

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 9/46 (2006.01)

G06F 17/30 (2006.01)



[12] 发明专利说明书

专利号 ZL 99811828.1

[45] 授权公告日 2006年5月10日

[11] 授权公告号 CN 1255728C

[22] 申请日 1999.10.8 [21] 申请号 99811828.1

[30] 优先权

[32] 1998.10.9 [33] US [31] 09/169,223

[86] 国际申请 PCT/GB1999/003360 1999.10.8

[87] 国际公布 WO2000/022526 英 2000.4.20

[85] 进入国家阶段日期 2001.4.6

[71] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 凯文·M·乔丹 吴坤龙 俞士纶

审查员 曲颖

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

代理人 朱海波

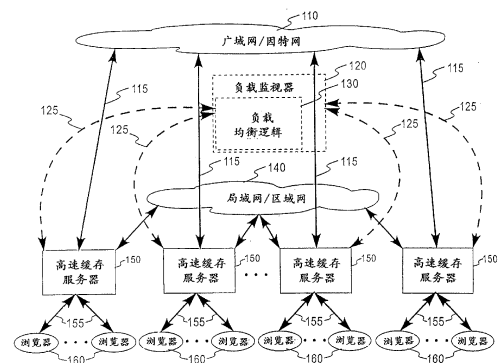
权利要求书 5 页 说明书 14 页 附图 5 页

[54] 发明名称

协同高速缓存服务器负载均衡方法

[57] 摘要

在一种包括例如代理高速缓存服务器的协同高速缓存服务器的集合的系统中，如果被请求对象没有在本地上找到，则一个请求可以被转发到一个协同高速缓存服务器。例如如果由于引用不平衡而检测到过载状况，一些对象由所有客户机所需，并且包含这些热点对象的高速缓存服务器由于被转发请求而过载。相应地，通过把一些或所有转发的请求从一个过载高速缓存服务器转移到一个较小负载的服务器而均衡负载。集中和分布的负载均衡环境都被描述。



1. 一种高速缓存服务器负载均衡方法，其中包括如下步骤：
接收响应来自多个协同高速缓存服务器之一上的一个对象的高速缓存不中而从该协同高速缓存服务器转发的转发请求；以及
根据动态保持的服务器负载状况和转发频率，把一个或多个要求该对象的所述转发请求在协同高速缓存服务器之间转移，所述转发频率包括对该对象的请求已经被转发的次数。
2. 根据权利要求1所述的方法，所述转移步骤还包括如下步骤：
定期监视一个拥有高速缓存服务器的负载状况和转发频率；以及
响应所述监视的结果，主动地把一个或多个用于被高速缓存对象的后续的转发的请求从该拥有高速缓存服务器转移到一个或多个所述协同高速缓存服务器。
3. 根据权利要求1或2所述的方法，所述转移步骤进一步包括响应转发的请求检查负载状况和转发频率的步骤。
4. 根据权利要求1或2所述的方法，其中所述转移步骤包括把对于该对象的拥有权改变为在两个或多个所述协同高速缓存服务器之间的共享拥有权。
5. 根据权利要求4所述的方法，其中进一步包括响应负载状况中的改变把所述共享拥有权合并的步骤。
6. 根据权利要求1所述的方法，其中进一步包括本地监视在每个协同高速缓存服务器上的负载的步骤。
7. 根据权利要求6所述的方法，其中进一步步骤包括如下步骤：
一个分布负载监视器监视并维护在所述每个协同高速缓存服务器上的本地负载状况、转发频率以及对于所高速缓存对象的拥有权信息。
8. 根据权利要求7所述的方法，其中进一步包括如下步骤：
所述协同高速缓存服务器定期交换和维护一个或多个：负载状况信息；转发频率；以及拥有权信息。

9. 根据权利要求7所述的方法，其中进一步包括如下步骤：

所述协同高速缓存服务器通过搭载一个或多个被转发请求和响应，而交换一个或多个：负载状况信息；转发频率；以及拥有权信息。

10. 根据权利要求1或2所述的方法，其中进一步包括如下步骤：
接收一个转发的请求并且更新转发频率。

11. 根据权利要求7、8或9所述的方法，其中进一步包括如下步骤：

识别一个较小负载的协同高速缓存服务器；以及

把一个或多个：转移请求；以及高速缓存对象的一个副本，传送到所述较小负载的协同高速缓存服务器。

12. 根据权利要求11所述的方法，其中进一步包括如下步骤：

所述较小负载的协同高速缓存服务器接收所述转移请求；以及
所述较小负载的协同高速缓存服务器响应所述转移请求，从一个源发对象服务器要求该对象的一个副本。

13. 根据权利要求11所述的方法，其中该副本是通过一个或多个内联网、广域网或因特网而获得。

14. 根据权利要求1或2所述的方法，其中进一步包括如下步骤：

把一个转移请求信息多点传送到一个或多个其它协同高速缓存服务器，使得后续的转发信息被转移。

15. 根据权利要求14所述的方法，其中进一步包括如下步骤：

协同高速缓存服务器维护一个高速缓存表的本地副本并且修改一个哈希函数；以及

该协同高速缓存服务器通过：更新一个高速缓存表的本地副本；以及修改一个哈希函数，而修改拥有权信息。

16. 根据权利要求15所述的方法，其中进一步包括如下步骤：

把对于该对象的拥有权改变为在至少两个所述协同高速缓存服务器之间的共享拥有权；以及

所述协同高速缓存服务器将后续对象请求转发到该对象的一个或多个较小负载的共享拥有者。

17. 根据权利要求 16 所述的方法，其中进一步包括如下步骤：
检测对于一个共享对象的负载状况的下降；以及
响应负载状况的下降，合并该共享的拥有权。

18. 根据权利要求 1 或 2 所述的方法，其中所述转移一个或多个所述转发的请求包括如下步骤：

把该对象的一个副本从拥有一个高速缓存服务器传送到一个或多个所述协同高速缓存服务器；以及

所述协同高速缓存服务器接收并高速缓存该对象的副本。

19. 根据权利要求 1 或 2 所述的方法，其中进一步包括如下步骤：
在过去的时间间隔中计算每个高速缓存服务器的负载状况；
在过去的时间间隔中计算所有高速缓存服务器的平均负载；以及
查找超过所述平均负载之上的一个阈值的高速缓存服务器。

20. 根据权利要求 1 或 2 所述的方法，其中所述协同高速缓存服务器的负载状况可以是到所述协同高速缓存服务器的所述转发的请求的计数值与直接请求的计数值的加权和。

21. 根据权利要求 1 或 2 所述的方法，其中进一步包括如下步骤：
在一个或多个：每个对象级；以及对象的一个分区级，维护高速缓存信息。

22. 根据权利要求 21 所述的方法，其中所述对象级或所述分区的所述高速缓存信息包括与该对象相关的转发频率。

23. 根据权利要求 22 所述的方法，其中进一步包括如下步骤：
一个分布的负载监视器监视并本地维护在每个高速缓存服务器上的负载状况、转发频率以及对于高速缓存对象的拥有权信息。

