观众、用户、人力资源、定时、语境、法规、
地理、IP 地址、来源、文件大小或类型或政治内容。

4. 一种由从第一计算机到第二计算机的数据流中的中间计算机
改变应用数据的方法，包括：

提取从至少一个 IP 分组接收的应用数据；

确定所述应用数据的一部分是否满足预定的用户条件；

响应于所述部分对预定的用户条件的满足，根据预定的用户规则
改变所述部分；

将所述经改变的部分与其它应用数据联合，并形成至少一个新的
IP 分组；以及

发送所述新的 IP 分组。

25

5. 如权利要求 4 的方法，其中所述应用数据是 HTML 信息。

6. 如权利要求 4 的方法，其中所述预定用户条件包括广告条
HTML 代码的存在，并且所述预定的用户规则是从由删除所述 HTML

30

代码、将广告条替换为另一产品的广告条以及添加公共通知所构成的一组规则中选择的。

7. 一种用于修改数据流中的应用数据元素的系统，包括：

5 第一超级模块，用于接收所述数据流中的至少一个应用数据元
素；

决策模块，用于根据一组预定的用户规则分析所述应用数据元素
并当满足预定条件时修改所述应用数据元素；

重分组模块，用于使用所述经修改的数据元素创建一信使分组；
10 和

第二超级模块，用于接收所述信使分组。

8. 如权利要求 7 的系统，其中所述预定用户规则集是基于内容、
广告、期望观众、用户、人力资源、定时、语境、法规、地理、IP 地
址、来源、文件大小或类型或政治内容。
15

用于使用元素解析来修改数据流的系统及方法

5 技术领域

本发明通常涉及通信领域，特别涉及在计算机网络上信息的有效传输。

背景技术

10 在过去的几十年中，因特网的范围从政府和大学之间的研究网络迅速成长为由企业和个人进行个人及商务交易的工具。因特网最初是设计为无组织的，使得在崩溃发生时实现通信的概率高。传输信息的方法是基于与通过邮件发送信件类似的概念。一条消息可以分解为多个 TCP/IP 分组（即，信件）并发送至收信人。与信件类似，每一分组可以采用不同的路径到达收信人。尽管许多路径上的许多小分组提供了相对低廉的用户对例如网站的访问，但是这比用户和网站之间的点对点连接慢许多。

20 图 1 是显示现有技术中用户到因特网的连接的框图。通常，用户 110 经由接入点 (PoP) 112 连接至因特网，在现有技术中，接入点 (PoP) 112 是由因特网服务提供者 (ISP) 运营的。PoP 连接至 ISP 的主干网 114，例如 ISP1。例如 ISP1 和 ISP2 的多个 ISP 主干网由例如 NAP 170 这样的网络接入点连接在一起，形成因特网“云” 160。

25 更特别的，在个人计算机 (PC) 120 处的单个用户具有多种选择以连接至 PoP 112，诸如在例如固定无线 PC 或移动电话上的直接用户线 (DSL) 调制解调器 122、TV 电缆调制解调器 124、标准拨号调制解调器 126 或无线收发器 128。这里使用术语个人计算机或 PC 以描述具有处理器和内存的任一设备，而不限于传统的桌面 PC。在 PoP 112 处，对每一种类型的调制解调器（或收发器）有一个对应的访问设备，

5 以从/向用户 110 接收/发送数据。对于 DSL 调制解调器 122，PoP 112 在其访问设备处有一数字用户线访问复用器（DSLAM）。对于电缆调制解调器 124，PoP 112 在其访问设备处有一电缆调制解调器终端系统（CTMS）头端器。DSL 和电缆调制解调器连接允许几百千比特每秒（Kbps）的速率，远远快于通过拨号远程访问服务器（RAS）134 在 PoP 112 接收数据的标准拨号调制解调器 126。无线收发器 128 可以是个人数字助理（PDA）或移动电话的一部分，并且在 PoP 112 处连接至例如基站这样的无线收发器 136。

10 企业用户（或具有家庭办公室的个人）可以具有局域网（LAN），例如通过以太网链路连接至 LAN 服务器 144 的 PC 140 和 142。企业用户可以具有到 PoP 112 的 T1（1.544 Mbps）、分数 T1 连接或更快的连接。来自 LAN 服务器 144 的数据通过路由器（未显示）发送至例如信道服务单元/数据服务单元（CSU/DSU）146 这样的数字连接设备，然后该数字连接设备通过 T1（或分数 T1）线 148 将数字数据发送至 PoP 112 处的 CSU/DSU。

15 20 PoP 112 可能包括连接了 DSLAM 130、CTMS 头端器 132、RAS 134、无线收发器 136 或 CSU/DSU 150 的 ISP 服务器 152。ISP 服务器 152 可以提供诸如电子邮件、新闻组或域名服务器（DNS）这样的用户服务。或者 DSLAM 130、CTMS 头端器 132、RAS 134、无线收发器 136 或 CSU/DSU 150 可以绕过 ISP 服务器 152，直接连接至路由器 154（虚线）。服务器 152 连接至路由器 154，路由器 154 将 PoP 112 连接至具有例如路由器 162、164、166 和 158 的 ISP1 的主干网。ISP2 具有提供诸如电子邮件和用户网络主机这样的竞争用户服务的 ISP2 服务器 180。向用户 110 提供在线内容的万维网服务器 182 和 184 连接至因特网“云”160。

25 30 尽管因特网为企业和个人提供了进行商业交易的基本功能，但是在例如万维网服务器和企业或个人用户之间传输信息的很长的时延是

一个大问题。例如 PC 120 处的用户需要来自万维网服务器 182 处的网站的信息。数据从万维网服务器 182 回到用户 PC 120 需要许多“路程段”。而且因为信息以分组“邮”回，所以分组通常通过不同路径返回。这些不同的路径由其它用户的分组共享，并且一些路径可能较慢。
5 于是即使在万维网服务器 182 和用户 120 之间的所有链路都有足够的容量，也将有显著的时延。然而，由于还有瓶颈，即流量超过容量的地方，所以甚至还会进一步的延迟。

两个主要的瓶颈是最后和倒数第二段距离。最后一段距离是从
10 PoP 112 至用户 110。当用户 120 使用最大速度为 56 Kbps 的拨号调制解调器时，这很明显。即使使用 512 Kbps 的 DSL 调制解调器，下载图片也很慢。倒数第二段距离在 ISP 之间。具有 PoP 112 的 ISP 可能通过其主干网 114 连接至更高层的 ISP（未显示），以达到地区/国家/全球覆盖。由于到更高层 ISP 的带宽增加提高了本地 ISP 的成本，所以具有例如 PoP 112 的本地 ISP 可能减少用户 110 的可用带宽量。效果是在万维网服务器 182 和 PC 120 之间的流量大于链路容量，于是有严重的延迟问题。在今天快节奏的世界中，这个问题大大妨碍了因特网作为商业工具的使用。
15

20 除了现有 TCP/IP 业务的瓶颈和低效，还有大量噪声流量。类似垃圾邮件使得业务通路变得堵塞并且用户被不需要的信息所淹没。由于网站和 ISP 可能接收来自广告商的资金，他们的兴趣会与寻求目标信息而不需要娱乐的商业用户不同。

25 用户的网络浏览器对例如弹出窗口、广告条和其他恼人的广告进行过滤或“广告阻塞”是已知的现有技术。尽管共同的服务器可能阻塞了选中的 URL 地址或 IP 地址，但是执行过滤的任务仍然在用户的浏览器上。

30 因此不仅需要提高在通信网络（例如因特网）上信息传输的效率，

而且需要减少对于用户不期望的数据业务。

发明内容

本发明提供了用于提高网络中信息传输效率并用于修改从服务器到用户的 data 流中的 data 的一种系统及方法。在一个示例性实施例中，来自用户请求的万维网服务器的应用数据，例如 HTML、XML、SGML、scripts 或其他软件代码，可以由位于用户的 PC 和万维网服务器之间的中间服务器解析为元素。该中间服务器可以基于一组用户定义的规则修改、删除、增加、查询、过滤或替换一个或多个元素并将修改的应用数据转发给用户。当中间服务器靠近万维网服务器并过滤出许多用户特定的不期望的数据（例如广告条）时，总有效网络带宽也增加了。

根据本发明的一个方面，提供了一种用于改变第一计算机系统经由通信网络发送至第二计算机系统的应用数据的方法，其中第二计算机具有用于显示该应用数据的浏览器。该方法包括：由第一计算机将应用数据解析为元素；如果多个元素中的一个元素满足预定的用户条件，则根据预定动作改变该元素，其中改变包括替换、修改和添加；以及将经改变的元素发送至该浏览器。

根据本发明的另一方面，提供了一种由从第一计算机到第二计算机的数据流中的中间计算机改变应用数据的方法。该方法包括：提取从至少一个 IP 分组接收的应用数据；确定该应用数据的一部分是否满足预定的用户条件；响应于该部分对预定的用户条件的满足，根据预定的用户规则改变该部分；将经改变的部分与其它应用数据联合，并形成至少一个新的 IP 分组；以及发送新的 IP 分组。

