



(12) 发明专利

(10) 授权公告号 CN 103827723 B

(45) 授权公告日 2016. 03. 02

(21) 申请号 201280012565. 0

(22) 申请日 2012. 01. 10

(30) 优先权数据

61/431, 079 2011. 01. 10 US

(85) PCT国际申请进入国家阶段日

2013. 09. 10

(86) PCT国际申请的申请数据

PCT/IL2012/050007 2012. 01. 10

(87) PCT国际申请的公布数据

W02012/095848 EN 2012. 07. 19

(73) 专利权人 斯托尔万有限公司

地址 以色列赖阿南纳

(72) 发明人 R·戈登

(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 11038

代理人 郭思宇

(51) Int. Cl.

G02B 21/18(2006. 01)

(56) 对比文件

US 2003/0188085 A1, 2003. 10. 02,

US 2003/0145038 A1, 2003. 07. 31,

US 7434002 B1, 2008. 10. 07,

CN 101868790 A, 2010. 10. 20,

US 2008/0091843 A1, 2008. 04. 17,

审查员 闫西章

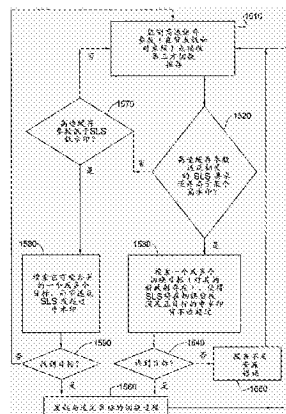
权利要求书4页 说明书48页 附图17页

(54) 发明名称

大规模存储系统

(57) 摘要

一种包括互连计算机节点的分布式存储系统;所述计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源;所述计算机节点中的至少一个包括一个或多个资源,所述一个或多个资源包括配置成高速缓存对象的至少一个高速缓存资源并具有相应的高速缓存相关参数;具有所述至少一个高速缓存资源的相应的计算机节点的至少一个UDSP代理配置成:监测连接到所述相应的计算机节点的所述至少一个高速缓存资源的高速缓存相关参数,用于确定所述高速缓存相关参数是否满足至少一个第一服务水平规范(SLS)标准;以及在所述至少一个第一SLS标准未被满足的情况下,发起所述至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到至少一个其它计算机节点的切换,所述至少一个其它计算机节点在接收到所述一个或多个高速缓存对象空间的所述至少部分之后,它的高速缓存相关参数满足至少一个第二SLS标准。



1. 一种包括互连计算机节点的分布式存储系统；

所述计算机节点中的每一个包括配置成执行统一分布式存储平台 (UDSP) 代理的至少一个处理资源；

所述计算机节点中的至少一个包括一个或多个资源,所述一个或多个资源包括配置成高速缓存对象的至少一个高速缓存资源并具有相应的高速缓存相关参数；

具有所述至少一个高速缓存资源的相应的计算机节点的至少一个 UDSP 代理配置成：

监测连接到所述相应的计算机节点的所述至少一个高速缓存资源的高速缓存相关参数,用于确定所述高速缓存相关参数是否满足至少一个第一服务水平规范 (SLS) 标准；以及

在所述至少一个第一 SLS 标准未被满足的情况下,发起所述至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到至少一个其它计算机节点的切换,所述至少一个其它计算机节点在接收到所述一个或多个高速缓存对象空间的所述至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准；

创建指示一个或多个高速缓存对象空间的所述至少部分的新位置的切换后高速缓存映射；

将所述切换后高速缓存映射发送到所述其它计算机节点；以及

将一个或多个高速缓存对象空间的所述至少部分的所有权传送到所述其它计算机节点。

2. 如权利要求 1 所述的分布式存储系统,其中具有所述至少一个高速缓存资源的所述相应的计算机节点的每个 UDSP 代理还配置成：

监测到所述高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准；以及

在满足所述至少一个第三 SLS 标准的情况下,发起所述至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到所述至少一个其它计算机节点的切换,所述至少一个其它计算机节点在接收到所述一个或多个高速缓存对象空间的所述至少部分之后,它的高速缓存相关参数满足所述至少一个第二 SLS 标准。

3. 如权利要求 1 所述的分布式存储系统,其中第一计算机节点具有相应的 UDSP 数据存储库,使得第一计算机节点的 UDSP 数据存储库包括关于相应于第二计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据；

所述第一计算机节点的所述 UDSP 代理配置成：

根据从关于至少一个第二计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测所述高速缓存相关参数,用于确定相应的第二计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准；以及

在至少一个第一 SLS 标准未被满足的情况下,向所述相应的第二计算机节点发起包括如从所述第一计算机节点的所述 UDSP 数据存储库提取的所述相应的第二计算机节点的所述高速缓存相关参数中的至少一个的消息,从而允许所述相应的第二计算机节点确定从所述第一计算机节点接收的所述高速缓存相关参数是否相对于所述第二计算机节点的所述高速缓存相关参数是优选的。

4. 如权利要求 3 所述的分布式存储系统,其中所述第一计算机节点具有连接到其的所

述至少一个高速缓存资源。

5. 如权利要求 4 所述的分布式存储系统, 其中所述第一计算机节点的所述 UDSP 数据存储库包括也关于相应于所述第一计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

6. 如权利要求 1 所述的分布式存储系统, 其中所述 UDSP 代理还配置成响应于发起所述切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐, 允许切换的成功发起。

7. 如权利要求 1 所述的分布式存储系统, 其中网关资源或客户端服务器具有相应的 UDSP 数据存储库, 使得网关资源或客户端服务器的 UDSP 数据存储库包括关于相应于计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;

所述网关资源或所述客户端服务器的所述 UDSP 代理还配置成:

根据从关于至少一个计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测所述高速缓存相关参数, 用于确定相应的计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准; 以及

在所述至少一个第一 SLS 标准未被满足的情况下, 向所述相应的计算机节点发起包括如从所述网关资源或所述客户端服务器的所述 UDSP 数据存储库提取的所述相应的计算机节点的所述高速缓存相关参数中的至少一个的消息, 从而允许所述相应的计算机节点确定从所述网关资源或所述客户端服务器接收的所述高速缓存相关参数是否相对于所述计算机节点的所述高速缓存相关参数是更优选的。

8. 如权利要求 1 所述的分布式存储系统, 其中所述相应的计算机节点的所述 UDSP 代理还配置成:

将所述切换后高速缓存映射发送到与一个或多个高速缓存对象空间的所述至少部分相关的一个或多个客户端服务器。

9. 如权利要求 1 所述的分布式存储系统, 其中所述相应的计算机节点的所述 UDSP 代理还配置成:

在所述切换期间接收与所述切换后高速缓存映射内的对象有关的请求;

如果根据所述切换后高速缓存映射所述对象不被所述相应的计算机节点拥有且所述相应的计算机节点没有所述对象的所有权, 则将所述请求转播到所述其它计算机节点。

10. 一种配置成连接到包括互连计算机节点的基础设施层的计算机节点, 所述计算机节点包括:

配置成高速缓存对象并具有相应的高速缓存相关参数的至少一个高速缓存资源;

配置成执行统一分布式存储平台 (UDSP) 代理的至少一个处理资源, 其中所述 UDSP 代理配置成:

监测连接到所述计算机节点的至少一个高速缓存资源的高速缓存相关参数, 用于确定所述高速缓存相关参数是否满足至少一个第一服务水平规范 (SLS) 标准; 以及

在所述至少一个第一 SLS 标准未被满足的情况下, 发起所述至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到所述互连计算机节点的至少一个其它计算机节点的切换, 所述至少一个其它计算机节点在接收到一个或多个高速缓存对象空间的所述至少部分之后, 它的高速缓存相关参数满足至少一个第二 SLS 标准;

创建指示一个或多个高速缓存对象空间的所述至少部分的新位置的切换后高速缓存映射；

将所述切换后高速缓存映射发送到所述其它计算机节点；以及

将一个或多个高速缓存对象空间的所述至少部分的所有权传送到所述其它计算机节点。

11. 如权利要求 10 所述的计算机节点, 其中所述 UDSP 代理还配置成：

监测到所述高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准；以及

在满足所述至少一个第三 SLS 标准的情况下, 发起所述至少一个高速缓存资源的一个或多个高速对象空间的至少部分到所述互连计算机节点的所述至少一个其它计算机节点的切换, 所述至少一个其它计算机节点在接收到一个或多个高速对象空间的所述至少部分之后, 它的高速缓存参数满足所述至少一个第二 SLS 标准。

12. 如权利要求 10 所述的计算机节点, 其中所述计算机节点还包括 UDSP 数据存储库, 所述 UDSP 数据存储库包括关于相应于所述互连计算机节点的至少一个其它计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据；且其中所述 UDSP 代理还配置成：

根据从关于至少一个其它计算机节点的所述一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测所述高速缓存相关参数, 用于确定相应的其它计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准；以及

在所述至少一个第一 SLS 标准未被满足的情况下, 向相应的其它计算机节点发起包括如所述 UDSP 数据存储库提取的所述相应的其它计算机节点的高速缓存相关参数中的至少一个的消息, 从而允许所述相应的其它计算机节点确定从所述计算机节点接收的所述高速缓存相关参数是否相对于所述其它计算机节点的高速缓存相关参数是优选的。

13. 如权利要求 12 所述的计算机节点, 其中所述 UDSP 数据存储库包括也关于相应于所述计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

14. 如权利要求 10 所述的计算机节点, 其中所述 UDSP 代理还配置成响应于发起所述切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐, 允许切换的成功发起。

15. 如权利要求 10 所述的计算机节点, 其中所述 UDSP 代理还配置成：

将所述切换后高速缓存映射发送到与一个或多个高速缓存对象空间的所述至少部分相关的一个或多个客户端服务器。

16. 如权利要求 10 所述的计算机节点, 其中所述 UDSP 代理还配置成：

在所述切换期间接收与所述切换后高速缓存映射内的对象有关的请求；

如果根据所述切换后高速缓存映射所述对象不被所述相应的计算机节点拥有且所述相应的计算机节点没有所述对象的所有权, 则将所述请求转播到所述其它计算机节点。

17. 一种操作计算机节点的方法, 所述计算机节点具有连接到其的至少一个高速缓存资源并配置成连接到包括互连计算机节点的基础设施层, 所述方法包括：

监测具有相应的高速缓存相关参数并配置成高速缓存对象的所述至少一个高速缓存资源的高速缓存相关参数, 用于确定所述高速缓存相关参数是否满足至少一个第一服务水平规范 (SLS) 标准；以及

在所述至少一个第一 SLS 标准未被满足的情况下,发起所述至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到所述互连计算机节点的至少一个其它计算机节点的切换,所述至少一个其它计算机节点在接收到一个或多个高速缓存对象空间的所述至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准;

创建指示一个或多个高速缓存对象空间的所述至少部分的新位置的切换后高速缓存映射;

将所述切换后高速缓存映射发送到所述其它计算机节点;以及

将一个或多个高速缓存对象空间的所述至少部分的所有权传送到所述其它计算机节点。

18. 如权利要求 17 所述的方法,还包括:

监测到所述高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准;以及

在满足所述至少一个第三 SLS 标准的情况下,发起所述至少一个高速缓存资源的一个或多个高速对象空间的至少部分到所述互连计算机节点的所述至少一个其它计算机节点的切换,所述至少一个其它计算机节点在接收到一个或多个高速对象空间的所述至少部分之后,它的高速缓存参数满足所述至少一个第二 SLS 标准。

19. 如权利要求 17 所述的方法,其中所述计算机节点还包括 UDSP 数据存储库,所述 UDSP 数据存储库包括关于相应于所述互连计算机节点的所述至少一个其它计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;所述方法还包括:

根据从关于至少一个其它计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测高速缓存相关参数,用于确定相应的其它计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准;以及

在所述至少一个第一 SLS 标准未被满足的情况下,向所述相应的其它计算机节点发起包括如从所述 UDSP 数据存储库提取的相应的其它计算机节点的高速缓存相关参数中的至少一个的消息,从而允许所述相应的其它计算机节点确定从所述计算机节点接收的高速缓存相关参数是否相对于所述其它计算机节点的高速缓存相关参数是优选的。

20. 如权利要求 19 所述的方法,其中所述 UDSP 数据存储库包括也关于相应于所述计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

21. 如权利要求 17 所述的方法,还包括响应于发起所述切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐,允许切换的成功发起。

22. 如权利要求 17 所述的方法,还包括:

