

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 17/30 (2006.01)

H04L 29/06 (2006.01)



[12] 发明专利说明书

专利号 ZL 98106098.6

[45] 授权公告日 2007 年 1 月 10 日

[11] 授权公告号 CN 1294512C

[22] 申请日 1998.3.10 [21] 申请号 98106098.6

[30] 优先权

[32] 1997. 3. 26 [33] US [31] 827075

[73] 专利权人 国际商业机器公司

地址 美国纽约州

[72] 发明人 V·伯斯蒂斯

[56] 参考文献

TW297194A 1997.2.1 H04L12/66

US5611049 1997.3.11 G06F15/163

US5581704 1996.12.3 G06F13/00

审查员 谢 静

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王 勇 王忠忠

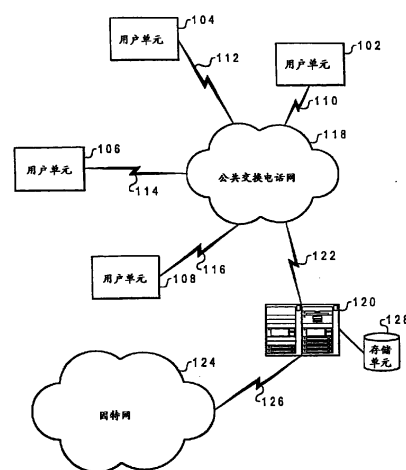
权利要求书 5 页 说明书 16 页 附图 7 页

[54] 发明名称

在服务器上预先高速缓存数据的方法和设备

[57] 摘要

在为多个用户提供访问环球网服务的服务器中，将 Web 站点上定期更新的选定页面预先高速缓存起来。有可能被用户访问的、与 Web 站点上主页链接的页面都被检索出来，存储在服务器上。当 Web 站点订户提出脱机浏览请求时，根据统计信息、链接关系以及/或者信息内容，将尚未出现在订户系统中的页面或页面部分，按照被订户访问的可能程度赋予优先级。最有可能被访问的页面或页面部分经压缩后被传送到该订户，使得所需的连接时间最少，能进行信息更新的用户数目最大。



1. 一种在数据处理系统中向多个用户有效地分配预先高速缓存的数据的方法，包括：

从某个系统接受数据请求；

确定尚未在所述某个系统中出现的被请求数据；

选择一部分要向所述某个系统传送的所述被确定的被请求数据来预先高速缓存，所述选择的根据是用户将访问所述被确定的被请求数据的所述被选择部分的可能性，并且所述被确定的被请求数据的所选择的部分受一个长度的限制；

将预先高速缓存的所述被确定的被请求数据的所述被选择部分传送给所述某个系统，其中，预先高速缓存数据可被有效地分配到多个用户。

2. 权利要求 1 的方法，其特征在于，所述从某个系统接受数据请求的步骤进一步包括：

接受对某个 Web 站点进行非高峰信息检索的请求。

3. 权利要求 1 的方法，其特征在于，所述从某个系统接受数据请求的步骤进一步包括：

从要求对多个 Web 站点进行浏览的请求的登记表中接受一个对某个 Web 站点的浏览请求。

4. 权利要求 1 的方法，其特征在于进一步包括：

将要求对多个 Web 站点进行浏览的请求收集成一个登记表。

5. 权利要求 4 的方法，其特征在于，所述将要求对多个 Web 站点进行浏览的请求收集成一个登记表的步骤进一步包括：

从所述登记表中剔除用户放弃的浏览请求。

6. 权利要求 1 的方法，其特征在于进一步包括：

为了向所述某个系统传送而将所述被确定的被请求数据的所述被选择部分进行压缩。

7. 权利要求 1 的方法，其特征在于，所述从某个系统接受数据请求的步骤进一步包括：

接受一个要求某个 Web 站点的数据页面的请求。

8. 权利要求 1 的方法，其特征在于：

所述从某个系统接受数据请求的步骤进一步包括，接受一个要求

某个 Web 站点的数据的请求;

所述为了向所述某个系统传送而选择所述被确定的被请求数据的一部分的步骤进一步包括, 为了向所述某个系统传送, 选择尚未出现在所述某个系统的所述 Web 站点的一部分所述数据。

9. 权利要求 8 的方法, 其特征在于, 所述选择所述 Web 站点的一部分所述数据的步骤进一步包括:

选择尚未出现在所述某个系统中、并且有可能被某个用户访问的所述 Web 站点的完整页面, 数据超过长度限制的完整页面除外。

10. 权利要求 8 的方法, 其特征在于, 所述选择所述 Web 站点的一部分所述数据的步骤进一步包括:

选择尚未出现在所述某个系统中、并且有可能被某个用户访问的所述 Web 站点的完整页面, 页面大小不超过从一组文件数量、字节数量和时间限制中选择的大小约束值。

11. 权利要求 1 的方法, 其特征在于:

所述接受某个系统的数据请求的步骤进一步包括, 接受一个对某个 Web 站点的至少一个页面的请求;

所述选择一部分所述被确定的被请求数据用于向所述某个系统传送的步骤进一步包括, 选择与所述至少一个数据页面链接的页面用于向所述某个系统传送。

12. 权利要求 1 的方法, 其特征在于,

所述从某个系统接受数据请求的步骤进一步包括, 接受一个对某个 Web 站点的一个页面的请求;

所述确定尚未出现在所述某个系统的被请求数据的步骤进一步包括, 确定尚未出现在所述某个系统的、与所述 Web 站点上被请求页面链接的页面;

所述选择一部分所述被确定的被请求数据用于向所述某个系统传送的步骤进一步包括, 选择有可能被用户访问的、与所述 Web 站点上被请求页面链接的页面。

13. 权利要求 1 的方法, 其特征在于进一步包括:

过一段时间后, 从所述某个系统接受第二个对所述数据的请求;

从所述被确定的被请求数据中确定尚未在所述某个系统中出现的其余数据;

从所述其余数据中选择一部分用于向所述某个系统传送；
以压缩形式将所述被选择的其余数据传送到所述某个系统。

14. 权利要求 1 的方法，其特征在于：

所述从某个系统接受数据请求的步骤进一步包括，接受一个对某个 Web 站点的至少一个页面的请求；

所述确定被请求数据的步骤进一步包括，确定尚未出现在所述某个系统的、与所述至少一个页面链接的页面的部分；

所述选择一部分所述被确定的被请求数据的步骤进一步包括，选择尚未出现在所述某个系统的、有可能被某个用户访问的与所述至少一个页面链接的页面的部分；以及

所述传送所述被选择的所述被确定的被请求数据部分的步骤进一步包括，将所述被选择的页面的部分传送到所述某个系统。

15. 一种向多个用户有效地分配预先高速缓存数据的设备，包括：

从某个系统接受数据请求的接收装置；

确定尚未在所述系统中出现的被请求数据的确定装置；

选择一部分要向所述系统传送的所述被确定的被请求数据来预先高速缓存的选择装置，所述选择的根据是用户将访问所述被确定的被请求数据的所述被选择部分的可能性，并且所述被确定的被请求数据的选择部分受一个长度的限制；