根据本发明的一个方面，提供了一种用于修改数据流中的应用数据元素的系统。该系统包括：第一超级模块，用于接收数据流中的至少一个应用数据元素；决策模块，用于根据一组预定的用户规则分析

该应用数据元素并当满足预定条件时修改该应用数据元素；重分组模块，用于使用经修改的数据元素创建一信使分组；和第二超级模块，用于接收该信使分组。

5 参考下面的说明书、所附权利要求和附图，可以更好的理解本发明的这些及其它实施例、特征、方面和优点。

附图说明

图 1 是显示现有技术中到因特网的用户连接的框图。

10 图 2 是图 1 的简化而扩展的框图，用于帮助解释本发明。

图 3 显示了 TCP/IP 协议栈以及各层的相关数据单元。

图 4 是根据本发明实施例的浏览器和万维网服务器之间的通信路径的框图。

15 图 5 是根据本发明实施例，插入到图 2 的现有系统中的超级模块的框图。

图 6 是根据本发明实施例，在超级用户处重分组多个应用数据单元的流程图。

图 7 是根据本发明实施例，在超级模块处重分组多个接收的 IP 分组的流程图。

20 图 8 详细解释了图 7 的步骤 922 和 924。在步骤 932，从 IP 分组中提取应用数据单元。

图 9 显示了根据本发明的一个方面，从超级用户到超级主机的信使分组的例子。

25 图 10 是根据本发明的一个实施例，在超级模块处改变应用数据单元的流程图。

具体实施方式

在下面的说明书中，给出了众多特定细节，以提供本发明特定实施例的更充分的说明。然后，对于本领域技术人员显然可以看出，本发明不限于所给出的所有特定细节。在其它示例中没有具体说明众所

周知的特征，从而更突出本发明。

为了使个人和企业将因特网作为有效的商业工具而使用，用户请求和接收信息所需的时间必须比现在典型的所需时间大为缩短。本发明提供了可以叠加在因特网基础设施的部件上的“超级”系统和技术，以增加网络中的信息流量，该系统和技术无论单独还是联合使用，都显著减少了用户对于来自例如万维网站点或其它用户的信息的等待时间。

图 2 是图 1 的简化而扩展的框图，用于帮助解释本发明。其中图 2 使用了与图 1 相同的标签。调制解调器 210 包括图 1 的 DSL 调制解调器 122、电缆调制解调器 124、拨号调制解调器 126 和无线收发器 128。类似地，访问设备 220 包括图 1 的相应 DSLAM 130、CMTS 头端器 132、RAS 134 和无线收发器 136。数字连接设备 212 和 222 包括 CSU/DSU 设备 146 和 150，另外还包括卫星、ISDN 或 ATM 连接设备。图 2 在 LAN 服务器 144 和调制解调器 210 之间具有附加连接，以说明 LAN 的另一选择，以连接至除数字连接设备 212 以外还连接到 PoP 112。图 2 中显示的多数计算机和网络系统使用标准传输通信协议/因特网协议（TCP/IP）协议进行通信。

图 3 显示了 TCP/IP 协议栈以及各层的相关数据单元。TCP/IP 协议栈 310 包括应用层 312、传输层 314、因特网层 316 和网络访问层 318。应用层接收应用程序或用户数据 320、数据的一个块或单元，我们可以称其为应用数据单元。例如，对于网页的用户请求可以是一个应用数据单元。在 TCP/IP 中有多种应用级协议，包括用于电子邮件的简单邮件传输协议（SMTP）和邮局协议（POP）、用于万维网的超文本传输协议（HTTP）、以及文件传输协议（FTP）。

传输层 314 包括传输控制协议（TCP）和用户数据报协议（UDP）。TCP 是基于连接的协议，在源和目的地之间提供可靠的虚电路。TCP

5

保证使用其的应用程序以字节发送的顺序投递字节流，即使 IP 分组投递服务不可靠也不会发生重复或数据损失。传输层通过 TCP 报头 322 向数据 320 添加控制信息，这称为 TCP 数据单元。UDP 不能保证分组投递，并且使用 UDP 的应用程序必须提供它们自己的保证投递的方式。

10

因特网层 316 是由于 TCP/IP 对网际互联的强调而命名的。这是发送并接收因特网协议（IP）分组的无连接层。尽管 IP 分组具有 IP 分组的原始源地址和最终目的地址，但是如果分组达到其最终目的地，则在特定节点的 IP 层无需任何知识便将 IP 分组路由至下一节点。IP 分组包括添加至 TCP 数据单元（TCP 报头 322 和数据 320）的 IP 报头 324。

15

网络访问层 318 是底层，处理 IP 分组的物理传输。网络访问层 318 将现存的数据链路和物理层标准集合，而不是定义其自身的标准。这一层定义了网络硬件和设备驱动。报头 326 和尾部（未显示）添加至 IP 分组，以允许 IP 分组在通信线路上的物理传输。

20

图 2 中 TCP/IP 协议的使用的一个例子是 PC 140 处的用户请求来自万维网服务器 182 的网页。用户通过其浏览器在应用层 312 创建对于网页的用户请求，即应用数据单元 320（图 3）。TCP/IP 栈 310 创建一个或多个 TCP 数据单元，其中每一 TCP 数据单元具有应用数据单元 320 的部分，其附有 TCP 报头 322。PC 140 处的传输层 314 与万维网服务器 182 处的传输层 314 的 TCP 建立对等连接，即虚电路。每一 TCP 数据单元被分为一个或多个 IP 分组。IP 分组被发送至 LAN 服务器 144，然后发送至 PoP 服务器 152，然后通过 PoP 路由器 154 发送至因特网 154。IP 分组通过因特网 160 上的多条路径到达万维网服务器 182。然后万维网服务器 182 处的传输层 314 从 IP 分组重组 TCP 数据单元，并将 TCP 数据单元传递至应用层 312，以重组用户请求。然后执行得到网页的用户请求。为了将网页发回到用户，可以在万维

25

30

网服务器 182 和 PC140 的传输层之间使用相同的 TCP 虚电路。然后网页被分解为 TCP 数据单元，然后 TCP 数据单元被分解为 IP 分组并通过因特网 160、PoP 路由器 154、PoP 服务器 152、LAN 服务器 144 发送至 PC 140。

5

图 4 是根据本发明实施例的浏览器和万维网服务器之间的通信路径的框图。上面说明了当用户使用浏览器 512 请求来自万维网服务器 182 的网页 514 时，在浏览器 512 和万维网服务器 182 之间的现有的交换。本发明的一个实施例创建了包括超级用户 540、超级应用 532、超级中央办公室（CO）服务器 534、超级 CO 集中器 536 和超级主机 538 在内的多个“超级”模块，提供了可选的超级高速公路路径，以在浏览器 512 和万维网服务器 182 之间交换数据。对于网页 514 的用户请求由 PC 140 上执行的浏览器 512 发送至同样运行在 PC 140 上的超级用户软件 530。然后超级用户 530 将用户请求发送至 LAN 服务器 144 上运行（或在另一实施例中，在其自己的服务器上执行）的超级应用 532。然后超级应用 532 将用户请求发送至超级 CO 服务器 534，后者将请求发送至超级 CO 集中器 536。超级 CO 服务器 534 和超级 CO 集中器 536 可以是单独的服务器也可以是在 PoP 服务器 152 上运行的软件。超级 CO 集中器 536 通过因特网 160 将用户请求发送至超级主机 538（或在另一实施例中，主机 538 是万维网服务器 182 上运行的软件）。用户请求从超级主机 548 前进至万维网服务器 182，后者从万维网服务器 182 上运行的网站中取回网页 154（万维网服务器 182 可以包括服务器的万维网群和多个网站）。然后网页 514 通过超级主机 538、超级 CO 集中器 536、超级 CO 服务器 534、超级应用 532 和超级用户 530 返回浏览器 512。

在另一实施例中，可能没有一个或多个超级模块，例如，超级应用 532。在没有超级应用 532 的情况下，超级 CO 服务器 534 通过 LAN 服务器 144 与超级用户 530 交换信息。另一个例子是，如果不存在超级主机 548，则万维网服务器 182 与超级 CO 集中器 536 交换信息。

30

5

于是如果缺少一个超级模块，使用相应的一般模块，例如 PC 140、LAN 服务器 144、PoP 服务器 150、PoP 路由器 154 和万维网服务器 182，作为替代。可以使用所有或一些超级模块，并且只要在至少两个不同的超级模块之间有至少一条通信链路，流过该链路的信息将会显著提高。另外，可以配置更多的超级模块，以在网络上延伸超级层的粒度。