将所述切换后高速缓存映射发送到与一个或多个高速缓存对象空间的所述至少部分或其部分相关的一个或多个客户端服务器。

23. 如权利要求 17 所述的方法,还包括:

在所述切换期间接收与所述切换后高速缓存映射内的对象有关的请求;

如果根据所述切换后高速缓存映射所述对象不被所述相应的计算机节点拥有且所述相应的计算机节点没有所述对象的所有权,则将所述请求转播到所述其它计算机节点。

大规模存储系统

[0001] 当前公开的主题的领域

[0002] 本发明涉及大规模存储系统,且特别是涉及用于实现这样的系统的装置和方法。

[0003] 背景

[0004] 分布式存储系统在最近十年快速地发展,因为网络在容量和速度方面增长。随着网络从局域网(LAN)扩展到全球广域网(WAN),商业变得更全球地分布,导致对分布式存储系统在远程地理位置上提供数据存储和访问的要求。因此在本领域中存在对用于通过通用网络分布数据存储的新方法和系统的需要。

[0005] 下面列出作为当前公开的主题的背景被认为是相关的现有技术参考文献。本文的参考文献的承认不应被推断为意味着这些参考文献以任何方式与当前公开的主题的专利性相关。

[0006] 美国专利公开号 2009/0070337 “Apparatus and Method for a Distributed Storage Global Database”涉及“用于管理数据元素的分布的地理上分布的存储系统,其中对给定数据元素的请求引起地理惯性。地理分布存储系统包括地理分布站点,每个站点包括用于在本地存储包括数据元素的全局连贯地分布的数据库的一部分的站点存储单元和用于接收与所述数据元素中的数据元素有关的请求的本地接入点。地理分布存储系统包括用于将至少一个所请求的数据元素转发到在地理分布站点(请求从地理分布站点接收)中的第一个处的本地接入点并在第一站点处存储至少一个所请求的数据元素的数据管理模块,从而为来自第一站点的未来请求提供对数据元素的本地可访问性,同时维持分布式数据库的全局连贯性”。

[0007] 美国专利号 5987505 “Remote Access and Geographically Distributed Computers in a Globally Addressable Storage Environment”涉及“计算机系统使用允许多个联网计算机通过寻址来访问数据的全局可寻址存储环境,即使数据存储持久性存储设备例如计算机硬盘和其它传统上非寻址数据存储设备上。计算机可位于单个计算机网络上或多个互连的计算机网络例如由广域网(WAN)耦合的两个局域网(LAN)上。全局可寻址存储环境允许数据由多个网络上的各种计算机访问和共享并在多个网络上的各种计算机当中被访问和共享”。

[0008] Ms. S. V. Patil 等人的“Unified Virtual Storage: Virtualization of Distributed Storage in a Network”(International Journal of Computer Applications 2010(0975 - 8887), Volume 1 - No. 22)描述了“有效地利用通过网络连接的桌上型机器上的自由磁盘空间的方法。在今天的很多网络中,仅偶尔使用客户端节点的本地磁盘。这是有效地管理网络中的数据存储并提供对共享 LAN 中的桌上型机器上的磁盘空间的软件支持的尝试。在当前情况中,在常规服务器上的存储扩展具有约束,如最大扩展限制、昂贵的事物,且在硬件更换、升级的情况下,数据的手动重新定位变得麻烦。UVS (统一虚拟存储)是有效地利用通过网络连接的桌上型机器上的自由磁盘空间的尝试。它的目的是减小网络服务器上的数据业务的负荷,有效地利用客户端节点上的空间,从而避免空间的浪费。它还消除了对存储器扩展的硬件限制,并提供数据存储库的位置透明性。UVS 的

主要优点是它可无缝地集成到现有基础设施(局域网系统)中。虚拟存储器实质上是无限支持可升级结构。客户端节点可使用统一虚拟驱动器作为对不同的服务器中的分布式存储的单点访问,从而消除了服务器的单独处理。在由网络连接的 UVS 服务器上实现的原型的性能比 n 个集中式系统更好,以及框架的开销即使在高负荷期间也是适度的”。

[0009] 美国专利公布号 2011/0153770 “Dynamic Structural Management of a Distributed Caching Infrastructure”涉及“用于 n 层分布式高速缓存基础设施的动态结构管理的方法、系统和计算机程序。在本发明的实施方案中, n 层分布式高速缓存基础设施的动态结构管理的方法包括建立到布置在 n 层高速缓存器中的相应层节点中的多个高速缓存服务器的通信连接,收集对 n 层高速缓存器中的相应层节点中的每个高速缓存服务器的性能度量,识别在越过阈值的 n 层的层节点中的相应层节点中的特定高速缓存资源的特征,并动态地构造包括特定的高速缓冲资源的一组高速缓存资源以解释所识别的特征”。

[0010] 概要

[0011] 根据当前公开的主题的方面,提供了配置成连接到包括互连计算机节点的基础设施层的计算机节点,互连计算机节点中的至少一个包括一个或多个存储相关资源,计算机节点包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更适当的计算机节点。

[0012] 根据当前公开的主题的某些例子,还提供了计算机节点,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个添加的互连计算机节点的计算。

[0013] 根据当前公开的主题的某些例子,还提供了计算机节点,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个互连计算机节点的计算。

[0014] 根据当前公开的主题的某些例子,还提供了计算机节点,其中更新的基础设施被动态地创建。

[0015] 根据当前公开的主题的某些例子,还提供了计算机节点,其中仅对未决的任务计算分数。

[0016] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0017] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一个或多个存储相关资源包括下列存储相关资源类别中的每个的至少一个存储相关资源:高速缓存资源、数

据存储资源和网络资源。

[0018] 根据当前公开的主题的某些例子,还提供了计算机节点,其中任务分数由优化引擎计算。

[0019] 根据当前公开的主题的某些例子,还提供了计算机节点,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0020] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一种或多种优化技术使用启发法或近似法。

[0021] 根据当前公开的主题的某些例子,还提供了计算机节点,其中更适当的计算机节点是基于所计算的分数的最适当的计算机节点。

[0022] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成基于所计算的任务分数计算总分数,且其中所述任务中的一个或多个的执行或将工作路由到更适当的计算机节点基于所计算的总分数。

[0023] 根据当前公开的主题的某些例子,还提供了计算机节点,其中工作从客户端服务器接收。

[0024] 根据当前公开的主题的某些例子,还提供了计算机节点,其中工作通过网关资源从客户端服务器接收。

[0025] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成:监测指示计算机节点的当前状态或连接到计算机节点的一个或多个存储相关资源的当前状态的至少一个参数,如果有的话;将指示对至少一个所监测的参数的变化的通知传播到互连计算机节点中的至少一个,

[0026] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成计算关于更新的基础设施的至少一个修改的互连计算机节点的下面的工作的任务的分数;更新的基础设施层包括至少一个修改的互连计算机节点;所述修改包括下列项中的至少一个;(i)至少一个新存储相关资源连接到相应的节点,(ii)至少一个现有的存储相关资源与相应的节点断开,(iii)至少一个现有的存储相关资源被修改;计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0027] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和存储相关资源参数数据计算分布式存储系统的配置;以及根据所计算的配置自动分配存储相关资源之一的至少部分。

[0028] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成:接收与分布式存储系统的动态行为有关的动态行为参数数据;当至少一个 SLS 被违反时,至少基于至少一个 SLS、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0029] 根据当前公开的主题的某些例子,还提供了操作配置成连接到包括互连计算机节点的基础设施层的计算机节点的方法,互连计算机节点中的至少一个包括一个或多个存储

相关资源,该方法包括:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更适当的计算机节点。

[0030] 根据当前公开的主题的某些例子,还提供了一种方法,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个添加的互连计算机节点的计算。

[0031] 根据当前公开的主题的某些例子,还提供了一种方法,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个互连计算机节点的计算。

[0032] 根据当前公开的主题的某些例子,还提供了一种方法,其中更新的基础设施被动态地创建。

[0033] 根据当前公开的主题的某些例子,还提供了一种方法,其中仅对未决的任务计算分数。

[0034] 根据当前公开的主题的某些例子,还提供了一种方法,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0035] 根据当前公开的主题的某些例子,还提供了一种方法,其中一个或多个存储相关资源包括下列存储相关资源类别中的每个的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0036] 根据当前公开的主题的某些例子,还提供了一种方法,其中计算分数由优化引擎执行。

[0037] 根据当前公开的主题的某些例子,还提供了一种方法,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0038] 根据当前公开的主题的某些例子,还提供了一种方法,其中一种或多种优化技术使用启发法或近似法。

[0039] 根据当前公开的主题的某些例子,还提供了一种方法,其中更适当的计算机节点是基于所计算的分数的最适当的计算机节点。

[0040] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:基于所计算的任务分数计算总分数,且其中执行所述任务中的一个或多个或将工作路由到更适当的计算机节点基于所计算的总分数。

[0041] 根据当前公开的主题的某些例子,还提供了一种方法,其中工作从客户端服务器接收。

[0042] 根据当前公开的主题的某些例子,还提供了一种方法,其中工作通过网关资源从

客户端服务器接收。

[0043] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:监测指示计算机节点的当前状态或连接到计算机节点的一个或多个存储相关资源的当前状态的至少一个参数,如果有的话;将指示对至少一个所监测的参数的变化的通知传播到互连计算机节点中的至少一个。

[0044] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:计算关于更新的基础设施的至少一个修改的互连计算机节点的下面的工作的任务的分数;更新的基础设施层包括至少一个修改的互连计算机节点;所述修改包括下列项中的至少一个;(i)至少一个新存储相关资源连接到相应的节点,(ii)至少一个现有的存储相关资源与相应的节点断开,(iii)至少一个现有的存储相关资源被修改;计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0045] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个SLS和存储相关资源参数数据计算分布式存储系统的配置;以及根据所计算的配置自动分配存储相关资源之一的至少部分。

[0046] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:接收与分布式存储系统的动态行为有关的动态行为参数数据;当至少一个SLS被违反时,至少基于至少一个SLS、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0047] 根据当前公开的主题的方面,还提供了分布式存储系统,其包括:包括互连计算机节点的基础设施层,其中互连计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源;互连计算机节点中的至少一个包括一个或多个存储相关资源;UDSP代理配置成接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到基础设施层的更适当的计算机节点;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0048] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个添加的互连计算机节点的计算。

[0049] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的

至少一个互连计算机节点的计算。

[0050] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中更新的基础设施被动态地创建。

[0051] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中仅对未决的任务计算分数。

[0052] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0053] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中一个或多个存储相关资源包括下列存储相关资源类别中的每个的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0054] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中任务分数由优化引擎计算。

[0055] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0056] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中一种或多种优化技术使用启发法或近似法。

[0057] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中更适当的计算机节点是基于所计算的分数最适当的计算机节点。

[0058] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成基于所计算的任务分数计算总分数,且其中所述任务中的一个或多个的执行或将工作路由到更适当的计算机节点基于所计算的总分数。

[0059] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中工作从客户端服务器接收。

[0060] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中工作通过网关资源从客户端服务器接收。

[0061] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成:监测指示计算机节点的当前状态或连接到计算机节点的一个或多个存储相关资源的当前状态的至少一个参数,如果有的话;将指示对至少一个所监测的参数的变化的通知传播到互连计算机节点中的至少一个。

[0062] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成:计算关于更新的基础设施的至少一个修改的互连计算机节点的下面的工作的任务的分数;更新的基础设施层包括至少一个修改的互连计算机节点;所述修改包括下列项中的至少一个;(i)至少一个新存储相关资源连接到相应的节点,(ii)至少一个现有的存储相关资源与相应的节点断开,(iii)至少一个现有的存储相关资源被修改;计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0063] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配

置成：接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)；至少基于至少一个 SLS 和存储相关资源参数数据计算分布式存储系统的配置；以及根据所计算的配置自动分配存储相关资源之一的至少部分。

