



(12) 发明专利

(10) 授权公告号 CN 101322116 B

(45) 授权公告日 2012. 06. 27

(21) 申请号 200680045183. 2

(51) Int. Cl.

(22) 申请日 2006. 12. 04

G06F 17/00 (2006. 01)

(30) 优先权数据

60/742, 508 2005. 12. 05 US

11/364, 234 2006. 02. 28 US

(56) 对比文件

US 6463501 B1, 2002. 10. 08,

WO 2005024552 A2, 2005. 03. 17,

(85) PCT申请进入国家阶段日

审查员 刘欢

2008. 06. 02

(86) PCT申请的申请数据

PCT/US2006/046222 2006. 12. 04

(87) PCT申请的公布数据

W02007/067480 EN 2007. 06. 14

(73) 专利权人 微软公司

地址 美国华盛顿州

(72) 发明人 G · 瑟里亚纳拉亚南 H · 刘

N · S · 伯杰纳

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 顾嘉运

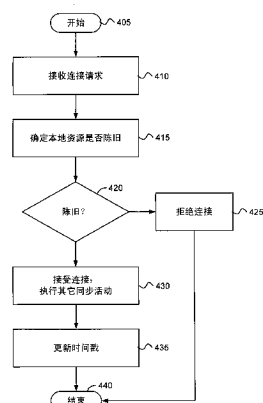
权利要求书 2 页 说明书 12 页 附图 6 页

(54) 发明名称

资源新鲜度和复制

(57) 摘要

此处所描述的主题的各方面涉及资源复制系统。在各方面中,描述了一种检测本地资源何时陈旧的机制。当最后一次成功同步活动与当前时间之间的时间超过一陈旧度值时,本地资源可以是陈旧的。如果本地资源被确定为陈旧的,则本地成员可制止特定的同步活动,直到该本地成员被允许继续这些特定的同步活动。如果本地资源不是陈旧的,则可执行包括在该成员与上游或下游合作伙伴之间同步资源的附加同步活动。



1. 一种计算机实现的用于资源更新的方法,包括:

经由一时间戳来检测主宿在本地机器上的本地资源是否是陈旧的,其中所述时间戳包括所述本地资源的先前同步的时间,所述先前同步包括将从上游合作伙伴接收的更新应用于所述本地资源,所述本地机器和上游合作伙伴参与一复制组,其中经由时间戳来检测本地资源是否是陈旧的包括确定所述时间戳与所述本地机器的本地系统时钟的时间之间的差异以及确定所述差异是否大于一陈旧度值,所述陈旧度值被设置成等于或小于一墓碑周期;

如果所述本地资源是陈旧的,则制止将来自所述本地机器的本地资源的更新发送到下游合作伙伴;以及

如果所述本地资源不是陈旧的,则将来自所述本地机器的本地资源的更新发送到下游合作伙伴。

2. 如权利要求1所述的方法,其特征在于所述陈旧度值能由系统管理员来选择。

3. 如权利要求1所述的方法,其特征在于,所述墓碑周期指示在与一资源相关联的资源元数据被删除之前该资源能在多长的时间内被标记为已删除。

4. 如权利要求1所述的方法,其特征在于,所述同步还包括:

获得所述上游伙伴上与所述本地机器上的相应资源不同的任何资源;

添加包含在所述上游伙伴上、但不包含在所述本地机器上的任何资源;以及

通过更新所述时间戳来将所述本地资源标记为新鲜。

5. 如权利要求4所述的方法,其特征在于,所述同步还包括删除在所述本地机器上、但不在所述上游合作伙伴上的任何资源。

6. 如权利要求1所述的方法,其特征在于,成功的同步包括维持与所述上游伙伴的连接。

7. 一种至少部分地由计算机实现的用于资源更新的方法,所述计算机包括至少一个处理器和存储器,所述方法包括:

储存包括了彼此复制资源的各成员之间的第一同步的先前时间的值,其中,所述第一同步包括将来自第一成员的更新应用到存储在第二成员上的资源;

接收参加第二同步的请求,其中所述请求是在各成员中的第一成员处接收的;

由所述第二成员通过计算当前时间和所述第一同步的先前时间之间的差值并确定所述差值是否超出一陈旧度值来确定在所述第二成员上存储的资源是否更新,所述陈旧度值能被设置成等于或小于一墓碑周期;以及

响应于所述确定步骤,由所述第二成员接受来自第三成员的连接请求,并拒绝传送资源更新给第三成员请求。

8. 如权利要求7所述的方法,其特征在于,所储存的值包括时间戳。

9. 如权利要求7所述的方法,其特征在于,所述第一同步包括在所述第一成员和所述第一成员的合作伙伴之间的成功的资源同步。

10. 如权利要求7所述的方法,其特征在于,所述第二同步包括在所述第一成员和所述第一成员的下游合作伙伴之间同步资源的请求。

11. 一种计算机实现的用于资源更新的系统,包括:

用于经由一时间戳来检测主宿在本地机器上的本地资源是否是陈旧的装置,其中所

述时间戳包括所述本地资源的先前同步的时间,所述先前同步包括将从上游合作伙伴接收的更新应用于所述本地资源,所述本地机器和上游合作伙伴参与一复制组,其中经由时间戳来检测本地资源是否是陈旧的包括确定所述时间戳与所述本地机器的本地系统时钟的时间之间的差异以及确定所述差异是否大于一陈旧度值,所述陈旧度值被设置成等于或小于一墓碑周期;

用于如果所述本地资源是陈旧的,则制止将来自所述本地机器的本地资源的更新发送到下游合作伙伴的装置;以及

用于如果所述本地资源不是陈旧的,则将来自所述本地机器的本地资源的更新发送到下游合作伙伴的装置。

资源新鲜度和复制

[0001] 背景

[0002] 用于复制资源的系统对于确保大型网络中的可用性和容错而言正变得越来越重要。复制包含域凭证和策略的文件的网络是其中需要可用性、可缩放性、一致性和可靠性的一个示例。然而,创建允许各种复制行为的灵活的资源复制系统是具有挑战性的。

[0003] 概述

[0004] 简言之,此处所描述的主题的各方面涉及资源复制系统。在各方面中,描述了一种检测本地资源何时为陈旧的机制。本地资源可能在最后一次成功的同步活动与当前时间之间的时间超过一陈旧性值时变得陈旧。如果本地资源被确定为陈旧的,则本地成员可制止特定的同步活动,直到本地成员被允许恢复这些特定的同步活动。如果本地资源不是陈旧的,则可执行包括在成员与上游或下游合作伙伴之间同步资源在内的其它同步活动。

[0005] 提供本概述以便简要地标识出将在以下详细描述中进一步描述的主题的某些方面。本概述并不旨在确定所要求保护的主题的关键或必要特征,也不旨在用于限制所要求保护的主体范围。

[0006] 除非上下文清楚地另外指明,否则短语“此处描述的主题”指的是在详细描述中描述的主题。术语“各方面”应当被理解为“一个或多个方面”。标识出在详细描述中所描述的主题的各方面并不旨在确定所要求保护的主题的关键或必要特征。

[0007] 当结合附图时,从以下详细描述中,可以清楚以上描述的各方面以及其它方面,附图中:

[0008] 附图简述

[0009] 图 1 是表示其中可包含此处描述的主题的各方面的计算机系统的框图;

[0010] 图 2 是一般表示根据此处描述的主题的各方面的参与复制组的机器的框图;

[0011] 图 3 是表示根据此处描述的主题的各方面的被配置成在资源复制系统中操作的机器的框图;以及

[0012] 图 4 是一般表示根据此处描述的主题的各方面的在确定本地机器上的内容是否足够新鲜来与复制组的其它成员同步时可能发生的动作的流程图;