24. 根据权利要求 23 所述的方法，其中进一步包括如下步骤：
所述协同高速缓存服务器定期的交换一个或多个负载状况，转发频率和拥有权信息。

25. 根据权利要求 22 所述的方法，其中进一步包括如下步骤：
所述协同高速缓存服务器通过搭载一个或多个被转发请求和响应，而交换一个或多个：负载状况；转发频率；以及拥有权信息。

26. 一种在协同高速缓存服务器的集合中均衡负载的方法，其中每个高速缓存服务器可以接收直接请求和转发请求，并且在一个高速缓存不中时，一个请求可以被转发到高速缓存所述对象的一个拥有高速缓存服务器，该方法包括如下步骤：

对所述协同高速缓存服务器监视负载状况和转发频率，所述转发频率包括对该对象的请求已经被转发的次数；以及

根据负载状况和转换频率中的改变，把已经被转发的一个或多个转发的请求从一个协同高速缓存服务器转移到第二个协同高速缓存服务器。

27. 根据权利要求 26 所述的方法，其中所述监视负载状况的步骤包括如下步骤：

在过去的时间间隔中计算每个高速缓存服务器的负载状况；

在过去的时间间隔中计算所有代理高速缓存服务器的平均负载；以及

查找超过高于所述平均的负载的一个阈值的那些代理高速缓存服务器。

28. 根据权利要求 26 或 27 所述的方法，其中所述转移步骤可以响应如下一种或多种情况而执行：来自所述协同高速缓存服务器的转发的请求；以及定期监视负载状况和转换频率。

29. 根据权利要求 26 或 27 所述的方法，其中进一步包括如下步骤：集中的逻辑负载监视器为该协同高速缓存服务器维护转发频率和负载状况。

30. 根据权利要求 26 或 27 所述的方法，其中所述高速缓存服务器的负载状况可以是到所述高速缓存服务器的转发的请求的计数值与直接请求的计数值的加权和。

31. 根据权利要求 26 或 27 所述的方法，其中进一步包括在每个对象级或在对象的一个分区级维护高速缓存信息的步骤。

32. 根据权利要求 31 所述的方法，其中所述对象级或分区级的高速缓存信息包括通过所述负载监视器到达所述对象的请求的转发频

率。

33. 根据权利要求 26 或 27 所述的方法，其中所述协同高速缓存服务器包括协同代理高速缓存服务器。

34. 根据权利要求 26 或 27 所述的方法，其中进一步包括如下步骤：

一个逻辑目录服务器维护一个高速缓存表和一个负载表；

所述高速缓存服务器向所述目录服务器查询对于一个本地不中的对象在另一个高速缓存服务器中的对象的位置；以及

响应对于对象位置的请求，所述目录服务器通过操纵所述高速缓存表在所述高速缓存服务器中负载均衡请求。

35. 根据权利要求 29 所述的方法，其中进一步包括如下步骤：

每个高速缓存服务器多点传送到一列表中的协同高速缓存服务器，以定位一个本地不中的对象的副本；以及

所述转移步骤包括从用于多点传送的相邻高速缓存服务器的子集中排除过载的高速缓存服务器的步骤。

协同高速缓存服务器负载均衡方法

本发明涉及在协同高速缓存服务器之间的负载均衡方法，特别涉及基于负载状态和从协同高速缓存服务器(cooperating cache server)发出请求的速率的负载均衡方法。

万维网的使用发展成指数增长。结果，用于访问 Web 对象的响应时间变得令人不满意地缓慢。用于改进 Web 访问时间的一个方法是在浏览器和源发 Web 服务器之间采用一个或多个代理高速缓存服务器。代理高速缓存服务器的例子是包括一组运行微软的 Windows NT_{TM} 的 PC 服务器，例如出自 IBM(国际商用机器公司)的 NETFINITY_{TM} 服务器；以及运行 IBM 的 AIX_{TM} 操作系统的工作站服务器，例如 IBM RS/6000_{TM} 或 SP/2_{TM}。实际上，越来越多的组织，例如因特网服务提供商(ISP)和公司使用一组协同代理高速缓存服务器来提高响应时间以及减小因特网的通信量。一组协同高速缓存服务器在可靠性和性能方面比单个高速缓存服务器具有明显优点。如果一个服务器发生故障，请求仍然可以由其它协同高速缓存服务器来提供服务。请求可以在服务器之间分配，因此增加可扩展性。最后，集合的高速缓存容量大得多，使得更加可能在其中一个高速缓存服务器中找到被请求的对象。

利用协同高速缓存服务器，由于高速缓存不中(cache miss)而造成不能够在本地对请求提供服务，则该请求可以被转发到另一个存储被请求的对象高速缓存服务器。结果，有两种请求可以到达高速缓存服务器：直接请求和转发请求。直接请求是直接来自客户接收的请求。转发请求是从用户的利益考虑由于在高速缓存服务器上高速缓存不中而来自其它协同高速缓存服务器的请求。如果一台服务器刚好包含大多数客户目前感兴趣的急需(或“热点”)对象，有了在高速缓存服务器之间转发的请求，该高速缓存服务器可能容易变得过载，造成高速缓

存服务器之间的不均衡的工作负载。由于许多高速缓存服务器等待同一个过载的高速缓存服务器来响应转发给它的请求，则不均衡的工作负载造成性能瓶颈。因此需要一种方式来执行在一组代理高速缓存服务器之间的动态负载均衡。本发明针对于这种需求。

负载均衡传统地是通过前端调度程序完成的。该调度程序在高速缓存服务器中“平均分配”输入的直接请求。例如，通过使用映射表可以在域名服务器(DNS)级完成负载均衡，如 Sun Microsystems 的 NETRA_{TM} 所做的 (“Sun Microsystems 的白皮书中 “Proxy Cache Server, Product Overview”，<http://www.sun.com/>)。在一簇服务器中的负载均衡也可以通过前端路由器完成，如 IBM 提供的 NETDISPATCHER_{TM} (例如，1997年5月，RC20853，IBM 研究报告，G.Goldszmidt和 G.Hunt所做的“Net Dispatcher:A TCP Connection Router”)。这里，输入的请求被 NETDISPATCHER_{TM} 分配给负载最小的服务器。然而，这些传统的方法只分配“直接请求”而不是针对由于太多的对热点对象的请求被同时转发到同一代理服务所造成的负载不均衡问题。本发明针对此需求。

协同高速缓存，或者远程高速缓存，已经被用在分布文件系统中，以提高系统性能(参见由 M.D.Dahlin 等人所著的“协同高速缓存：使用远程客户机存储器来提高文件系统性能”，关于操作系统设计和应用的第一次讨论会的会议录，pp.1-14，1994)。在此，分布于局域网上的工作站的集合的文件高速缓存协同工作，以形成一个更加有效的整体文件高速缓存。每个工作站文件高速缓存不但是由本地请求所引用的对象，而且还是可能由来自远程工作站的请求所引用的对象。在本地高速缓存不中时，一个本地请求可能被发送到其它可以获得一个副本的客户工作站，如果能找到这种工作站的话。否则，该对象从对象服务器获得。在此强调的主要是如何在更新时保持高速缓存的一致性，以及如何通过把局部替换的对象转移到另一个工作站的高速缓存而保持高-速缓存命中率。在此没有动态负载均衡。