10

图 5 是根据本发明实施例，插入到图 2 的现有系统中的超级模块的框图。对于相同或相似的设备，在图 4 中使用了与图 2 相同的标签。超级用户 540 通过调制解调器 210 经由访问设备 220 连接至 PoP 服务器 152。具有超级用户 530、超级用户 542 和超级应用 532 的局域网连接至调制解调器 210 或数字连接设备 212，其中数字连接设备 212 由数字连接设备 222 连接至 PoP 服务器 152。超级应用 532 包括在 LAN 服务器 144 上执行的软件。服务器 152 通过交换机 420 连接至路由器 154，交换机 420 使分组业务转向超级 CO 服务器 534 和超级 CO 集中器 536。路由器 154 连接至因特网云 160。从因特网 160 开始，业务可以到达连接至万维网服务器 182 的超级主机 538 或者连接至万维网服务器 184 的超级主机 550 或者连接至 ISP 服务器 180 的超级主机 552。

15

20

超级系统部件

下面说明的是图 5 的超级系统的各个部件的一个实施例，包括超级用户 540、超级应用 532、超级 CO 服务器 534、超级 CO 集中器 536 和超级主机 538。

25

30

超级用户 530 包括位于用户 PC，例如 PC 140，中的软件。浏览器，例如 Microsoft 的 Internet Explorer，被设置为到超级用户 530 的代理，以使所有浏览器数据请求都由超级用户 530 提供。另外，通过浏览器的所有用户请求都被发送至超级用户 530。于是浏览器由超级用户与网络的其它部分隔离。超级用户在用户的 PC 上的本地高速缓存中对用户请求的所有数据进行高速缓存，使得当用户再次请求数据

时，如果可用即从本地高速缓存中本地取回。如果被高速缓存的数据超过预定的文件大小，则超级用户分析本地高速缓存中的所有数据，并删除最不常使用的数据。例如，可以使用一种现有的最近不常使用算法丢弃旧数据。超级用户 540 的一些软件功能有：

5 1. 高速缓存：如果浏览器请求在本地高速缓存中存在的数据并且数据满足高速缓存寿命需求，则从本地高速缓存提供该数据。否则从更新的数据是可访问的最近的超级模块高速缓存，例如超级应用 532 或超级 CO 服务器 534、超级 CO 集中器 536 或超级主机 538，中取回该数据，或者如果对于任一超级高速缓存都是不可访问的，则可从万维网服务器取回该数据。每一数据元素具有一高速缓存寿命，即它在需要刷新之前可以被高速缓存使用多久。

10 2. 刷新高速缓存：当超级用户 PC 空闲时（没有活跃地取回来自因特网的数据），超级用户检查本地高速缓存并自动刷新达到高速缓存寿命的数据。超级用户使用人工智能（AI）或其它技术，基于其确定的用户最可能请求哪些而把刷新区分优先次序。例如，超级用户可以对用户访问一网页的频率进行计数。较高的计数表示用户更可能在将来请求该网页，并且超级用户将自动刷新该页。

15 3. 预取：超级用户在空闲时间期间，使用 AI 或其它技术预取被用户需要的可能性高的网页（即，取回用户还没有请求的网页）。例如，如果用户正在浏览目录站点上的某些网页，则用户将浏览同一目录中的站点上的其它页的概率高。超级用户将预取这些页。预取增加了用户从本地高速缓存获得数据的概率。

20 4. 信使分组（后面描述）被分组，并且经分组的数据在发送至超级应用或超级 CO 服务器之前由超级用户压缩。信使分组被拆包，并且被拆包的数据在发送至浏览器之前由超级用户解压缩。

25 超级应用 532 包括在 LAN 服务器 144 上执行的软件。超级应用 532 执行的一些功能包括防火墙安全、全局高速缓存、分群、智能主机以及电子邮件管理。超级应用软件执行的功能还包括：

30 1. 如果超级应用连接到超级 CO 服务器，则它们之间传输的所

有数据被压缩并分组为信使分组，否则使用标准因特网请求并且响应在被发送至超级用户之前被分组为信使分组。

2. 超级应用还自动复制并保留其用户频繁使用的网站。

5 3. 如果超级应用连接到超级 CO 服务器，则它只有当被通知来自超级 CO 服务器的改变时，才更新网站的拷贝。如果超级应用没有连接到超级 CO 服务器，则它在空闲时间期间和/或在周期性的预定间隔期间检查网站的更新。

10 4. 如果超级用户连接到超级应用，则所有数据响应以压缩的格式传输至超级用户。如果固定用户连接到超级应用，则数据响应在超级应用中被解压缩并发送至用户。如果超级用户正在保留网站，则只要网页在超级应用上被更新时，就将一条提醒发送至超级用户，使得超级用户可以请求该变化。

15 5. 超级用户还向超级应用提醒有关用户 PC 监视器密度的信息，使得可以对局域网上传输的图片进行调整。向不能显示图片的监视器发送高密度图片是网络资源的浪费。超级应用中的软件在传输数据之前调整图片密度。

20 6. 如果多于一个超级用户请求同一数据，则超级应用发出请求并仅发送一个请求至下一超级模块，例如，超级 CO 服务器。如果在超级应用和网站之间没有另一超级模块，则请求仍然内爆（implode）并且作出标准 TCP/IP 请求。当接收到对于内爆的请求的响应时，超级应用发出该数据并且数据被发送至合适的超级用户。

25 超级应用保留的网站越多，网页的访问速度越接近局域网的速度。超级用户处保留的网页越多，网络访问速度越接近硬盘访问速度。超级应用和超级用户上能复制和保留的网页越多，最后一英里距离越不会成为响应时间的瓶颈。

30 超级 CO 服务器 534 是因特网主干网 114 和用户 110 之间的桥梁。超级 CO 服务器 534 的一个目的是将用户和因特网之间的流量最小化。超级 CO 服务器通过复制超级用户或一般用户经由超级 CO 服务器访

间的网站而实现这一点。超级 CO 服务器上的网站越多，通过减少数据在网络上的移动而对网络的优化越大。如果超级 CO 服务器上的网站来自超级 CO 集中器 536 存储的网站，则无论何时超级 CO 集中器 536 向超级 CO 服务器 534 提醒网页已经变化，超级 CO 服务器 534 都请求更新的网页。来自超级 CO 集中器 536 的网页以压缩且重分组的格式存储。如果超级 CO 服务器上的网站没有存储在超级 CO 集中器中，则超级 CO 服务器以预定的间隔检查在主万维网服务器的网站中的变化。超级 CO 服务器保留在每一超级应用 532 高速缓存上的网站的日志。当超级应用高速缓存上存在的网站发生变化时，向超级应用发送发生变化以及超级应用应请求改变的网页的更新拷贝的提醒。
当数据从非超级 CO 集中器站点接收时，它被压缩、分组并保存在超级 CO 服务器上。超级 CO 服务器从其请求日志确定被其用户访问的网站，并确定复制并在超级 CO 服务器 534 高速缓存保留哪一网站。超级 CO 服务器还将删除未被使用的站点。如果一个网站没有被存储和保留，则该网页保留在单独的全局高速缓存中，使得如果被再次请求则可以从全局高速缓存中提供。在全局高速缓存和网络主机中应维持正确的平衡。全局高速缓存和超级 CO 服务器可以作为一个高速缓存而实现并分别管理，或者作为两个单独的高速缓存而实现。如果超级应用请求一网页，则该网页以超级压缩且重分组的格式发送，否则该网页被解压缩并发送至请求用户。最靠近用户的超级模块拆包任何重分组格式并解压缩数据，使得它以原始形式发送至用户。最靠近用户的超级模块还高速缓存未压缩且未分组格式的信息。使用的优化涉及应用于可变数据（通常为文本）的压缩量和网页上可变数据的量。因特网上使用的 Rich 数据格式越多，得到的优化越大。Flash 软件、文件、java 程序、java script 等都存储在超级 CO 服务器。

来自超级应用的没有被超级 CO 服务器高速缓存满足的数据请求发送至负责为请求的 URL（网站）服务的超级 CO 集中器 536。请求被根据优化机制而分组压缩并内爆。在一个实施例中，在超级 CO 服务器发生数据内爆的第一级。在另一实施例中，内爆由超级应用完成。

超级 CO 服务器由 ISP 布局组织，使得区域定向的复制使用特性可以在请求上内爆并在响应上外爆（explode）。不能由超级 CO 服务器的高速缓存中的数据响应的所有请求和内爆的请求被传递至超级 CO 集中器。

5

超级 CO 集中器 536 由网站（URL 的）组织。这增加了用户需要的网站数据在 CO 集中器的高速缓存中的概率。它还增加了请求能被内爆以及网络流量能得到减小的概率。每一超级 CO 集中器负责高速缓存以及与超级主机，例如 538，和其它非超级主机网站接口。对于非超级主机网站，超级 CO 集中器 536 是遇到的第一个超级模块，并且发生了所有图片的初始重分组、第一压缩、最终内爆、第一外爆、转化至优化的压缩格式，例如 PGN 或合适的压缩算法，以及发生超级高速缓存的第一级。这里也是其它超级模块发生所有检查和刷新的地点。当来自网站的数据被刷新并更新时，通知超级 CO 服务器，使得所有高速缓存都可以被更新和刷新。