[0013] 图 5 是一般表示根据此处描述的主题的各方面的在一成员被确定为具有陈旧内容并且被设置为从设备来恢复内容时可能发生的动作的流程图;以及

[0014] 图 6 是一般表示根据此处描述的主题的各方面的在一成员被确定为具有陈旧内容并且被设为将内容与上游合作伙伴合并时可能发生的动作的流程图。

[0015] 详细描述

[0016] 示例性操作环境

[0017] 图 1 示出了适合在其中实现此处所描述的主题的各方面的计算系统环境 100 的一个示例。计算系统环境 100 仅为合适的计算环境的一个示例,并非对此处所描述的主题的各方面的使用范围或功能提出任何局限。也不应将计算环境 100 解释为对示例性操作环境 100 中示出的任一组件或其组合具有任何依赖或需求。

[0018] 此处所描述的主题的各方面可以使用众多其它通用或专用计算系统环境或配置

来操作。适用于此处所描述的主题的各方面的众所周知的计算系统、环境和 / 或配置的示例包括但不限于,个人计算机、服务器计算机、手持式或膝上型设备、多处理器系统、基于微处理器的系统、机顶盒、可编程消费电子产品、网络 PC、小型机、大型计算机机、包括任一上述系统或设备的分布式计算环境等等。

[0019] 此处所描述的各方面可在诸如程序模块等由计算机执行的计算机可执行指令的一般上下文环境中描述。一般而言,程序模块包括执行特定的任务或实现特定的抽象数据类型的例程、程序、对象、组件、数据结构等等。此处所描述的主题的各方面也可以在其中任务由通过通信网络链接的远程处理设备来执行的分布式计算环境中实践。在分布式计算环境中,程序模块可以位于包括存储器存储设备的本地和远程计算机存储介质中。

[0020] 参考图 1,用于实现此处所描述的主题的各方面的示例性系统包括计算机 110 形式的通用计算设备。计算机 110 的组件可包括但不限于,处理单元 120、系统存储器 130 以及将包括系统存储器 130 的各类系统组件耦合至处理单元 120 的系统总线 121。系统总线 121 可以是若干种总线结构的任一种,包括存储器总线或存储器控制器、外围总线以及使用各类总线体系结构的任一种的局部总线。作为示例而非局限,这类体系结构包括工业标准体系结构 (ISA) 总线、微通道体系结构 (MCA) 总线、增强型 ISA (EISA) 总线、视频电子技术标准协会 (VESA) 局部总线以及外围部件互连 (PCI) 总线 (也称为小背板 (Mezzanine) 总线)。

[0021] 计算机 110 通常包括各种计算机可读介质。计算机可读介质可以是可由计算机 110 访问的任一可用介质,包括易失性和非易失性介质、可移动和不可移动介质。作为示例而非局限,计算机可读介质包括计算机存储介质和通信介质。计算机存储介质包括以用于储存诸如计算机可读指令、数据结构、程序模块或其它数据等信息的任一方法或技术实现的易失性和非易失性,可移动和不可移动介质。计算机存储介质包括但不限于, RAM、ROM、EEPROM、闪存或其它存储器技术、CD-ROM、数字多功能盘 (DVD) 或其它光盘存储、磁盒、磁带、磁盘存储或其它磁存储设备、或可以用来储存所期望的信息并可由计算机 110 访问的任一其它介质。通信介质通常以诸如载波或其它传输机制等已调制数据信号来体现计算机可读指令、数据结构、程序模块或其它数据,并包括任一信息传送介质。术语“已调制数据信号”指以对信号中的信息进行编码的方式设置或改变其一个或多个特征的信号。作为示例而非局限,通信介质包括有线介质,如有线网络或直接连线连接,以及无线介质,如声学、RF、红外和其它无线介质。上述任一的组合也应当包括在计算机可读介质的范围之内。

[0022] 系统存储器 130 包括易失性和 / 或非易失性存储器形式的计算机存储介质,如只读存储器 (ROM) 131 和随机存取存储器 (RAM) 132。基本输入 / 输出系统 133 (BIOS) 包括如在启动时帮助在计算机 110 内的元件之间传输信息的基本例程,它通常储存在 ROM 131 中。RAM 132 通常包含处理单元 120 立即可访问和 / 或当前正在操作的数据和 / 或程序模块。作为示例而非局限,图 1 示出了操作系统 134、应用程序 135、其它程序模块 136 和程序数据 137。

[0023] 计算机 110 也可包括其它可移动 / 不可移动、易失性 / 非易失性计算机存储介质。仅作示例,图 1 示出了对不可移动、非易失性磁介质进行读写的硬盘驱动器 141,对可移动、非易失性磁盘 152 进行读写的磁盘驱动器 151,以及对可移动、非易失性光盘 156,如 CD ROM 或其它光介质进行读写的光盘驱动器 155。可以在示例性操作环境中使用的其它可移动 /

不可移动、易失性 / 非易失性计算机存储介质包括但不限于, 磁带盒、闪存卡、数字多功能盘、数字录像带、固态 RAM、固态 ROM 等等。硬盘驱动器 141 通常通过不可移动存储器接口, 如接口 140 连接到系统总线 121, 磁盘驱动器 151 和光盘驱动器 155 通常通过可移动存储器接口, 如接口 150 连接到系统总线 121。

[0024] 上文讨论并在图 1 示出的驱动器及其关联的计算机存储介质为计算机 110 提供了计算机可读指令、数据结构、程序模块和其它数据的存储。例如, 在图 1 中, 示出硬盘驱动器 141 储存操作系统 144、应用程序 145、其它程序模块 146 和程序数据 147。注意, 这些组件可以与操作系统 134、应用程序 135、其它程序模块 136 和程序数据 137 相同, 也可以与它们不同。这里对操作系统 144、应用程序 145、其它程序模块 146 和程序数据 147 给予不同的标号来说明至少它们是不同的副本。用户可以通过输入设备, 如键盘 162 和定位设备 161 (通常指鼠标、跟踪球或触摸垫) 向计算机 110 输入命令和信息。其它输入设备 (未示出) 可包括话筒、操纵杆、游戏手柄、圆盘式卫星天线、扫描仪等等。这些和其它输入设备通常通过耦合至系统总线的用户输入接口 160 连接至处理单元 120, 但是也可以通过其它接口和总线结构连接, 如并行端口、游戏端口或通用串行总线 (USB)。监视器 191 或其它类型的显示设备也通过接口, 如视频接口 190 连接至系统总线 121。除监视器之外, 计算机也可包括其它外围输出设备, 如扬声器 197 和打印机 196, 它们通过输出外围接口 195 连接。

[0025] 计算机 110 可以使用到一个或多个远程计算机, 如远程计算机 180 的逻辑连接在网络化环境中操作。远程计算机 180 可以是个人计算机、服务器、路由器、网络 PC、对等设备或其它常见的网络节点, 并通常包括许多或所有相对于计算机 110 所描述的元件, 尽管在图 1 中仅示出了存储器存储设备 181。图 1 描述的逻辑连接包括局域网 (LAN) 171 和广域网 (WAN) 173, 但也可包括其它网络。这类网络环境常见于办公室、企业范围计算机网络、内联网以及因特网。

[0026] 当在 LAN 网络环境中使用时, 计算机 110 通过网络接口或适配器 170 连接至 LAN 171。当在 WAN 网络环境中使用时, 计算机 110 通常包括调制解调器 172 或用于通过 WAN 173, 如因特网建立通信的其它装置。调制解调器 172 可以是内置或外置的, 它通过用户输入接口 160 或其它适当的机制连接至系统总线 121。在网络化环境中, 相对于计算机 110 所描述的模块或其部分可储存在远程存储器存储设备中。作为示例而非局限, 图 1 示出远程应用程序 185 驻留在存储器设备 181 上。可以理解, 示出的网络连接是示例性的, 也可以使用在计算机之间建立通信链路的其它手段。