协同高速缓存还被用于集合的代理高速缓存服务器中，以减少存

取时间。在一个高速缓存不命中时，不直接通过广域网到达源发 Web 服务器，而是一个同速高速缓存服务器把该请求转发以从一个局域网或区域网中的协同高速缓存服务器获得该对象。例如，在 SQUID 系统中一个本地高速缓存不命中时，一个高速缓存服务器把一个请求多点传送(使用因特网高速缓存协议(ICP))到一组其它高速缓存服务器(参见由 D.Wessels 等人所著的“Squid 因特网对象高速缓存”，<http://squid.nlanr.net/>)。如果它们的高速缓存中包含所请求的对象，则这些协同高速缓存服务器用一个消息表明这种情况。所请求的对象然后从第一个响应该请求的协同高速缓存服务器获得，而不是从因特网上的源发 Web 服务器获得。但是，如果在超时周期之后没有响应，则所请求对象将从源发 Web 服务器获得。由于转发请求而可能在一个高速缓存服务器中出现负载不均衡。

代替多点传送，CRISP 系统使用一种逻辑中央目录，来定位在另一个代理服务器上被高速缓存的对象(参见由 S.Gadde 等人所著的“用于可缩放的因特网高速缓存的目录结构”，杜克大学计算机科学系，技术报告 CS-1997-18, 1997)。在此，在一个高速缓存不命中时，高速缓存服务器向目录服务器请求该对象。利用高速缓存对象存储器的中心知识，目录服务器把这种请求发送到其高速缓存中包含该对象的服务器。如果找到该对象的话，则该对象被发送到发出请求的服务器，而源发服务器继续高速缓存该对象。如果没有一个高速缓存具有所请求对象的副本，则发出请求的服务器通过因特网(可能通过一个广域网)从源发 Web 服务器获得该对象。由于后续请求被转发到这个高速缓存服务器，因此这也可能在该高速缓存服务器中造成负载不均衡。

在一个协同高速缓存服务器上定位一个对象的另一种方法是通过哈希函数(hash function)。一个例子是高速缓存阵列路由协议(CARP)(参见 V.Valloppillil 和 K.W.Ross 所著的“高速缓存阵列路由协议 V1.0”，因特网草案，<http://ircache.nlanr.net/Cache/ICP/draft-vinod-carp-v1-03.txt>, 1998 年 2 月)。在 CARP 中，整个对象空间在协同高速缓存服务器中分区，每个高速缓存服务器使用一个分区。当一个请

求由高速缓存服务器从配置的客户浏览器接收时，一个哈希函数被应用到来自该请求的一个关键字，例如 URL 或者目的 IP 地址，以识别该分区。如果该哈希分区被分配给发出请求的高速缓存服务器，则该请求在本地解决。否则，它被转发送到在被识别分区中的适当的高速缓存服务器。

SQUID、CRISP 和 CARP 使用其它代理服务器的高速缓存来减少必须通过广域网来寻找不中的对象的可能性。它们在用于定位高速缓存中包含被请求对象的副本的一个协同高速缓存服务器的机制上是不同的。每个高速缓存服务器处理两种请求：直接请求和转发请求。直接请求是直接由连接到该代理服务器的浏览器所发出的请求。转发请求是由高速缓存中没有被请求对象的协同高速缓存服务器所发出的请求。在任何情况下，根据在给定时刻一个代理服务器高速缓存的对象类型，其 CPU 可能由于忙于处理直接和转发请求而过载。

根据上述需要，本发明的一个方面是提供一种高速缓存服务器负载均衡方法，其中包括如下步骤：响应在一个协同高速缓存服务器上的一个对象的高速缓存不中，从该协同高速缓存服务器接收被转发请求；以及根据负载状况和要求该对象的转发频率把一个或多个要求该对象的所述转发请求在协同高速缓存服务器之间转移。

本发明另一个方面提供一种在协同高速缓存服务器的集合中均衡负载的方法，其中每个高速缓存服务器可以接收直接请求和转发请求，并且在一个高速缓存不中时，一个请求可以转发到一个高速缓存所述对象的拥有高速缓存服务器，该方法包括如下步骤：对于所述协同高速缓存服务器监视负载状况以及转发频率；以及根据负载状况和转发频率中的变化，把一个或多个转发请求从一个协同高速缓存服务器转移到第二个高速缓存服务器。

例如，在一种包括协同代理高速缓存服务器的集合的系统中，如果在本地没有找到被请求对象时，一个请求可以被转发到另一个协同服务器。不是通过因特网从源发 Web 服务器获得该对象，而是一个高速缓存服务器可以在局域网或内部网中从一个协同高速缓存服务器获

得一个副本。访问一个对象的平均响应时间可以通过该协同高速缓存服务器而大大改进。但是，由于引用不对称，一些对象可能被所有客户所要求。结果，包含这些热点对象的代理高速缓存服务器可能由于来自其它代理高速缓存服务器的转发请求而变得过载，产生一个性能瓶颈。根据本发明，我们提出一种用于协同代理高速缓存服务器的集合的负载均衡方法，它是通过把一些或所有转发请求从一个过载的高速缓存服务器转移到另一个负载较小的服务器而实现的。

根据本发明的高速缓存服务器负载均衡方法的一个例子包括如下步骤：响应在一个协同高速缓存服务器上的一个对象的高速缓存不中，从该协同高速缓存服务器接收被转发的请求；以及根据负载状况和要求该对象的转发频率把一个或多个要求该对象的所述转发请求在协同高速缓存服务器之间转移。

本发明还包括如下特点：定期监视该拥有高速缓存服务器的负载状况以及转发频率；以及响应该监视情况，主动把一个或多个要求被高速缓存的对象的后续转发请求从该拥有高速缓存服务器转移到一个或多个协同高速缓存服务器。另外，转移步骤还包括响应一个被转发请求的接收，检查负载状况和转发频率。在一个例子中，协同高速缓存服务器的负载状况是到达所述协同高速缓存服务器的所述被转发请求的计数值和直接请求的计数值的加权和。在另一个例子中，高速缓存信息被保持在：每个对象级；或者对象的分区级。

本实施例包括用于执行负载均衡的各种实现方式，包括集中和分布环境以及其各种混合形式。例如，分布的负载监视器可以用于监视并保持局部负载状况，转发频率和在每个协同高速缓存服务器上对于被高速缓存对象的拥有信息。该协同高速缓存服务器可以定期交换和保持一个或多个：负载状况信息；转发频率；以及拥有信息。例如，协同高速缓存服务器可以通过随着一个或多个转发请求和响应，搭载(piggyback)负载状况信息、转发频率以及拥有信息中的一种或多种来交换信息。