将预先高速缓存的所述被确定的被请求数据的所述被选择部分传送给所述系统的传送装置，其中，预先高速缓存数据可被所述设备有效地分配到多个用户。

16. 权利要求 15 的设备，其特征在于进一步包括：

对要向所述系统传送的所述被确定的被请求数据的被选择部分进行压缩的压缩装置。

17. 权利要求 15 的设备，其特征在于，所述接收装置进一步包括：接受对某个 Web 站点的一个数据页面的请求的接收装置。

18. 权利要求 17 的设备，其特征在于：

所述接收装置进一步包括接受要求某个 Web 站点的数据的请求的接收装置。

所述选择装置进一步包括选择用于向所述系统传送的、尚未在所述系统中出现的所述 Web 站点的所述数据的一部分的装置。

19. 权利要求 18 的设备，其特征在于，所述选择装置进一步包括：选择尚未在所述系统中出现、并且有可能被某个用户访问的所述 Web 站点的页面的选择装置。

20. 权利要求 18 的设备，其特征在于：

所述接收装置进一步包括，接受要求某个 Web 站点至少一个页面的请求的装置；以及

所述选择装置进一步包括选择用于向所述系统传送的、与所述至少一个数据页面链接的页面的装置。

21. 权利要求 18 的设备，其特征在于：

所述接收装置进一步包括，接受要求某个 Web 站点的一个页面的请求的装置；

所述确定装置进一步包括，确定尚未出现在所述系统的、与所述 Web 站点上所述被请求页面链接的页面的装置；

所述选择装置进一步包括，选择有可能被用户访问的、与所述 Web 站点上被请求页面链接的页面的装置。

22. 权利要求 18 的设备，其特征在于进一步包括：

过一段时间后，从所述系统接受第二个对所述数据的请求的接收装置；

从所述被确定的被请求数据中确定尚未在所述系统出现的其余数据的确定装置；

从所述其余数据中选择一部分用于向所述系统传送的选择装置；

以压缩形式将所述被选择的其余数据传送到所述系统的传送装置。

23. 权利要求 18 的设备，其特征在于：

所述接受装置进一步包括，接受一个对某个 Web 站点的至少一个页面的请求的装置；

所述确定装置进一步包括，确定尚未在所述系统中出现的、与所述至少一个页面链接的页面的部分的装置；

所述选择装置进一步包括，选择尚未在所述系统中出现的、有可能被某个用户访问的与所述至少一个页面链接的页面的部分的装置；以及

所述传送装置进一步包括传送所述被确定的被请求数据的所述被

选择部分的装置，进一步包括将所述被选择的页面的部分传送到所述系统的装置。

24. 一种向多个用户有效地分配数据的方法，包括：

接受多个用户的数据请求；

根据所收到的请求，从较大范围的可用数据中选择有可能被该多个用户访问的数据；

将所选择的数据预先高速缓存到服务器系统中；

当该多个用户中有用户连接到该服务器系统时，将该预先高速缓存的数据的一部分传送到该用户，选择该部分预先高速缓存的数据的依据，是事先收到的来自该用户的请求。

在服务器上预先高速缓存数据的方法和设备

技术领域

本发明总体来说涉及数据处理系统，特别是数据处理系统的数据分配机制。更具体来说，本发明涉及一种从因特网向许多数据处理系统进行信息的数据分配的机制。

背景技术

因特网已经变成作为一种信息和娱乐来源的文化事物。许多商家正在建立 Web 站点，将其作为整个市场营销努力中的一个部分，向消费者传播商家所提供的产品或劳务信息，或者提供其它信息，以期培养消费者的品牌忠诚度。许多联邦、州、地方的政府机构也正在将因特网站点用于信息目的，特别是那些几乎必须与社会各方面打交道的机构，诸如国内税收服务（Internal Revenue Services）和州秘书部门。通过提供联机公开文件信息指导和可检索数据库，可以节省运作费用。

当前在因特网上传输数据最普遍采用的方法是采用环球网（World Wide Web，简称 Web）环境。其它用于信息传输的因特网资源，例如有文件传输协议（File Transfre Protocol，简称 FTP）和 Gopher，但它们都不及 Web 用得普遍。在 Web 环境中，服务器和客户机使用超级文本传输协议（HTTP）进行数据事务处理。HTTP 是一种有名的处理各种不同数据文件（例如文本、静态图像、声音、动态画面等等）传输的协议。信息是由一种叫超文本标记语言（Hypertext Markup Language，简称 HTML）的标准页面描述语言格式化后向用户表示的。除了进行基本的表示格式化外，HTML 还允许开发商规定“链接”，连接到用统一资源定位器（Uniform Resource Locator，简称 URL）标识的其它 Web 资源。URL 是一种特殊的语义标识符，它定义通往特定信息的通讯路径。用户可访问的被称为“页面”的各逻辑信息块由 URL 标识。

Web 上的信息检索一般是用客户机上与 HTML 兼容的“浏览器”完成的。浏览器是个程序，它能提交对由 URL 标识的信息的信息请求。请求被提交给连接客户机的服务器，可以由一系列服务器进行处理，进行对所请求信息的检索。检索到的信息按照 HTML 格式化后被提供给客

户。

目前大多数公众消费者还不能访问这些 Web 资源。一般来说，要这部分消费者为利用这些资源去购买必须的硬件和软件，掌握必要的计算机使用技术，他们要么是没有能力，要么是没有兴趣。虽然目前多数销售的计算机都预装了因特网入网设施，包括 Web 浏览器，还有相当数量的家庭没有个人计算机。市场需求成本低廉、操作简单的数据处理系统，使对计算机不熟练的用户也有访问因特网的机会。在某种程度上，例如“Web 电视”（WebTV）等“顶置”（“set-top”）系统满足了这种需求。这些系统能迅速转换电视机的功能，即接收常规无线或有线电视节目功能和作为访问因特网的用户接口功能。用户电视因此变成 Web 设备的组成部分。

然而，在为 Web 设备设计成本低廉、操作简单的数据处理系统时，必须假设目标用户是知识肤浅、无经验的用户。因此，数据处理系统的操作必须既简单又有指导性，用户无需具有或只需较少的技术知识。为此，常规 Web 浏览器的许多性能必须改变，使其在 Web 设备上实现时对用户透明。

Web 浏览器在结合 Web 设备实现时特别有益的一个性能是脱机浏览（off-line browsing）。如果对特定 Web 站点需求的通信量很大，对这些站点的访问就可能有困难。脱机浏览能使客户机在没有同时用户交互作用的情况下在非高峰期间检索出站点的信息，供用户随后进行脱机浏览。脱机浏览过程就是浏览被高速缓存在例如是硬盘的本地存储器中的 Web 页面，期间无需与这些页面的源 Web 站点连接。页面一般是在非高峰期间从源 Web 站点检索的，或是在该站点的通信量最小的期间检索的。

一般来说，用户可以用调度实用程序来检索特定的 Web 页面，将其存储到用户的硬盘驱动器中供以后浏览。脱机浏览器对单个用户而言可能是有益的，但是它不能支持在一组客户与 Web 之间通信的最优化。各自采用非高峰期间检索的单个用户，如果连接的是相同的服务器或服务器群组，仍然会对通信资源造成压力。尤其是当有相当数量的 Web 设备通过一个服务提供商访问因特网时，就会发生这种情况。在这种环境中，脱机信息检索实际上受到的约束会使脱机浏览变得困难，除此之外，服务提供商也要对为脱机浏览进行的非高峰信息检索的时间分配加以限

制。

因此，就有必要提供一种自动的且更高效的从受欢迎的因特网站点向特定的用户群组下载信息的性能。多用户对非高峰信息检索的使用，即使有时间间隔，也会在服务器和因特网之间产生瓶颈，因而需要额外的资源来满足带宽条件。因此，还有必要提供一种机制来消除这种引发的带宽条件。这种机制的另外优点是尽量缩短了从信息源到单个用户的数据传输时间，并且对服务器要求的资源量最小。

发明内容

因此，本发明的一个目的是提供一种改进的数据处理系统。

本发明的另一个目的是提供一种用于数据处理系统的改进的数据分配机制。

本发明的另一个目的是提供一种改进的、从因特网向大量数据处理系统进行信息的数据分配的方法和机制。

现在描述上述目的的实现方法。在一个向多个用户提供访问环球网服务的服务器中，从定期更新的 Web 站点选择的页面被预先快速存储起来。有可能被用户访问的与 Web 站点主页面链接的诸页面被检索出来储存在服务器上。当该 Web 站点的预订用户提出脱机浏览请求时，服务器利用统计数据、链接关系和/或内容，将尚未在该预订用户系统出现的这些页面或这些页面的一部分，按被访问可能性赋予优先权。这些最有可能被访问的页面或部分页面经压缩后被传送到预订用户，这样就使所需的连接时间最小化，使可能被更新的预订用户的数量最大化。

本发明的上述目的以及其它目的、特点、优点，明显地表现在以下的详细说明中。

根据本发明第一方面提供一种在数据处理系统中向多个用户有效地分配预先高速缓存的数据的方法，包括：

从某个系统接受数据请求；

确定尚未在所述系统中出现的被请求数据；

选择一部分要向所述系统传送的所述被确定的被请求数据，所述选择的根据是用户将访问所述被确定的被请求数据的所述被选择部分的可能性，并且所述被确定的被请求数据的所选择的部分受一个长度的限制；

将所述被确定的被请求数据的所述被选择部分传送给所述系统，其中，预先高速缓存数据可被有效地分配到多个用户。

根据本发明第二个方面提供一种向多个用户有效地分配预先高速缓存数据的设备，包括：

从某个系统接受数据请求的接收装置；

确定尚未在所述系统中出现的被请求数据的确定装置；

选择一部分要向所述系统传送的所述被确定的被请求数据的选择装置，所述选择的根据是用户将访问所述被确定的被请求数据的所述被选择部分的可能性，并且所述被确定的被请求数据的选择部分受一个长度的限制；

将所述被确定的被请求数据的所述被选择部分传送给所述系统的传送装置，其中，预先高速缓存数据可被所述设备有效地分配到多个用户。

根据本发明第三方面提供一种预先高速缓存 Web 页面的方法，包括：

保存一个由被请求 Web 站点的主页面组成的登记表；

对于每个被请求 Web 站点的主页面建立一个高速缓存的 Web 站点；

当有要求所述登记表内某个 Web 站点的主页面的请求时，确定对应的高速缓存的 Web 站点中发生变化的部分；

传送对应的高速缓存的 Web 站点的所述发生变化的部分，其大小不超过一个大小限制条件。

根据本发明第四方面提供一种向多个用户有效地分配数据的方法，包括：

接受多个用户的数据请求；

根据所收到的请求，从较大范围的可用数据中选择有可能被该多个用户访问的数据；

将所选择的数据预先高速缓存到服务器系统中；

当该多个用户中有用户连接到该服务器系统时，将该预先高速缓存的数据的一部分传送到该用户，选择该部分预先高速缓存的数据的依据，是事先收到的来自该用户的请求。

附图说明

说明书后所附的权利要求中列举了被认为是本发明特点的新颖之处。然而，要最完整准确地理解本发明本身以及较佳实施方案及其其它目的和优点，要结合附图阅读以下对示例的详细描述。对各附图简要说明如下：