[0027] 资源复制

[0028] 可以理解, 现代机器可以在相当短的一段时间内处理上千次资源改变。复制这些资源以及跨经由具有变化的可靠性和带宽的各种网络连接的上百或上千台机器保持它们同步提出了一个重大的挑战。

[0029] 乐观地, 多主设备复制系统允许在参与给定复制组的任何机器上对复制内容的不受限制的改变。复制组包括在参与该复制组的机器上复制的一组资源。复制组的该组资源可跨各卷。例如, 一复制组可包括与 C:\DATA、D:\APPS 和 E:\DOCS 相关联的资源, 这些资源可以在参与该复制组的一组机器上复制。潜在的冲突改变在复制系统的控制下使用一组冲突解决准则来调和, 该组冲突解决准则对每一冲突的情况定义了哪一冲突改变将优先于其它改变。

[0030] 一复制组成员或一复制组的成员包括通过复制其上包含的资源来参与该复制组的机器。内容集（有时简称为“内容”）包括驻留在复制组成员上的资源。

[0031] 术语“机器”不仅限于物理机器。相反，单个物理机器可以包括多个虚拟机器。如此处所使用的，从一个机器到另一机器的复制意味着同一复制组的一个或多个成员从虚拟或物理的一台机器到虚拟或物理的另一台机器的复制。单个物理机器可包括同一复制组的多个成员。由此，复制一复制组的成员可涉及将包括同一复制组的两个或多个成员的单个物理机器的成员同步。

[0032] 资源可被认为是对象。每一资源与资源数据和资源元数据相关联。资源数据可包括内容以及与该内容相关联的属性，而资源元数据包括在协商同步以及在冲突解决时可能相关的其它属性。资源数据和元数据可被储存在数据库或其它合适的存储中；在另一实施例中，可使用单独的存储来储存资源数据和元数据。

[0033] 在包括基于文件系统命名文件的数据存储的复制系统中，资源数据可包括文件内容，以及储存在该文件系统上的与文件内容相关联的任何文件属性。文件属性可包括访问控制列表（ACL）、创建/修改时间、以及与文件相关联的其它数据。如此处所使用的，文件可包括目录（即，可以与后代文件及目录以及祖先目录相关联的文件系统对象）以及非目录。除非上下文明确另外指明，否则术语文件可被理解为“资源数据”。

[0034] 在包括不基于文件系统命名文件的数据存储的复制系统（例如，其中资源被储存在数据库或基于对象的数据存储的复制系统）中，储存对数据存储适当的资源数据。贯穿全文，有时使用基于文件系统中的文件的复制系统来进行说明，但是可以认识到，可使用任何能够储存内容的数据存储，而不脱离此处所描述的主题的精神或范围。

[0035] 对每一资源，资源元数据可包括全局唯一标识符（GUID）、该资源是否已被删除、版本序列号加上改变的原作者、反映发生改变的时间的时钟值、以及其它字段，诸如概括资源数据的值并且可包括用于资源内容的签名的摘要。摘要例如可用于快速比较以便绕过复制同步期间的数据传输。如果目的地机器上的资源与源机器上的内容同步（例如，如由摘要所指示的），则可通过仅发送资源元数据而不发送资源数据本身来最小化网络开销。进行发送资源元数据，使得目的地机器可以在其后续的复制活动中反映包括在源机器上的元数据。这可允许目的地机器例如在后续的复制活动中成为源机器。资源元数据可以与资源数据一起或分开储存，而不脱离此处所描述的主题的精神或范围。

[0036] 在复制资源时可以使用版本向量。版本向量可被示为参与复制组的机器的一组全局计数器或时钟。参与复制组的每个机器维护表示该机器的当前最新版本和该机器接收到的关于其他机器的最新版本的版本向量。每次从机器创建、修改或删除资源时，该资源的版本就被设置成等于该机器的当前版本号加一的版本号。该机器的版本向量也被更新，以便反映该机器的版本号被递增了。

[0037] 版本向量维护有关知识状态的信息。版本向量可以将机器标识符映射到时间戳。参与复制组的每个机器可以维护其自身的时钟向量以便记录它所知道的其他机器已获得的时钟值。时钟向量可以包括版本序列号。在一个实施例中，可以维护一组时钟值，而非为每个机器维护单个时钟值，以便适应于处理无序更新并利用对版本向量的基本集合操作。

[0038] 在一个实施例中，在同步期间，可发送版本向量以便在同步资源时使用。例如，如果机器 A 和 B 参加诸如联合等同步活动，则机器 B 可将其版本向量发送给 A。在接收到 B 的

版本向量之后,A 然后可发送对具有未被 B 的版本向量包含(即,不被支配)的版本的资源(如果有)的改变。短语“一个版本向量支配另一版本向量”是对版本向量之间对应的包含关系的同义词。

[0039] 在题为“Interval Vector Based Knowledge Synchronization for ResourceVersioning(用于资源版本的基于内部向量的知识同步)”的美国专利申请序列号 10/791,041、题为“Garbage Collection of Tombstones for OptimisticReplication Systems(用于乐观复制系统的墓碑的无用信息收集)”的美国专利申请序列号第 10/779,030、以及题为“Granular Control Over the Authority of Replicated Information via Fencing and UnFencing(经由围栅或解围栅对复制信息的授权的粒度控制)”的美国专利申请序列号第 10/733,459 中描述了在同步中使用版本向量的示例。

[0040] 图 2 是根据此处所描述的主题的各方面一般地示出参与复制组的机器的框图。作为一个示例,上游机器 201 和下游机器 202 可以参与包括两个资源的复制组。这两个资源可以包括例如文档目录 205 和 215 和帮助目录 210 和 220(在两台机器上给予它们不同的号码以表示在一特定时刻,这些资源可能不包括相同的资源数据 – 即它们可以是非同步的)。

[0041] 周期性地,下游机器 202 可以向上游机器 201 请求更新,并且可基于该更新来更新其资源。虽然仅在图 2 中示出了两台机器,但是上游机器 201 和下游机器 202 可以是包括许多其他机器的复制系统的一部分。在一个交互中为源的机器(有时称为上游机器)可以稍后成为另一个交互中的目的地(有时称为下游机器),反之亦然。

[0042] 下游机器 202 可以包括确定何时从上游机器 201 获取更新的更新机制 225。如果下游机器 202 也是另一机器(未示出)的上游合作伙伴,那么更新机制 225 也可以确定要将哪些更新发送给其他机器。更新机制将在以下结合图 3 来更详细描述。

[0043] 图 3 是表示根据此处所描述的主题的各方面的被配置成在资源复制系统中运作的机器的框图。机器 305 包括更新机制 225、资源 322 和通信机制 340。更新机制包括出现时间逻辑 345、协议逻辑 320 和其他同步逻辑 320。

[0044] 资源 322 包括用于存储资源数据的资源数据存储 325 和资源元数据存储 330。虽然在相同的框中示出,但是资源数据存储 325 可以相对于资源元数据存储 330 一起存储或被存储分开的存储中。特别地,资源元数据存储 330 可以包括存储在资源存储 325 中的每个资源数据记录的版本,并且还可以包括内部向量(框 335)。

[0045] 资源元数据存储 330 还可包括与复制文件夹相关联的资源,其中每一记录指示(例如,经由时间戳)其各自的复制文件夹何时最后一次参与与另一成员上的相应复制文件夹的成功同步活动。如此处所使用的,复制文件夹与内容集同义。

[0046] 通信机制 340 允许更新机制 225 与其他机器上的其他更新机制(未示出)通信。通信机制 340 可以是网络接口或适配器 170、调制解调器 172 或者结合图 1 所描述的用于建立通信的任何其他装置。

[0047] 更新机制 225 包括用于与其他复制合作伙伴通信的协议逻辑 315,以及可用于确定一机器是否参与与其它机器的同步活动的出现时间逻辑 345,如以下更详细描述的。其他同步逻辑 320 包括除了协议逻辑之外的同步逻辑(例如在冲突更新的情况下做什么,如何确定要获取哪些更新等)。虽然将协议逻辑 315、出现时间逻辑 345 和其他同步逻辑 320 示为分开的框,但是在其他实施例中,可以全部或部分地将它们组合在一起。