在另一个例子中，一个过载协同高速缓存服务器可以识别一个较

小负载的协同高速缓存服务器；并且把一个转移请求和高速缓存对象的副本传送到该较小负载的协同高速缓存服务器(然后该服务器高速缓存该对象)，使得要求该对象的后续请求将不被转发。另外，一个过载的协同高速缓存服务器可以把该转移请求传送到该较小负载的协同高速缓存服务器，然后响应该转移请求，该较小负载的服务器从源发对象服务器获得该对象的副本。在另一个例子中，该拥有高速缓存服务器可以把转移请求消息多点传送到一个或多个其它协同高速缓存服务器，使得后续的转发请求被转移。

在本发明的一个完整分布应用中，每个协同高速缓存服务器包括一个分布的负载监视器，用于监视和本地维护负载状况，并且还可以在一个高速缓存表的本地副本中或者通过一个哈希函数保持转发频率和拥有信息。该协同高速缓存服务器可以通过高速缓存表的本地副本或哈希函数的方式而改变拥有信息。

本发明还包括其它特征，把对该对象的拥有权改变为在至少两个协同高速缓存服务器之间的共享拥有权，并且把后续的对象请求转发到一个或多个该对象的较小负载的共享拥有者。如果检测到对于一个共享对象的负载状况减小，则共享拥有权可以合并，对应于在负载状况中的减小。

在另一个例子中，根据负载状况和转发频率转移一个或多个转发请求的转移可以通过把该对象的副本从该拥有高速缓存服务器传送到一个或多个协同高速缓存服务器而实现，使得后续请求将不被转发(只要该对象保留在该接收方的高速缓存中)。

根据本实施例的集中环境的一个例子包括：用于为协同高速缓存服务器维护转发频率和负载状况的中心逻辑负载监视器。该负载监视器可以包括一个逻辑目录服务器，用于维护一个用于监控在该高速缓存服务器上的负载的一个负载表以及一个用于监视转发频率和定位对象的高速缓存表(或者哈希函数)。目录服务器接收对于一个在其它高速缓存服务器中本地不中的对象的对象定位请求，并且把要求本地不中的对象的请求转发。目录服务器通过操纵基于一个给定对象的负载

和转换频率的高速缓存表，响应要求对象位置的请求，在协同高速缓存服务器之间对请求进行负载均衡。

参照下文的描述、所附权利要求以及结合附图，本发明的这些和其它特点、方面和优点将得到更好的理解，其中：

图 1a 以方框图的形式示出采用代理高速缓存服务器的集合的一种系统的例子，其中可以应用根据本发明的集中负载均衡逻辑；

图 1b 以方框图的形式示出采用代理高速缓存服务器的集合的一种系统的例子，其中可以应用根据本发明的分布负载均衡逻辑；

图 2a-b 示出用于对由图 1a-b 中所示的负载监视器所维护的两个表的数据格式；

图 3 示出用于负载监视器响应由于一个高速缓存不中而来自一个高速缓存服务器的请求的逻辑流程的一个例子；以及

图 4 示出用于高速缓存服务器响应要求一个对象的请求的逻辑流程的一个例子。

下面将对用于集中和分布结构的本发明的负载均衡逻辑的例子进行描述。图 1a 示出采用代理高速缓存服务器的集合的一个系统的方框图的例子，其中可以应用本发明中提出的集中负载均衡逻辑。如图所示，该系统包括代理高速缓存服务器 150 的集合。尽管仅仅示出一级的高速缓存服务器，但是在此可以有多级高速缓存服务器 150。通常，这些代理高速缓存服务器可以通过局域网(LAN)或者区域网或者内部网 140 相互连接。每个高速缓存服务器还连接到广域网(WAN)或者因特网 110。通过 WAN，这些代理高速缓存服务器可以到达 115 源发 Web 服务器，要求不能在它们的高速缓存中本地地找到的对象。

根据本实施例，一个逻辑负载监视器 120 包括一个负载均衡逻辑 130，用于监视协同高速缓存服务器 150 的负载状况和转发频率(图 2a)并且为它们提供负载均衡。如下文所述，各种负载监视器 120 的特征可以：驻留在一个或多个高速缓存服务器中；在高速缓存服务器之间复制和分布；或者驻留在另一个专用系统中，例如个人计算机(PC)服务器或工作站。在集中系统结构中，负载监视器 120 在指导请求 125

转发到高速缓存服务器中，可以执行中央目录功能。一个或多个浏览器 160 可以连接到每个高速缓存服务器 150。直接请求 155 从例如运行常规浏览器 160 的计算机这样的客户机发送到配置的高速缓存服务器 150。如果可以在本地找到所请求的对象，则它被返回到该浏览器。否则，该高速缓存服务器 150 把一条消息传送到负载监视器 120。负载监视器 120 的各种实现例子将在下文中更加具体描述。如果不存在负载不均衡状况或趋势，则负载监视器 120 把该请求 125 转发到拥有所请求对象的高速缓存服务器 150。该拥有高速缓存服务器然后把所请求对象例如通过 LAN 140 发送到发出请求的高速缓存服务器。

如果实际负载不均衡情况被识别，或者根据负载趋势而预测到，则负载监视器 120 开始一个从过载高速缓存服务器向一个或多个低负载(或者较小负载)的服务器转发请求的转移。如下文更具体的描述，拥有权的转移可以根据服务器 150 的负载状况和转发频率以及其它因素。

图 2a-b 示出由负载监视器所维护的两个表的数据格式。如图所示，该表包括一个负载表 102 和一个高速缓存表 101。本领域内的技术人员将认识到可以使用一个表格或者各种其它数据结构来代替或等价替换。负载表 102 包括每个(A、B、C...1022)高速缓存服务器 150 的负载状况 1021，使得过载或低负载服务器可被识别。通常，负载状况 1021 可以通过检测每个高速缓存服务器而定期更新。一个高速缓存服务器的负载可以是转发请求数与直接请求数的加权和。过载高速缓存服务器 150 可以用任何常规技术而识别。例如，负载监视器可以在过去的时间间隔中计算所有代理高速缓存服务器的平均负载。过载的高速缓存服务器可以是具有超过上述平均负载之上的一个阈值的负载的服务器。根据本实施例，负载均衡考虑到过载量，以及由于被高速缓存对象的转发频率 1011 而造成的负载。这样，负载监视器可以确定是否要继续把一些或所有转发请求从一个过载高速缓存服务器 C 10213 转移到一个低负载服务器 A 10211。高速缓存表 1011 包括一个对象或一部分对象的转发频率 1011 和拥有权 1012 信息。如下文所述，

