



[12] 发明专利申请公布说明书

[21] 申请号 200480042399.4

[43] 公开日 2007 年 6 月 13 日

[11] 公开号 CN 1981501A

[22] 申请日 2004.3.12

[21] 申请号 200480042399.4

[86] 国际申请 PCT/US2004/007652 2004.3.12

[87] 国际公布 WO2005/099223 英 2005.10.20

[85] 进入国家阶段日期 2006.9.12

[71] 申请人 汤姆森许可公司

地址 法国布洛涅 - 比扬古

[72] 发明人 J·李 J·张 S·维马

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 程天正 刘杰

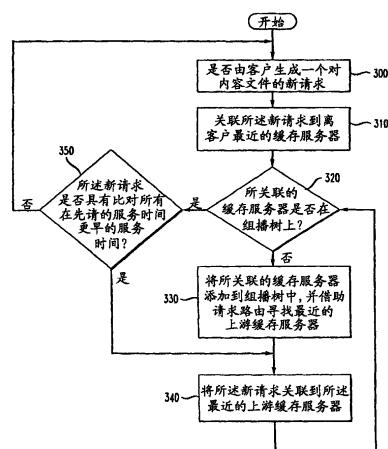
权利要求书 3 页 说明书 5 页 附图 2 页

[54] 发明名称

调度内容文件的分发的方法和缓存服务器网络

[57] 摘要

公开了一种用于在内容递送网络中调度内容文件的分发的技术，和一种适于执行这种操作的内容递送网络。该技术包括基于递送的位置、内容请求的服务时间和缓存服务器的分层，来调度内容文件的分发。优选地，基于位置和服务时间的因素，在内容递送网络中动态建立一用于递送每个内容的组播树。



1.一种用于从内容递送网络系统处理对内容文件的请求的方法，该方法包括：

接收一在服务时间递送内容的请求；

将该内容文件和一特定缓存服务器关联；

动态建立缓存服务器的组播树；以及

在服务时间，从缓存服务器的组播树递送所请求的内容。

2.如权利要求 1 所述的方法，其中所述关联的步骤进一步包括将所述请求与一最近的缓存服务器相关联的步骤。

3.如权利要求 1 所述的方法，其中所述动态建立步骤进一步包括以下步骤：

如果所述缓存服务器还没有与所述组播树关联，就增加一个与一请求关联的缓存服务器。

4.如权利要求 1 所述的方法，其中所述关联步骤进一步包括以下步骤：

如果所述请求具有早于在前请求的服务时间，将所述请求与一最近的缓存服务器关联。

5.一种用于从内容递送网络系统处理对内容文件的请求的方法，该方法包括：

(a) 接收对一对于具有第一服务时间的内容文件的第一请求；

(b) 将所述第一请求与一获取用的缓存服务器相关联；

(c) 确定所述关联的缓存服务器是否在一个组播树上，该组播树的根处于作为内容文件源的内容服务器；

(d) 当确定所述关联的缓存服务器不在所述组播树上时，就将关联的缓存服务器添加到所述组播树上，查找一朝向内容服务器的上游缓存服务器，将所述第一请求与所找到的上游缓存服务器相关联，使得该上游缓存服务器成为关联的缓存服务器，并重复步骤 (c)，直到到达所述内容服务器以及第一请求被关联到该内容服务器，其中当该第一请求与所述内容服务器关联时，就从步骤 (a) 开始处理对内容文件的下一个请求；

(e) 当确定所关联的缓存服务器在所述组播树上时，就确定所述第一服务时间是否早于所有对已经存在于该关联缓存服务器上的内容文件的请求的服

务时间；

(f) 当确定所述第一服务时间不早于所有对已经存在于该关联缓存服务器上的内容文件的请求的其它服务时间，就关联该第一请求与该内容服务器并从步骤(a)开始处理对内容文件的下一个请求；和

(g) 当确定所述第一服务时间早于所有对已经存在于该关联缓存服务器上的请求的其它服务时间，就将该第一请求与已被确定是朝向组播树的内容服务器的上游缓存服务器的缓存服务器相关联，使得这个缓存服务器成为关联的缓存服务器，并返回到步骤(c)，直到所述第一请求被关联到内容服务器，其中当所述第一请求与所述内容服务器关联时，就从步骤(a)开始处理对内容文件的下一个请求。