10

万维网服务器作为连接到因特网的一个或多个网站的主机。超级主机，即超级主机 538，响应超级 CO 集中器例如 536 作出的请求。每次作出下载万维网服务器上任一网站的请求时，超级主机 538 从万维网服务器中取回网页并在将内容发送至请求的超级 CO 集中器之前将内容压缩并分组。这通过有效的压缩率以及通过对所有被请求的网页数据发送单个数据块，提高了网络传输的效率。使用对于数据的特定类型执行的最好的技术，分析并压缩每一条信息。当接收到每一超级 CO 集中器请求时，超级主机记录超级 CO 集中器的 IP 地址。超级主机检查在万维网服务器中包含的网站，并将任何变化的网页通知曾经请求来自该网站的数据的任一超级 CO 集中器。这使得超级 CO 集中器能够知道何时需要刷新其网站版本，并通过允许超级 CO 集中器从超级 CO 集中器的高速缓存中的网页的版本直接为用户对于网页的请求提供服务而最小化万维网流量。网页的超级 CO 集中器版本需要刷新的时刻仅仅是其发生改变时。这使得可以最小化从万维网主机站

15

20

25

30

点到 ISP 站点的流量。在每一网站处有许多 ISP 站点访问数据。这是将网站移至因特网的外缘并将压缩和分组带至因特网的内部工作中的一个步骤。将网站移至因特网的外缘的困难是确保数据是当前的，超级模块高速缓存的互锁保证了这一点。

5

重分组

如今典型的网页包括一超文本链接标示语言（HTML）文档以及许多内嵌的图象。浏览器的现有行为是取回基础 HTML 文档，然后，在接收到基础 HTML 文档之后，浏览器对许多内嵌的对象进行二次取回，这些内嵌的对象通常位于同一万维网服务器上。每一内嵌对象，即应用数据单元，被放入一 TCP 数据单元，并且每一 TCP 数据单元被分割为一个或多个 IP 分组。为许多内嵌对象发送许多 TCP/IP 分组而不是例如一个大的 TCP/IP 分组，意味着网络耗费的时间比发送控制数据所需的时间更多，换言之，控制数据/时间对应用数据/时间的比率太大。将许多内嵌对象联合为一个大的应用数据单元然后创建一个（或至少是最小数目的）大的 TCP 数据单元是更为有效的。对于一个大的 TCP 数据单元，该发送器超级模块和下一接收器超级模块之间的链路的最大传输单元（MTU）用于 IP 分组。发送器超级模块将通过尝试使各个 IP 分组尽可能实际地靠近 MTU，而试图将发送的 IP 分组的数目最小化。对于在超级模块发送器和超级模块接收器之间的各个链路，对于该链路确定 MTU 并且 IP 分组的大小可以改变。现有技术中，通常使用在用户和万维网服务器之间的通信链路的所有 MTU 中的最小公分母 MTU，与现有技术不同，在本实施例中，使用各个链路的 MTU。

10

15

20

25

在本发明的一个实施例中，当需要时，在每一超级模块，例如超级用户、超级应用、超级中央办公室（CO）服务器、超级 CO 集中器和超级主机，将例如用户请求及万维网服务器响应这样的应用数据单元重分组（或拆包）为一较大（或多个较小）的经修改的应用数据单元。例如，将两个 IP 分组联合为一个 IP 分组，这是“信使”分组的

30

一个例子。第一 IP 分组具有第一 IP 报头、第一 TCP 报头和第一应用数据单元。第二 IP 分组具有第二 IP 报头、第二 TCP 报头和第二应用数据单元。创建第一修改应用数据单元，其具有第一应用数据单元和第一伪报头，第一伪报头具有来自第一 IP 报头和第一 TCP 报头的控制数据，诸如重构第一 IP 分组所需的源地址、源及目的端口以及其它控制信息。创建第二修改应用数据单元，其具有第二应用数据单元和第二伪报头，第二伪报头具有来自第二 IP 报头和第二 TCP 报头的控制数据，诸如重构第二 IP 分组所需的源地址、源及目的端口以及其它控制信息。令联合应用数据单元具有连接至第二修改应用数据单元的第一修改应用数据单元。向联合应用数据单元添加一新的 TCP 报头和 IP 报头，并且形成信使分组。于是必需的控制信息嵌入联合应用数据单元，并且 TCP/IP 协议用于在超级模块发送器和超级模块接收器之间移动联合应用数据单元。当接收器不是超级模块时，将联合应用数据单元拆离，并且第一 IP 分组和第二 IP 分组被重建并由超级模块发送器发送至一般接收器。

图 6 是根据本发明的实施例，在超级用户处对多个应用数据单元进行重分组的流程图。在步骤 910，超级用户将具有同一目的地的多个应用数据单元联合为一个联合数据单元。例如，对于一个万维网服务器的多个用户请求被联合。在步骤 912，从这一个应用数据单元形成一个 TCP 数据单元（或最小数目的 TCP 数据单元）。在步骤 914，创建一个 IP 分组（或最小数目的 IP 分组），即信使分组，其中每一 IP 分组被填充以尽可能接近链路的字节的 MTU 数目或直至转发计时器 T 过期。在步骤 916，信使被发送至目的路径中的下一超级模块，例如超级应用或超级 CO 服务器。

图 7 是根据本发明的另一实施例，在超级用户处对多个应用数据单元进行重分组的流程图。在步骤 920，超级模块接收具有同一目的地的多个 IP 分组。在步骤 922，从多个 IP 分组提取应用信息。在步骤 924，提取的应用用于形成重分组的分组（即，信使分组）。在步骤

926，重分组的分组被发送至到公共目的地的路径上的下一超级模块。

图 8 具体解释了图 7 的步骤 922 和 924。在步骤 932，从 IP 分组中提取应用数据单元。对于每一应用数据单元，检查相关的 TCP 报头和 IP 报头控制信息。并且将应用控制信息，例如源、源及目的端口和数据长度，添加至响应的应用数据单元以形成修改应用数据单元（步骤 934）。在步骤 936，集合修改应用数据单元以形成一个 TCP 数据单元（或最小数目的 TCP 数据单元）。在步骤 938，使用发送器和接收器超级模块之间的链路的 MTU，从 TCP 数据单元中形成新的重分组的 IP 分组。
5
10

是否在步骤 936 形成一个大 TCP 数据单元或者多个小 TCP 数据单元的判定是根据在离开发送者超级模块的链路上的业务负载而动态确定的。例如，如果链接几乎满载，则发送多个小 TCP 数据单元更加有效，并因此是小 IP 分组，而一个（或者几个）大 IP 分组将不得不等待。
15

图 9 显示了根据本发明的一个方面，从超级用户到超级主机的信使分组的例子。超级用户 530 根据图 6 的流程图，将用户请求 1020 和 1022，即应用数据单元 D1 和 D2，联合为信使分组 1024。超级用户 1010 具有 IP 分组 1026 中的其用户请求 D3，超级用户具有 IP 分组 1028 中的用户请求 D5。这两个单超级用户请求都被重分组为信使分组并发送至合适的超级应用。在第一超级应用 530，信使分组 1024 和 IP 分组 1026 被根据图 7 的流程图而接收并重分组以形成更大的应用信使分组 1030。应用信使分组 1030 具有例如已经过修改以包括来自 IP 分组 1024 的 TCP 和 IP 报头 H1 的控制信息的应用数据单元 D1。第二超级应用 1014 接收信使分组 1028-1，不改变它（1028-2）并将其转发至超级 CO 服务器 534。超级 CO 服务器 534 接收来自超级应用 532 的应用信使分组 1030 和来自超级应用 1014 的信使分组 1028-2。根据图 7 的流程图对信使分组 1030 和 1028-2 重分组，以形成被发送至超
20
25
30

5

级 CO 集中器 536 的 CO 信使分组 1034。超级 CO 服务器 1036 具有也被发送至超级 CO 集中器 536 的 CO 信使分组 1038。超级 CO 集中器 536 对 CO 信使分组 1034 和 1038 重分组为 CO 集中器信使分组 1040，CO 集中器信使分组 1040 被发送至超级主机 538。超级主机拆包 CO 集中器信使分组 1040 以得到用户请求 D1、D2、D3、D4、D5、D6 和 D7（例如，HTTP 或 FTP 请求），并且这些请求被发送至万维网服务器。对于从万维网服务器到超级用户 538 再经由超级 CO 集中器 536、超级 CO 服务器 534 和超级应用 532 返回超级用户 530 的数据响应，也发生根据图 6、7 和 8 的拆包。