拥有权可以单独在 A 中 10122, 或者在两个或更多的协同高速缓存服务器之间共享 10121、10123。转发频率 1011 表示用于一个对象的请求已经通过负载监视器转发的次数。除了转发频率 1011 之外, 高速缓存表 101 还可以维护时间戳 1013, 而时间戳表示用于一个对象的请求被转发的最近时间。另外, 用于一个对象或分区 1010 的高速缓存信息可以包括在给定时间段(计数/时间)上对于通过负载监视器 120 的对象标识或分区标识 1010 的转发频率。对象的分区 1010 另外可以根据关于对象标识符的哈希函数, 或者可以根据在 Web 服务器上组织对象的目录结构。在一个分区的情况下, 属于一个分区的任何对象将被负载监视器所转发。拥有权的转移可以根据服务器的负载状况、转发频率 1011 和其它信息, 例如时间戳信息。

图 3 示出响应由于一个高速缓存不中而来自高速缓存服务器 150 的请求由负载监视器 120 所执行的步骤的逻辑流程的一个例子。如图所示, 在图 201 中, 检查所请求对象/分区可以在高速缓存表中找到。如果为否, 在步骤 202 中, 为该对象/分区创建一个新的条目, 并且一个高速缓存服务器被指定为拥有者。在该条目位于高速缓存表中之后, 在步骤 203 中, 转发频率 1011 被更新, 例如, 加 1。然后, 负载监视器检查负载表 102, 以在步骤 204 中察看该拥有者当前是否过载(以及看转发频率 1011 是否为一个重要的因素)。如果为是, 在步骤 205 中, 负载监视器找到一个低负载(或者较小负载)的高速缓存服务器, 并且把其指定为所请求对象的新的(或共享的)拥有者 10122。在高速缓存表 101 中对于该对象的拥有权信息 1012 被相应地更新。本领域内的技术人员将认识到该逻辑流程可以在该高速缓存表 101 中包括共享的 10123 或分级别的拥有权 1012 或采用其它数据结构。该请求(可能带有被请求对象的副本)然后可以在步骤 206 中被转发 125 到一个新的单独的 10122(或者共享的 10123)的拥有者。另外, 新的拥有者可以被请求从源发对象服务器获得 115 一个对象的副本, 例如, 通过因特网 110。本领域内的技术人员将认识到, 负载检查步骤 204 可以被主动执行, 即, 定期地或者响应一个对于给定对象标识/分区标识 1010 以及高速

缓存服务器(拥有权 1012)的被识别的至少部分由于较高转发频率 1011 所造成的过载或过载倾向 1021。如果是这样的话,则在步骤 205,负载监视器找到一个低负载(或者较小负载的)高速缓存服务器,把其指定为所请求对象的新的(或者共享的)拥有者,并且可能如上文所述把该对象的一个副本发送到该新的(或者共享的)拥有者。相反,如果在步骤 208 中一个共享的拥有权模型被使用,当对于共享拥有的对象(p 10101)的负载状况 10211 和转发步骤 10111 下降到预定阈值,在步骤 209 中,共享的拥有权(B,A 10121)可以合并到单个拥有权,并且其中一个副本被从该高速缓存服务器 A 10121 中清除,例如,为另一个热点对象腾出空间。

图 4 示出当对于一个对象的请求被直接 155 从浏览器 160 接收或者从负载监视器 120 转发 125 接收时用于一个缓冲服务器的逻辑流程的例子。如图所示,在步骤 301 中,首先检查被请求对象是否可以在其本地高速缓存中找到。如果为是,在步骤 302 中,它返回该对象并且该过程结束,在步骤 306 中。否则,在步骤 303 中,检查该请求是一个直接请求还是一个转发请求。如果它是一个直接请求,则在步骤 304 中,该请求被发送到负载监视器并且该过程结束,在步骤 306 中。另一方面,如果该请求是一个转发请求,在步骤 305 中,高速缓存服务器将从源发 Web 服务器取得该对象并且返回该对象。然后该过程结束,在步骤 306。

现在参照图 1a 和 2a-b,例如假设,一个连接到高速缓存服务器 C 10223 的浏览器 160 请求 155 一个对象 p 10101。从高速缓存表 101 中,可以看出对象 p 10101 没有被高速缓存在服务器 C 上,但是它高速缓存在(“被拥有”)高速缓存服务器 B 上(假设 B, A 10121 最初仅仅由 B 所指定)。响应对象 p 的高速缓存不中,服务器 C 10223 把一个请求发送到负载监视器 120 要求对象 p。根据在服务器 B 上要求 p 10101 的请求的负载状况 10212 以及转发频率 1011,负载监视器可以把该请求转发到服务器 B,请求它把对象 p 的副本发送到服务器 C。或者,如果服务器 B 当前为过载或者它趋向过载,则负载监视器可以通过找到

一个低负载(或者较小负载)的服务器来作为一个新的(或者如在 B, A 10121 中共享的)对象 p 的拥有者, 而把转发请求转移。然后, 该请求被转发到对于该对象的新的(或者共享的, 例如 A)拥有服务器。请注意, 即使在拥有权转移之后, 对象 p 的副本仍然在服务器 B 的高速缓存中, 并且仍然可以处理来到服务器 B 的直接请求。但是, 在该例子中, 将来对于对象 p 的所有转发请求(或者在共享拥有权的情况下可能是一些请求)将被转移到服务器 A。另外, 在共享拥有权 B, A 10121 的情况下, 对于对象 p 10101 的将来的转发的请求可以被发送到较小负载的服务器。

现在, 根据本实施例的负载均衡方法用于代理高速缓存服务器的集合已经进行了描述, 其中一个逻辑中央目录被用于定位一个对象, 但是各种变型将被考虑。本发明可以适用于对这些系统获得负载均衡。

例如, 本发明可以用于对协同代理高速缓存服务器的集合执行负载均衡, 其中每个高速缓存服务器 150 多点传送到一协同高速缓存服务器列表, 以定位本地不中的对象的副本。在这种情况下, 没有任何特定的拥有权信息需要被保持在该系统中的任何地方。但是, 在此也不能保证从协同高速缓存服务器找到一个对象。假设一个逻辑负载监视器 120 被用于维护所有代理高速缓存服务器的负载状况 1021 并且与每个高速缓存服务器共享此信息。负载均衡可以通过从该协同服务器列表中排除过载的服务器而实现, 其中一个高速缓存服务器把其请求(也称为转移请求)多点传送到该高速缓存服务器的列表上。结果, 仅仅较小负载的高速缓存服务器将接收被转发请求 125。

另一个替换是一种用于协同代理高速缓存服务器的集合的负载均衡方法, 其中一个哈希函数被用于定位本地不中的对象的一个副本。在这种情况下, 对象空间可以被在该协同代理高速缓存服务器 150 中分区, 一个分区用于每个高速缓存服务器。为了通过转移被转发请求而实现负载均衡, 可以改变哈希函数, 使得被转发请求不能到达过载服务器。一个优选的方法是把对象空间散列为大量存储段(bucket), 比代理高速缓存服务器的总数大的多。这些散列存储段然后被分配给

高速缓存服务器，目的是在它们之间均衡负载。可以定期地把一个或多个散列存储段从一个过载服务器转移到一个低负载服务器，有效地改变哈希函数。

在任何一种情况中，协同高速缓存服务器的负载状况可以把转发频率作为要素直接包括到所计算的负载状况中。例如，负载状况可以是到所述协同高速缓存服务器的所述转发请求的计数值以及直接请求的计数值的加权和。另外，负载监视器可以分别维护对于每个协同高速缓存服务器的整体转发频率。