附图 1 是在其中实现本发明的一个分布式数据处理系统。

附图 2 是可以按照本发明一个较佳实施方案实现服务器功能的一个数据处理系统的方框图。

附图 3 是可以按照本发明较佳实施方案实现用户单元功能的一个数据处理系统的表示图。

附图 4 是可以按照本发明一个较佳实施方案实现用户单元功能的数

据处理系统的主要部件的方框图。

附图 5 是按照本发明一个较佳实施方案在服务器高速缓存数据过程的高级流程图。

附图 6 是按照本发明一个较佳实施方案向用户单元传送被高速缓存的下载数据过程的高级流程图。

附图 7 是按照本发明一个较佳实施方案在用户单元处理从服务器接收的被高速缓存的下载数据过程的高级流程图。

附图 8 是按照本发明一个较佳实施方案从 Web 站点或服务器高速缓存存储器检索数据过程的高级流程图。

具体实施方式

现在参考各附图，特别是附图 1，这是本发明的一个分布式数据处理系统在其中实现的示意图。用户单元 102、104、106、108 有通信链路 110、112、114、116，使这些用户单元能够访问公共交换电话网（PSTN）118。用户单元通过这些通信链路与服务服务器 120 通信，后者通过通信链路 122 与 PSTN 118 连接。服务器 120 通过通信链路 126 使用户单元 102 ~ 108 能够访问因特网 124。服务器 120 除了提供用户单元 102 ~ 108 对因特网 124 的访问之外，还能在存储设备（SD）128 上储存各种配置信息、口令、电子邮件报文和备份数据。用户单元 102 ~ 108 可以位于远程地理位置，例如加利福尼亚或纽约。此外，用户单元 102 ~ 108 还可以位于地球上的其它大陆。

参照附图 2，这是可以按照本发明实现例如附图 1 中服务器 120 的服务器功能的一个数据处理系统的方框图。数据处理系统 200 可以是一个包括了连接到系统总线 206 的多个处理器 202 和 204 的对称多处理器（SMP）系统。与系统总线 206 相连的还有存储器控制器/高速缓存 208，它提供与本地存储器 209 连接的接口。输入/输出（I/O）总线桥 210 与系统总线 206 相连，提供到 I/O 总线 212 的接口。如图所示，存储器控制器/高速缓存 208 与 I/O 总线桥 210 可以被集成在一起。

与 I/O 总线 212 连接的外围部件互连（PCI）总线桥 214 提供到 PCI 总线 216 的接口。可以有多个调制解调器 218 ~ 220 与 PCI 总线 216 相连。典型的 PCI 总线设备支持四个 PCI 扩展槽或增加（add-in）连接器。用调制解调器 218 ~ 220 通过增加板连接到 PCI 本地总线 216，可以实现连接到附图 1 中 PSTN 118 的通信链路。图内示例中的调制解调器 218 ~ 220 也提供了向附图 1 中所示的因特网 124 的连接。

其它 PCI 总线桥 222、224 为其它 PCI 总线 226、228 提供接口，从这些总线也可以支持其它调制解调器。这种方式下，服务器 200 允许多个用户单元同时拨号。如图所示，内存映射的图形适配器 230 和硬盘 232 也可以直接或间接地与 I/O 总线 212 连接。

本领域的普通技术人员知道，附图 2 中描绘的硬件是可以有所变化的。例如，可以增添光盘驱动器之类的外围设备，或以其代替图中的硬件。图内所示例子并非意味着对本发明的体系结构作出限制。

附图 2 中的数据处理系统例如可以是国际商用机器 (IBM) 公司 (该公司位于纽约 Armonk) 的生产的 IBM RISC/System 6000 系统，该系统运行高级交互执行 (Advanced Interactive Executive, 简称 AIX) 操作系统。该数据处理系统提供一个 Web 服务器平台，并且可以作为因特网服务提供商使用的一组互连服务器中的一个，为 Web 客户或用户单元提供对因特网的访问。

本发明较佳实施例中的服务器数据处理系统含有一个 Web 服务器程序，例如 (美国加州 Mountain View 的) 网景通讯公司 (Netscape Communications Corporation) 出品的 Netscape Enterprise Server (2.0 版)，该程序支持接口扩展。服务器因此含有一组已知的服务器应用功能软件 (SAF)，它们接受客户请求以及配置信息，并返回一个响应。服务器也可以包括一个提供扩展功能的应用程序接口 (API)，使应用程序开发人员能够通过通常称作“插件” (plug-ins) 的软件程序来扩充或定制服务器应用功能软件 (SAF)。服务器支持客户的脱机浏览，提供预先高速缓存 Web 页面的存储器。服务器还执行和/或支持下文描述的、由客户对预先高速缓存和脱机下载的 Web 页面进行选择的过程。

现在参见附图 3，图示的数据处理系统可以实现用户单元功能，例如附图 1 中用户单元 102 - 108 的功能。附图 3 是数据处理系统的整体图示。图例中的数据处理系统 300，以对用户花费最少的硬件，提供对因特网的访问。数据处理系统 300 包括一个数据处理单元。数据处理单元 302 的大小与典型的娱乐中心环境最好恰当匹配，具有一般个人计算机均具有的、使用户能浏览因特网所需的全部功能。此外，数据处理单元 302 也可以有其它的一般功能，例如用作应答机、发送或接收传真、提供语音信箱功能。

作为 Web 设备, 数据处理单元 302 与用于显示图形信息的电视机 304 连接。电视机 304 可以是任何合适的电视机, 不过带 S-视频输入的彩色电视机会更好地表现图形信息。数据处理单元 302 可以通过一根标准同轴电缆线与电视机 304 相连。用户可以用遥控器 306 与数据处理单元 302 交互作用, 对其进行控制。遥控器 306 发出的红外 (IR) 信号, 最好以与通常的电视机、音响设备、录像机的红外遥控频率不同的频率进行调制, 以避免干扰。遥控器 306 具有常规个人计算机中的指点设备的功能, 能移动显示器上的光标, 选择项目。

现在参见附图 4, 该图是本发明较佳实施例中数据处理单元 302 中主要部件的方框图。与普通个人电脑一样, 数据处理单元 302 有一块主板 402, 主板上的处理器 404 和内存 406 与系统总线 408 相连。处理器 404 最好是一个主频不低于 100MHz 的 486 以上的处理器。内存 406 中有一个包括基本输入/输出服务 (BIOS) 例行程序的只读存储器 (ROM) 406a, 还可以有高速缓存和/或视频 RAM。

主板 402 上连接系统总线 408 的视频/TV 转换器 410 产生计算机监视器用的计算机视频信号、复合电视信号以及 S-视频信号。视频/TV 转换器 410 的功能可以用市售的视频和转换器芯片来实现。主板 402 上的键盘/遥控器控制接口单元 412 通过控制器 414 接收键盘代码, 无论使用的是有线的键盘或指点设备还是红外键盘或遥控器。红外遥控器 306 发送的信号, 如同普通鼠标器或指点设备的移动所产生的控制信号, 最终被传送到串行口。遥控器 306 上有两个按钮具有与普通鼠标器上的两个按钮完全相同的功能, 其它按钮则发送与红外键盘的按键对应的信号。这样, 遥控器 306 就具有红外键盘提供的功能的子集。主板 402 上的连接器/指示器 416 提供上述数据处理单元 302 上的连接和指示器。

图内示例中主板 402 外部连接有电源 418、硬盘驱动器 420、调制解调器 422 和扬声器 424。电源 418 能从控制器 414 接受控制信号, 使主板 402、硬盘驱动器 420 和调制解调器 422 断电, 除此之外和普通电源一样。在一些恢复情况中, 关闭电源后再重新启动是将这些设备复位到已知状态的唯一有效的办法。因此, 电源 418 接到控制器 414 的信号时, 能关闭电源, 重新启动数据处理单元 302。

硬驱 420 中含有操作系统和数据处理单元 302 的应用软件, 最好包括 IBM 公司的 IBM DOS 7.0、Microsoft 公司的 Windows 3.1 和

Netscape 通讯公司的 Netscape Navigator。硬驱 420 上也可以存储数据。插在主板 402 边侧插槽内的调制解调器 422 最好是 33.6kbps 的调制解调器，支持 V.42bis、v34bis、v.34、v.17 传真、MNP 1-5 和 AT 指令集。硬驱 420 也可以存储数据，例如用户喜欢的因特网站点清单或来自因特网站点的尚未阅览的下载信息。此外，硬驱 420 还含有为建立与服务提供商的通信链路和启动数据处理系统的配置过程所必须有的指令。

控制器 414 最好是一个或多个 805x 系列的控制器。控制器 414 的电源不间断，当数据处理单元 302 被启动时，控制器监测系统是否定时发出一个表明数据处理单元 302 正常工作的信号(ping)。如果控制器在一定的时间内收不到这种信号(ping)，控制器就使系统断电，并重新启动系统。例如在系统发生一般性的保护错误时，这种控制是必要的。如果多次重新启动系统都不成功，控制器 414 就要关闭数据处理单元 302，并通过指示器 416 发出要求维修的信号。这样，数据处理单元 302 在某些情况下无需用户参与就能进行故障的自我恢复。