10

改变应用数据

15

从上面对于重分组的讨论可以看出，应用数据单元在作为信使分组从超级主机 533 返回到超级用户（例如 530）时被多次检查。由于信使分组在任两个超级模块之间使用，如图 7 和 8 的流程图用于接收信使分组并且如果必要则将它们重分组为新的信使分组。在各个情况下都提取应用数据。如果应用数据是 HTML，则应用数据可以被解析为程序元素并且应用 IF-THEN 规则（即，如果满足一条件则执行预定动作）。另一实施例可以使用诸如 Perl（实用摘录与报告语言）这样的脚本语言，其寻找应用数据中的模式并执行一定的动作，诸如满足该模式的数据的删除、修改、替换或增加。删除、增加、修改或替换的规则可以基于任一期望的用户准则，包括内容、广告、期望观众、用户、人力资源、定时、内容、法规、地理、IP 地址、来源、文件尺寸或类型或政治内容。

20

25 为了说明，使用在若干画面中旋转的广告条的 HTML 代码作为可以作为对于用户对网页的请求的响应而返回的应用数据的一个例子。这里是事件处理机基于某种事件（诸如到达网页）而触发条幅显示的代码：

```
<body on Load = “rotate Banner” (‘images/Banner1.jpg’)”>
```

30

接下来是显示第一条幅图像（即 Banner1.jpg）的代码：

```
5      <table>
          <tr><td><img name = “banner” src = “images/Banner1.jpg”></td>
          </tr>
</table>
```

最后是函数 rotateBanner()，它每 5 秒钟循环调用自身并改变上面的“src”参数，从而显示新的条幅图像：

```
10     Function rotateBanner (BannerSrc)
          { var Timer ID
              // swap the picture
              Document.banner.src = BannerSrc;
              // wait for timeout and call self to swap next picture
              If (BannerSrc == images/Banner1.jpg")
                  TimerID = setTimeout ("rotateBanner ('images/Banner2.jpg')",
15                  5000);
                  Else If (BannerSrc == images/Banner2.jpg")
                      TimerID = setTimeout ("rotateBanner ('images/Banner3.jpg')",
                  5000);
20                  .....
          }
```

上面的例子显示了标准程序构造，并且对于本领域技术人员可以以多种软件程序解析。一旦解析了程序构造和参数，这些元素就可以由用户定义的规则操纵。于是用户能够使用在从万维网服务器到浏览器的路径中的任一超级模块过滤或修改他/她请求的数据。

尽管应用数据流修改软件可以是任一超级模块的一部分，但是在优选实施例中，其位于超级 CO 服务器 534、超级 CO 集中器 536 或超级主机 538（图 4），即，在最后一段的因特网侧（POP 服务器和 LAN

5

服务器 144 之间)。通过去除广告条和多余的图片，减少了在最后一段上的流量。通过在超级主机 538 处去除广告条，例如将节省在因特网 160 上不必要的数据流量。另外，超级 CO 服务器 534 能够使用除用户之外的其它改变规则。例如，政府组织可能希望以公共服务通知代替广告条。

10

图 10 是根据本发明的一个实施例，在超级模块处改变应用数据单元的流程图。在步骤 1010，从收到的 IP 分组中提取应用数据单元(见图 8 的步骤 932)，这些 IP 分组可能是来自一般模块的一般 IP 分组或者来自其它超级模块的信使分组。在步骤 1012，分析一个应用数据单元并且检查用户的 IF-THEN 规则集。当应用数据单元满足用户规则的“IF”条件时，则可以删除、修改或替换该数据(步骤 1016)。如果在信使分组中还有其它应用数据单元，则重复步骤 1012。否则，在步骤 1022，前面提取的伪 TCP 和 IP 报头信息添加至各个应用数据单元。然后将这些应用数据单元集合，并且添加一新的 TCP 和 IP 报头以形成一新的信使分组(步骤 1024)。新的信使分组被发送至下一超级模块。

15

于是，在信使分组中的应用数据通过各个超级模块的同时，它们被根据用户定义的规则而动态评估并改变。在另一实施例中，在一个或多个超级模块的超级高速缓存中根据用户定义的规则检查并改变应用数据。

20

结论

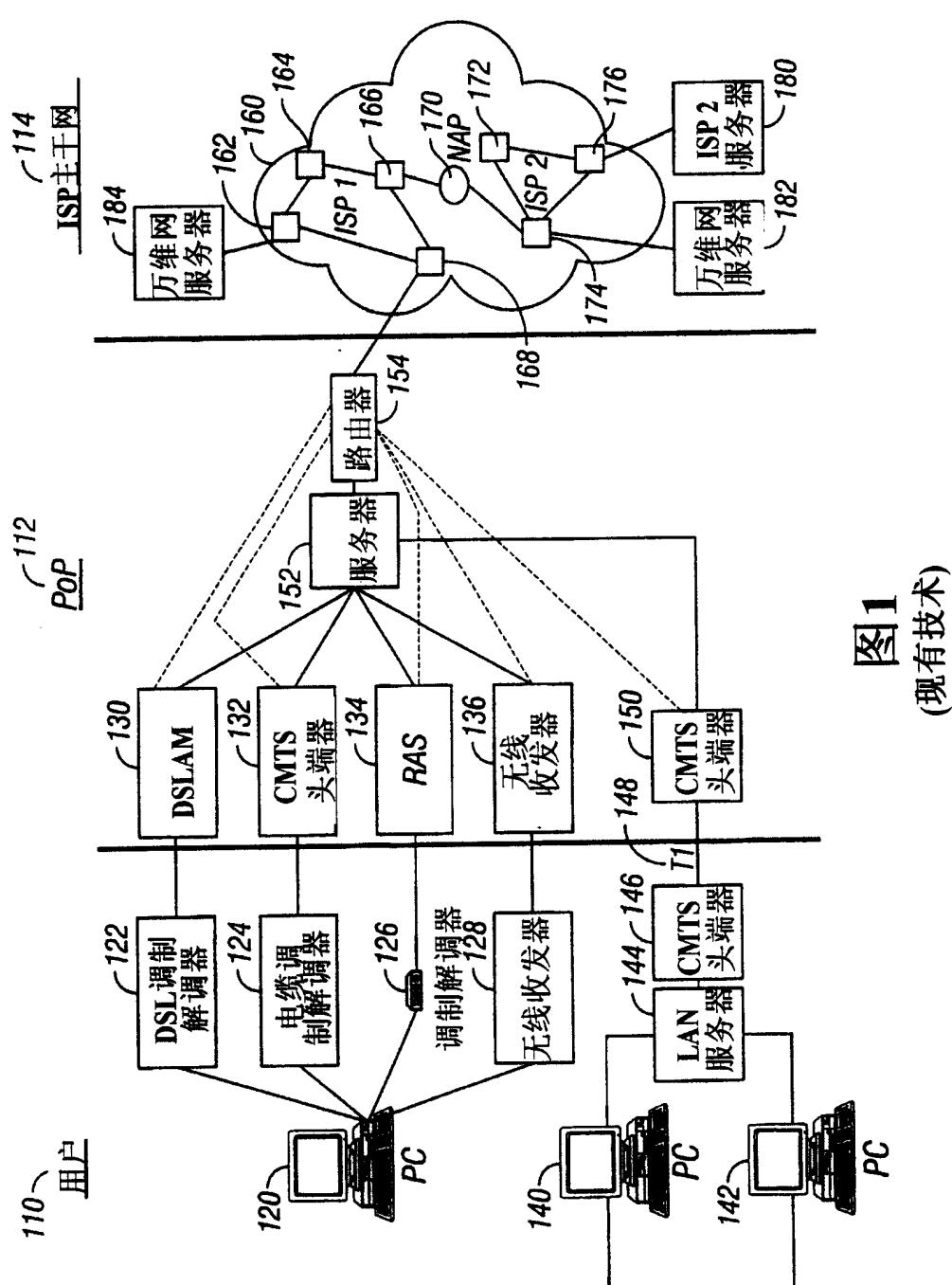
25

尽管已经描述了本发明的特定实施例，但是在本发明的范围也可以涵盖多种改进、变化、其它构造和等同物。所述的发明不限于在某一特定的数据处理环境中的操作，而是在多种数据处理环境中自由操作。另外，尽管使用特定的事务和步骤顺序描述本发明，对于本领域熟练技术人员可以理解，本发明不限于所述的事务和步骤的顺序。

30

进一步，尽管本发明是使用硬件和软件的特定组合描述的，但是应当理解，硬件和软件的其它组合也在本发明的范围之内。本发明可以仅以硬件或仅以软件或使用二者的结合而实现。

5 因此，说明书和附图应看作描述性的而不是限制性的。然而，很显然可以在不脱离权利要求所给出的本发明的更宽的精神和范围的前提下，对本发明作出增加、删剪和其它修改及变化。