现在参照图 1b 和 2a-b, 另一个变型是分布的负载监视器 120, 即, 其中一些或所有负载监视器被在高速缓存服务器 150 上重复。在一个例子中, 分布的负载监视器包括协同高速缓存服务器 150 的本地负载状况信息 1021(如下文所述, 可能是所有(A, B, C...1022)的负载状况)。分布的负载监视器 120'最好还包括具有对于每个对象标识/分区标识 1010 的转发频率 1011 和拥有权 1012 的高速缓存表 101。另外, 一个哈希函数, 如上文所述, 可以分布并存储在该高速缓存服务器中。负载状况信息 1021 和/或高速缓存信息 101: 可以被定期交换; 当存在状态改变时(拥有权或者在负载状况中的重大改变); 或者搭载高速缓存转发请求和响应。负载状况 1021 信息还可以具有一个与用于跟踪或其它目的相关的时间戳(未示出)。

在此, 如果一个高速缓存服务器 150 具有一个高速缓存不中, 则本地负载监视器 120'在其本地高速缓存表 101 中查找被请求对象的拥有权, 并且把该请求转发到该拥有高速缓存服务器。另外, 哈希函数可以应用于来自该请求的一个关键字, 例如 URL(统一资源定位符)或者目的 IP 地址, 以识别该分区, 然后该请求被转发到正确的高速缓存服务器。当被转发请求(即, 来自具有高速缓存不中的一个高速缓存服务器)被接收时, 该拥有高速缓存服务器把它识别为一个转发请求(例如, 通过把它识别为来自与一个客户机相对的另一个高速缓存服务器), 并且把其转发频率 1011 信息更新为可应用(图 3, 步骤 203)。如果表明一个过载倾向或状况(步骤 204), 该拥有高速缓存服务器可以用

一个转移请求和所高速缓存对象的副本来响应发出请求的高速缓存服务器。另外，发出请求的高速缓存服务器可以通过内部网、广域网或因特网 110 从源发对象服务器获得一个副本。在任何情况中，当转发服务器高速缓存该对象的一个副本时，只要该对象保留在高速缓存中，该服务器不再发出转发请求(步骤 301, 302)，从而成比例地减小该拥有服务器上的负载。另外，该拥有高速缓存服务器可以把一个转移请求消息多点传送到一个或多个其它协同高速缓存服务器 150，使得后续的转发请求被转移，例如，通过更新高速缓存表的本地副本或者修改哈希函数(步骤 205)。在这一点，其它高速缓存服务器可以把它们的请求转发到新的拥有者(或者转发到两个或多个高速缓存服务器 150 的最小负载拥有者，如果拥有权为共享的话)，如高速缓存表 101 的本地副本中所示。当源发高速缓存拥有者的负载下降到一个可接受的水平(步骤 204)，例如，由一个阈值所示，该共享拥有信息可以被合并到其原始状态(例如，B, A 10121--> B)。

在用于所有高速缓存服务器(A, B, C...1022)的负载状况信息 1021 被完全分布的情况下，发出请求的高速缓存服务器可以主动检查拥有服务器的负载状况(以及相关时间戳)(步骤 204)，即，在转发该请求之前。如果过载，则发出请求的服务器可以从拥有服务器(或者通过内联网或因特网 110 从源发服务器)要求该对象的一个副本，并且可能要求一个负载状况确认。该拥有高速缓存服务器可以更新其高速缓存表 101 或者修改哈希函数，以指示新的共享拥有权(步骤 205)。然后，发出请求的服务器(或者拥有服务器)可以把一个消息多点传送到所有其它高速缓存服务器 150，指示该对象的新的共享拥有权，并且可能包括一个更新的负载状况。在这一点，其它高速缓存服务器将更新它们的高速缓存表，或者修改哈希函数，以指示新的共享拥有权(步骤 202)，并且可以把它们的请求(步骤 206)转发到共享拥有权的两个或多个高速缓存服务器 150 的较小负载的拥有者，如高速缓存表 101 的本地副本中所示。当共享的高速缓存拥有者的负载已经降低到一个可接受的水平(步骤 204 和 208)，例如，如一个阈值所示，则拥有权信息可以合

并到其原始状态，在步骤 209 中。

本发明的一个优选实施例包括如下特征，即可以作为实际体现在计算机程序产品或者程序存储设备上，用于在具有高速缓存服务器 150 的处理器(未示出)或者实现负载监视器 120 的其它计算机上执行的软件而实现，例如在上述集中模式中。例如，以流行的面向对象的计算机可执行代码实现的软件，例如 JAVA 语言，提供跨越不同平台的可移植性。本领域内的技术人员将认识到还可以采用许多其它编译或解释，面向过程和/或面向对象(OO)编程环境，包括 REXX、C、C++ 和 Smalltalk 语言，但不限于此。

本领域内的技术人员还认识到本实施例的方法可以是体现在磁性、电、光或者其它持久编程和/或数据存储设备上的软件，其中包括但不限于：磁盘、直接存取存储设备(DASD)、磁泡存储器；磁带；光盘格式，例如 CD-ROM 和 DVD；以及其它持久(也称为非易失)存储设备，例如磁芯存储器、ROM、PROM、闪存器、或者电池供电的 RAM。本领域内的技术人员将认识到在本发明的精神和范围内，在服务器 120' 的存储器中的一个或多个部分可以直接通过磁盘(未示出)、网络、其它服务器而存取和保存，或者可以在多个服务器上分布。

总而言之，在包括协同高速缓存服务器，例如代理高速缓存服务器，的集合的系统中，如果所请求对象不能在本地找到，则该请求可以被转发到协同高速缓存服务器。如果例如由于引用不均衡而造成过载状况被检测，即所有客户机所要求一些对象，则包含这些热点对象的高速缓存服务器由于转发请求而过载。相应地，通过把一些或所有的转发请求从过载的高速缓存服务器转移到较小负载的服务器而均衡负载。集中和分布的负载均衡环境都已描述。

尽管我们已经描述了本发明的优选实施例及其替代，但是本领域内的技术人员应当知道，现在和将来都可以进行各种改进和增强，而这些改进落在下文的权利要求的范围中。这些权利要求应当被认为是保持对第一次公开的本发明的适当保护。