控制器 414 也接收并处理来自红外遥控器 306、红外键盘、有线键盘或有线鼠标器的输入信号。当有一个键盘或指点设备被使用时，所有其它输入设备就被锁定（不予响应），直到在一定的时间段内没有输入设备工作。然后，最先动作的键盘或指点设备将其它输入设备锁定。控制器 414 还直接控制除了用于指示调制解调器的使用情况的所有发光二集管指示器和指示在电源关-开期间内自举分区选择。

本发明的技术人员知道，以上描述的附图 3 和 4 内的部件，在具体应用或实施中会有所变化。实施本发明中的这种变化可以认为不超出本发明的精神和范围。

现在参见附图 5，附图 5 是本发明一个较佳实施例中在服务器处预先高速缓存数据过程的高级流程图。该过程可以由 Web 服务器在非高峰期间执行，也可以作为后台应用程序连续一日或数日地运行。

图中该过程从步骤 502 开始。步骤 502 是根据服务器系统时钟或服务单元内的其它事件来定时启动或者说自动启动过程。下载启动的时序可以按用户单元进行自动下载的时间来协调安排。例如，如果该服务器的用户单元都设置了在凌晨 3:00 ~ 5:00 进行自动下载，就要将服务器启动该过程的时间设置在凌晨 1:00 左右，以便在用户单元进行下载之前，向服务器的必要下载能按时完成。一种备选方案是，各个自动下载可以重叠

进行，其中向服务器的下载在用户单元发出请求之前完成。该备选方案可能适于按后台应用程序运行的过程。

该过程接着到达步骤 504，从待被服务器高速缓存的数据登记表中选择一项。登记表的内容是客户或用户单元定期（例如每晚一次或每周一次）下载的信息的标识，以及当前登记的对应每个标识信息的用户数。例如，登记表可以包含例如被客户或客户频繁请求的各种 Web 页面的 URL 的列表。

登记表可以通过要求向登记表添加条目的特定用户请求来生成，可以通过监控从相同数据源定期传输的用户传输数据来生成，或者同时采用上述两种方法来生成。例如，可以利用 Web 数据要被预先高速缓存的服务器通过编辑客户发出的特定非高峰信息检索请求，或通过检查客户的 Web 浏览器的“书签”或“偏爱”列表，来生成登记表。另一种备选方案可以是，对客户的 Web 浏览器进行设置，使得向书签列表增加一个 URL 时，启动一次对用户关于将该 URL 添加到非高峰检索列表中的询问。

登记表内容的维护通过监测实际的用户传输并将它们与登记请求比较而进行。对于某数据项，当请求该项的用户超过一个门限时间（例如一个月）后还没有进行下载，就将该项的请求数减一。用这种方法可以将被请求用户放弃或不用的请求项从登记表中剔除。例如，用户单元可以监测下载信息是否被浏览过，如果过了一段时间，特定的下载信息还未被浏览，用户单元就可以停止对这种数据的自动下载。

从步骤 504，过程到达步骤 506。该步骤要确定登记定期下载选定项目的客户或用户单元数是否超过一个阈值。该阈值是根据预先高速缓存选定项目的下载信息的效率而确定的。例如，如果对于特定下载只有一个用户登记，那么就让用户直接下载所请求的数据，不要在服务器中预先高速缓存下载数据，这样做的效率可能会更高一些。登记表上各数据项的阈值的设定，可以取决于用户特别请求非高峰信息检索该数据项还是仅仅按照对该数据项请求的频率将数据项添加到登记表这两种不同的情况。

如果对于选定项，登记的用户数小于阈值，过程继续到步骤 507，确定是否全部条目都被检查完毕。如果是，过程继续到步骤 514（下文将要描述）；如果不是，过程返回到 504 去选择另一个数据项。如果对

于选定项，登记的用户数不小于阈值，过程继续到步骤 508，该步骤提取由该选定项标识的数据，压缩该数据并将其在服务器中存储起来。可以用任何合适的压缩应用程序来进行数据压缩。

在较佳实施例中，被提取的数据包括因特网上某 Web 站点出的 Web 页面。检索出的 Web 页面数用规则或试探的方法确定，决定哪些 Web 页面或子页面最有可能被用户访问。下载的 Web 页面，每个由多个文件或部分（components）组成。因此，作为该步骤的一部分，可以将以前下载的、存储在服务器上、但被更新过或不再在 Web 页面中被引述的下载 Web 页部分删除。可以将从不同站点获得的 Web 页进行压缩，并分别存储，以方便将每个用户所请求的页仅分发到该用户。因此，对于利用该服务器进行非高峰检索的用户来讲，每一个服务器对于所请求的特定 Web 站点起到一个镜像站点的作用。

然后该过程进到步骤 510，该步骤对于包含为客户下载的预先高速缓存在服务器中的项列表的下载表进行更新。该下载表可以包含其它信息，比如每个所下载的的部分的时间和日期。该过程下一步转移到步骤 512，确定是否登记表中的全部条目或项都被检查完毕。如果不是，该过程返回到步骤 504 去选择登记表中的另一项；但如果是，该过程就继续到步骤 514，表示过程进入空闲状态，等待下个定时的预先高速缓存下载的启动。

参见附图 6，该图是按照本发明较佳实施方案向用户单元（或 Web 客户）传送被预先高速缓存的下载数据过程的高级流程图。过程从步骤 602 开始。步骤 602 表示服务器接收下载请求，例如对特定 Web 页面的自动、定时的非高峰检索请求，或者在拨号时进行的联机请求。该请求可以含有一个特定 Web 站点的 URL，下文对此将有所叙述。过程下一步转移到步骤 604，该步骤表示检查服务器中预先高速缓存下载信息的下载表。然后过程转移到步骤 606，确定被请求的下载信息是否已经被在服务器中预高速缓存。如果不是，该过程转移到步骤 608，将被请求的下载信息传送到适当的地点。然后，该过程转移到步骤 616，表示过程进入空闲状态，直到接收下一个下载请求。

如果被请求的下载信息已经被预高速缓存在服务器中，则该过程转移到步骤 610，服务器截取该下载请求。该过程下一步转移到步骤 612，从服务器向发出请求的用户单元或 Web 客户传送被请求的下载数据。被

请求数据存储在用户单元的本地存储器中，如硬盘驱动器。然后，该过程转移到步骤 614，用户单元自动解压所下载的数据。然后，该过程转移到步骤 616，表示该过程进入空闲状态，直到接收下一个下载请求。

现在参见附图 7，附图 7 是按照本发明较佳实施方案在用户单元处理从服务器接收的预先高速缓存的下载数据过程的高级流程图。该过程自步骤 702 开始，用户单元从服务器接收预先高速缓存的下载数据。然后，该过程转移到步骤 704，删除以前从服务器接收的性质类似的下载信息（即所下载的 Web 页部分的旧版本）。该过程下一步转移到步骤 706，在接收下载信息数据时对其进行数据解压，以便提高浏览速度。然后，该过程转移到步骤 708，表示该过程进入空闲状态，直到接收另一个下载。

参见附图 8，这是按照本发明较佳实施方案从 Web 站点或服务器高速缓冲存储器检索数据的过程的高级流程图。图中的全部或部分过程，用于从 Web 站点检索如附图 5 中步骤 508 所示在 Web 服务器预先高速缓存的信息，或者从 Web 服务器检索如附图 6 中步骤 612 所示在 Web 客户的硬驱中存储的信息。

从服务器预先高速缓存的信息的检索过程需要的不止是一次简单提取，而必须使其与非高峰信息检索的实际情况相适应。一般认为，用户的非高峰检索请求远远超出了为满足这种请求所具有的带宽、时间和资源条件。就是说，假设真要满足这些请求，那势必要耗费比实际上所具备的更多的资源。所以，检索预先高速缓存的信息的过程就必须分析请求，分析被请求的数据，智能化地提取数据。

Web 页面或者超文本文件，是通过标识着页面的通信源的 URL 被检索的。URL 的典型格式是“www.域名.扩展名/文件路径/文件名”（www.domainname.ext/filepath/filename）。域名和扩展名标识着含有被请求信息的特定 Web 站点（Web 域或服务器）。被请求信息中含有一个或一组在作为请求对象的 Web 站点上的目录内组织的文件。所以，URL 中必须有包含被请求信息的文件的路径，也可以要求一个文件名。无需为文件名规定扩展名，这是因为只有 HTML 文件才能被浏览器显示，它们的缺省扩展名为“.html”或“.htm”。如果请求中不指定文件名，浏览器就在指定的 Web 域和路径检索一个名为“default”或“index”的 HTML 文件。