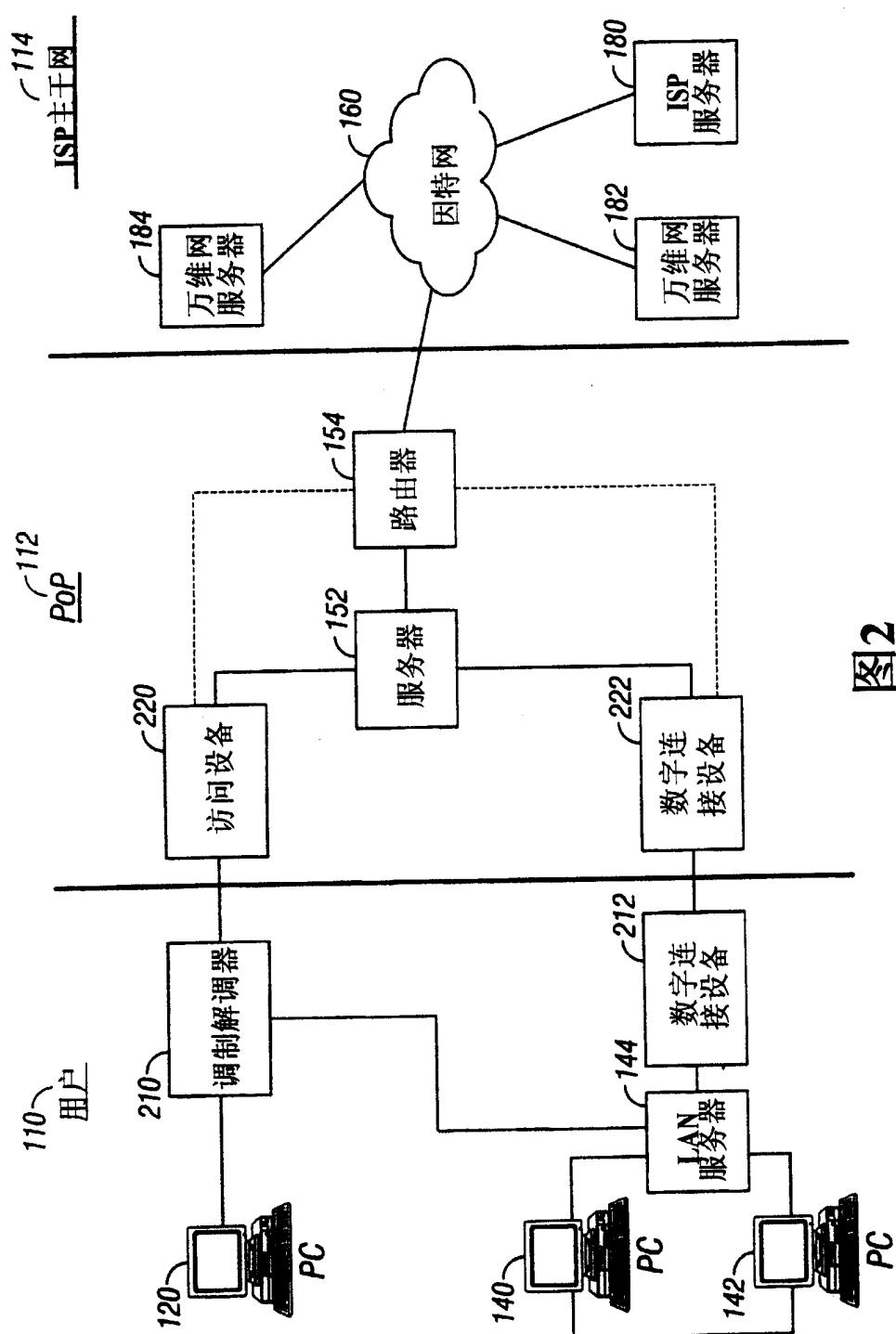


图2

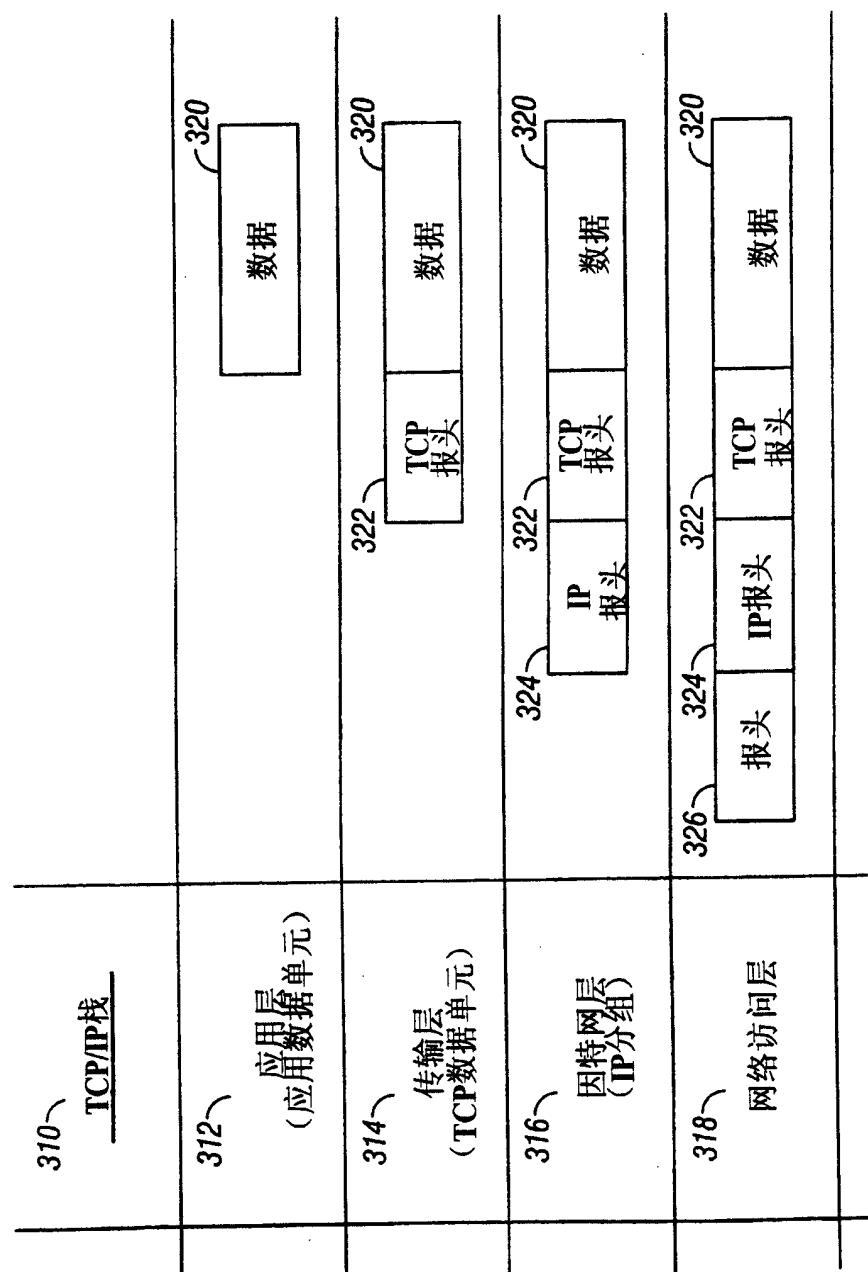


图3

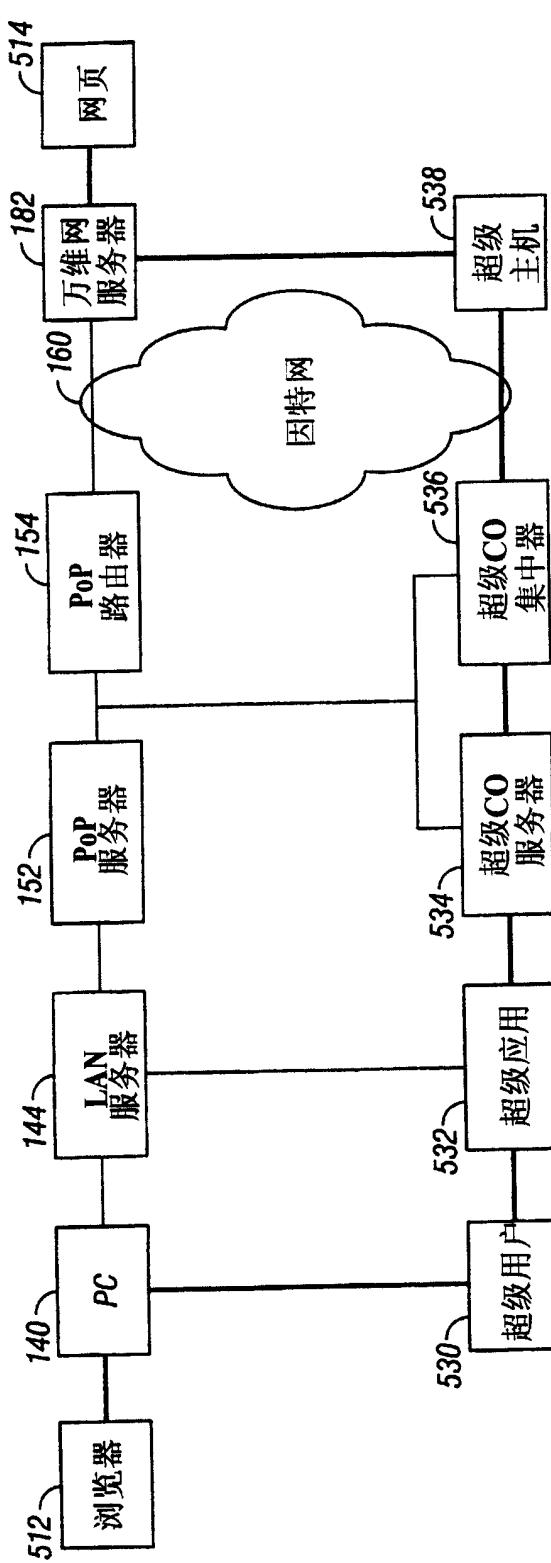


图4