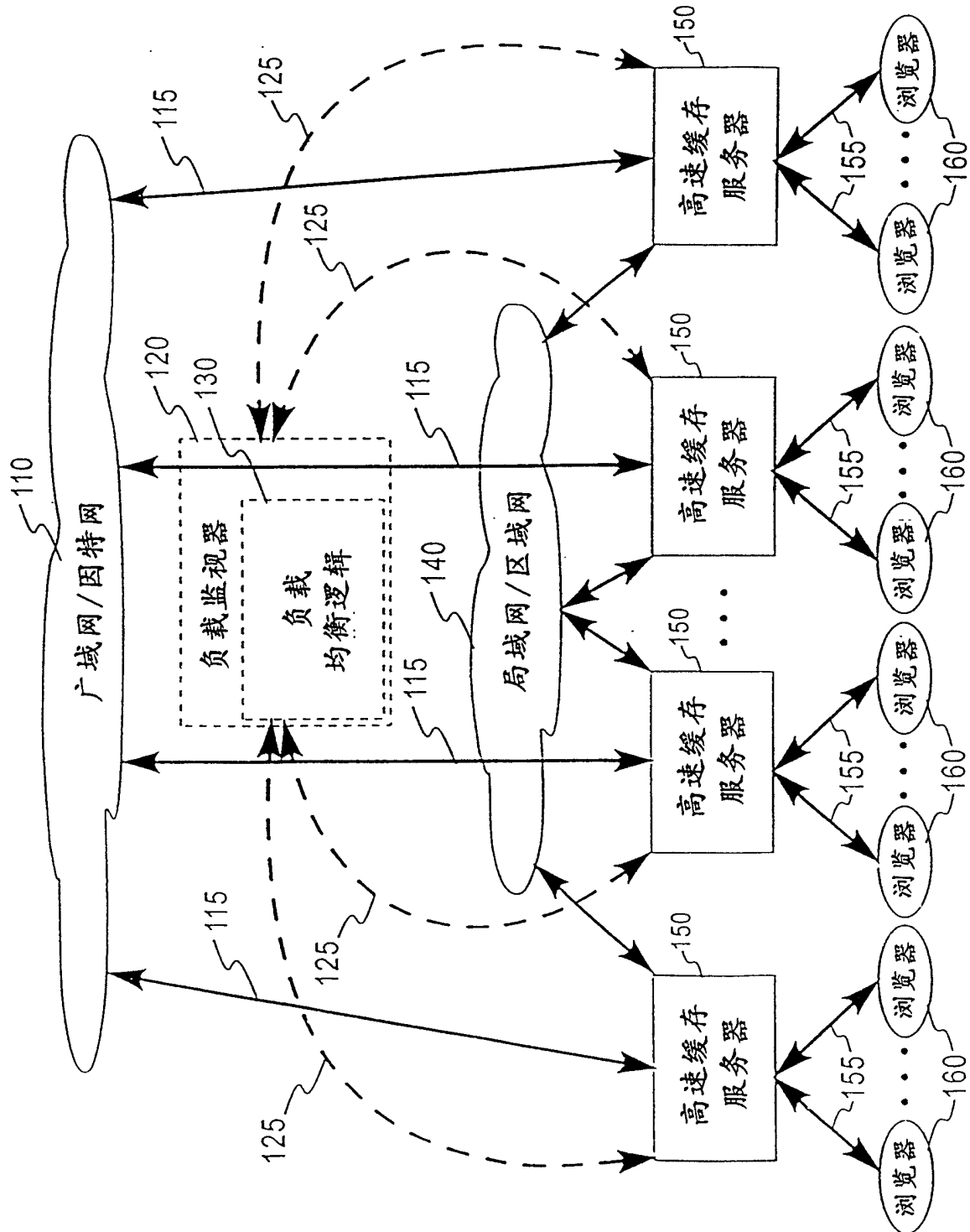


图 1a

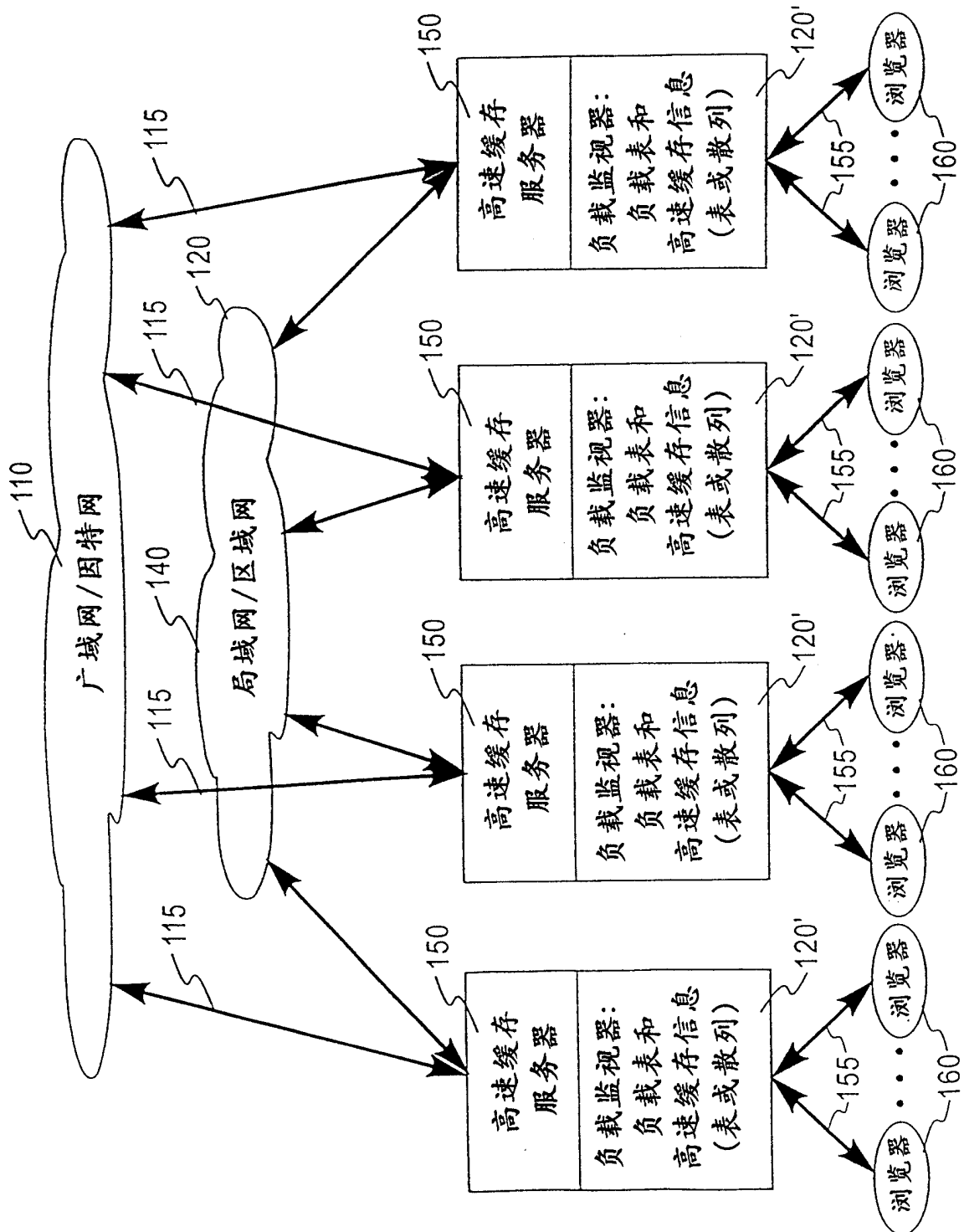


图 1b

101

高速缓存表			
对象标识/ 分区标识	时间戳	转发频率 (计数或速率)	拥有权
p	hh:mm	25,2	B,A
q	hh:mm	2	A
r	hh:mm	10,50	C,A