特定位置（或域）的 Web 页面含有一个或多个 HTML 文件以及与该 HTML 文件关联的图形、声音、动画和可执行的脚本文件（script file）。一个 HTML 文件构成了 Web 页面的一部分，它经常包含各种引用（reference），用于联系到图形文件 - 诸如 JPEG 或 GIF 格式的图形，声音文件 - 诸如 WAV 或 MID 格式的音频信息，动画文件 - 诸如 MPEG 格式的视频信息，以及/或者可执行的脚本文件 - 诸如 JAVA、JAVASCRIPT 或公共网关接口（CGI）等脚本文件。更为重要的是，HTML 文件一般含有包括 URL 在内的“链路”或内置的引用，用于“跳转”到（实际上是检索）其它 HTML 文件。这些其它 HTML 文件可以是本地文件（位于相同的 Web 域，不过路径可能不同），也可以是远程文件（位于不同的 Web 域或服务）。

用浏览器显示 Web 页面，需要检索出至少一个 HTML 文件，来对欲显示的 Web 页面和 HTML 文件中引用的每个图形、声音、动画和脚本文件进行格式化。此外，要有效地高速缓存一个 Web 页面，要求分解出该页内的链接，检索出用于显示的数据。链接中涉及的页面，可能本身也含有通往其它页面等的链接。按照这种方式，一次脱机浏览请求，可能会请求一个含有连接“树”的根部的页面，如果将该根部充分扩展开来，可以包含当前已经开放的每一个 Web 页面。此外，鉴于以上叙述的原因，客户的非高峰检索时间可能会受到限制。种种这些原因表明，从服务器的高速缓存中检索数据或向服务器的高速缓存存储检索数据的过程，必须采取智能方法来进行。

过程从步骤 802 开始，其接受一个信息提取的请求。提取请求可以由服务器在客户的非高峰检索请求中接收，或者作为检索服务器高速缓存数据过程的一部分来接收。然后，该过程转移到步骤 804，确定根据该请求所要检索或传输的是什么信息。对从 Web 站点检索的高速缓存到服务器的信息或者从服务器的高速缓存向客户传输的信息进行选择时采用的是一组规则或试探方法，以确定最可能满足用户需要的文件，而不会过分地增加可用资源的负担。

在服务器高速缓存情形中，确定要检索哪些在服务器中要被高速缓存的 Web 站点文件的规则根据两个目标而制订，目标之一是可能满足用户需求的页面的跨度，目标之二是对于一总的高速缓存空间，将高速缓存数据的长度限制在恰当的范围。检索与某 Web 页面关联的范

围较广的文件，对于沿着链接的特定序列或路径检索不要投入不合理数量的系统资源。

例如，如果用户对特定的新闻页面站点感兴趣，系统一开始会检索该新闻站点的首页或缺省 HTML 文件以及该文件内引用的图形、脚本等等。包含缺省 HTML 文件及其密切关联的图形、脚本等的文件组有时被称作“主页”。一般来说，可以将 HTML 文件连同该 HTML 文件的文本内引用的图形、声音、动画和/或脚本文件称作“页面”。相同的图形、声音、动画和/或脚本文件可以被一个以上的 HTML 文件关联或引用，因此会在一个以上的页面中出现。页面内对独立的 HTML 文件或链接的引用与对图形、声音、动画和/或脚本文件的引用不同，是对不相同的页面的引用。一般来说，从 Web 站点检索用于高速缓存的信息时，最好检索完整的页面而不是一部分页面（例如忽略声音文件），除非数据长度会超出限制范围（例如页面中有长度异乎寻常大的动画文件）。

下一步，系统就要分解新闻站点的主页内的链接，顺着这些通往其它页面的链接去检索这些页面。顺着这些二级页面内的链接，过程继续按递归方式运行，直到将从主页开始的连接树穷尽检索完，或者直到检索量达到阈值。阈值可以被规定为检索文件时的文件数量、字节检索量或者检索时花费的时间量。因为大多数情况下，在连接树被搜寻完之前就会达到所选阈值，所以有必要提供一种能在连接树内检索最有用的页面的方法。

许多页面都有能监测访问该特定页面的用户数量的功能。因此可以利用这类统计数据来识别站点上人们喜欢检索的最受欢迎区域。有一种“宽度优先”的检索系统，它可以结合统计数据或代替统计数据，检索最受欢迎的页面。“宽度优先”的系统在连接树的一定层次上，首先检索该层次上的所有页面或者一定数量的页面，然后搜寻连接树的下一层次。这与“深度优先”的检索系统不同，“深度优先”的系统在特定路径上每一层，先搜寻尽连接树内指定路径上的所有层次，然后再从该路径寻找处于不同层次的分支。读者可以参考 J.Thompson 和 V.Berstis 发明的共同未决美国专利申请“对非高峰数据高速缓存进行最优化的方法”中介绍了一种高速缓存的“宽度优先”方法。该专利的受让人与本发明的受让人相同，系列号：08/797,902，申请日：97年2月10日，审查号：AT997048。

还有一个可以与上述系统结合使用的关于页面检索的规则，即在检索位于各个不同站点的页面之前，优先检索具有相同站点（即页面 URL 内相同域名标识的站点）的页面。该规则还可以延伸为，在给定站点上，在检索各个不同目录内的页面之前，优先检索位于相同目录内的页面。基于站点的优先规则能在检索中过滤掉所谓的“超链接”页面。超链接页面中含有连接 Web 站点的大量链接，经常起着搜寻特定感兴趣区域的资源定位器的作用。按照超链接页面内的所有链接检索会很快消耗可用的系统资源。也可以结合基于站点的优先，使用页面连接计数，忽略掉含有超过一个阈值数（例如 100）的页面链接。要是将如此众多的链路分解开来，还是会很快消耗可用的资源。

在客户检索的情形中，确定在非高峰信息检索期间内应向客户传送哪些信息的规则，是建立在向用户提供感兴趣的页面同时不超过被分配资源的阈值这些类似目标之上的。非高峰检索的用户非常重视连接时间的最小化，在向用户下载高速缓存信息的情况下也是如此。这就得以公平地为大量客户提供高质量的服务。从这个角度来说，对非高峰检索的时间和带宽可以限制得更严一些。

由于上下文之间的不同，确定应向客户传送哪些信息的规则也有所不同。例如，如果客户以前从未下载过特定的高速缓存的站点，就必须下载该站点的全部内容（每一页）。如果限制用户每晚只能有一个小时的非高峰信息检索时间，则不可能在一个晚上将所有页面下载到客户。下载一个特定站点可能需要几个晚上。因此，必须为下载的内容建立一种优先系统。

必须记住，要将用户请求的站点的 Web 页面有效地反映在用户单元供脱机浏览。要区分第一次下载新站点的用户和只要求进行内容更新的老客户。对于大型 Web 站点，有可能不能在一个晚上向一个新用户下载其请求的站点上的所有请求文件。相反，向该新用户下载该 Web 站点文件可能需要数个晚上。在这种情况下，最好不要下载完整的页面，使用户在等待非高峰检索用于脱机浏览信息的同时，还能阅览该请求 Web 站点的部分内容。例如，可以将与页面关联的声音文件的优先级，视为低于 HTML 及关联的图形文件的优先级。这样，可以将给定站点上所有声音文件赋予最低的下级优先级，它们在最后一个晚上才被下载。此外，可以根据统计，选择采用以上叙述的宽度优先、基于站点优先或基于目

录优先, 来确定选择哪些页面进行下载。

在客户下载的情形中, 连接时间最小化的目标要求, 也增加了对信息重叠的考虑。例如, 如果特定站点上许多页面都使用相同的图形文件, 比如标识图 (logo image), 则可以赋予该图形文件一个比该站点上其它文件高的下载优先级。采用这种做法的另一种情况是, 客户已经下载了给定站点上的信息, 仅仅需要进行更新, 例如有相同广告内容的每日新闻站点。只要将服务器高速缓存中的文件的日期及数据长度与客户系统的内容进行比较, 就能知道哪些信息发生变化, 必须为客户进行更新。另一种方法是, 保留以前下载到客户的文件清单。这种做法更好, 这是因为不管怎样都要保留为客户所提供服务的清单。这种清单也可作为更新要高速缓存的站点登记表、挑选以前向客户下载过的文件的依据。

附图 8 给出了客户下载情形时步骤 804 过程的扩展图。在客户下载情形中, 对应传送什么数据的确定始于步骤 804a, 该步骤首先确定被请求 Web 站点上已经被下载到客户的数据。图中的示例中假设, 给定 Web 站点上已经被下载到客户的文件清单, 要么保存在客户机中, 要么保存在服务器中。该清单可以与该 Web 站点上当前文件的清单加以比较, 根据文件名、日期、长度的不同确定内容的变化。然后, 该过程转移到步骤 804b, 按照上文叙述的规则或类似规则, 对有待下载的数据规定优先级。

图中所示过程可用作多投射 (multicasting) 过程的一部分, 即将相同的信息流同时提供给不同的接收者。例如, 可能有许多用户预订了相同新闻站点的非高峰检索, 因此, 他们需要的更新也相同。如果各个下载过程被协调好去接收该相同的信息流, 就可以用一个服务器过程去对每个用户的信息进行更新。在一户中有多个用户单元的情形中, 多投射过程特别有用。