[0048] 如此处所使用的,复制服务可以包括以上结合图 3 所描述的任何一个或多个组件。

[0049] 在操作中,机器 305 可以用作下游机器和 / 或上游机器。下游机器可以建立与上游和下游机器都参与的复制组的上游机器的连接。这可以通过例如通信机制 340 来完成。在建立连接的过程中,每个合作伙伴 (即上游和下游机器) 可以将其版本向量发送给其他合作伙伴。接着,建立会话以便将更新从上游机器发送到下游机器。

[0050] 可以将会话用于将上游机器的复制文件夹与下游机器的其相应复制文件夹绑定。可以为每一复制组的每个复制文件夹建立会话。可以在上游和下游机器之间的单个连接上建立多个文件夹的会话。

[0051] 在来自会话的所有更新被处理或丢弃之后,下游机器可关闭会话。

[0052] 下游机器可请求 (例如经由通信介质 340) 上游机器在与会话相关联的任何资源的更新发生时通知下游机器。当上游机器通知下游机器更新可用时,下游机器可以请求更新的版本向量。作为响应,上游机器发送其版本向量 (有时称为“VVup”)。注意 VVup 可以包括完整的版本向量或者包括自从发送上一个版本向量以来改变的版本向量。通知下游机器更新可用并且等待下游机器请求更新可以用两个步骤执行,这样使得下游机器不会被来自多个上游合作伙伴的版本向量意外地溢满。

[0053] 下游机器使用它接收到的上游版本向量 (即“VVup”),并计算与其自身的版本向量的设置差异,以便计算出下游机器不知道的驻留在上游机器上的版本。下游机器接着可以请求有关版本的元数据。在请求更新时,下游机器可以包括增量版本向量,该增量版本向量指示下游机器需要哪些更新。

[0054] 下游机器可以分开或一起请求墓碑 (tombstone) 或活更新。墓碑表示资源被删除,而活更新表示不删除资源的更新。在某些实现中,下游机器会在请求活更新之前请求墓碑。这可以用于提高效率,因为被修改接着被删除的资源无需在它在复制合作伙伴上被删除之前被修改。此外,在活更新之前处理墓碑可清除下游机器的数据存储 (例如在文件系统中) 的名称空间,以便为处理活替换更新作准备。

[0055] 在接收到更新之后,下游机器可以开始处理更新,以便确定从下游机器请求与更新相关联的哪些资源或其部分。可以例如通过其他同步逻辑 340 来执行该处理,该同步逻辑 340 可以被分成多个组件 (未示出)。例如,更新可以指示一特定资源的源数据或其部分被改变。在一个实施例中,可以由下游机器请求与资源相关联的所有资源数据。在另一实施例中,下游机器可以请求包括改变的资源数据的一部分。如此处所使用的,涉及资源数据的交互 (例如请求、响应、更新等) 应该被理解为意指涉及与资源相关联的资源数据的一部分或全部的交互。例如,对资源数据的请求可以意味着对于与资源相关联的资源数据的一部分或全部的请求。

[0056] 在确定需要被请求的资源数据之后,下游机器可以请求资源数据。作为对于资源数据的请求的响应,下游机器可以发送与资源相关联的资源数据。请求和响应可以继续直至下游机器确定需要更新的所有资源数据已经被请求。注意,并非所有的资源数据可以被发送,因为例如如果资源已经被删除,上游机器可能不再具有所请求的资源数据。其中可能不发送资源数据的另一示例是如果相对于下游机器的唯一有效的改变是资源被重命名或元数据属性被更新。在这种情况下,接收更新并重命名本地资源或更新本地元数据是将下

游资源与上游资源同步所需的全部。

[0057] 例如,如果复制文件夹被删除、如果在复制期间发生不可恢复的错误、或者如果复制系统被关闭,那么可以关闭会话。否则,所建立的会话可用于涉及上述事件的全部或一部分的后续同步动作。

[0058] 更新机制 225 可以使用数据库(例如资源元数据存储 330)来跟踪资源数据存储 325 中的资源数据的文件和目录状态。每次资源被更新时,该资源可以在数据库中被分配一个新的版本。

[0059] 更新机制 225 可以从上游合作伙伴接收更新并且更新本地资源,使得它们与上游合作伙伴上的复制资源同步。对资源元数据的更新可以被存储在资源元数据存储 330 中,而对资源数据的更新可以被存储在资源数据存储 325 中。在一个实施例中,资源数据存储 325 和资源元数据存储 330 可以被包括在相同的卷和物理设备(例如硬盘驱动器)上。

[0060] 也可以通过与资源数据相关的本地文件活动来改变资源数据。例如,复制合作伙伴可以用作许多客户机机器(未示出)的文件服务器。客户机机器可以访问和修改复制合作伙伴上的源数据(例如文件)。更新机制 225 可以包括可以监视对复制资源的改变(例如对文件的改变)的资源数据改变监视器(未示出)。当复制资源改变时,资源改变监视器可以更新资源元数据存储 330 以便指示更新。这可用于保持资源元数据存储 330 与资源数据 325 同步以及复制组成员之间的同步。

[0061] 内容新鲜度

[0062] 出现时间逻辑 345 可以在确定一复制组的成员的内容集的内容新鲜度时使用。当该成员与具有该复制组的成员的其它机器断开连接达一延长的时间段时,该成员的资源可变为非同步且陈旧。如果该成员重新加入到复制组并且没有使用此处所描述的出现时间逻辑 345,则该机器可开始进行与复制组的其它成员的复制。如果在该成员上没有更新,则该成员可从其它成员接收更新,并且可以不复制出任何陈旧的资源。然而,如果在该成员上有不是有帮助的新近更新,则这些更新可被复制出去并且盖写其它成员上的好的最新的版本。由此,在一种意义上,内容新鲜度可被示为改变相关度的一种度量。

[0063] 何处会发生这一情况的一个示例是当机器由于硬件问题而被关闭了达延长的一段时间(例如,几个月)的时候。在该机器被修复之后,该机器可回到在线并且主宿在该机器上的复制组的成员可以开始与其它机器上的合作伙伴的同步。在开始与合作伙伴同步时或刚好在这之前,某些应用程序(例如,反病毒应用程序)可遍历一个或多个陈旧文件并接触它们(例如,更新其最后修改时间)。如果没有此处所描述的方面,则这可导致该成员上的陈旧文件看似为最近被更新并且被复制出去到其它机器上的其它成员,从而盖写了其它机器上的更好的版本。

[0064] 在一个实施例中,复制陈旧内容(以上示例为其典型)可以通过使用出现时间来避免或减少。在一个实施例中,一机器可被指定为主设备,并且可在该机器上恒定地更新一出现时间值并将其传播到其它机器。在同步期间,如果两台机器之间的出现时间具有大于容许的差异,则从具有陈旧内容的机器(例如,具有较小出现时间的机器)的更新将不被复制回到最新的机器。

[0065] 在各方面中,出现时间值可以在机器、复制文件夹或个别资源上设置。如果出现时间值被设置在机器上,则该机器上的复制文件夹和资源可被示为具有相同的出现时间值。

如果出现时间值被设置在复制文件夹上,则与该文件夹相关联的所有复制资源可以被示为具有相同的出现时间值。该机器上的另一复制文件夹可具有不同的出现时间值。如果每一资源具有其自己的出现时间值,则每一复制文件夹中的不同资源可具有不同的出现时间值。