6.如权利要求5所述的方法，其中查找上游缓存服务器的步骤包括使用请求路由程序查找最近的上游缓存服务器。

7.如权利要求6所述的方法，其中使用从地理距离、缓存的占有率和网络链路的负载均衡所组成的组中选择的至少一个因素来确定接近程度。

8.如权利要求5所述的方法，其中查找上游缓存服务器的步骤包括使用一种分层关系来查找所述上游缓存服务器。

9.一种用于处理对内容文件的请求的内容递送网络系统，包括一内容服务器和一具有至少一个缓存服务器的CDN网络，适于(a)从客户接收对于内容文件的第一请求；(b)将所述第一请求与获取用的缓存服务器关联，(c)确定所关联的缓存服务器是否在一个组播树上，并且如果不在，将所述内容服务器关联到所述组播树。

10.如权利要求9所述的系统，具有确定第一服务时间是否早于所有对已存在于所述组播树中所关联的缓存服务器上的请求的其它服务时间的装置，并且如果不早于，该装置就将所述第一请求与组播树上的缓存服务器关联，以及如果早于，该装置就查找一个上游缓存服务器并将所述第一请求与上游缓存服务器关联，直到所述第一服务时间不早于所有对已存在于所述组播树上的所关联的缓存服务器上的请求的其它服务时间或者所述第一请求被关联到内容服务器为止。

11.如权利要求9所述的系统，具有用于如果所述第一服务时间不早于所有对已存在于所述树上所关联的缓存服务器上的请求的其它服务时间，将所述

第一请求与已经确定是在组播树上的缓存服务器关联的装置。

12.权利要求 9 所述的系统，具有通过路由请求的方式查找最近的上游缓存服务器的装置。

13.如权利要求 12 所述的系统，具有使用从地理距离、缓存的占有率和网络链路的负载均衡所组成的组中选择至少一个因素来查找最近的上游缓存服务器的装置。

14.如权利要求 9 所述的系统，具有使用一种分层关系来查找一上游缓存服务器的装置。

15.如权利要求 9 所述的系统，进一步包括一内容递送网络代理，该代理适于为一个请求路由程序提供信息，该请求路由程序的结果将由所述内容服务器使用，并且该信息是关于内容递送网络中一个或多个缓存服务器上所请求的内容文件的可用性。

16.如权利要求 9 所述的系统，进一步包括一内容递送网络代理，该代理适于为一个请求路由程序提供信息，该请求路由程序的结果将由所述内容服务器使用，并且所述消息是关于在内容递送网络中一个或多个缓存服务器上所请求的内容文件的可用性，或者用于调度一个或多个缓存服务器 13 上所述内容文件的未来可用性。

17.如权利要求 9 所述的系统，还包括一个或多个缓存服务器和内容递送网络代理及用于代理、缓存服务器的装置，和/或用于确定未来时间段的内容服务器和客户可以向其请求文件的服务器。

调度内容文件的分发的方法和缓存服务器网络

技术领域

本发明一般涉及数据通信和缓存服务器网络领域，特别是用于在内容递送网络中调度内容文件的组播分发的系统和方法。

背景技术

对于较大的内容（例如电影）来说，内容客户通常能够容忍用一些延迟来换取更好的质量。客户可能宁愿在将来预定的时间观看高质量的下载视频，也不愿意即刻观看低质量的数据流视频。例如，一移动用户当处于他/她在蜂窝移动网络中时可预先订购一个视频，并且在稍后的时间当他/她处于热点无线 LAN 中时下载该视频。这就是众所周知的远程站点下载。这样，移动用户可以低成本享受高质量的内容。