一旦在步骤 804 中确定了待被传送或检索的数据, 过程就转移到步骤 806, 该步骤检索被确定的信息并为在服务器上存储而对其进行压缩, 或者将被确定的信息从压缩的高速缓存文件中抽取出来, 压缩后向客户传送。在客户下载的情形中, 用户关心如何最小化连接时间, 这在传送高速缓存数据中进行数据压缩更加重要。对服务器上单个文件中的 Web 站点信息进行数据压缩, 并不严重妨碍从该 Web 站点中选择页面或文件向客户传送。已知的算法允许从压缩的文档中提取文件, 在传

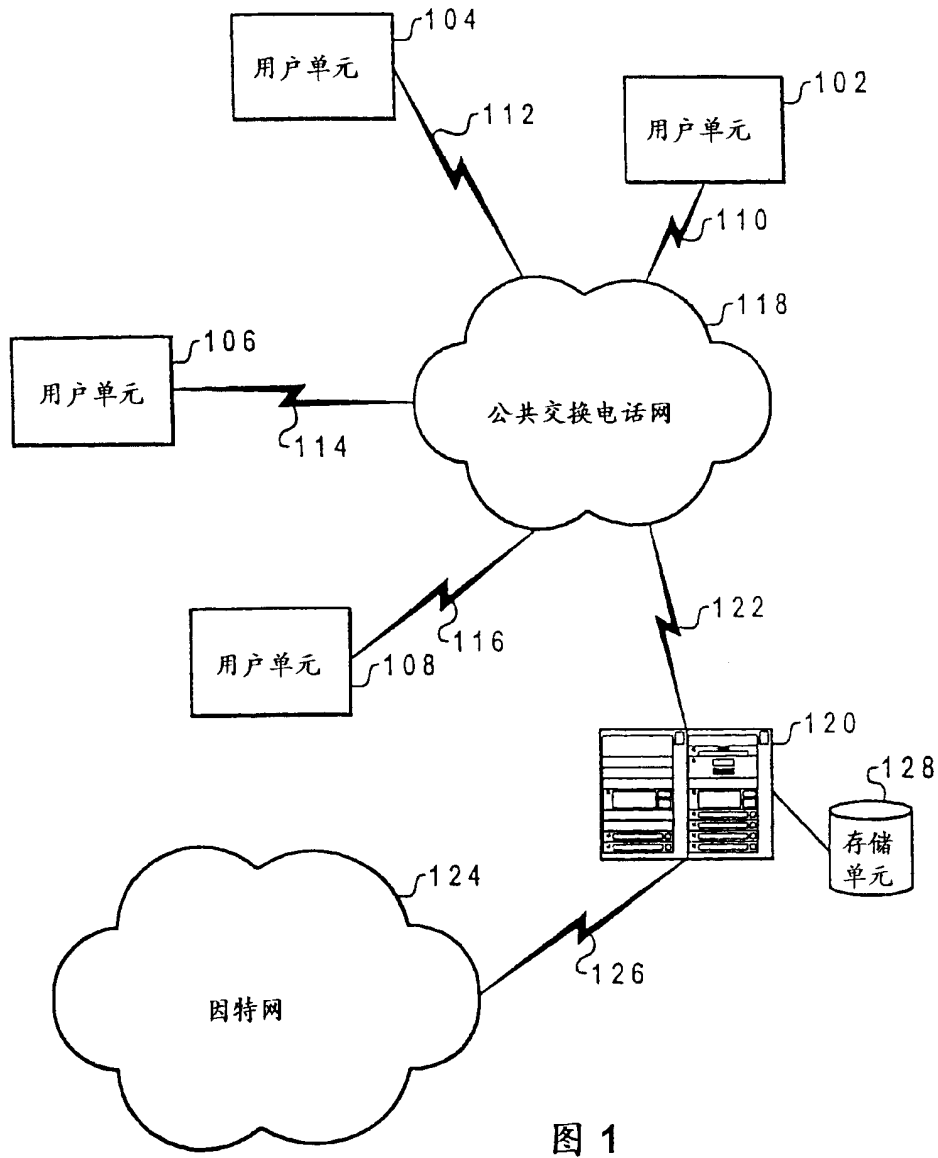
送中同时对文件进行压缩，使得数据中只有哪些被选择的页面或文件才需要被传送，并且可以以压缩的方式被传送，以节省连接时间。

人们预期 Web 站点最终将实现最优化高速缓存的页面。例如，Web 站点会将用于优先进行高速缓存和下载信息的规则的知识，用于设置页面，建立最优化高速缓存的页面，用于照顾高速缓存及向客户下载。Web 出版者也可以通过注解来标明哪些页面和/或文件属于同一个 Web “出版物”，哪些链接引用不同出版物的页面。对页面定期更新的 Web 站点，可以将 Web 站点的更新或改变后的文件压缩成一个数据组，以提高检索效率。对于提供向这种 Web 站点脱机预定服务的服务器来说，服务器只需充当改变后的压缩文件的一个镜面。

值得注意的是，尽管本文描述本发明所针对的环境是功能齐备的数据处理系统，本领域的熟练人员应知道，本发明的机制能用载有多种不同形式的指令的计算机可读介质进行分配，并且，不管实际用于进行分配的载有信号的介质是什么具体类型，本发明均能同等地应用。计算机可读介质的例子包括：软盘、CD - ROM 之类的可记录式媒体，数字和模拟通信线路之类的传输式介质。

本发明提供了一种从多个信息源向多个目的地有效地进行数据分配的方法，用于当单独进行数据传送时所有数据必定要经过相同节点の場合。通过对服务器的多个用户请求的数据进行预先高速缓存或“镜象”的方法，降低了从服务器到因特网的带宽要求。对被请求数据进行的压缩，会减少占用服务器的存储空间，加快向用户的数据传递。