[0064] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中 UDSP 代理还配置成：接收与分布式存储系统的动态行为有关的动态行为参数数据；当至少一个 SLS 被违反时，至少基于至少一个 SLS、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置；以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0065] 根据当前公开的主题的方面，还提供了分布式存储系统，其包括：包括互连计算机节点的基础设施层，其中：互连计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源；互连计算机节点中的至少一个包括一个或多个存储相关资源；UDSP 代理配置成：接收包括至少一个任务的工作；计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数，使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性，同时满足至少一个服务水平规范(SLS)要求，该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算；基于所计算的分数执行任务中的一个或多个或将工作路由到基础设施层的更适当的计算机节点；计算关于更新的基础设施层的至少一个修改的互连计算机节点的下面的工作的任务的分数；更新的基础设施层响应于至少一个互连计算机节点的修改而创建；所述修改包括下列项中的至少一个；(i)至少一个新存储相关资源连接到相应的节点，(ii)至少一个现有的存储相关资源与相应的节点断开，(iii)至少一个现有的存储相关资源被修改；计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数，产生更新的基础设施层；以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0066] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中 UDSP 代理还配置成：接收与分布式存储系统的动态行为有关的动态行为参数数据；当至少一个 SLS 被违反时，至少基于至少一个 SLS 要求、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置；以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0067] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中当计算下面的工作的任务的分数时，创建更新的基础设施层，且执行关于所创建的更新的基础设施层的至少一个修改的互连计算机节点的计算。

[0068] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中当计算下面的工作的任务的分数时，创建更新的基础设施层，且执行关于所创建的更新的基础设施层的至少一个互连计算机节点的计算。

[0069] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中更新的基础设施被动态地创建。

[0070] 根据当前公开的主题的某些例子，还提供了分布式存储系统，其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源：高速缓存资源、数据存储资源和网络资源。

[0071] 根据当前公开的主题的方面，还提供了配置成连接到包括互连计算机节点的基础

设施层的计算机节点,互连计算机节点中的至少一个包括一个或多个存储相关资源,计算机节点包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到基础设施层的更适当的计算机节点;计算关于更新的基础设施层的至少一个修改的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于至少一个互连计算机节点的修改而创建;所述修改包括下列项中的至少一个;(i)至少一个新存储相关资源连接到相应的节点,(ii)至少一个现有的存储相关资源与相应的节点断开,(iii)至少一个现有的存储相关资源被修改;计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数,产生更新的基础设施层;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0072] 根据当前公开的主题的某些例子,还提供了计算机节点系统,其中 UDSP 代理还配置成:接收与存储系统的动态行为有关的动态行为参数数据;当至少一个 SLS 要求被违反时,至少基于至少一个 SLS 要求、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0073] 根据当前公开的主题的某些例子,还提供了计算机节点系统,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个修改的互连计算机节点的计算。

[0074] 根据当前公开的主题的某些例子,还提供了计算机节点系统,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个互连计算机节点的计算。

[0075] 根据当前公开的主题的某些例子,还提供了计算机节点系统,其中更新的基础设施被动态地创建。

[0076] 根据当前公开的主题的某些例子,还提供了计算机节点系统,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0077] 根据当前公开的主题的某些例子,还提供了操作配置成连接到包括互连计算机节点的基础设施层的计算机节点的方法,互连计算机节点中的至少一个包括一个或多个存储相关资源,该方法包括:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到基础设施层的更适当的计算机节点;计算关于更新的基础设施层的至少一个修改的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于至少一个互连计算机节

点的修改而创建 ;所述修改包括下列项中的至少一个 ;(i) 至少一个新存储相关资源连接到相应的节点,(ii) 至少一个现有的存储相关资源与相应的节点断开,(iii) 至少一个现有的存储相关资源被修改 ;计算关于具有连接到其的新存储相关资源或具有从其断开的存储相关资源的至少一个互连计算机节点的下面的工作的任务的分数,产生更新的基础设施层 ;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0078] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括 :接收与存储系统的动态行为有关的动态行为参数数据 ;当至少一个 SLS 要求被违反时,至少基于至少一个 SLS 要求、存储相关资源参数数据和动态行为参数数据计算对分布式存储系统的重新配置 ;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0079] 根据当前公开的主题的某些例子,还提供了一种方法,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个修改的互连计算机节点的计算。

[0080] 根据当前公开的主题的某些例子,还提供了一种方法,其中当计算下面的工作的任务的分数时,创建更新的基础设施层,且执行关于所创建的更新的基础设施层的至少一个互连计算机节点的计算。

[0081] 根据当前公开的主题的某些例子,还提供了一种方法,其中更新的基础设施被动态地创建。

[0082] 根据当前公开的主题的某些例子,还提供了一种方法,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源 :高速缓存资源、数据存储资源和网络资源。

[0083] 根据当前公开的主题的方面,还提供了配置成连接到包括分布式存储系统的基础设施层的计算机节点,基础设施层包括互连计算机节点,计算机节点包括 :配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成 :接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据 ;至少基于至少一个 SLS 和所述存储相关资源参数数据计算分布式存储系统的配置 ;根据所计算的配置自动分配存储相关资源之一的至少部分 ;接收与分布式存储系统的动态行为有关的动态行为参数数据 ;当至少一个 SLS 要求被违反时,至少基于至少一个 SLS、存储相关资源参数数据和动态行为参数数据计算对分布式存储系统的重新配置 ;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0084] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成 :接收包括至少一个任务的工作 ;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足 SLS 的至少一个用户定义的存储相关要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算 ;基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点 ;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数 ;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建 ;

以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0085] 根据当前公开的主题的某些例子,还提供了计算机节点,其中配置由优化引擎计算。

[0086] 根据当前公开的主题的某些例子,还提供了计算机节点,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0087] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一种或多种优化技术使用启发法或近似法。

[0088] 根据当前公开的主题的某些例子,还提供了计算机节点,其中存储要求包括下列项中的至少一个:位置、本地保护水平、备份保留政策、远程保护水平、性能水平、加密水平、重复删除、压缩、存储方法。

[0089] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0090] 根据当前公开的主题的某些例子,还提供了计算机节点,其中存储相关资源分布在至少两个位置当中。

[0091] 根据当前公开的主题的某些例子,还提供了计算机节点,其中配置包括过度调拨存储系统资源供使用。

[0092] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成响应于故障来确定配置,以给用户提供故障通知或指示将允许配置的成功计算的至少一个额外的存储相关资源的添加的推荐。

[0093] 根据当前公开的主题的方面,还提供了操作配置成连接到包括互连计算机节点的基础设施层的计算机节点的方法,互连计算机节点中的至少一个包括一个或多个存储相关资源,该方法包括:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和存储相关资源参数数据计算分布式存储系统的配置;根据所计算的配置自动分配存储相关资源之一的至少部分;接收与分布式存储系统的动态行为有关的动态行为参数数据;当至少一个 SLS 被违反时,至少基于至少一个 SLS、存储相关资源参数数据和动态行为参数数据计算对存储系统的重新配置;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0094] 根据当前公开的主题的某些例子,还提供了一种方法,其包括:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;以及基于所计算的分数执行下面的工作的任务

中的一个或多个或将下面的工作路由到更新的基础设施层的更适当的计算机节点。

[0095] 根据当前公开的主题的某些例子,还提供了一种方法,其中计算配置由优化引擎执行。

[0096] 根据当前公开的主题的某些例子,还提供了一种方法,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0097] 根据当前公开的主题的某些例子,还提供了计算机一种方法,其中一种或多种优化技术使用启发法或近似法。

[0098] 根据当前公开的主题的某些例子,还提供了一种方法,其中存储要求包括下列项中的至少一个:位置、本地保护水平、备份保留政策、远程保护水平、性能水平、加密水平、重复删除、压缩、存储方法。

[0099] 根据当前公开的主题的某些例子,还提供了一种方法,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0100] 根据当前公开的主题的某些例子,还提供了一种方法,其中存储相关资源分布在至少两个位置当中。

[0101] 根据当前公开的主题的某些例子,还提供了一种方法,其中配置包括过度调拨存储系统资源供使用。

[0102] 根据当前公开的主题的某些例子,还提供了一种方法,还包括:响应于故障来确定配置,以给用户故障通知或指示将允许配置的成功计算的至少一个额外的存储相关资源的添加的推荐。

[0103] 根据当前公开的主题的方面,还提供了分布式存储系统,其包括:包括互连计算机节点的基础设施层,其中,计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和所述存储相关资源参数数据计算分布式存储系统的配置;根据所计算的配置自动分配存储相关资源之一的至少部分;接收与分布式存储系统的动态行为有关的动态行为参数数据;当至少一个 SLS 被违反时,至少基于至少一个 SLS、存储相关资源参数数据和动态行为参数数据计算对分布式存储系统的重新配置;以及根据所计算的重新配置自动分配存储相关资源之一的至少部分。

[0104] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足 SLS 的至少一个用户定义的存储相关要求,该分数也基于与连接到相应的计算机节点的一个或多个存储相关资源有关的参数数据——如果有的话——来计算;基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点;计算关于更新的基础设施层的至少一个添加的互连计算机节点的下面的工作的任务的分数;更新的基础设施层响应于向其添加至少一个额外的互连计算机节点而创建;以及基于所计算的分数执行下面的工作的任务中的一个或多个或将下面的工作路由到

更新的基础设施层的更适当的计算机节点。

[0105] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中配置由优化引擎计算。

[0106] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中优化引擎使用下列优化技术中的一个或多个:线性编程;模拟退火;遗传算法。

[0107] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中一种或多种优化技术使用启发法或近似法。

[0108] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中存储要求包括下列项中的至少一个:位置、本地保护水平、备份保留政策、远程保护水平、性能水平、加密水平、重复删除、压缩、存储方法。

[0109] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中一个或多个存储相关资源包括下列存储相关资源类别的至少一个存储相关资源:高速缓存资源、数据存储资源和网络资源。

[0110] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中存储相关资源分布在至少两个位置当中。

[0111] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中配置包括过度调拨存储系统资源供使用。

[0112] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成响应于故障来确定配置,以给用户提供故障通知或指示将允许配置的成功计算的至少一个额外的存储相关资源的添加的推荐。

[0113] 根据当前公开的主题的方面,还提供了分布式存储系统,其包括:包括互连计算机节点的基础设施层,其中,互连计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和所述存储相关资源参数数据计算分布式存储系统的配置;以及根据所计算的配置自动分配存储相关资源之一的至少部分。

[0114] 根据当前公开的主题的方面,还提供了配置成连接到分布式存储系统的基础设施层的计算机节点,基础设施层包括互连计算机节点,计算机节点包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和所述存储相关资源参数数据计算分布式存储系统的配置;以及根据所计算的配置自动分配存储相关资源之一的至少部分。

[0115] 根据当前公开的主题的方面,还提供了操作配置成连接到包括互连计算机节点的基础设施层的计算机节点的方法,基础设施层包括互连计算机节点,该方法包括:接收包括涉及至少一个逻辑存储实体的用户定义的存储要求的至少一个服务水平规范(SLS)和与连接到互连计算机节点的一个或多个存储相关资源有关的存储相关资源参数数据;至少基于至少一个 SLS 和所述存储相关资源参数数据计算分布式存储系统的配置;以及根据所计算

的配置自动分配存储相关资源之一的至少部分。

[0116] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其包括:包括互连计算机节点的基础设施层,其中:互连计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源;互连计算机节点中的至少一个包括一个或多个其它存储相关资源;UDSP 代理配置成:接收包括至少一个任务的工作;计算关于所述互连计算机节点中的至少一个的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求;以及基于所计算的分数执行任务中的一个或多个或将工作路由到基础设施层的更适当的计算机节点。

[0117] 根据当前公开的主题的方面,还提供了配置成连接到包括互连计算机节点的基础设施层的计算机节点,计算机节点包括:配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源,其中 UDSP 代理配置成:接收包括至少一个任务的工作;计算关于连接到基础设施层的至少一个其它计算机节点并包括一个或多个存储相关资源的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求;以及基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点。

[0118] 根据当前公开的主题的方面,还提供了操作配置成连接到包括互连计算机节点的基础设施层的计算机节点的方法,该方法包括:接收包括至少一个任务的工作;计算关于连接到基础设施层的至少一个其它计算机节点并包括一个或多个存储相关资源的所述任务中的至少一个的分数,使得每个分数指示互连计算机节点的相应计算机节点执行所述任务中的相应任务的适合性,同时满足至少一个服务水平规范(SLS)要求;以及基于所计算的分数执行任务中的一个或多个或将工作路由到更适当的计算机节点。

[0119] 根据当前公开的主题的方面,还提供了分布式存储系统,其包括互连计算机节点;计算机节点中的每一个包括配置成执行统一分布式存储平台(UDSP)代理的至少一个处理资源;计算机节点中的至少一个包括一个或多个资源,所述一个或多个资源包括配置成高速缓存对象并具有相应的高速缓存相关参数的至少一个高速缓存资源;具有至少一个高速缓存资源的相应计算机节点的至少一个 UDSP 代理配置成:监测连接到相应的计算机节点的至少一个高速缓存资源的高速缓存相关参数,用于确定高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,发起至少一个高速缓存资源的一个或多个高速对象空间的至少部分到至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速缓存对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0120] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中具有至少一个高速缓存资源的相应计算机节点的每个 UDSP 代理还配置成:监测到高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准;以及在满足至少一个第三 SLS 标准的情况下,发起至少一个高速缓存资源的一个或多个高速对象空间的至少部分到至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0121] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中第一计算机节