近年来，内容递送网络（CDN）技术的使用已扩展到因特网上，来改善网页的下载。内容递送网络（CDN）由多个处于不同地理位置的缓存服务器，即具有存储和传输能力的网络节点组成。CDN 技术的基本前提是缓存服务器和客户之间的链路具有低成本和高带宽。如果在客户请求内容文件的时候，该内容文件存储在一个邻近的缓存服务器的缓存中，那么该下载将进行得很快。否则，该客户可能经历一个较长的延迟。因此，优选地客户从最近的缓存服务器中下载内容文件。为客户查找邻近的缓存服务器的技术被称为请求路由。这是将内容请求重定向到更近的缓存服务器的一种程序。例如修改 URL，将原始的 URL 修改成由缓存服务器加上前缀的 URL。在我们的另一种应用中，提供了一种具有内容定时的常规请求路由的扩展，来基于缓存服务器上所请求内容的未来可用性将该请求重定向到更接近的缓存服务器。

一般地，客户可以容忍对较大尺寸的内容文件的延迟，最多达到客户指定作为他/她希望获得该内容文件的时间的期望的服务时间。这样，即使请求的内容文件当前没有存储在邻近客户的缓存服务器中，只要在所期望的服务时间之前下载系统将内容文件转送到邻近的缓存服务器，客户就将不会经历延迟。通过恰当地调度请求的内容文件下载到适当的缓存服务器供客户提取，来减少这

些延迟也是这个行业的目标。

可以在不同的缓存服务器上请求组播内容递送。由于在内容下载之前在先内容请求信息的可用性，所以通过组播技术在 CDN 中优化内容分发是可能的。典型地，下载服务要求 CDN 将内容文件的分发提供给离产生对那个内容文件的客户请求的地方最近的缓存服务器。该内容文件必须存储在那个缓存服务器上，并且准备在不晚于客户所指定的期望服务时间将其下载到客户端。因此，存在一种需求，希望有一种将内容文件的分发调度到与对那个内容文件的请求关联的缓存服务器的改进系统和方法。

发明内容

简要地，本发明涉及一种在缓存的网络环境下调度内容文件的分发的方法。该方法包括以下步骤：接收一个在服务时间递送内容的请求，将该内容文件与一特定缓存服务器相关联，动态建立缓存服务器的组播树，和在服务时间从缓存服务器的组播树递送所请求的内容。

附图说明

图 1 是根据本发明一个实施例的一种内容递送网络的示意图。

图 2 是考虑缓存服务器的接近程度根据本发明一个实施例的图 1 内容递送网络的示意图，示出了由对内容文件的许多用户请求所产生的下载和关联结构。

图 3 是考虑缓存服务器的接近程度根据本发明一个实施例的一种在内容递送网络中调度递送和下载内容文件的方法流程图。

具体实施方式

组播分发可以在传输层或应用层实现。因为存在许多与传输层组播关联的缺点，所以本发明仅考虑应用层组播。传输层组播需要一种能够组播的传输网络。因特网典型地不具有这种传输网络。另外，即使有一种可用的可组播的传输网络，组播树的所有分支上的传输也要必须是同时的。在任一组播会话期间，要是组播树上任何网络节点（也就是缓存服务器）不具有传输或缓存能力，这可能就不可行了。然而，在组播树上安排点到点的传输，应用层组播能有更好的灵活性。对于在不同的期望服务时间都有很多下载请求的下载服务，应用层组播可能更合适。在这里所使用的，应用层组播定义为在组播树上的每一个网络节点处的存储/转发行为（forward action）。存储意味着在中间节点上缓存，，

而转发意味着同时或在不同时间向多个端口传输。

现在参照图 1, 示出了本发明的一种内容递送系统 100 的实施例。系统 100 包括内容服务器 S, 和 CDN, 该 CDN 包括缓存服务器 A、缓存服务器 B 和缓存服务器 C。虽然为仅具有三个缓存服务器的 CDN 示出了一个特定的结构, 但本发明不限于任何特定的网络结构或配置。不管 CDN 网络的结构(平面的或分层的)如何, 组播树都能根据本发明而建立起来以供内容文件分发之用。只要对相同内容文件的请求不具有相同的期望服务时间, 下载和关联结构(也就是组播树结构)就将不仅取决于源和目的地之间的距离, 还取决于不同的期望服务时间之间的间隙。