[0066] 使用出现时间值可涉及在同步期间发送出现时间值,使得陈旧度可被确定。因此,在逐个资源的基础上实现出现时间值可以施加比在复制文件夹或在逐个机器的基础上实现出现时间值更多的开销。出现时间值可以在一台机器上以一种方式(例如,按照复制文件夹)来设置,而在另一台机器上以另一种方式(例如,按照资源)来设置,而不脱离此处所描述的主题的精神或范围。

[0067] 在另一实施例中,成员可储存使用本地系统时钟生成的时间戳,而非使用出现时间值。该本地系统时钟可以从时间服务器周期性地更新以保持其最新。该时间戳也可以取决于各种同步活动而被更新。在一个实施例中,用于复制文件夹的时间戳可以在每次该复制文件夹成功地与另一成员的复制文件夹同步时被更新。在另一实施例中,该时间戳可以在每次建立了到复制组的另一成员的良好连接时被更新。

[0068] 在又一实施例中,该时间戳可以每天一次地或者取决于同步活动以某一其它所选频率来更新。如果在该周期期间发生了成功的同步活动,则可更新时间戳。成功的同步活动可以涉及将一复制文件夹与另一成员的复制文件夹成功同步、建立与领域成员的良好连接、维持与另一成员的良好连接(例如,如由“心跳”消息或以其它方式指示的)、从上游合作伙伴获得版本向量、向下游合作伙伴发送版本向量、或任何其它同步活动。

[0069] 采用时间戳,每一复制文件夹可在其元数据存储中具有相关联的记录(例如,CONTENT_SET_RECORD)。除了包括在该记录中的其它字段,该记录还可包括例如反映该记录的相应复制文件夹的新鲜度的时间戳。

[0070] 注意,情况可能是一台机器上主宿的一个复制文件夹是最新的,而该机器上主宿的另一复制文件夹是陈旧的。会发生这一情况是因为一个复制文件夹已经能够与其复制组中的其它成员同步,而另一复制文件夹尚不能够与其复制组中的其它成员同步。

[0071] 在主宿复制文件夹的成员试图连接到另一成员或接受来自另一成员的连接请求时,该成员可在参加任何其它复制活动之前检查该复制文件夹的时间戳是否是最新的(例如,在从当前时刻开始的可设置的天数之内)。如果其时间戳是最新的,则该成员可连接到另一成员或允许该连接请求,并继续同步活动。否则,该成员可以不尝试连接到另一成员,或可拒绝任何连接请求并制止启动与该其它成员的任何另外的复制活动。

[0072] 在另一实施例中,如果一成员的内容是陈旧的,则该成员可接受该连接请求而非拒绝连接请求,但是拒绝参加同步会话。在一个实施例中,可改为发生对参加同步会话的拒绝,并且该拒绝可以在此处关于拒绝连接请求所描述的任何时刻发生。

[0073] 在另一实施例中,具有陈旧内容的成员可能无法初始化,使得以后该成员将从不尝试连接到其它成员,也不能接受任何连接请求。

[0074] 拒绝接受连接请求的成员可返回一公知的出错码,使得其它成员知道该成员的内容是陈旧的。另外,该成员可将事件(例如,EVENT_DFSR_CS_STALE)记入日志。

[0075] 复制组中具有旧的时间戳的成员可被认为是具有陈旧内容。在复制成员的内容被认为陈旧之前本地系统时钟和时间戳之间可以容许的最大差异可由系统管理员、用户等

（以下有时称为“系统管理员”）来设置。该最大差异有时被称为“陈旧度值”。

[0076] 为避免对陈旧度的假指示，响应于成功同步活动更新时间戳的频率可以被选为在对应于该陈旧度值的一段时间期间发生至少一次。例如，如果该成员维持了与另一成员的良好连接，则该时间戳可在该成员维持良好连接期间每天更新。如果陈旧度值是 30 天，并且时间戳是每天更新的，则不可能出现对陈旧度的假指示。

[0077] 选择不保证对于在对应于陈旧度值的时间期间的更新的至少一个可能性的更新频率可能会导致对陈旧度值的假指示。例如，如果该时间戳仅每 30 天或更长时间更新一次，并且陈旧度值是 30，则可能会发生成功的同步活动（即，良好连接）而在时间戳中没有更新。

[0078] 可以认识到，在与陈旧度值相比非常频繁地更新时间戳与较不频繁地更新时间戳之间可以有一折衷（例如，开销）。在一个实施例中，更新时间戳的频率是每天一次（如果在该天中发生成功的同步活动），而陈旧度值被设为 30 天。在其它实施例中，可取决于例如实现和系统配置来选择其它频率和陈旧度值。

[0079] 另外，通过将陈旧度值设为等于或小于墓碑收集周期，则可避免或减轻各种问题。墓碑收集周期是一成员在删除指示一资源已被删除的元数据资源记录之前等待的周期。通常，成员等待删除记录，使得墓碑可被传播到复制组中的其它成员。如果一成员与该复制组的其它成员断开连接达足够的一段时间，则可能已经创建了该资源的墓碑，并且然后删除该元数据资源记录。没有此处所描述的机制，如果该资源存在于陈旧成员上（并且如果发生结合图 3 所描述的同步），则该资源可能从该陈旧成员复制出去并被重新引入到该复制组的其它成员。

[0080] 为避免或减少这一行为，陈旧度值可被设为等于或小于墓碑周期。较小的陈旧度值与墓碑周期相比，有较小的可能性使得在对一已删除资源的墓碑被传播到所有非陈旧成员之前也将被删除。只要一成员在墓碑周期中成功地与其它成员同步，则该成员将更有可能接收到墓碑并且标记该成员的相应资源用于删除。如果该成员在该周期内没有成功地与另一成员同步，则该成员上的内容可被标记为陈旧，并且可采取以下描述的其它机制来确保该陈旧内容不被重新引入到复制组的其它成员中。

[0081] 当一复制文件夹变为陈旧时，系统管理员可禁用它，然后重新启用它，使得当一入站连接可用时该复制文件夹与其它成员同步。重新启用复制文件夹可使得该复制文件夹经历一初始同步。在该初始同步期间，可使用三种同步策略中的至少一种来如下所述地同步复制文件夹。

[0082] 在一种策略中，可将复制合作伙伴（例如，具有陈旧内容的合作伙伴）之一设为只读（或从设备）。简言之，只读成员上的任何更新被丢弃。如果一文件是在该成员上创建的，则它被删除。如果一现有文件被删除，则它被重新激活（例如，取消删除）。如果一文件被更新，则该更新被丢弃，并且重新安装原始版本或来自另一合作伙伴的较新版本。一般而言，这使得来自非只读合作伙伴（例如，主设备）的所有更新被复制到只读合作伙伴。

[0083] 在第二种策略中，发生资源的合并。在合并时，对在陈旧合作伙伴上不同的所有资源，来自上游合作伙伴的资源复制到陈旧合作伙伴。如果陈旧合作伙伴不包括资源，则该资源被复制到陈旧合作伙伴。如果该陈旧合作伙伴包括上游合作伙伴不包括的资源，则该资源不在陈旧合作伙伴上删除，并且可在陈旧合作伙伴是另一成员的上游合作伙伴的情况下

从该陈旧合作伙伴中复制出来。

[0084] 在第三种策略中,可发生如以上结合图 3 所描述的同步。在该策略中,一般具有更新近修改日期的资源被复制出来,并且盖写其它成员上具有较不新近的修改日期的资源。

[0085] 有可能一复制组的所有成员都是陈旧的。在这一情况下,复制成员之一可被设置为主成员。如果一成员被指定为主成员,则该成员可将资源复制到其下游合作伙伴。如果一成员没有被指定为主成员,则该成员可以不将资源复制到其下游合作伙伴,直到它成功地将资源与上游合作伙伴同步。在一成员成功地与另一成员进行了复制之后,该成员可恢复如结合图 3 所描述的标准复制活动。