点具有相应的 UDSP 数据存储库,使得第一计算机节点的 UDSP 数据存储库包括关于相应于第二计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;第一计算机节点的 UDSP 代理配置成:根据从关于至少一个第二计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测高速缓存相关参数,用于确定相应的第二计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,向相应的第二计算机节点发起包括如从第一计算机节点的 UDSP 数据存储库提取的相应的第二计算机节点的高速缓存相关参数中的至少一个的消息,从而允许相应的第二计算机节点确定从第一计算机节点接收的高速缓存相关参数是否相对于第二计算机节点的高速缓存相关参数是优选的。

[0122] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中第一计算机节点具有连接到其的至少一个高速缓存资源。

[0123] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中第一计算机节点的 UDSP 数据存储库包括也关于相应于第一计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

[0124] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中 UDSP 代理还配置成响应于发起切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐,允许切换的成功发起。

[0125] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中网关资源或客户端服务器具有相应的 UDSP 数据存储库,使得网关资源或客户端服务器的 UDSP 数据存储库包括关于相应于计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;网关资源或客户端服务器的 UDSP 代理还配置成:根据从关于至少一个计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测高速缓存相关参数,用于确定相应的计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,向相应的第二计算机节点发起包括如从网关资源或客户端服务器的 UDSP 数据存储库提取的相应的计算机节点的高速缓存相关参数中的至少一个的消息,从而允许相应的计算机节点确定从网关资源或客户端服务器接收的高速缓存相关参数是否相对于计算机节点的高速缓存相关参数是更优选的。

[0126] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中相应的计算机节点的 UDSP 代理还配置成:创建指示一个或多个高速缓存对象空间的至少部分的新位置的切换后高速缓存映射;将切换后高速缓存映射发送到其它计算机节点或与一个或多个高速缓存对象空间的至少部分相关的一个或多个客户端服务器,如果有的话;以及将一个或多个高速缓存对象空间的至少部分的所有权传送到其它计算机节点。

[0127] 根据当前公开的主题的某些例子,还提供了分布式存储系统,其中相应的计算机节点的 UDSP 代理还配置成:在切换期间接收与切换后高速缓存映射内的对象有关的请求;如果根据切换后高速缓存映射对象不被相应的计算机节点拥有且相应的计算机节点没有对象的所有权,则将请求转播到其它计算机节点。

[0128] 根据当前公开的主题的某些例子,还提供了配置成连接到包括互连计算机节点的基础设施层的计算机节点,计算机节点包括:配置成高速缓存对象并具有相应的高速缓存相关参数的至少一个高速缓存资源;配置成执行统一分布式存储平台(UDSP)代理的至少

一个处理资源,其中 UDSP 代理配置成:监测连接到计算机节点的至少一个高速缓存资源的高速缓存相关参数,用于确定高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,发起至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到互连计算机节点的至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速缓存对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0129] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理配置成:监测到高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准;以及在满足至少一个第三 SLS 标准的情况下,发起至少一个高速缓存资源的一个或多个高速对象空间的至少部分到互连计算机节点的至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0130] 根据当前公开的主题的某些例子,还提供了计算机节点,其中计算机节点还包括 UDSP 数据存储库,UDSP 数据存储库包括关于相应于互连计算机节点的至少一个其它计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;且其中 UDSP 代理还配置成:根据从关于至少一个其它计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测高速缓存相关参数,用于确定相应的其它计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,向相应的其它计算机节点发起包括如从 UDSP 数据存储库提取的相应的其它计算机节点的高速缓存相关参数中的至少一个的消息,从而允许相应的其它计算机节点确定从计算机节点接收的高速缓存相关参数是否相对于其它计算机节点的高速缓存相关参数是优选的。

[0131] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 数据存储库包括也关于相应于计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

[0132] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成响应于发起切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐,允许切换的成功发起。

[0133] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成:创建指示一个或多个高速缓存对象空间的至少部分的新位置的切换后高速缓存映射;将切换后高速缓存映射发送到其它计算机节点和与一个或多个高速缓存对象空间的至少部分相关的一个或多个客户端服务器,如果有的话;以及将一个或多个高速缓存对象空间的至少部分的所有权传送到其它计算机节点。

[0134] 根据当前公开的主题的某些例子,还提供了计算机节点,其中 UDSP 代理还配置成:在切换期间接收与切换后高速缓存映射内的对象有关的请求;如果根据切换后高速缓存映射对象不被相应的计算机节点拥有且相应的计算机节点没有对象的所有权,则将请求转播到其它计算机节点。

[0135] 根据当前公开的主题的某些例子,还提供了操作计算机节点的方法,计算机节点具有连接到其的至少一个高速缓存资源并配置成连接到包括互连计算机节点的基础设施层,该方法包括:监测具有相应的高速缓存相关参数的至少一个高速缓存资源的高速缓存相关参数,用于确定高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个

第一 SLS 标准未被满足的情况下,发起至少一个高速缓存资源的一个或多个高速缓存对象空间的至少部分到互连计算机节点的至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速缓存对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0136] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:监测到高速缓存相关参数满足指示相应的高速缓存资源的利用不足的至少一个第三 SLS 标准;以及在满足至少一个第三 SLS 标准的情况下,发起至少一个高速缓存资源的一个或多个高速对象空间的至少部分到互连计算机节点的至少一个其它计算机节点的切换,至少一个其它计算机节点在接收到一个或多个高速对象空间的至少部分之后,它的高速缓存相关参数满足至少一个第二 SLS 标准。

[0137] 根据当前公开的主题的某些例子,还提供了一种方法,其中计算机节点还包括 UDSP 数据存储库,UDSP 数据存储库包括关于相应于互连计算机节点的至少一个其它计算机节点的一个或多个高速缓存资源的高速缓存相关参数的数据;该方法还包括:根据从关于至少一个第二计算机节点的一个或多个高速缓存资源的 UDSP 数据存储库提取的数据来监测高速缓存相关参数,用于确定相应的其它计算机节点的高速缓存相关参数是否满足至少一个第一 SLS 标准;以及在至少一个第一 SLS 标准未被满足的情况下,向相应的其它计算机节点发起包括如从 UDSP 数据存储库提取的相应的其它计算机节点的高速缓存相关参数中的至少一个的消息,从而允许相应的其它计算机节点确定从计算机节点接收的高速缓存相关参数是否相对于其它计算机节点的高速缓存相关参数是优选的。

[0138] 根据当前公开的主题的某些例子,还提供了一种方法,其中 UDSP 数据存储库包括也关于相应于计算机节点的高速缓存资源的自我高速缓存相关参数的数据。

[0139] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括响应于发起切换的故障而给用户提供故障通知或指示至少一个额外的高速缓存资源的添加的推荐,允许切换的成功发起。

[0140] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:创建指示一个或多个高速缓存对象空间的至少部分的新位置的切换后高速缓存映射;将切换后高速缓存映射发送到其它计算机节点和与一个或多个高速缓存对象空间的至少部分或其部分相关的一个或多个客户端服务器,如果有的话;以及将一个或多个高速缓存对象空间的至少部分的所有权传送到其它计算机节点。

[0141] 根据当前公开的主题的某些例子,还提供了一种方法,其还包括:在切换期间接收与切换后高速缓存映射内的对象有关的请求;如果根据切换后高速缓存映射对象不被相应的计算机节点拥有且相应的计算机节点没有对象的所有权,则将请求转播到其它计算机节点。

[0142] 附图简述

[0143] 为了理解当前公开的主题并查看它可实际上如何被执行,现在将仅作为非限制性的例子参考附图描述该主题,其中:

[0144] 图 1 示意性示出根据本发明的示例性实施方案的包括基础设施层的分布式存储系统的顶级结构;

[0145] 图 2 示意性示出根据当前公开的主题的用于配置分布式存储系统的简化示例性

系统；

[0146] 图 3 示意性示出根据当前公开的主题的由基于目标的管理系统执行的优化过程的简化和示例性流程图；

[0147] 图 4 示意性示出根据当前公开的主题的由基于目标的管理系统执行的配置过程的示例性操作算法的简化流程图；

[0148] 图 5 是示意性示出根据当前公开的主题的某些例子的连接到分布式存储系统的示例性计算机节点的方框图；

[0149] 图 6 是示出根据当前公开的主题的某些例子的为了创建工作而执行的一序列操作的流程图；

[0150] 图 7 是示出根据当前公开的主题的某些例子的为了创建示例性存储块-写工作而执行的一序列操作的流程图；

[0151] 图 8 是示出根据当前公开的主题的某些例子的为了管理 UDSP 代理所接收的工作而执行的一序列操作的流程图；

[0152] 图 9 是示出根据当前公开的主题的某些例子的为了给节点适合性评分以执行未决的工作任务而执行的一序列操作的流程图；

[0153] 图 10 是示出根据当前公开的主题的某些例子的为了在计算机节点上执行未决的任务而执行的一序列操作的流程图；

[0154] 图 11 是示出根据当前公开的主题的某些例子的为了管理分布式存储系统(DSS)的重新配置而执行的一序列操作的流程图；

[0155] 图 12 是示出根据当前公开的主题的某些例子的为了监测计算机节点和连接到其的资源的本地参数而执行的一序列操作的流程图；

[0156] 图 13 是示出根据当前公开的主题的某些例子的为了探测和管理连接到计算机节点的资源而执行的一序列操作的流程图；

[0157] 图 14 是示出根据当前公开的主题的某些例子的为了将新计算机节点连接到分布式存储系统(DSS)而执行的一序列操作的流程图；

[0158] 图 15 是示出根据当前公开的主题的某些例子的为了从远程计算机节点接收通知并相应地更新统一分布式存储平台(UDSP)数据库而执行的一序列操作的流程图；

[0159] 图 16 是示意性示出根据当前公开的主题的某些例子的高速缓存管理模块的方框图；

[0160] 图 17 是示出根据当前公开的主题的某些例子的为了管理计算机节点的本地高速缓存资源而执行的一序列操作的流程图；

[0161] 图 18 是示出根据当前公开的主题的某些例子的为了管理远程计算机节点的远程高速缓存资源而执行的一序列操作的流程图；

[0162] 图 19 示意性示出根据本发明的示意性实施方案的分配高速缓存资源的各种场景；

[0163] 图 20 是示出根据当前公开的主题的某些例子的为了通过切换发起者执行高速缓存切换而执行的一序列操作的流程图；

[0164] 图 21 是示出根据当前公开的主题的某些例子的为了由切换目标执行高速缓存切换而执行的一序列操作的流程图；

[0165] 图 22 是示出根据当前公开的主题的某些例子的为了处理在切换期间处理切换发起者所接收的对象相关请求而执行的一序列操作的流程图；以及

[0166] 图 23 是示出根据当前公开的主题的某些例子的为了处理在切换期间处理切换目标所接收的对象相关请求而执行的一序列操作的流程图。

[0167] 相关描述

[0168] 在附图和所阐明的描述中，相同的参考数字指示不同的实施方案或配置所共有的那些部件。

[0169] 除非另有特别说明，如从下面的讨论中明显的，应认识到，在整个说明书讨论中，利用术语例如“接收”、“计算”、“执行”、“路由”、“监测”、“传播”、“分配”、“提供”等包括将数据操纵和 / 或转换成其它数据的行动和 / 或过程，所述数据被表示为物理量，例如电子量，和 / 或所述数据表示物理对象。术语“计算机”应被广泛地解释为涵盖具有数据处理能力的任何种类的电子设备，作为非限制性的例子包括个人计算机、服务器、计算系统、通信设备、处理器（例如，数字信号处理器（DSP）、微控制器、现场可编程门阵列（FPGA）、专用集成电路（ASIC）等）、任何其它电子计算设备和 / 或其任何组合。