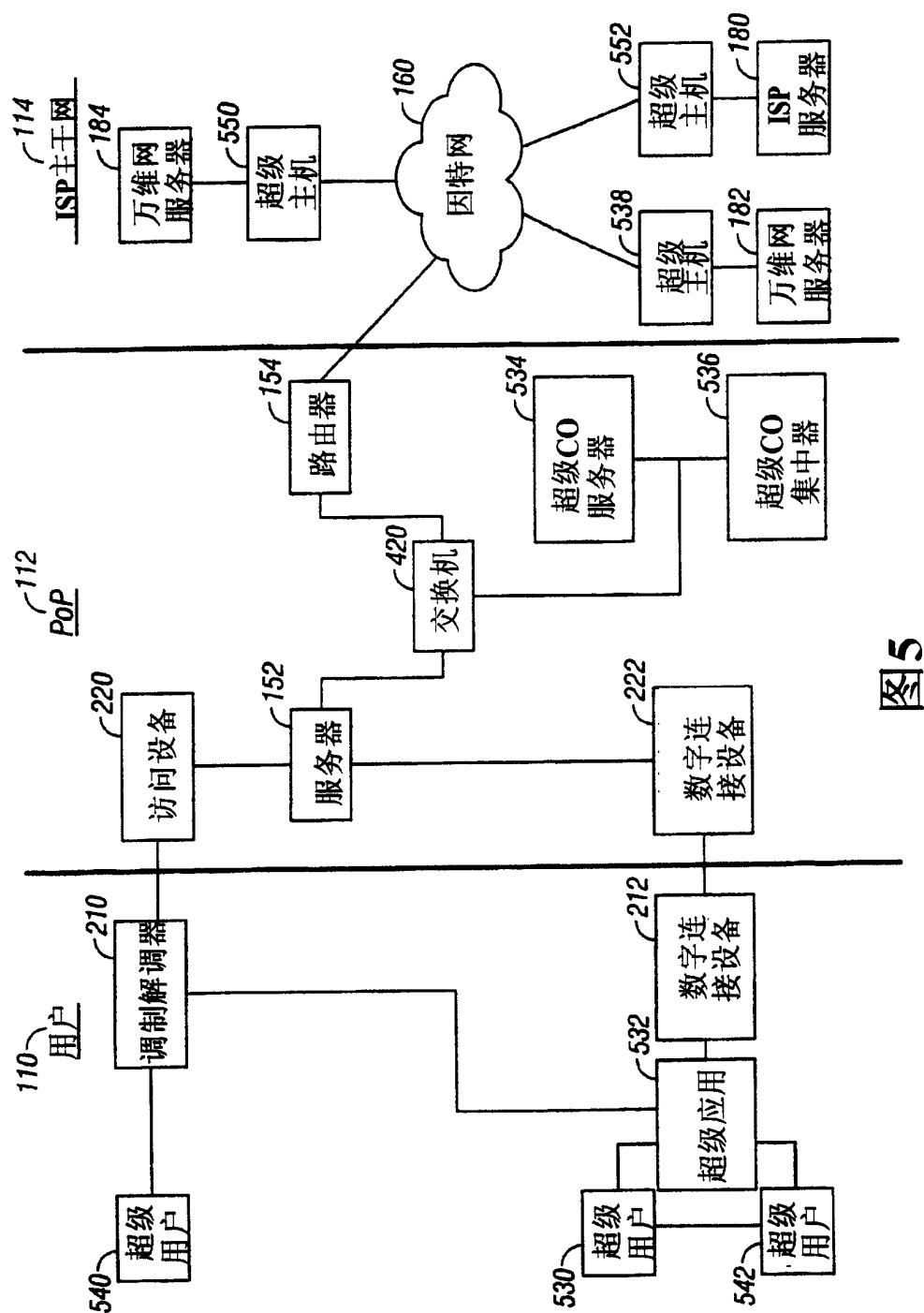


图5

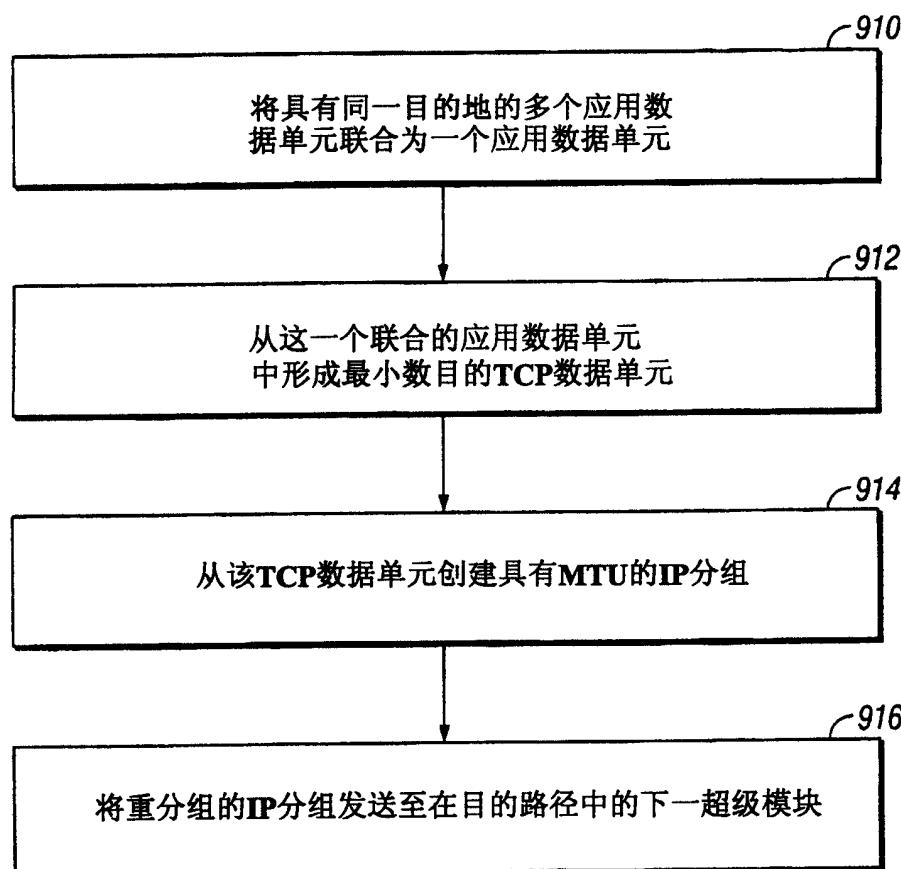


图6

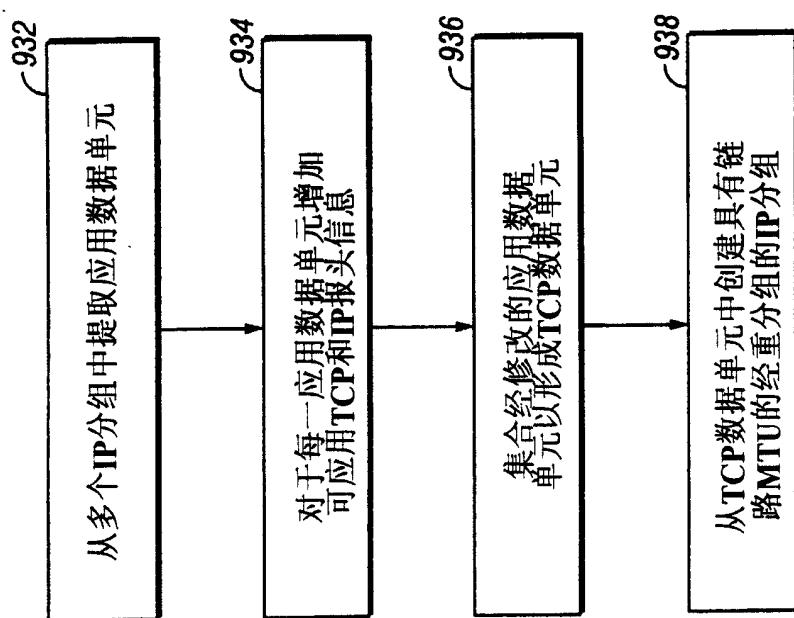


图8