[0086] 可设置控制在同步期间哪些资源获胜(并且被复制出去)的栅栏(fence)值。栅栏值允许系统管理员指示一成员的被复制到复制组的其它成员的资源,而不管资源的其它特性如何。例如,已被分配了较高的栅栏值的资源可被复制出去,即使它比另一成员上的相应资源旧。

[0087] 在一个实施例中,当一复制组的所有成员都是陈旧的时候执行自动恢复。主成员可能已由系统管理员预先指定以在所有成员都变为陈旧的情况下成为主成员。或者,可使用成员的其它特性(例如,位置、分层次序(如果有)、最新时间戳等等)来确定哪一机器被指定为主机器。

[0088] 在另一实施例中,当一复制组的所有成员都是陈旧的时候可以执行手动干预。在这一干预中,可通知系统管理员所有成员都是陈旧的,并且一个成员需要被设为主成员以使同步继续。系统管理员然后可指示哪一成员是主成员。

[0089] 如果一复制组成员具有权威资源或仅有出站连接(即,该成员没有上游合作伙伴),则如果该成员与其它成员断开连接达延长的一段时间,它也可被认为是陈旧的。权威资源是被指定来盖写复制组的所有其它成员上的相应资源(除了被指定为权威的其它资源之外)的资源。在一个实现中,这可通过向每一权威资源分配一特殊栅栏值来完成。内容集的资源(例如,特定成员上的资源)可在例如对一成员执行备份恢复并且期望所有其它成员也都具有该相同的内容集的时候被指定为权威的。

[0090] 内容集的资源可在一成员从资源丢失中恢复时被指定为非权威的。当为非权威的时候,该成员的资源败给来自上游合作伙伴的资源。如果对于一特定资源,该资源数据与包含在上游合作伙伴上的资源数据相同,则该资源数据不需要从上游合作伙伴下载,即使该资源的资源元数据不同。内容集中的资源可在例如当一复制组的成员的资源已丢失或变得破坏时、该资源已从备份恢复时、以及期望该资源不被从该成员中复制出去时被指定为非权威的。

[0091] 有时候,一成员的资源元数据将变得被破坏或丢失或删除。当发生这一情况时,可删除储存资源元数据的数据库,并且从资源数据重建该数据库。在这一情况下,该成员的内容集的资源也可被指定为非权威的。

[0092] 当一资源及其在上游合作伙伴上的相应资源没有被指定为权威或非权威的时候,可使用普通的冲突解决来确定哪些资源获胜并被复制出去。

[0093] 在期望使一陈旧成员变为新鲜的情况下,可为系统管理员展示一迫使陈旧的复制文件夹变为最新的接口。在一个实施例中,该接口可更新时间戳以使得陈旧成员变为新鲜的。

[0094] 在其中监控系统的健康状态的系统中,健康状态模型可将陈旧状态下的复制文件夹认为是警告状态,并且可在报告中显示警告。如上所述,陈旧成员可能不需要系统管理员干预,因此在一个实施例中,陈旧成员可被注释为警告而非关键错误。

[0095] 图 4 是一般地表示根据此处所描述的主题的各方面的在确定本地机器上的内容是否足够新鲜来与复制组的其它成员同步时可能发生的动作的流程图。在框 405 处,该动作开始。

[0096] 在框 410 处,从合作伙伴(或者是上游或者是下游)接收连接请求。在框 415 处,接收到该请求的成员如上所述地确定其本地资源是否是陈旧的。在框 420 处,如果本地资源是陈旧的,则该动作在框 425 处继续;否则,该动作在框 430 处继续。在框 425 处,拒绝连接,并且可向请求者返回指示本地资源陈旧的消息。

[0097] 在框 430 处,接受连接请求,并且可执行其它同步活动(例如,结合图 3 所描述的那些活动)。在框 435 处,可更新时间戳以指示最近发生了同步。

[0098] 在框 440 处,该动作结束。

[0099] 尽管未示出,但是在一成员请求与另一成员的连接之前,该成员可确定其本地资源是否是陈旧的。如果是,则该成员可制止请求连接;否则,该成员可请求连接、执行其它同步活动、并更新其时间戳。

[0100] 图 5 是一般表示根据此处所描述的主题的各方面的在一成员被确定为具有陈旧内容并且被设定为一从设备以恢复内容时可能发生的动作的流程图。在框 505 处,该动作在初始同步阶段中开始。

[0101] 在框 510 处,删除本地包含在一成员上但是没有包含在该本地合作伙伴正在与其同步的上游合作伙伴上的任何资源。

[0102] 在框 515 处,添加没有本地地包含的但被包含在上游合作伙伴上的任何资源。

[0103] 在框 520 处,获得与上游合作伙伴不同的任何资源并相应地更新本地资源。

[0104] 在框 525 处,该动作结束。在该动作结束之后,本地合作伙伴可被认为是通过了初始同步阶段,并且之后可如结合图 4 所描述地同步资源。

[0105] 图 6 是一般地示出根据此处所描述的主题的各方面的在一成员被确定为具有陈旧内容并且被设置来将内容与上游合作伙伴合并时可能发生的动作的流程图。在框 605 处,该动作在初始同步阶段中开始。

[0106] 在框 610 处,获得与上游合作伙伴不同的任何资源并相应地更新本地资源。

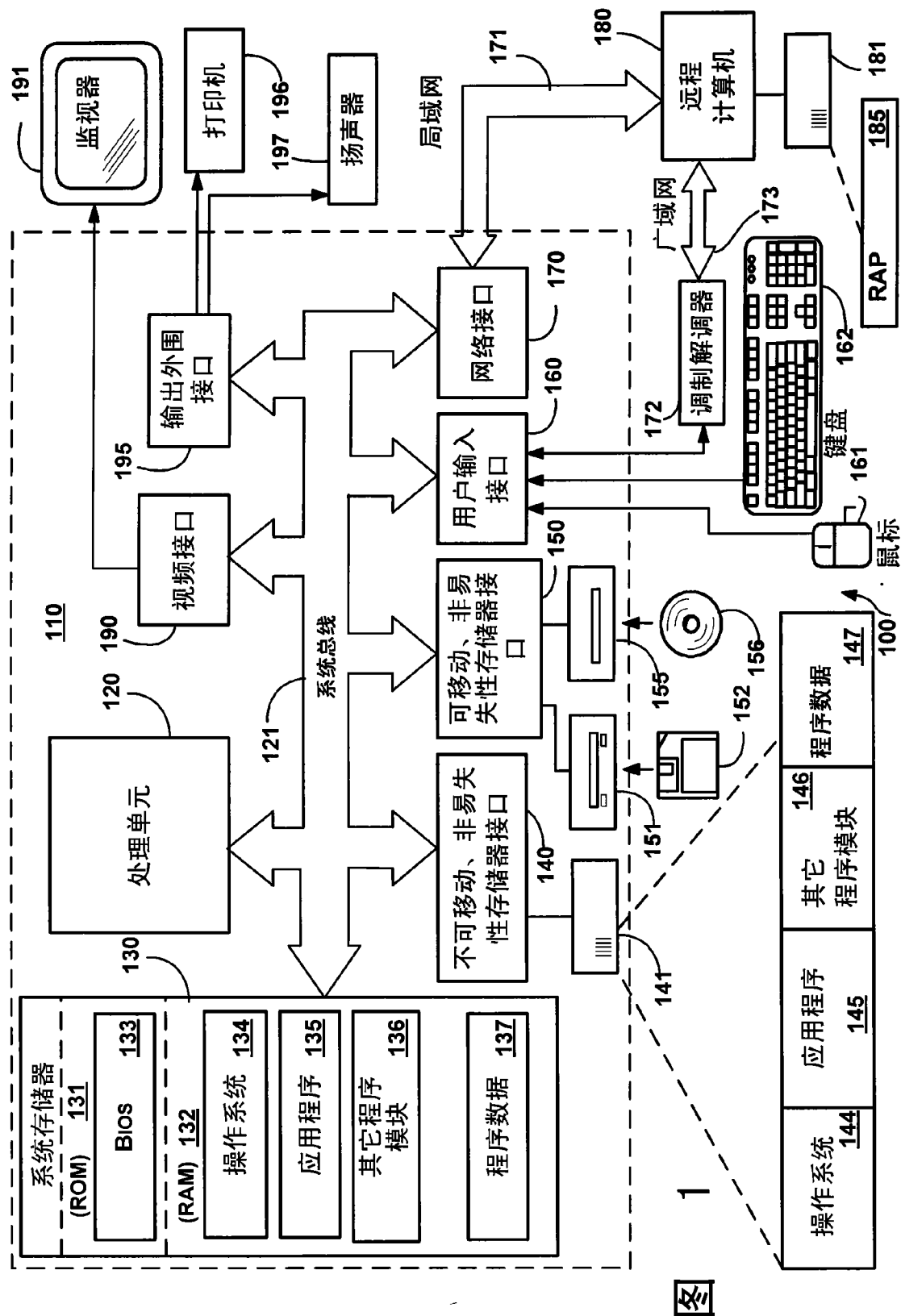
[0107] 在框 615 处,添加没有被本地包含的但被包含在上游合作伙伴上的任何资源。