[0170] 根据本文的教导的操作可由为了期望目的而特别构造的计算机或为了期望目的通过存储在计算机可读存储介质中的计算机程序特别配置的通用计算机执行。

[0171] 如本文所使用的，术语“例如”、“诸如”、“比如”和其变形描述当前公开的主题的非限制性实施方案。在说明书中对“一种情况”、“一些情况”、“其它情况”或其变形的提及意味着关于实施方案描述的特定特征、结构或特性包括在当前公开的主题的至少一个实施方案中。因此，短语“一种情况”、“一些情况”、“其它情况”或其变形的出现不一定指同一实施方案。

[0172] 应认识到，在单个实施方案中也可组合地提供为了清楚起见在单独实施方案的上下文中描述的当前公开的主题的某些特征。相反，也可单独地或以任何适当的子组合提供为了简洁起见在单个实施方案的上下文中描述的当前公开的主题的各种特征。

[0173] 根据当前公开的主题的实施方案，比图 3、4、6-18 和 20-23 中所示的步骤更多和 / 或不同的步骤可被执行。在当前公开的主题的实施方案中，图 3、4、6-18 和 20-23 中所示的一个或多个步骤可以按不同的顺序执行，和 / 或一组或多组步骤可同时被执行。图 1、2、5 和 19 示出根据当前公开的主题的实施方案的系统结构的一般示意图。在图 1、2、5 和 19 中的每个模块可由执行如在本文中定义和解释的功能的软件、硬件和 / 或固件构成。在图 1、2、5 和 19 中的每个模块可集中在一个位置中或分散在多于一个的位置上。在当前公开的主题的其它实施方案中，系统可包括比图 1、2、5 和 19 所示的那些模块更少、更多和 / 或不同的模块。

[0174] 记住此，注意图 1，其示意性示出根据当前公开的主题的包括基础设施层的分布式存储系统的顶级结构。根据当前公开的主题的例子，分布式存储系统（DSS）200 可包括下列层中的一个或多个：基础设施层 201、同一分布式存储平台（UDSP）层 202 和 API/ 框架层 203。

[0175] 根据当前公开的主题的一些例子，基础设施层 201 可包括一个或多个互连计算机节点 205（例如，任何类型的计算机，特别包括一个或多个处理资源例如一个或多个处理单元、一个或多个存储器资源例如存储器、以及一个或多个网络接口）和在一些情况下包括两

个或多个互连计算机节点 205,在本文特别参考图 5 提供了关于互连计算机节点 205 的更详细的描述。基础设施层 201 还可包括下面的存储相关资源中的一个或多个:(a) 数据存储资源(例如,数据存储设备 204、RAID (独立冗余磁盘阵列) 206、DAS (直接附加存储器) 208、JBOD (只是一堆驱动器) 210、网络存储器具 207 (例如 SAN、NAS 等)、SSD 213 等);(b) 在一些情况下可另外或可选地用作高速缓存资源的高速缓存资源 212 例如存储器资源(例如 RAM、DRAM 等)、易失性和 / 或非易失性和 / 或数据存储资源(例如 SSD 213)等;(c) 网络资源 214 ;以及(d)向 DSS 200 提供另外的功能和 / 或增强其性能(例如压缩加速器、加密加速器 209、实现与 SAN 资源的通信的主机总线适配器(HBA)等)的额外资源。

[0176] 在一些情况下,资源可包括多于一个同一类型的设备和 / 或多于一个不同类型的设备。本文接下来是一些资源的更详细的描述。

[0177] 根据当前公开的主题的一些例子,计算机节点 205 可由网络(例如,通用网络)互连。

[0178] 在一些情况下,基础设施层 201 的一个或多个资源可直接连接到一个或多个计算机节点 205。在一些实施方案中,基础设施层 201 的一个或多个资源可包括在计算机节点 205 中并形成其一部分。在一些情况下,基础设施层 201 的一个或多个资源可通过网络(例如通用网络)连接(例如由逻辑连接例如 iSCSI 222 等)到一个或多个计算机节点 205。

[0179] 可选地,网络可以是通用网络。可选地,网络可包括 WAN。可选地,WAN 可以是全球 WAN,例如互联网。可选地,网络资源可使用 IP 网络基础设施互连。可选地,网络可以是存储区域网络(SAN)。可选地,网络可包括存储虚拟化。可选地,网络可包括 LAN。可选地,网络基础设施可包括以太网、Infiniband、FC (光纤通道) 217、FCoE (以太网上光纤通道)等或一个或两个网络基础设施的任何组合。可选地,网络可以是本领域中已知的任何类型的网络,包括通用网络和 / 或存储器网络。可选地,网络可以是适合于应用用于分配和管理网络内的资源的基于目标的管理系统的任何网络,如本文进一步描述的。可选地,网络可以是任何两个或多个网络类型的组合(特别包括本文公开的网络类型)。

[0180] 根据当前公开的主题的一些例子,基础设施层 201 (特别包括计算机节点 205、数据存储资源、高速缓存资源、网络资源、连接到计算机节点 205 的额外的资源或任何其它资源)的至少一个资源可以是现成的商品,不是连接到网络和 / 或一个或多个计算机节点 205 的故意构建的资源。应注意,这样的资源可以如本文详述的被互连,而不考虑资源特征,例如制造商、尺寸、计算能力、容量等。因此,可与计算机节点 205 通信的任何资源(特别包括计算机节点 205)——而不考虑其制造商——可连接到基础设施层 201,并由 DSS 200 利用,如本文详细描述。在一些情况下,任何数量的资源(特别包括计算机节点 205)可连接到网络和 / 或一个或多个计算机节点 205 并由 DSS 200 利用,因而实现 DSS 200 的可量测性。在一些情况下,任何数量的计算机节点 205 可连接到网络,且任何数量的资源可连接到一个或多个计算机节点 205 并由 DSS 200 利用,因而实现 DSS 200 的可量测性。应注意,在本文特别关于图 5 更详细描述了关于将新资源(特别包括计算机节点 205)连接到 DSS 200 的过程的更详细的解释。

[0181] 转到 UDSP 层 202,根据当前公开的主题的一些例子,它可包括可安装在一个或多个计算机节点 205 上(或以另外方式与一个或多个计算机节点 205 相关或包括在一个或多个计算机节点 205 内)的一个或多个 UDSP 代理 220。在一些情况下,UDSP 代理 220 可安装

在每个计算机节点 205 上(或以另外方式与每个计算机节点 205 相关)。在一些情况下,UDSP 代理 220 可此外安装在一个或多个网关资源 216 (其特别可充当如本文进一步描述的协议转换器)上(或以另外方式与一个或多个网关资源 216 相关),或在一些情况下在每个网关资源 216 上。在一些情况下,UDSP 代理 220 此外安装在一个或多个客户端服务器 218 (例如,服务器或连接到 DSS 200 作为客户端的其它设备)上(或以另外方式与一个或多个客户端服务器 218 相关),或在一些情况下在每个客户端服务器 218 上。应注意,在一些情况下,客户端服务器 218 可与 DSS 200 直接交互作用,而不需要可选的任何网关资源 216。应进一步注意,在一些情况下,在 UDSP 代理 220 中根据其安装位置或其关联可能有差异(例如,在其功能和 / 或其容量等方面的差异)(例如,在安装在计算机节点 205 上或以另外方式与计算机节点 205 相关的 UDSP 代理 220、安装在网关资源 216 上或以另外方式与网关资源 216 相关的 UDSP 代理 220、安装在客户端服务器 218 上或以另外方式与客户端服务器 218 相关的 UDSP 代理 220 之间可能有差异,等等)。

[0182] 应注意,在本文特别关于图 5 提供了 UDSP 代理 220 的详细描述。据说,应注意,根据当前公开的主题的一些例子,UDSP 代理 220 可配置成控制并管理 DSS 200 的各种操作(特别包括自动分配和管理基础设施层 201 的资源,处理数据路径操作,等等)。在一些情况下,UDSP 代理 220 可配置成管理新计算机节点 205 到 DSS 200 的基础设施层 201 的连接。在一些情况下,UDSP 代理 220 可配置成探测连接到它们被安装于的计算机节点 205 的资源并管理这样的资源。如上所述,在本文特别关于图 5 提供了 UDSP 代理 220 的更详细的描述。

[0183] 在一些情况下,UDSP 层 202 可包括 UDSP 225,其包括 DSS 200 的管理系统。可选地,可通过安装在基础设施层 201 中的计算机节点 205 上的一个或多个 UDSP 代理 220 或通过安装在网关资源 216 上或在访问 DSS 200 (例如,直接地和 / 或通过网关资源 216)的客户端服务器 218 上的一个或多个 UDSP 代理 220 或其任何组合来实现管理系统处理。

[0184] 测量系统可使用户能够执行与 DSS 200 有关的各种管理工作(特别包括监测和报告工作),例如创建可与服务水平规范(SLS)相关的新逻辑存储实体(例如,逻辑单元、对象存储库、文件系统实例等)(在一些情况下,每个逻辑存储实体与单个 SLS 相关),更新逻辑存储实体,准予逻辑存储实体对网关资源 216 和 / 或客户端服务器 218 的访问允许,创建快照,创建备份,到远程站点的故障转移,对主要站点的故障自动恢复,监测 DSS 200 的动态行为,监测 SLS 顺应性,涉及 DSS 200 的不同范围(例如,在整个 DSS 200 的分辨率、某些站点、某些类型的使用例如对于某些 SLS、某些资源等中)的各种(例如,预先定义的和 / 或用户定义的,等等)报告(例如,性能报告、资源可用性报告、库存报告、指示计算机节点 205 和其它资源之间的关系的关系报告、趋势报告和各種参数——包括关键性能指示器——的预测报告等)的产生,管理 DSS 200 所提供的警告(例如,出故障的硬件的警告等),等等。应注意,上述管理工作仅作为非限制性的例子被提供。应注意,在一些情况下,逻辑存储实体可由 DSS 200 根据 SLS 自动创建,如在本文进一步详述的。应注意,每个存储实体可与一个或多个数据存储资源相关。

[0185] 应注意,在整个说明书中,当提到用户时,这可以指操作人员,例如系统管理员或任何类型的辅助实体。辅助实体例如可以指外部应用,例如外部管理系统,包括不需要任何人干预的辅助实体,等等。

[0186] 在一些情况下,管理系统可使用户能够给 DSS 200 提供用户定义的存储要求,其

定义指定用户要求 DSS 200 满足的各种要求的服务水平规范(SLS)。在一些情况下,SLS 可与逻辑存储实体相关。可选地,SLS 可包括信息,例如数据将被存储和 / 或处理的一个或多个地理位置的规范;定义可用性、保留、恢复参数(例如,RPO——恢复点目标,RTO——恢复时间目标)的本地保护水平;定义多长信息应被保留的备份保留政策;定义一个或多个远程地理位置以便在各种灾难情况下实现规定的可用性、保留和恢复目标的灾难恢复(DR)的远程保护水平;本地和 / 或远程复制政策;使用度量例如 IOPS (每秒输入 / 输出操作)、响应时间和吞吐量定义(可选地,调拨)的性能水平;加密要求;重复删除要求;压缩要求;存储方法(物理容量、稀少的容量 / 供应)等。

[0187] 在一些情况下,管理系统可实现各种服务水平组(SLG)的管理(包括创建、更新和删除)。SLG 是可在多个逻辑存储实体当中被共享的模板 SLS。SLG 可以是部分 SLS (其需要增强)和 / 或包含可被覆盖的设置。因此,例如,SLG 可定义可被各种 SLS 继承的各种恢复参数,每个 SLS 可添加和 / 或覆盖 SLS 参数。

[0188] 根据当前公开的主题的一些例子,UDSP 225 可包括用于分配资源和管理 DSS 200 中的资源的自动管理系统。可选地,自动管理系统是基于目标的管理系统(OBMS)100,其可配置成特别基于一个或多个服务水平规范(SLS)所定义的用户定义的要求、与计算机节点 205 和 / 或连接到其的资源有关的各种参数的数据、涉及 DSS 200 或其部分的各种参数的数据(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)和涉及 DSS 200 和环境(例如,客户端服务器 218、网关资源 216 等)的动态行为的各种参数的数据中的任一个或任何组合来分配和管理网络中的资源,如在本文特别关于图 2 和图 5 进一步描述的。可选地,OBMS 100 处理可通过安装在基础设施层 201 中的一个或多个计算机节点 205 上的一个或多个 UDSP 代理 220 或通过安装在网关资源 216 上或在访问 DSS 200(例如,直接地或通过网关资源 216)的代理服务器 218 上的一个或多个 UDSP 代理 220 或其任何组合来实现。