下面参照图 2 和图 3, 图 2 示出了使用数据流箭头连接所示各部件的内容递送系统 100, 图 3 示出了本发明方法的一种实施例, 在步骤 300, 客户/用户 A1、B1 和 C1 每个都生成一个对相同内容文件的请求。每一个对内容文件的请求都具有一个预计的服务时间。如在这里所使用的, 请求的服务时间是客户请求内容文件可用来下载到客户电子设备上的时间, 所述客户电子设备可以是计算机、蜂窝电话或个人数据助理。当产生请求时, 服务时间由客户 A1、B1 和 C1 指定。在这个实施例中, 由客户 A1、B1 和 C1 生成的请求的服务时间分别是 7PM、5PM 和 8PM。尽管服务时间具有 B1-A1-C1 的时间次序, 但对这些请求可能已经由客户生成和/或由系统 100 接收的次序则没有限制。

由客户 A1、B1 和 C1 所生成的请求(为了简单理解将称之为请求 A1、B1 和 C1)分别与缓存服务器 A、B 和 C 相关联, 这样完成步骤 310。在图 2 中, 请求 A1、B1 和 C1 与缓存服务器 A、B 和 C 的关联分别由线 1、2 和 3 所指定。关联的决定由 CDN 通过静态分层结构或动态请求路由过程来做出。每个请求与哪个缓存服务器相关联的选择是由该缓存服务器与客户的接近程度来确定的。优选地, 一个请求与作出对内容文件请求的客户最近的缓存服务器相关联。如这里所使用的, 接近程度可考虑带宽成本和/或拥塞, 使用物理位置和/或网络所在地来确定。用来完成这种关联的技术就是众所周知的请求路由技术。哪个缓存服务器关联到一个请求的确定可以被静态学习(也就是说, 同在多代理服务器中一样, 存在预先建立的分层结构), 或者被动态地学习。可选择地, 客户可以指定一个他/她的请求应该要关联到的一个特定的缓存服务器, 例如指定一个具有热点远程站点下载功能的缓存服务器。

当一个请求与缓存服务器的关联是通过请求路由技术被动态地确定时，应该使用扩展的请求路由技术。在这种情况下，即使所请求的内容文件当前在缓存服务器上不可得到，请求路由仍能将该请求关联到那个缓存服务器，因为这个关联意味着要在一个未来的时间把该内容文件递送到那个缓存服务器。

按照 B1、C1 和 A1 的次序将请求 A1、B1 和 C1 发送到内容服务器 S。组播树最初仅有一个节点，内容服务器 S。因为请求 B1 是发送到内容服务器 S 的第一个请求，所以首先为缓存服务器 B 执行步骤 320。在步骤 320，确定该缓存服务器 B 是否在组播树上。如果答案是否（在这种情况下确实如此），那么该系统就将节点 B 添加到组播树上并继续到步骤 330。在步骤 330，该系统检查最近的上游缓存服务器的存在，并且在这种情况下，找到游缓存服务器 C。这是通过静态分层或请求路由来进行的。举例说明了请求路由。然后将请求 B1 关联到缓存服务器 C，这样完成步骤 340。请求 B1 到缓存服务器 C 的关联在图 2 中由线 4 示出。

然后为缓存服务器 C 执行步骤 320。根据步骤 320，接着确定缓存服务器 C 是否在组播树上。如果答案是否，在这种情况确实如此，该系统就将节点 C 添加到组播树上。然后缓存服务器 C 查找一个离它最近的上游节点，该节点就是内容服务器 S，这样完成步骤 330。然后在步骤 340 中请求 C1 关联到所述内容服务器 S，这在图 2 中由线 5 示出。

然后为内容服务器 S 执行步骤 320。根据步骤 320，由于所述内容服务器 S 在组播树上，所以答案就是是并且进入步骤 350。由于当前服务器是该内容服务器，所以在步骤 350 答案是否并且继续进行处理下一个请求。

现在转向请求 C1，请求 C1 在步骤 300（继请求 B1 之后）生成，并且在步骤 310 与缓存服务器 C 关联。由于在执行对请求 B1 的过程时，节点 C 就已经被添加到组播树上了，所以在步骤 320 的答案是是，并且该过程继续进行到步骤 350。由于 C1 的服务时间（8PM）晚于 B1 的服务时间（5PM），所以步骤 350 的答案是否。然后，该过程重新开始并处理下一个请求。