[0108] 在框 620 处,该动作结束。在该动作结束之后,本地合作伙伴可被认为是通过了初始同步阶段,并且之后可如结合图 4 所描述地同步资源。这可包括如上所述同步在其它成员上被删除的对其墓碑也被删除的资源。

[0109] 应当理解,结合图 4-6 所描述的动作并没有全部包括在检查复制系统中的资源的新鲜度时可采取的所有动作。此外,尽管这些动作被描述为以特定的次序发生,但是在其它实施例中,某些动作可以并行发生或可以按其它次序执行而不脱离此处所描述的主题的精神或范围。

[0110] 如可以从以上详细描述中见到的,描述了与内容新鲜度以及资源复制系统有关的各方面。尽管此处所描述的主题的各方面容许各种修改和替换构造,但是其某些所示实施

例在附图中示出并且已在上文中详细描述。然而,应当理解,并没有任何意图将所要求保护的主题的各方面限于所公开的具体形式,而是相反,意在覆盖落入此处所描述的主题的各方面的精神和范围之内内的所有修改、替换构造和等效技术方案。



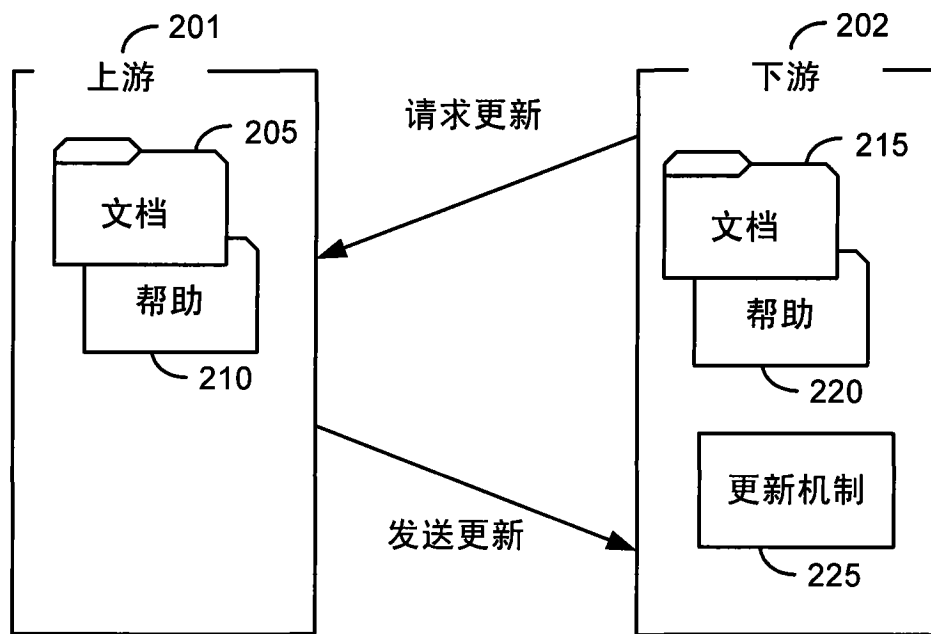


图 2

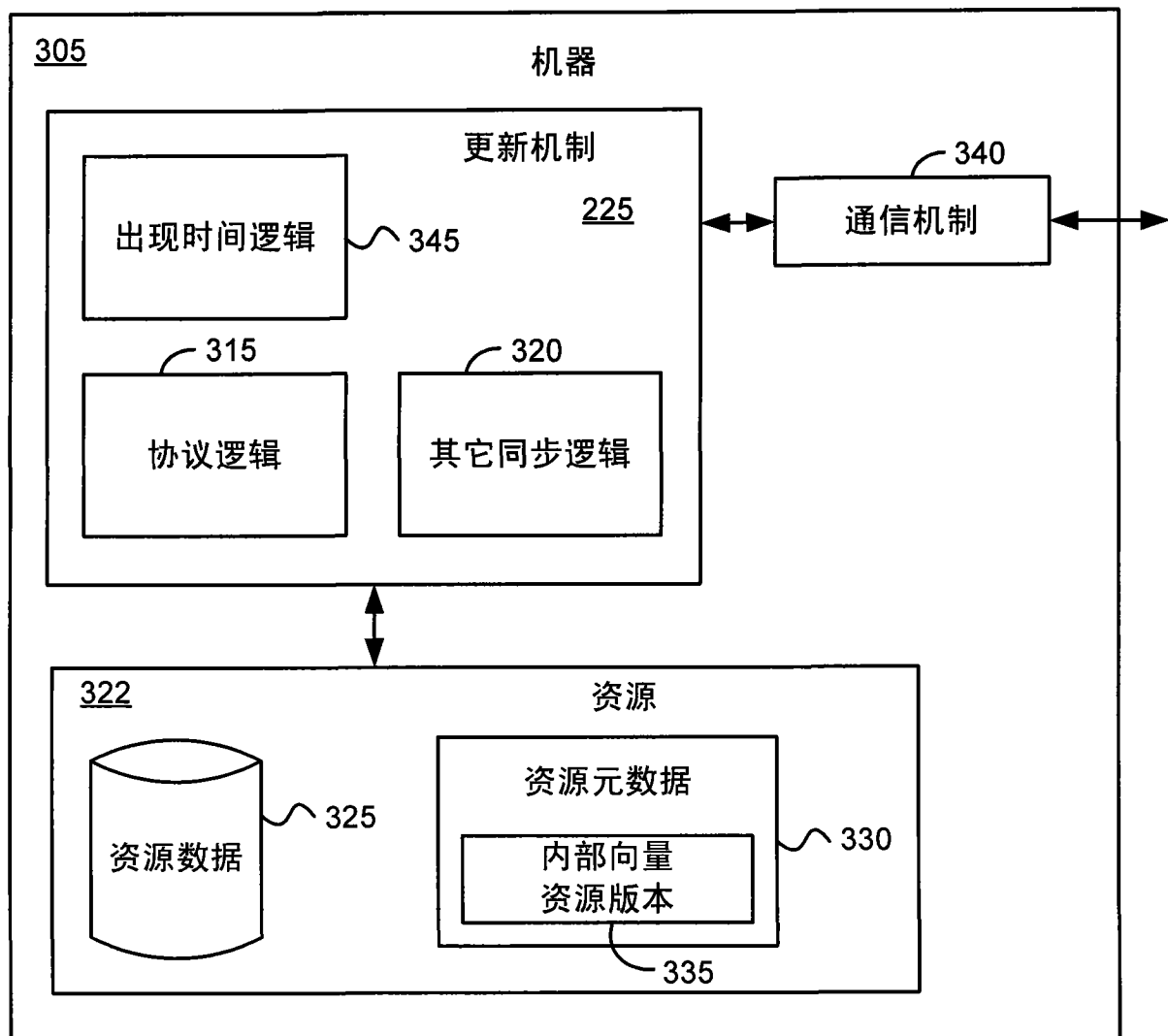


图 3

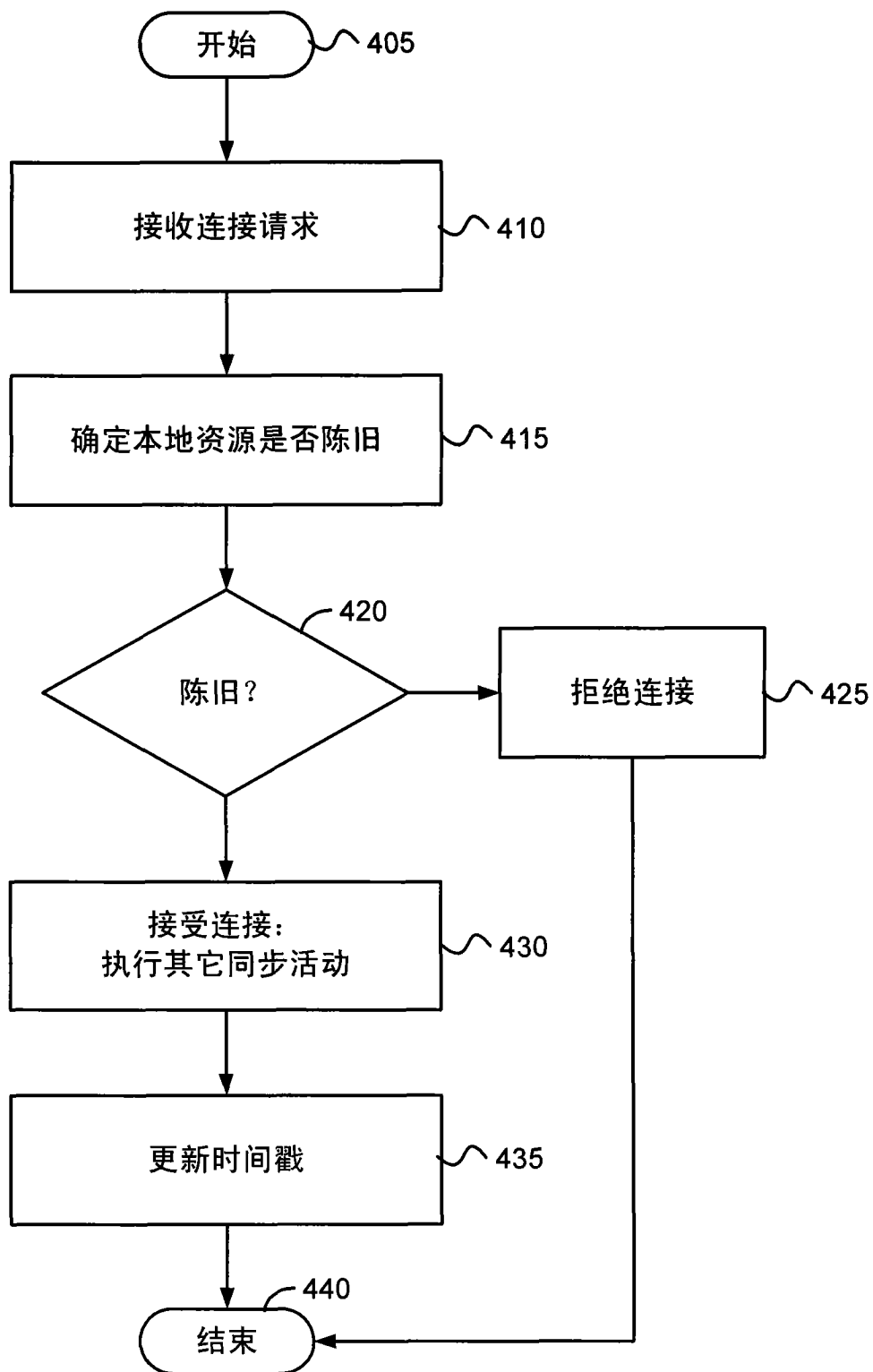


图 4

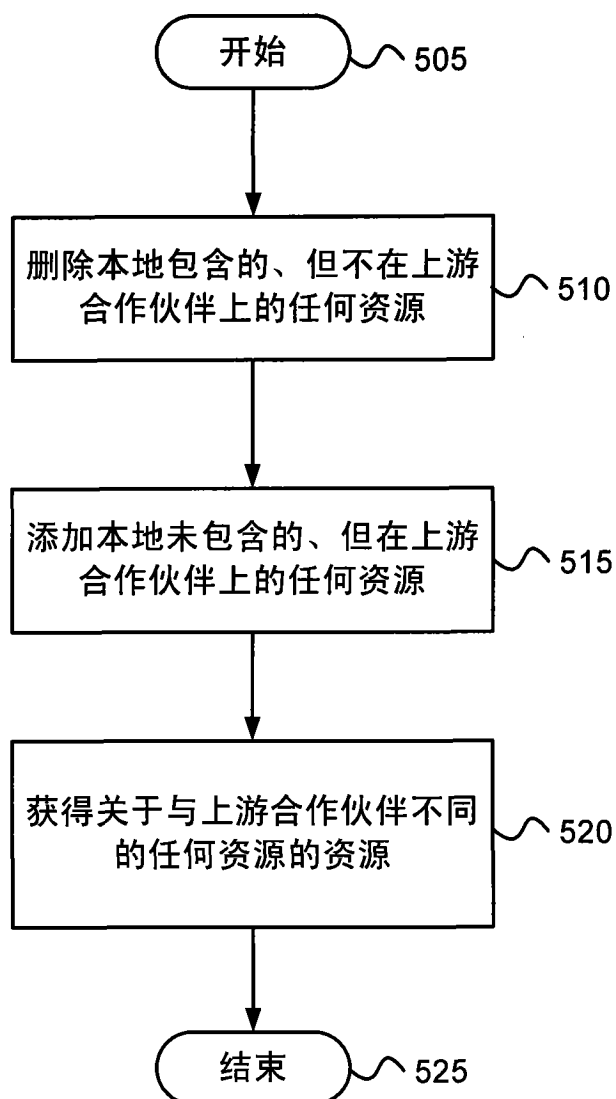


图 5

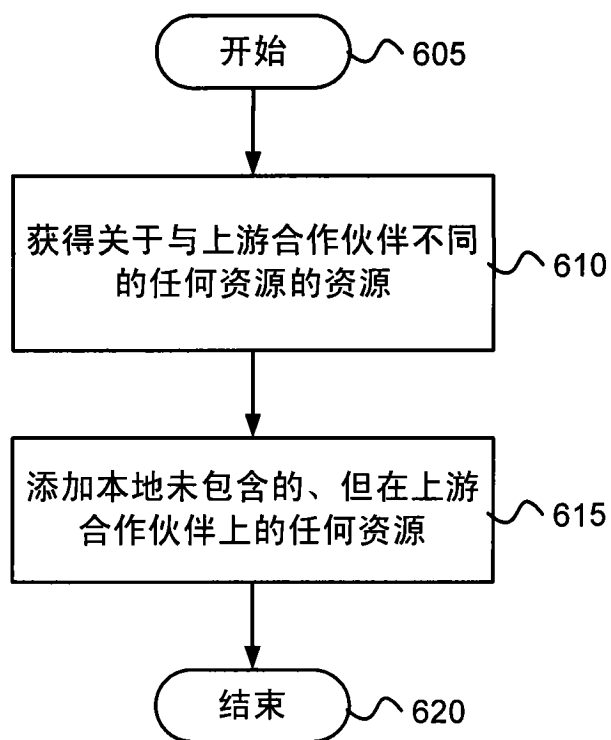


图 6