1010 10101 10102 10103

1011 1012 10121 10122 10123

图 2a

102

负载表	
高速缓存服务器	负载状况
A	0.15
B	0.45
C	0.80

10221 10222 10223

1021 10211 10212 10213

图 2b

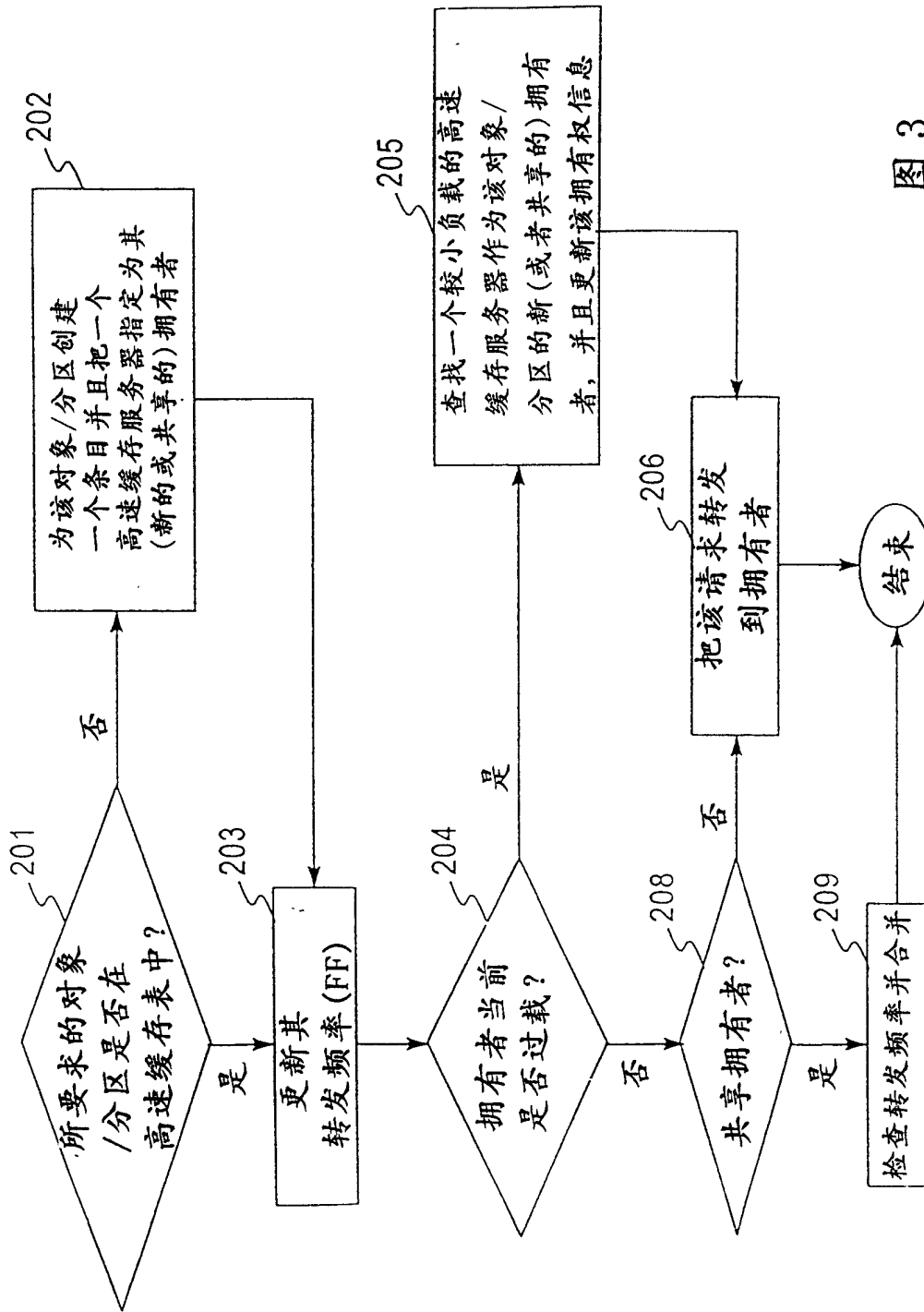


图 3

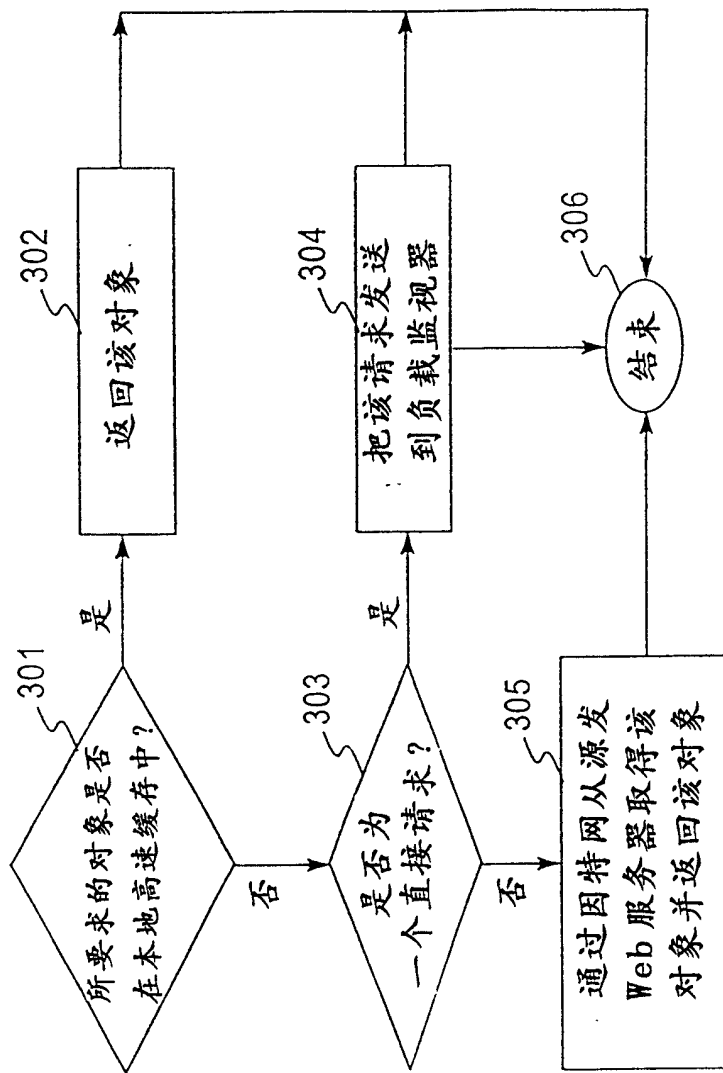


图 4