现在转向请求 A1，请求 A1 是继请求 C1 之后接收到的，请求 A1 是在步骤 300 生成的并在步骤 310 与缓存服务器 A 相关联。根据步骤 320，确定缓存服务器 A 是否在组播树上。在这种情况下，如果答案是否，然后该过程继续到步骤 330。在这一点，节点 A 首先被添加到所述组播树上，然后缓存服务器 A

就查找其上游缓存服务器 B，这样完成步骤 330。然后，在步骤 340 中，请求 A1 被关联到缓存服务器 B。此关联在图 2 中由线 6 示出。由于节点 B 已经在组播树上并且 A1 的服务时间晚于节点 B 的服务时间，所以步骤 350 的答案是否。该过程然后重新开始并处理下一个请求。

在确定缓存服务器之间距离中使用的算法不仅基于地理距离还基于其它因素，例如缓存容量、网络链路的负载均衡等。例如节点 A 可能发现节点 C 是其上游节点，因为在节点 B 把内容从 5PM 缓存到 7PM 的成本可能大于链路 7 和链路 6 之间成本差。

尽管本发明已经以充分的细节进行了描述和举例说明，但是本领域技术人员将容易地制造和使用本发明，并且在不脱离本发明的精神和范围的情况下，各种可选择的替代、修改和改进都变的很明显。