[0189] 根据当前公开的主题的一些例子,API/ 框架层 203 包括便于软件扩展(插件)添加到 DSS 200 的插件层。这样的插件可用于例如将过程应用于数据,将新功能和特征引入到 DSS 200,使 DSS 200 与特定的应用通过接口连接并实现应用特定的任务(例如,存储相关工作等),实现各种资源特定的驱动器,引入新的 SLS 参数和 / 或参数组(例如,与插件功能和 / 或目标有关),实现管理功能,等等。在一些情况下,插件层也可包括与各种硬件部件(例如,加密卡等)相关的驱动器。

[0190] 在一些情况下,插件可部署在一个或多个 UDSP 代理 220 上。在一些情况下,例如根据插件规范(例如软件加密插件可安装在任何 UDSP 代理 220 上),根据连接到计算机节点 205 和 / 或连接到网关资源 216 和 / 或连接到 UDSP 代理 220 安装于的客户端服务器 218 的各种资源(例如,硬件加速器插件可自动部署在与计算机节点 205 相关的每个 UDSP 代理 220 上,计算机节点 205 与这样的硬件加速器相关),根据自动管理系统(例如,OBMS 100)的决定,或根据系统管理员的选择,等等,插件可部署在一个或多个 UDSP 代理 220 上。在一些情况下,可例如通过自动管理系统(例如,OBMS 100)和 / 或通过计算机节点 205 来自动部署插件。可选地,软件扩展可包括数据处理插件 226,例如启用例如存储在 DSS 200 上的数据的重复删除的数据重复删除插件、启用例如存储在 DSS 200 上的数据的加密 / 解密的数据加密插件、启用例如存储在 DSS 200 上的数据的压缩 / 解压缩的数据压缩插件,等等。可选

地,软件扩展可包括存储特征插件 228,例如实现例如存储在 DSS 200 上的数据的编索引的内容编索引插件、实现例如存储在 DSS 200 上的数据的快照的管理的快照管理插件、实现例如存储在 DSS 200 上的数据的分层布置的分层布置管理插件、实现例如与灾难恢复有关的过程、政策、程序的管理的灾难恢复插件、实现例如存储在 DSS 200 上的数据的连续或实时备份的管理的连续数据保护插件等。可选地,软件扩展可包括应用插件 230,例如实现例如使查询处理加速的数据库插件、实现例如各种 DSS 200 的管理工作的性能和与用户、客户端服务器 218 和连接到 DSS 200 的其它实体的其它互动的管理插件 233、以及其它适当的应用插件。

[0191] 如本文指示的,在一些情况下,插件可引入新的 SLS 参数和 / 或参数组(例如,与插件功能和 / 或目标有关)。在这样的情况下,根据插件功能,相应的 SLS 参数和 / 或参数组可被引入到 DSS 200。这样引入的 SLS 参数可被使用,以便由用户和 / 或自动地由自动管理系统(例如,OBMS 100)等设置插件相关的要求。

[0192] 在一些情况下,软件扩展可存储在计算机节点 205 之一上或分布在多于一个计算机节点 205 上。在一些情况下,软件扩展可存储在连接到一个或多个计算机节点 205 的一个或多个数据存储资源上。在一些情况下,软件扩展可存储在可被 UDSP 代理 220 共享的虚拟软件扩展库上。

[0193] 在一些情况下,软件扩展可被自动和 / 或手动地(例如,通过系统管理员)管理。有时可通过利用管理插件 233 来执行这样的管理。在这样的情况下,管理插件 233 可实现软件扩展到 DSS 200 的添加 / 从 DSS 200 的移除、各种软件扩展到一个或多个 UDSP 代理 220 的添加 / 从一个或多个 UDSP 代理 220 的移除,等等。

[0194] 在 DSS 200 的顶级结构的描述之后,特此提供可由基于目标的管理系统(OBMS) 100 执行的 DSS 200 的配置过程的详细描述。为了这个目的,现在注意图 2,其示出根据当前公开的主题的用于配置分布式存储系统 200 的简化示例性系统。为了这个目的,OBMS 100 可特别配置成自动分配和管理基础设施层 201 中的资源。OBMS 100 可包括输入模块 102、一个或多个处理器 104 和输出模块 106。

[0195] 在一些情况下,输入模块 102 可配置成接收输入数据。这样的输入数据可特别包括由一个或多个服务水平规范(SLS)定义的用户定义的存储要求、一个或多个逻辑存储实体的定义、与计算机节点 205 和 / 或连接到其的资源(包括也被称为存储相关资源数据的存储相关资源)有关的各种参数的数据、涉及 DSS 200 或其部分的各种参数的数据(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)、与 DSS 200 和环境(例如,客户端服务器 218、网关资源 216 等)有关的动态行为(动态行为参数数据)的各种参数的数据等中的任一个或任何组合。

[0196] 在一些情况下,用户定义的要求可定义一个或多个服务水平规范(SLS),其规定一个或多个用户要求 DSS 200 和 / 或一个或多个逻辑存储实体满足的各种要求。

[0197] 在一些情况下,与 DSS 200 和环境(动态行为参数数据)有关的动态行为的各种参数的数据可包括指示 DSS 200 的部件(包括计算机节点 205 和连接到其的资源)中的一个或多个的当前状态的各种参数数据。这样的数据可包括存在和 / 或负荷和 / 或可用性和 / 或故障和 / 或能力和 / 或响应时间和 / 或连接性和 / 或成本(例如,网络链路的成本,不同类型的数据存储资源)的数据和 / 或与一个或多个资源有关的任何其它数据,包括与一个或多

个计算机节点 205、一个或多个网关资源 216、一个或多个客户端服务器 218 等有关的数据。在一些情况下,这样的数据可特别包括各种统计数据。

[0198] 在一些情况下,与计算机节点 205 和 / 或连接到其的资源(包括也被称为存储相关资源数据的存储相关资源)有关的各种参数的数据可包括指示 DSS 200 的资源的各种参数的数据,所述资源例如包括硬件资源,包括存储相关资源,例如:

[0199] a. 与数据存储资源有关的参数(例如,对于其硬盘驱动器中的每个):

[0200] 1. 硬盘驱动器类别参数(例如,硬盘驱动器尺寸、接口(例如, SAS、SATA、FC、超 SCSI 等)、高速缓存器尺寸、特殊特征(例如,驱动加密等)等;

[0201] 2. 硬盘驱动器性能参数(例如,响应时间、平均时延、随机查找时间、数据传输率等);

[0202] 3. 硬盘驱动器功率消耗;

[0203] 4. 硬盘驱动器可靠性参数(例如,在故障之间的平均时间(MTBF)、年故障率(AFR)等)。

[0204] b. 计算机节点 205 的参数:

[0205] 1. CPU 和每 CPU 核心的数量;

[0206] 2. 每个 CPU 和 / 或核心的性能参数,例如频率、L2 和 L3 高速缓存器尺寸;

[0207] 3. 结构(例如,CPU 和 / 或核心支持 64 位计算?它是小 endian 还是大 endian?);

[0208] 4. 对某些指令集(例如, AES-NI、用于加速 AES 加密的新指令集)的支持;

[0209] 5. 可用的硬盘驱动器插槽的数量;

[0210] 6. 可用的存储接口(SATA、SAS 等);

[0211] 7. 最大存储量;

[0212] 8. 所支持的存储器配置;

[0213] c. 高速缓存资源参数:

[0214] 1. 高速缓存资源类型(例如, DRAM、SSD)、尺寸和性能;

[0215] 2. 高速缓存存储空间是本地的还是远程的?

[0216] 3. NUMA 参数。

[0217] d. 网关资源参数:

[0218] 1. CPU 和每 CPU 核心的数量;

[0219] 2. 每个 CPU 和 / 或核心的性能参数,例如频率、L2 和 L3 高速缓存器尺寸;

[0220] 3. 结构(例如,CPU 和 / 或核心支持 64 位计算?它是小 endian 还是大 endian?);

[0221] 4. 对某些指令集(例如, AES-NI、用于加速 AES 加密的新指令集)的支持;

[0222] 5. 外壳中的可用硬盘驱动器插槽的数量;

[0223] 6. 可用的存储接口(SATA、SAS 等);

[0224] 7. 最大存储量;

[0225] 8. 所支持的存储器配置;

[0226] 9. 与网关有关的联网参数(端口的数量、每个端口的速度和类型等)。

[0227] e. 网络资源参数:

[0228] 1. 切换和路由能力;

[0229] 2. 网络类型;

[0230] 3. 安全参数。

[0231] 应注意,这些仅仅是例子,且可使用额外的和 / 或可选的各种参数。

[0232] 在一些情况下,与 DSS 200 和环境的动态行为有关的数据(动态行为参数数据)可包括指示 DSS 200 的资源的各种参数,包括例如硬件资源,例如:

[0233] a. 与数据存储资源有关的参数(例如,对于其硬盘驱动器中的每个):

[0234] 1. 硬盘驱动器自由空间。

[0235] 2. 硬盘驱动器的 S. M. A. R. T. 参数。

[0236] 3. 硬盘驱动器的功率状态(关闭、在启动阶段中、准备等)。

[0237] 4. 硬盘驱动器上的最近和当前负荷。

[0238] 5. 现有的分配和保留。

[0239] b. 计算机节点 205 的参数:

[0240] 1. 对每个核心的最近和当前的负荷统计。

[0241] 2. 现有的分配和保留。

[0242] 3. 当前存储量。

[0243] c. 高速缓存资源参数:

[0244] 1. 可用尺寸。

[0245] 2. 高速缓存器的占用水平。

[0246] 3. 最近和当前的交换 / 页面故障统计。

[0247] 4. 现有的分配和保留。

[0248] d. 网关资源参数:

[0249] 1. 最近和当前的网络连接统计。

[0250] 2. 最近和当前的节点负荷统计。

[0251] 3. 最近和当前的时延统计。

[0252] 4. 最近和当前的路由成本统计(对于通过网关路由到 DSS 中的命令)。

[0253] 5. 现有的分配和保留。

[0254] e. 网络资源参数:

[0255] 1. 网络段的最近和当前负荷。

[0256] 2. 网络段的最近和当前的可靠性和质量参数。

[0257] 3. 现有的分配和保留。

[0258] 应注意,这些仅仅是例子,且额外的和 / 或可选的各种参数可被使用。

[0259] 在一些情况下,输入模块 102 可配置成将输入数据转移到一个或多个处理器 104。如所示的,通过一个或多个 UDSP 代理 220 (例如,当利用如特别参考图 5 在本文进一步详述的基于目标的配置模块 380 时),例如通过安装在基础层 201 中的一个或多个计算机节点 205 上的 UDSP 代理 220、或通过安装在一个或多个网关资源 216 上的 UDSP 代理 220、或通过安装在访问 DSS 200 (例如直接地或通过网关资源 216)的一个或多个客户端服务器 218 上的 UDSP 代理 220 或其任何组合,可实现 OBMS 100 的处理。在这样的情况下,一个或多个处理器 104 可以是与这样的 UDSP 代理 220 相关的一个或多个处理资源(例如,处理单元)(例如,如果处理通过安装在计算机节点 205 上的 UDSP 代理 220 实现,则处理器可以是该计算机节点 205 的处理单元,等等)。应注意,多于一个处理资源(例如,处理单元)可例如在并行

和 / 或分布式处理的情况下被使用。

[0260] 一个或多个处理器 104 可配置成从输入模块 102 接收输入数据并基于输入数据执行优化过程用于确定配置要求,配置要求满足特别关于它们影响的实体(例如与这样的 SLS 相关的逻辑存储实体)由 DSS 200 的一个或多个用户提供的所有用户定义的存储要求(例如, SLS)的。在本文特别关于图 3 提供了优化过程和所确定的配置要求的更详细的描述。

[0261] 配置要求可传送到输出模块 106,其在一些情况下可确定当前 DSS 200 的资源是否足以满足所确定的配置要求。因此,输出模块 106 可配置成执行解推动的行动,这些行动可包括资源的分配、保留、调拨或过度调拨(例如,实际上分配比在基础设施层 201 中可用的实际资源多的资源)——如果配置要求可被系统满足,或发出用户按照来行事的改进推荐,其可包括添加资源和 / 或添加插件和 / 或用于使系统能够满足配置要求的任何其它推荐。这样的改进推荐可包括例如添加一个或多个资源、添加或升级一个或多个插件、使基础设施跨越额外的和 / 或不同的位置(本地和 / 或远程)等的推荐。