以上结合较佳实施例对本发明作了特别的表示和描述，本领域的熟练人员知道，在本发明的思想和范围内，其中的许多细节都可以作出各种改进。



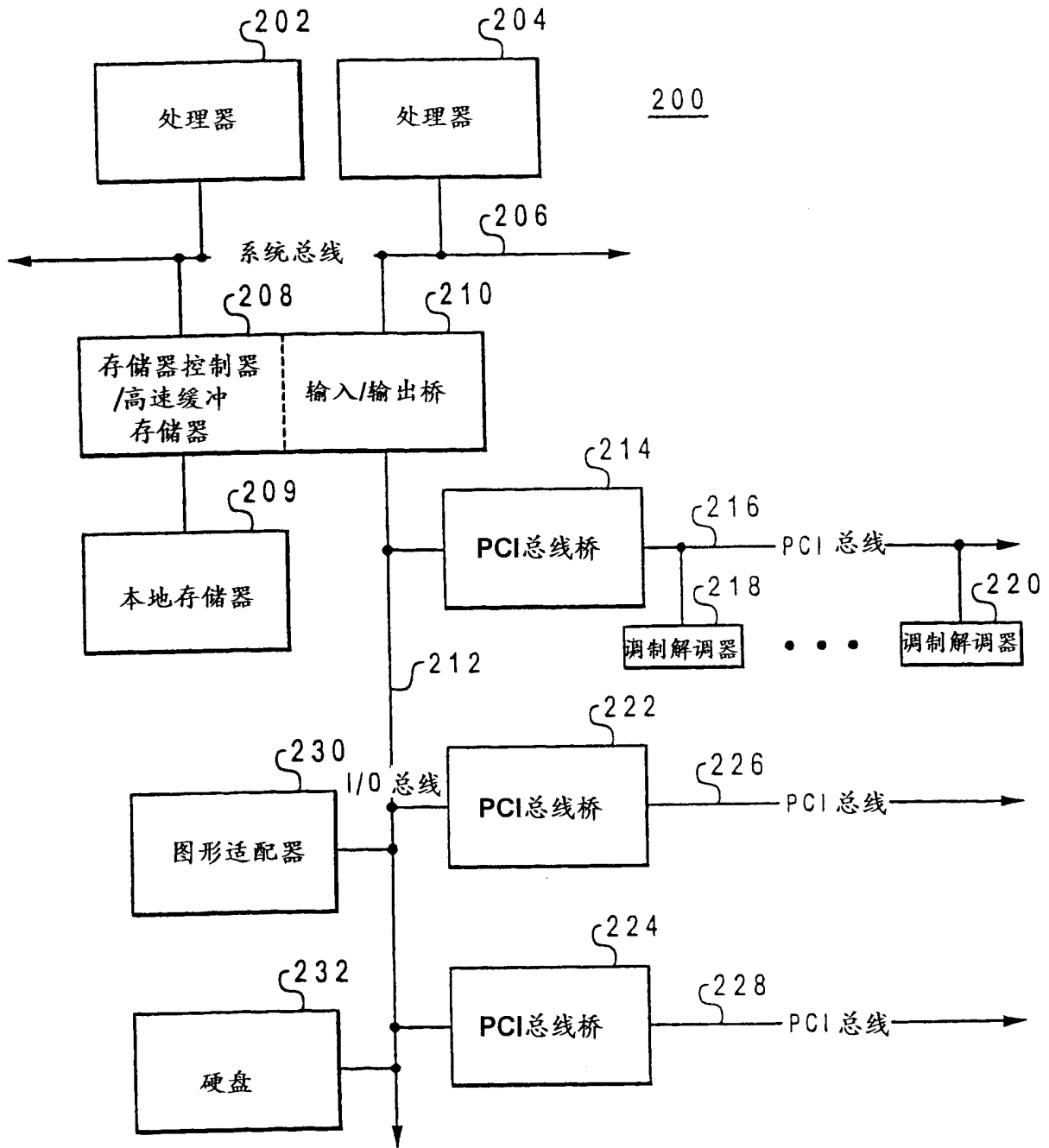


图 2

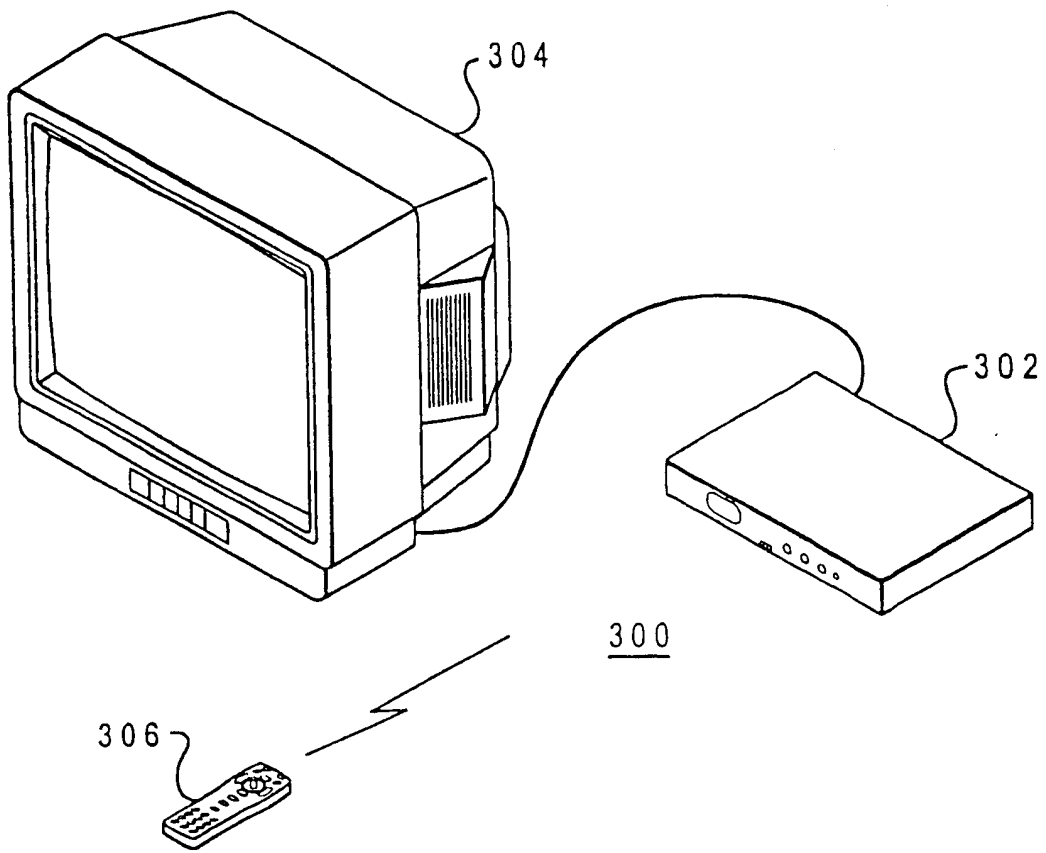


图 3

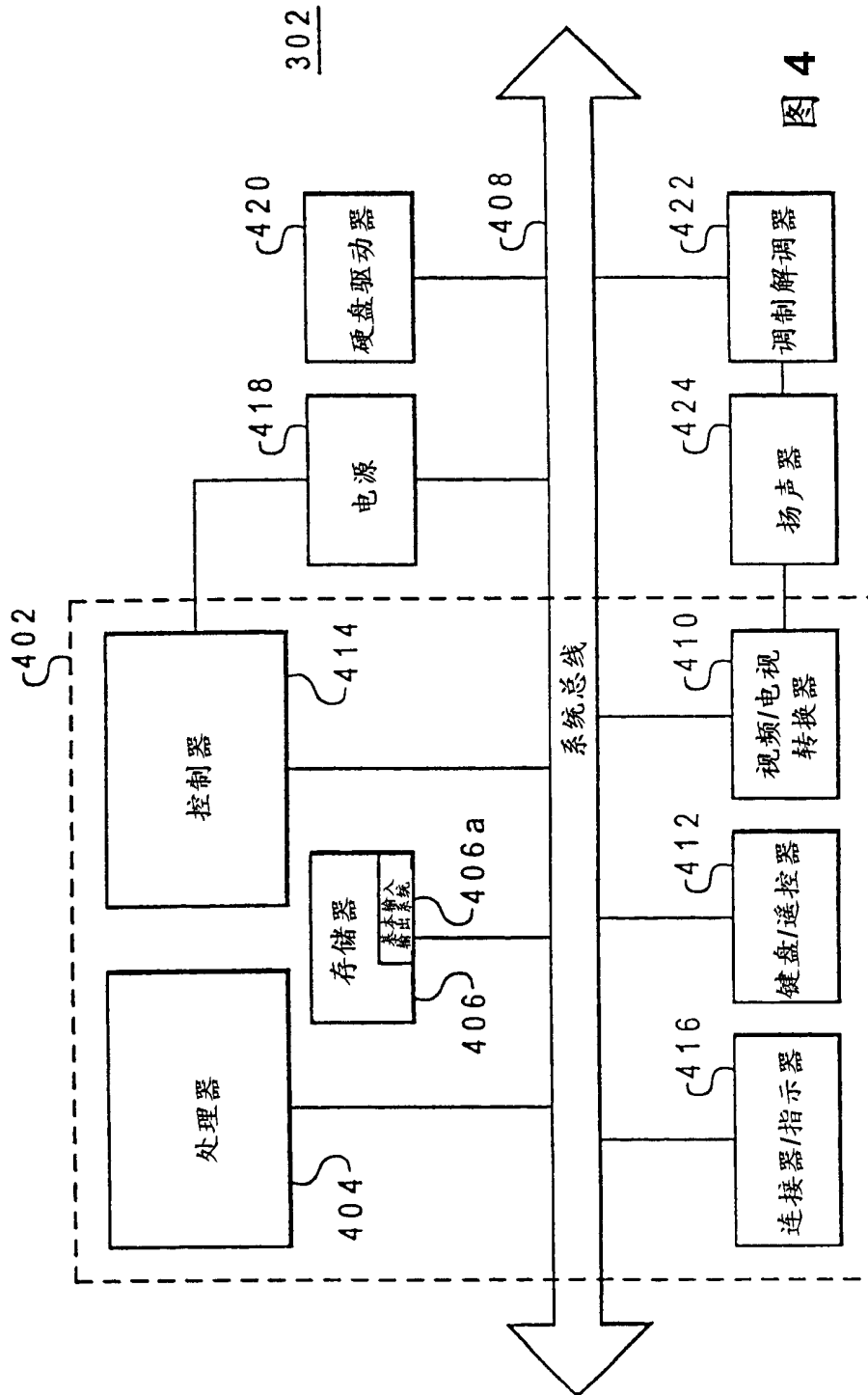