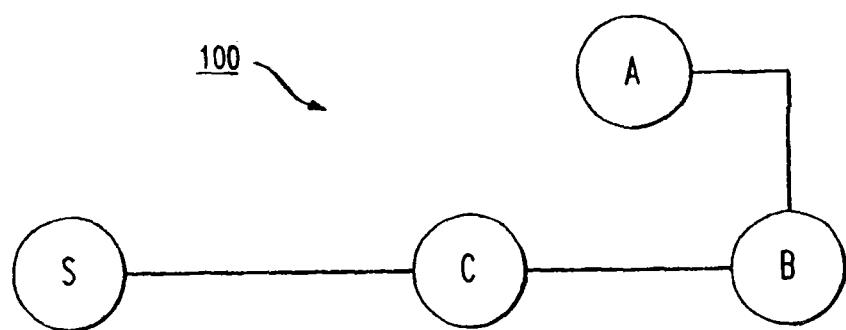


图 1

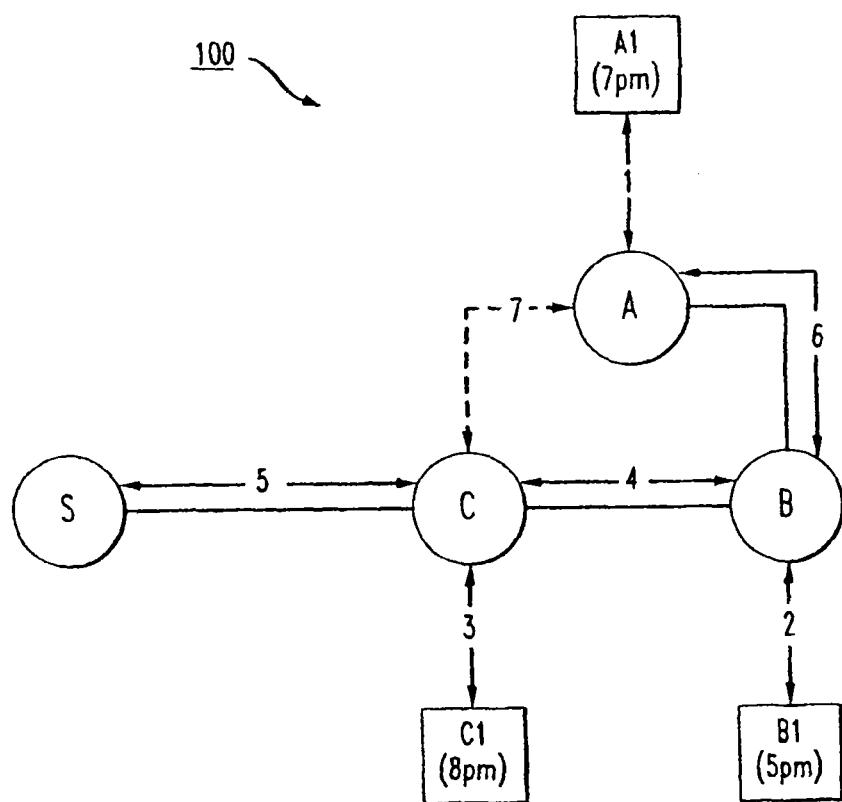


图 2

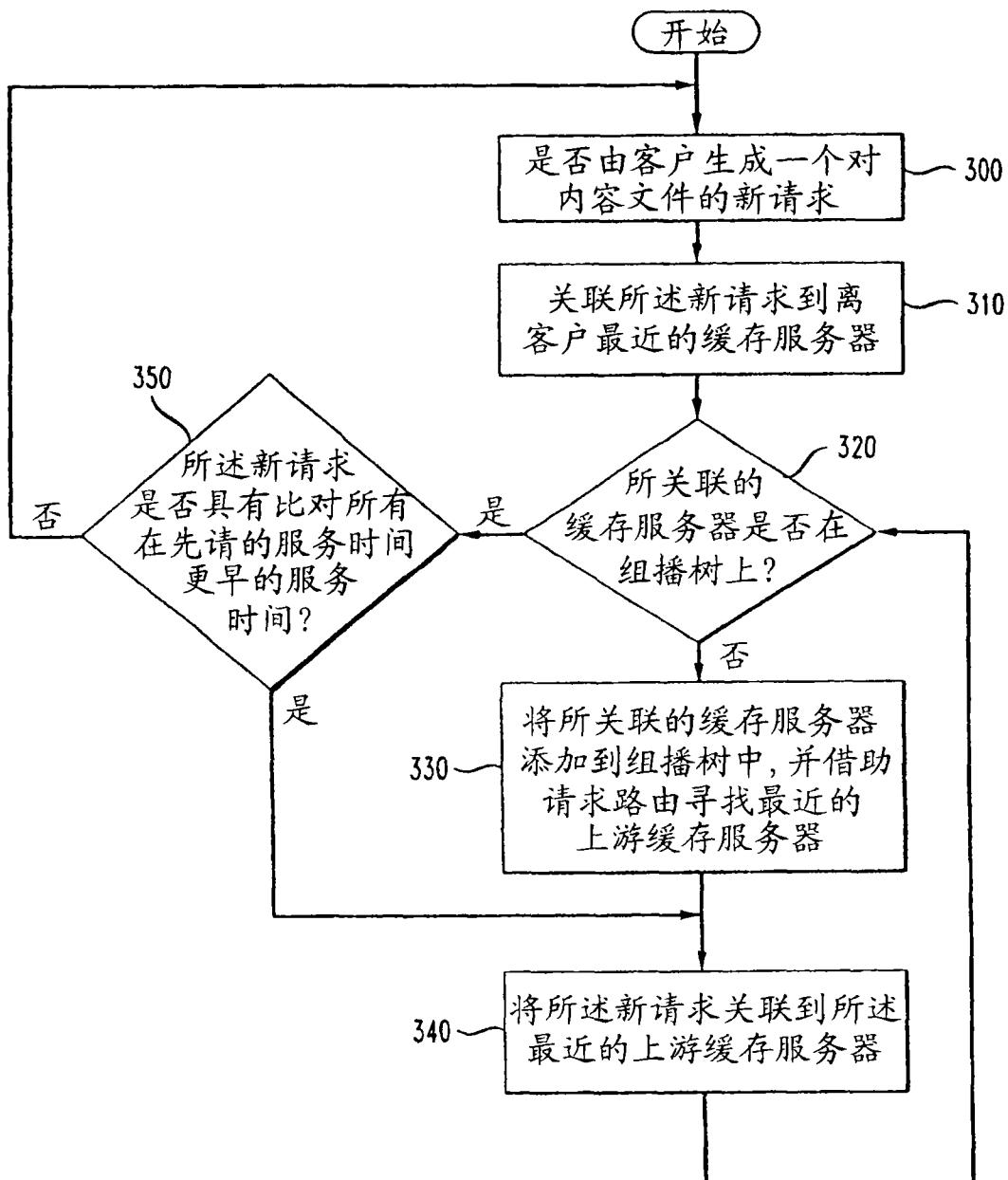


图 3