[0262] 应注意,在一些情况下,当首次部署 DSS 200 和 / 或一个或多个逻辑存储实体时和 / 或在施加到 DSS 200 和 / 或一个或多个逻辑存储实体的一个或多个变化(例如,预先定义的变化)(例如,资源例如计算机节点 205、高速缓存资源、数据存储资源、网络资源、插件或任何其它资源到 DSS 200 的添加 / 移除;一个或多个用户定义的存储要求中的变化等)之后,和 / 或根据 DSS 200 的动态行为(如下面特别关于图 5 和图 11 进一步详述的)等,可发起配置过程或其部分。此外或可选地,可以用半连续的方式(例如,以预先确定的时间间隔等)发起配置过程或其部分。此外或可选地,可连续地执行配置过程或其部分。

[0263] 应进一步注意,参考图 2,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0264] 现在注意图 3,其示意性示出根据当前公开的主题的由基于目标的管理系统执行的优化过程的简化和示例性流程图。在一些情况下,一个或多个处理器 104 可配置成接收输入数据(例如,从输入模块 102),且在一些情况下将所接收的输入数据转换成适合于由优化引擎处理的格式(例如,转换成优化问题表示)(块 112)。

[0265] 与一个或多个处理器 104 相关的优化引擎可配置成基于原始和 / 或转换的输入数据执行优化过程以达到满足如输入数据所定义的要求的所需配置(如在本文特别关于图 2 进一步详述的)(块 114)。应注意,在一些情况下,可指示优化过程返回它找到的第一个有效解,而在其它情况下,可指示优化过程从一组计算出的有效解搜索最佳解。可选地,在优化过程中的使用的优化技术可包括线性编程、模拟退火、遗传算法或本领域中已知的任何其它适当的优化技术中的任一个或任何组合。可选地,优化技术可利用启发法和 / 或近似法。可选地,可基于部分的和 / 或不是最新的信息来采取优化决定。

[0266] 在一些情况下,优化引擎的输出可由一个或多个处理器 104 从优化解表示转换成配置要求表示(块 116)。

[0267] 在一些情况下,配置要求由一个或多个处理器 104 输出,例如作为下列项中的任一个或任何组合:位置要求(例如,至少一个额外的站点的可用性、在额外的站点中的某个量的存储空间的可性、在站点之间的最大时延、在例如用于灾难恢复目的站点之间的最

小地理距离,等等)、高速缓存资源要求(例如,所需的高速缓存器尺寸、所需的高速缓存器类型、所需的高速缓存器位置、所需的高速缓存器性能参数等)、网关资源要求(例如,所需的光纤通道带宽、所需的处理性能参数等)、网络资源要求(例如,所需的网络带宽、所需的网络类型等)、计算资源要求(例如,计算机节点处理性能参数、CPU 核心的计算机节点数量等)、数据存储资源要求(例如,所需的存储空间、所需的存储类型等)、额外的资源要求(例如,所需的压缩性能、所需的加密性能等)、插件要求(例如,所需的数据插件等)、环境要求(例如,所需的物理安全水平等)等(块 117)。

[0268] 应注意,参考图 3,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0269] 转到图 4,其示出了根据当前公开的主题的由基于目标的管理系统执行的配置过程的示例性操作算法的简化流程图的示意图。在一些情况下,如上所示,输入模块 102 可接收输入数据,并将数据转移到一个或多个处理器 104 (块 110)。如上面进一步指示的,一个或多个处理器 104 在一些情况下可将输入数据转换成适合于由优化引擎处理的格式(例如,转换成优化问题表示)(块 112)。

[0270] 与一个或多个处理器 104 相关的优化引擎可配置成基于原始和 / 或转换的输入数据执行优化过程以达到满足如输入数据所定义的要求的所需配置(如在本文特别关于图 2 进一步详述的)(块 114)。在一些情况下,优化引擎的输出可由一个或多个处理器 104 从优化解表示转换成配置要求表示(块 116)。

[0271] 在一些情况下,输出模块可比较所需的配置与 DSS 200 的资源(例如,计算机节点 205、存储相关资源等)和 / 或环境的实际数据用于确定 DSS 200 是否可满足所需配置(块 118)。应注意,在一些情况下,实际 DSS 200 的资源可以涉及 DSS 200 的资源的当前可用的那些部分。如果实际 DSS 200 的资源和 / 或环境可满足所需的配置,则 OBMS 100 可配置成根据所需的配置来保留和 / 或分配资源(块 126)。在一些情况下,OBMS 100 可配置成建立 DSS 200 的配置和 / 或执行任何引起的部署行动(块 128)。在一些情况下,建立和 / 或部署行动可特别包括自动创建与 SLS 相关的新逻辑存储实体(例如逻辑单元、对象存储库、文件系统实例等)。在一些情况下,每个逻辑存储实体与单个 SLS 相关。

[0272] 作为建立存储配置和 / 或执行任何引起的部署行动的部分,相关的建立和 / 或部署行动请求可被发送到 UDSP 代理 205;在一些情况下,这样的请求被发送到与存储相关资源相关的 UDSP 代理 205,存储相关资源与所请求的建立和 / 或部署行动相关。在一些情况下,接收这样的请求的 UDSP 代理 205 可配置成更新与其相关的关于被请求来由 DSS 200 使用的建立和 / 或部署的数据存储库,如下面特别关于图 5 进一步详述的。在一些情况下,在部署之后,部署 DSS 200 的过程成功地结束(块 130)。

[0273] 如果实际 DSS 200 的资源和 / 或环境不能满足所需的配置,OBMS 100 可配置成将消息发送给用户(例如系统管理员),给用户提供故障通知和 / 或关于由用户采取的校正行动的推荐,用于允许所需的基础设施配置的实现(块 120)。可选地,行动可包括添加允许配置的成功计算的基础设施资源。可选地,行动可包括添加相关的插件。可选地,行动可涉及使基础设施资源跨越额外的和 / 或可选的位置。应注意,本文公开的推荐仅仅是例子,且其

它推荐可此外或可选地被发给用户。在一些情况下,OBMS 100 可配置成做出关于在某个间隔 / 延迟之后所需的基础设施配置是否应被重新评估的决定(块 122)。如果是,OBMS 100 可配置成返回到块 112。可选地,输出模块 106 可选地在某个间隔 / 延迟之后自动转到 112,如果被设置到连续模式。可选地,是否重试的决定基于重试指令的用户输入。如果否,部署 DSS 200 的过程失败。在一些情况下,OBMS 100 可配置成报告故障。

[0274] 应注意,参考图 4,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0275] 现在注意图 5,其中示出了方框图,其示意性示出根据当前公开的主题的某些例子的连接到分布式存储系统的示例性计算机节点。

[0276] 根据当前公开的主题的一些例子,计算机节点 205 可包括一个或多个处理资源 310。一个或多个处理资源 310 可以是处理单元、微处理器、微控制器或适合于独立地或协作地处理数据的任何其它计算设备或模块,包括多个和 / 或并行和 / 或分布式处理单元,其用于控制相关的计算机节点 205 的资源 and / 或连接到计算机节点 205 的存储相关资源和用于实现与计算机节点 205 的资源和 / 或连接到计算机节点 205 的存储相关资源有关的操作。

[0277] 计算机节点 205 还可包括用于使计算机节点 205 能够特别与其它计算机节点和 / 或连接到 DSS 200 的其它资源通信的一个或多个网络接口 320(例如,网络接口卡或任何其它适当的设备)。

[0278] 根据当前公开的主题的一些例子,计算机节点 205 可与配置成存储数据的 UDSP 数据存储库 330 相关,该数据特别包括定义 SLS 的各种用户定义的存储要求的数据和 / 或与每个 SLS 相关的逻辑存储实体的数据和 / 或与计算机节点 205 和 / 或连接到其的存储相关资源有关的各种参数的数据和 / 或与涉及 DSS 200 或其部分的各种参数有关的数据和 / 或与 DSS 200 和环境(例如,客户端服务器 218、网关资源 216 等)的动态行为相关的数据和 / 或与 DSS 200 的建立和 / 或部署相关的数据和 / 或任何其它数据。在一些情况下,UDSP 数据存储库 330 还可配置成实现所存储的数据的检索、更新和删除。应注意,在一些情况下,UDSP 数据存储库 330 可在本地位于计算机节点 205 上、连接到计算机节点 205 的存储相关资源(例如,数据存储资源、高速缓存资源或任何其它适当的资源)上、客户端服务器 218 上、网关资源 216 上、或任何其它适当的位置。在一些情况下,UDSP 数据存储库 330 可分布在两个或多个位置之间。在一些情况下,UDSP 数据存储库 330 可此外或可选地存储在 DSS 200 内的一个或多个存储实体上。在一些情况下,此外或可选地,UDSP 数据存储库 330 可在多个计算机节点之间被共享。

[0279] 根据当前公开的主题的一些例子,计算机节点 205 还可包括可例如由一个或多个处理资源 310 执行的 UDSP 代理 220。如上所示,UDSP 代理 220 可特别配置成控制和管理计算机节点 205 和 / 或 DSS 200 的各种操作。UDSP 代理 220 可包括下列模块中的一个或多个:工作管理模块 335、多播模块 340、工作创建模块 345、执行模块 350、本地参数监测模块 360、远程节点参数监测模块 370、云即插即用模块 380、资源探测和管理模块 385、基于目标的配置模块 390、高速缓存管理模块 397 和基于目标的路由模块 395。

[0280] 根据当前公开的主题的一些例子,工作管理模块 335 可配置成管理所接收的工作,例如数据路径操作(例如,读 / 写操作),如特别关于图 8 进一步详述的。

[0281] 多播模块 340 可配置成将各种通知传播(例如,通过单播 / 多播 / 重播传输)到各种 UDSP 代理 220 (例如,安装在其它计算机节点、网关资源 216、客户端服务器 218 等上的 UDSP 代理)。这样的通知可包括例如资源状态改变的通知、新资源的添加的通知、资源的断开的通知、本地参数中的变化的通知等。此外,多播模块 340 可配置成处理在各种 UDSP 代理 220 和 DSS 200 的其它实体以及外部实体(例如,外部管理系统等)之间的任何协议。

[0282] 工作创建模块 345 可配置成创建用于在 DSS 200 中执行的新工作,如特别关于图 8 和 9 进一步详述的。

[0283] 执行模块 350 可配置成在本地执行与所接收的工作相关的一个或多个任务,如在本文特别关于图 10 进一步详述的。

[0284] 本地参数监测模块 360 可配置成监测各种本地参数,例如指示计算机节点 205 和 / 或连接到其的任何资源的动态行为的参数,并传播指示对一个或多个本地参数的变化的通知(例如,当利用多播模块 340 时),如特别关于图 12 进一步详述的。应注意,在一些情况下,本地参数是与特定的计算机节点 205 (或网关资源 216 或客户端服务器 218,加以必要的变更)和 / 或连接到其的资源有关的参数,监测在计算机节点 205 上被执行。

[0285] 远程节点参数监测模块 370 可配置成接收指示一个或多个远程计算机节点 205 和 / 或连接到其的资源的一个或多个参数中的变化的通知,并相应地更新 UDSP 数据存储库 330,如特别关于图 15 进一步详述的。在一些情况下,远程节点参数监测模块 370 可配置成向另一计算机节点 205 注册(例如,向与其它计算机节点 205 相关的 UDSP 代理 220 注册)以从其接收选择通知。应注意,在一些情况下,远程节点参数监测模块 370 可配置成独立地和 / 或主动查询远程计算机节点 205 以找到任何所需的信息。

[0286] 云即插即用模块 380 可配置成实现计算机节点 205 到 DSS 200 的自主和 / 或自动连接,如特别关于图 14 进一步详述的。

[0287] 资源探测和管理模块 385 可配置成探测和管理连接到计算机节点 205 的资源,如特别关于图 13 进一步详述的。

[0288] 基于目标的配置模块 390 可配置成配置和 / 或重新配置 DSS 200,如特别关于图 2-4 和 11 进一步详述的。

[0289] 基于目标的路由模块 395 可配置成将所接收的工作路由到计算机节点 205,如特别关于图 6 和 8 进一步详述的。

[0290] 高速缓存管理模块 397 可特别配置成监测与高速缓存资源有关的参数,并管理连接到计算机节点的高速缓存资源(特别包括执行高速缓存切换),如特别关于图 16-22 在本文进一步详述的。