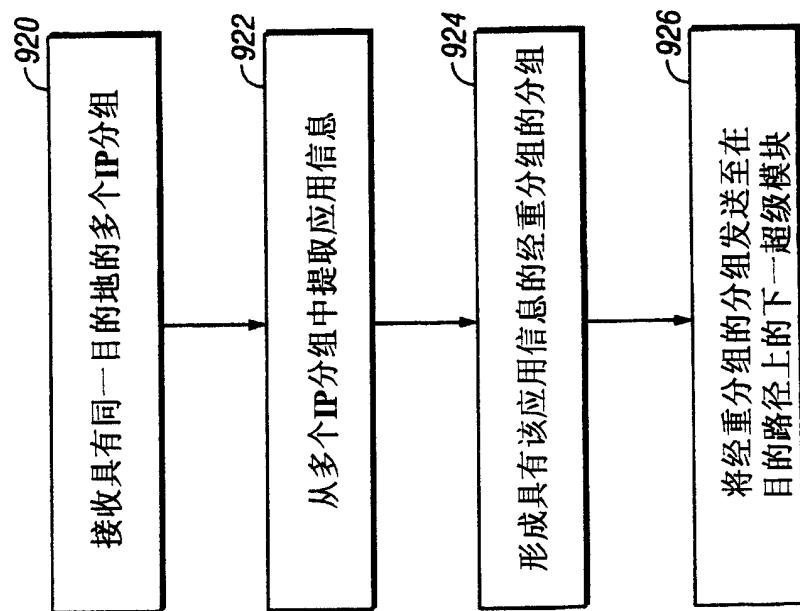


图7

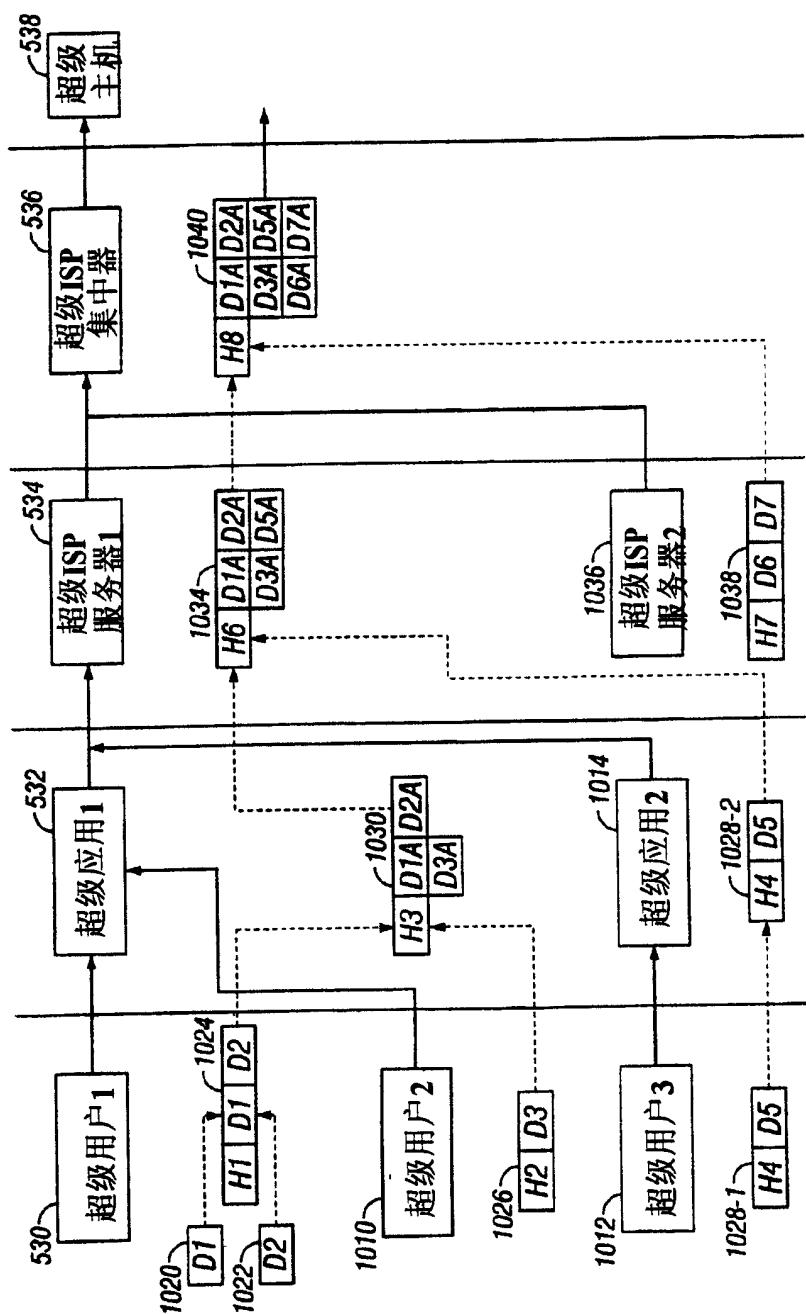


图9

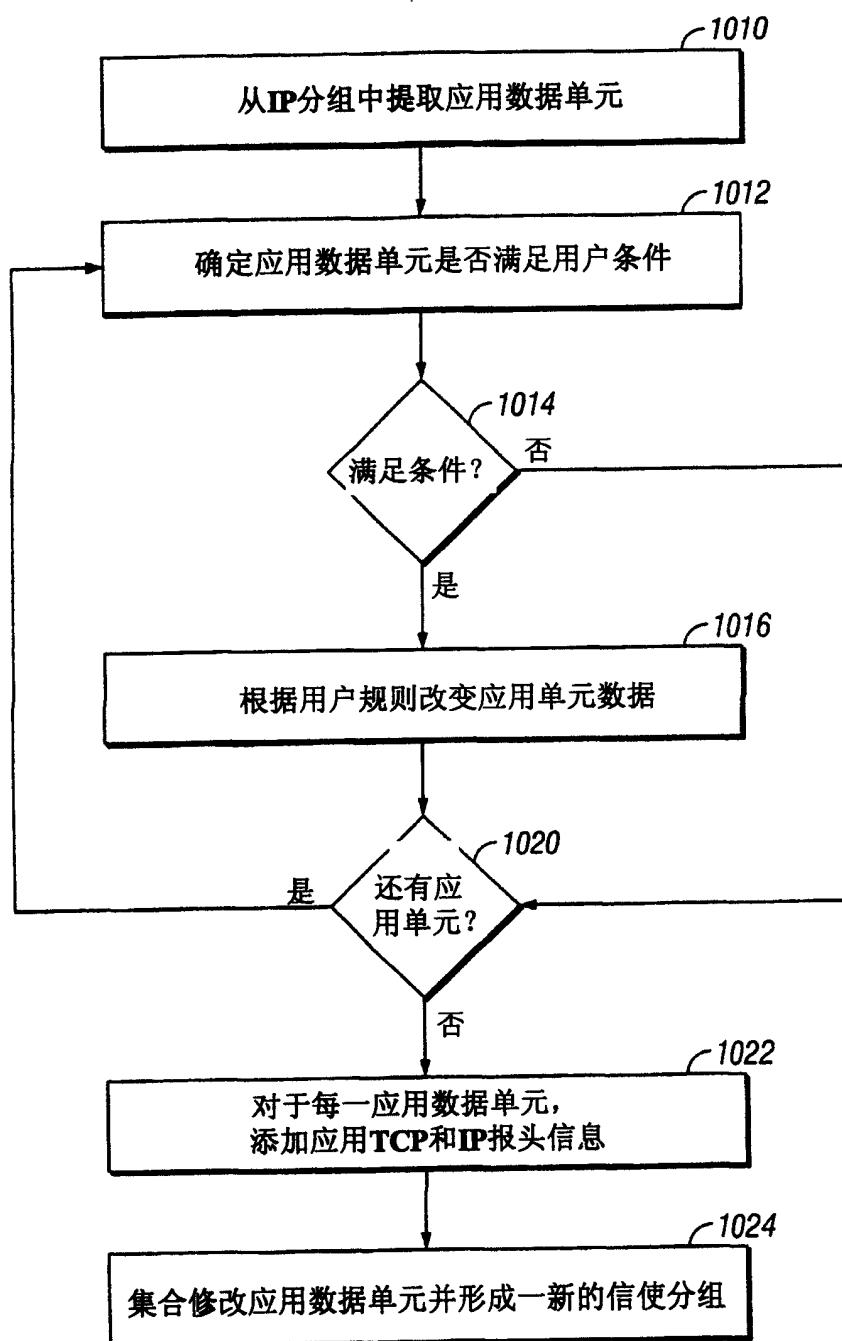


图10