图 4

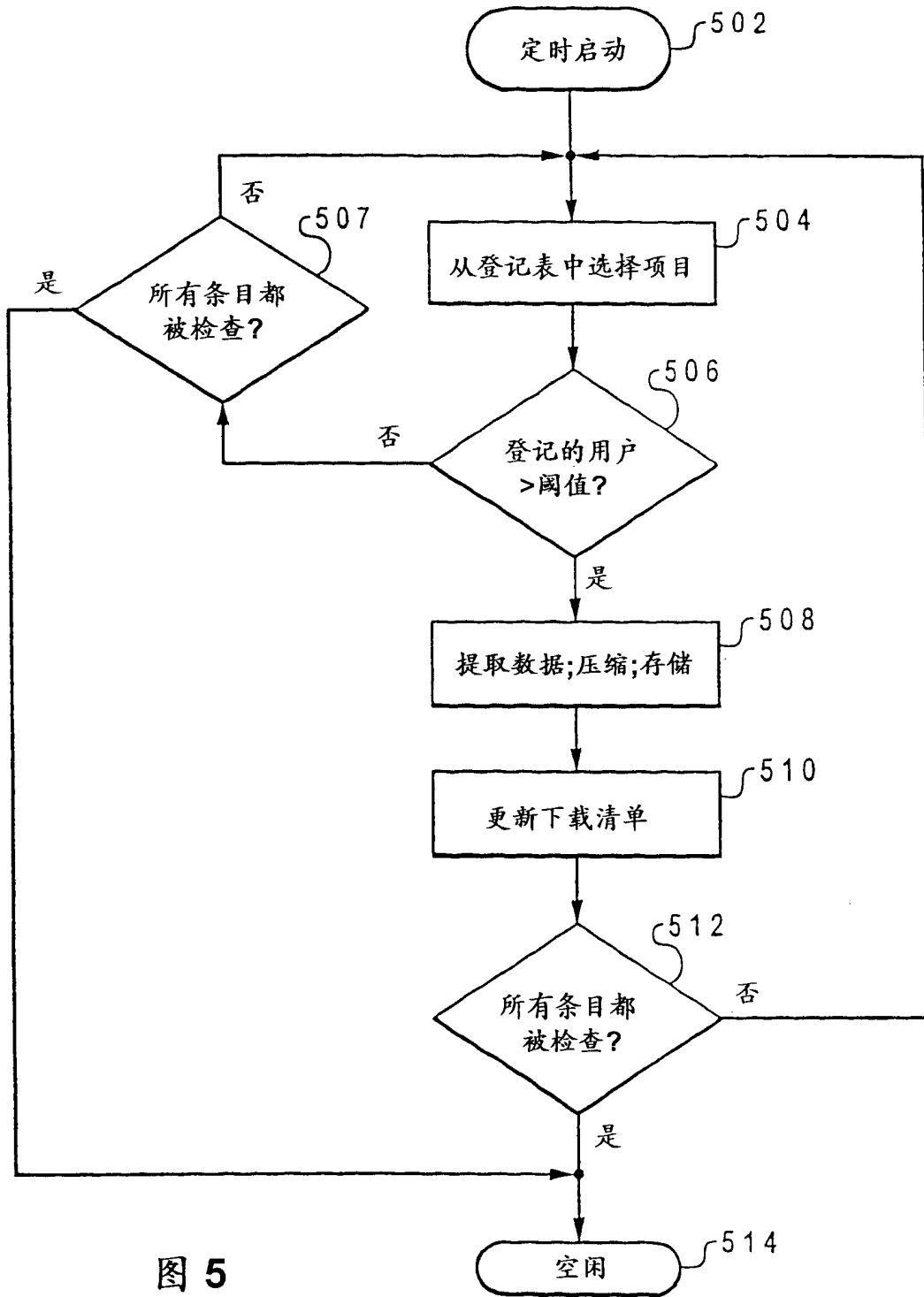


图 5

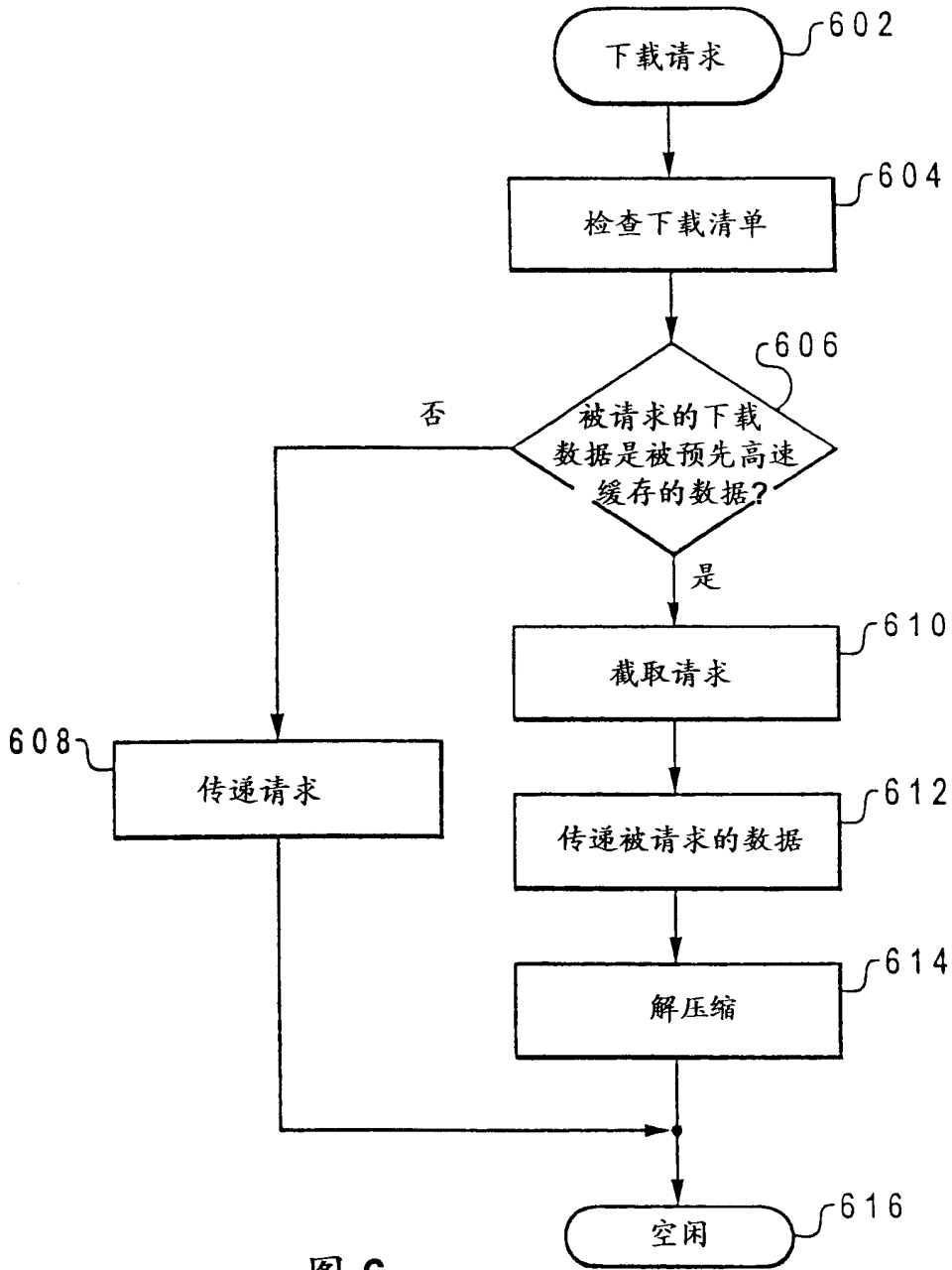


图 6

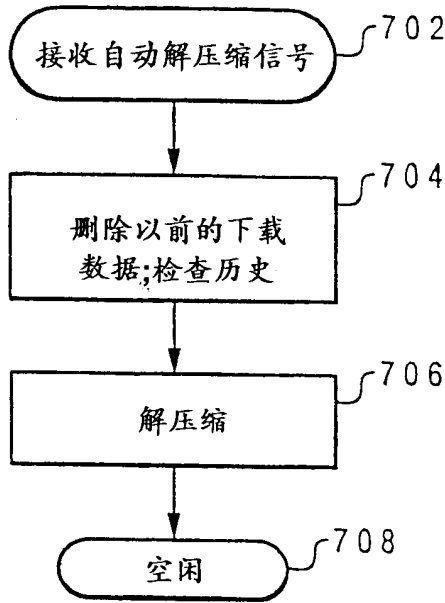


图 7

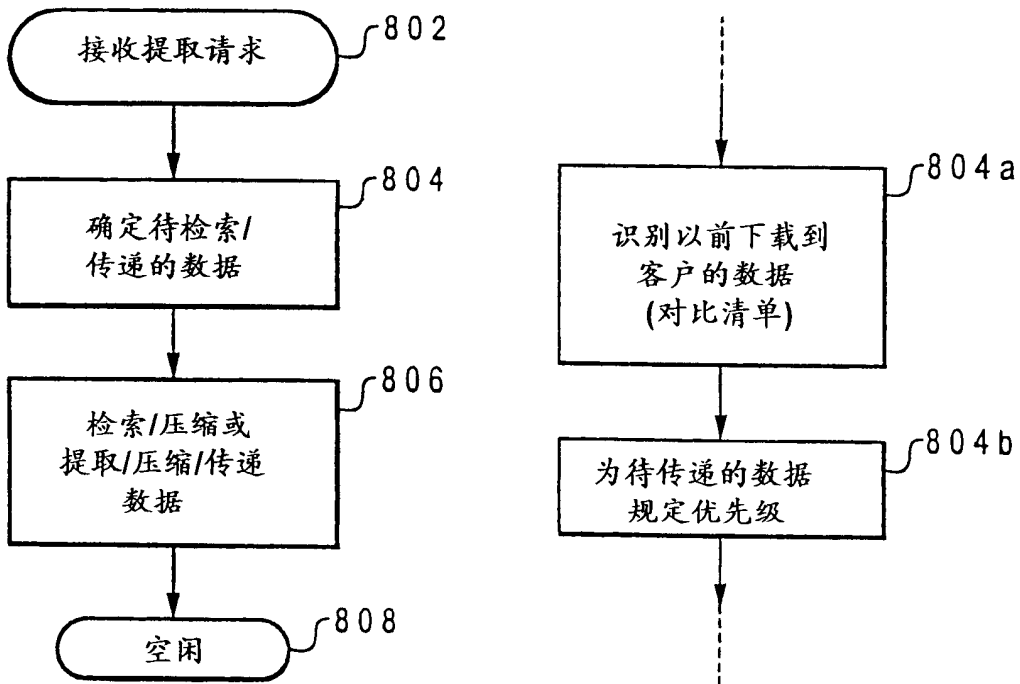


图 8