[0291] 应注意,一个或多个处理资源 310 可配置成执行 UDSP 代理 220 和其中包括的任何模块。

[0292] 应注意,根据当前公开的主题的一些例子,UDSP 代理 220 中的一些或全部可被组合和设置为单个模块,或作为例子,它们中的至少一个可以以两个或多个模块的形式实现。应进一步注意,在一些情况下,UDSP 代理 220 可此外或可选地安装在一个或多个网关资源 216 和 / 或客户端服务器 218 等上。在这样的情况下,UDSP 代理 220 的部分或修改版本可

安装在一个或多个网关资源 216 和 / 或客户端服务器 218 上和 / 或由一个或多个网关资源 216 和 / 或客户端服务器 218 使用。

[0293] 转到图 6, 其示出了说明根据当前公开的主题的某些例子的为了创建工作而执行的一序列操作的流程图。工作可产生, 以便执行由 DSS 200 接收的所请求的操作 (例如, 读 / 写操作、管理操作等)。在一些情况下, 工作可包括作为所请求的操作而被执行的一个或多个任务的列表。

[0294] 在一些情况下, 工作创建模块 345 可执行工作创建过程 500。为了这个目的, 在一些情况下, 工作创建模块 345 可接收例如起源于客户端服务器 218、网关资源 216、计算机节点 205 或任何其它源的所请求的操作 (块 510)。接收的所请求的操作可包括指示操作的类型 (例如, 读、写、管理等) 的数据和 / 或与所请求的操作有关的任何其它数据 (例如, 在写请求中, 指示相关逻辑存储实体 (在其上, 操作将被执行, 块将被写, 等等) 的数据)。

[0295] 工作创建模块 345 可配置成创建工作容器 (块 520)。工作容器可特别包括下列项中的一个或多个: 指示所请求的操作发起者的数据 (例如, 其网络标识符)、指示操作将被执行的相关逻辑存储实体的数据、操作特定的数据 (例如, 在块写操作 - 待写的块的情况下) 和空任务列表。

[0296] 在一些情况下, 例如当请求与逻辑存储实体相关时, 工作创建模块 345 可配置成取回与逻辑存储实体相关的 SLS, 并根据 SLS 创建待执行的一个或多个任务 (例如, 如果 SLS 要求数据被加密, 加密任务可被自动创建, 等等) (块 530)。

[0297] 应注意, 工作创建过程 500 可由与计算机节点 205 相关的 UDSP 代理 220 的工作创建模块 345 执行。然而, 应注意, 此外和 / 或可选地, 工作创建过程 500 可由与客户端服务器 218 和 / 或网关资源 216 或具有工作创建模块 345 的任何其它源相关的 UDSP 代理 220 的工作创建模块 345 执行。因此, 在一些情况下, 计算机节点 205 可接收已经由客户端服务器 218 和 / 或网关资源 216 等创建的一个或多个工作。

[0298] 应注意, 参考图 6, 一些块可整合成合并的块或可分成几个块, 和 / 或其它块可被添加。此外, 在一些情况下, 可以用与本文所述的不同的顺序执行块。也应注意, 虽然也参考实现它们的系统元件描述了流程图, 这决不是限制性的, 且块可由除了本文所述的元件以外的元件执行。

[0299] 为了更好地理解工作创建的过程, 注意图 7, 其示出说明根据当前公开的主题的某些例子的为了创建示例性存储块 - 写工作而执行的一序列操作的流程图。在本文提供的例子中, 工作创建模块 345 可接收将被写在 DSS 200 中的块数据和指示块将被写的相关逻辑存储实体的数据 (块 605)。

[0300] 在一些情况下, 工作创建模块 345 可配置成创建新工作容器。工作容器可特别包括指示操作起源于的发起者的数据 (例如, 其网络标识符)、指示块将被写的相关逻辑存储实体的数据、将被写在逻辑存储实体中的存储块数据和空任务列表 (块 610)。

[0301] 在一些情况下, 每个工作可被分配世代号。这样的世代号可以是唯一的顺序 (或任何其它排序的值) 标识符, 其可由各种插件和资源使用, 以便解决冲突并处理有故障的情况。例如, 可假设第一工作 (FT) 在第二冲突工作 (ST) 之前被发出, 以及 ST 被接收用于首先处理。在这样的情况下, 执行模块 350 可配置成检查 FT 的世代号是否比 ST 的世代号早, 且在这样的情况下, 执行模块 350 可配置成不盖写以前根据 ST 更新的数据。

[0302] 工作创建模块 345 还可配置成取回与操作将被执行的逻辑存储实体相关的 SLS (块 615), 并相应地将相关的任务引入到与工作相关的任务列表。因此, 工作创建模块 345 可配置成根据 SLS 来检查压缩是否是需要的 (块 620), 且如果是, 工作创建模块 345 可配置成将相关的任务 (例如, 压缩数据) 添加到任务列表 (块 625)。工作创建模块 345 还可配置成根据 SLS 来检查加密是否是需要的 (块 630), 且如果是, 工作创建模块 345 可配置成将相关的任务 (例如, 加密数据) 添加到任务列表 (块 635)。

[0303] 假设根据 SLS 只有两个待执行的任务, 工作创建模块 345 成功地创建新工作且新工作准备执行 (块 640)。

[0304] 应注意, 参考图 7, 一些块可整合成合并的块或可分成几个块, 和 / 或其它块可被添加。此外, 在一些情况下, 可以用与本文所述的不同的顺序执行块。也应注意, 虽然也参考实现它们的系统元件描述了流程图, 这决不是限制性的, 且块可由除了本文所述的元件以外的元件执行。

[0305] 在关于工作及其创建的简要解释之后, 注意图 8, 其示出了说明根据当前公开的主题的某些例子的为了管理 UDSP 代理所接收的工作而执行的一序列操作的流程图。

[0306] 在一些情况下, UDSP 代理 220 的工作管理模块 335 可配置成接收工作 (块 405)。应注意, 工作可从客户端服务器 218 (例如, 直接地或通过可特别充当协议转换器的网关资源 216)、从网关资源 216、从另一计算机节点 205、从外部实体 (例如, 应用等) 或从任何其它源接收。

[0307] 在接收到工作之后, 工作管理模块 335 可配置成取回指示 DSS200 的资源 (例如, 计算机节点和 / 或存储相关资源等) 的全部或部分的动态行为的数据的全部或部分 (块 410)。

[0308] 在一些情况下, 工作管理模块 335 可配置检查工作是否与 SLS 相关 (例如, 工作与特定的逻辑存储实体有关, 等等) (块 412), 且如果是, 则取回与相关于工作的逻辑存储实体相关的 SLS (例如, 从 UDSP 数据存储库 330, 或者如果在 UDSP 数据存储库 330 中不是可用的则从另一计算机节点的 UDSP 数据存储库, 等等) (块 413)。

[0309] 工作管理模块 335 可配置成利用基于目标的路由模块 395 以为 DSS 200 的计算机节点 205 中的一个或多个的适合性评分以执行一个或多个未决的工作任务 (块 415)。

[0310] 未决的工作任务是在其执行之前没有未完成的先决条件的任务。例如, 压缩任务可取决于重复删除任务的以前执行, 加密任务可取决于压缩任务的以前执行, 等等。

[0311] 计算机节点 205 执行未决的工作任务的适合性以及因而其分数可例如取决于其资源 (例如, 其处理能力), 包括其存储相关资源, 和 / 或在工作涉及逻辑存储实体的情况下, 取决于它们满足一个或多个 SLS 要求 (例如, 具有能够用于执行在这样的逻辑存储实体的范围中的一个或多个工作任务的资源) 的能力, 如果这样的要求存在, 和 / 或取决于它们的动态行为和当前状态, 等等。关于图 9 提供了分级过程的更详细的描述。

[0312] 基于所计算的分数, 工作管理模块 335 可配置成利用基于目标的路由模块 395 来按照评分结果将工作路由到例如更合适的计算机节点 205, 且有时路由到最合适的计算机节点 (例如, 工作可被路由到具有最高分数的计算机节点 205) (块 420)。

[0313] 任务管理模块 335 可配置成检查任务是否被路由到另一计算机节点 (块 425)。如果任务被路由到另一计算机节点, 则与本地计算机节点 205 (例如, 运行该过程的计算机节点 205) 有关的过程结束 (块 440)。然而, 如果本地计算机节点 205 是最适当的计算机节点,

则一个或多个未决的工作任务可在本地计算机节点 205 上例如通过利用 UDSP 代理 220 的执行模块 350 来执行(块 430)。

[0314] 应注意,在一些情况下,不是本地计算机节点 205 能够执行的所有未决的工作任务都被它执行,而是仅仅它被选择为最适当的计算机节点的未决的工作任务。因此,例如,如果工作包括未决的工作任务,则其中两个工作任务可由本地计算机节点 205 执行:它有最高分数的工作任务和它没有最高分数的工作任务——与本地计算机节点 205 相关的 UDSP 代理 220 可配置成只执行本地计算机节点 205 具有最高分数的任务。应进一步注意,本地计算机节点 205 的 UDSP 代理 220 在一些情况下可利用本地计算机节点 205 的多于一个处理资源(如果这样的资源存在),用于一个或多个任务的并行和 / 或同时处理。在一些情况下,对于多于一个任务的这样的并行和 / 或同时处理,本地计算机节点 205 可利用远程处理资源(例如,与一个或多个远程计算机节点 205 相关的处理资源)。特别关于图 10 提供了任务执行的更详细的描述。

[0315] 工作任务模块 335 还可配置成检查在任务的执行之后在本地计算机节点 205 上额外的任务是否存在和 / 或在本地计算机节点 205 上的任务的执行是否触发一个或多个新工作和 / 或任务的创建(例如,复制任务可导致多个写工作的产生,每个写工作预定在不同的位置)(块 435)。如果否,过程结束(块 440)。如果是,过程返回到块 405,其中具有剩余任务的工作和 / 或一个或多个新工作由与本地计算机节点 205 相关的 UDSP 代理 220 接收,且管理每个工作的过程开始。

[0316] 在一些情况下,例如通过将一个或多个互连计算机节点 205 添加到基础设施层,通过从基础设施层移除一个或多个计算机节点 205,通过修改基础设施层的一个或多个现有的计算机节点 205(例如,添加处理资源 310 和 / 或连接到其的其它存储相关资源,移除处理资源 310 和 / 或连接到其的其它存储相关资源)等,可更新基础设施层。在一些情况下,对基础设施层的这样的变化可被动态地执行(例如,每当用户期望时),包括在 DSS 200 的操作期间。

[0317] 工作管理模块 335 在一些情况下可配置成利用基于目标的路由模块 395 来对被添加或修改的一个或多个更新的基础设施层计算机节点 205 的适合性评分,以执行下面的工作的一个或多个未决的工作任务。在一些情况下,更新的基础设施层可在这样的评分计算期间被创建,且计算可关于更新的基础设施层的一个或多个计算机节点 205 来执行。在一些情况下,计算可关于更新的基础设施层的一个或多个计算机节点 205 来执行。

[0318] 工作管理模块 335 在一些情况下可配置成执行下面的工作的所述未决的任务中的一个或多个,或基于所计算的分数将所述下面的工作路由到更新的基础设施层的更适当的计算机节点 205 (且在一些情况下路由到最适当的计算机节点 205)。

[0319] 应注意,参考图 8,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0320] 如本文详述的,工作管理模块 335 可配置成利用基于目标的路由模块 395 来对 DSS 200 的计算机节点 205 中的一个或多个的稳定性评分以执行未决的工作任务。注意图 9,其示出根据当前公开的主题的某些例子的为了给节点适合性评分以执行未决的工作工作而

执行的一序列操作的流程图。

[0321] 评分过程 700 可例如通过基于目标的路由模块 395 接收下列项中的至少一个来开始：待执行的任务、指示 DSS 200 的资源（包括计算机节点和 / 或存储相关资源等）的全部或部分的动态行为的数据、或可由评分过程使用的任何其它数据（块 710）。在一些情况下，当工作与特定的逻辑存储实体相关时，基于目标的路由模块 395 也可接收与相关于工作的逻辑存储实体相关的 SLS。

[0322] 基于目标的路由模块 395 可配置成对一个或多个计算机节点 205 的稳定性评分以执行每个未决的工作任务（块 720）。评分可特别基于所接收的数据来执行。

[0323] 应注意，可对连接到 DSS 200 的每个计算机节点 205 或只对计算机节点 205 中的一些（例如根据网络拓扑、离本地计算机节点 205 的地理距离、随机地和 / 或确定地选择计算机节点 205，直到适合于执行一个或多个未决的工作任务的足够数量的计算机节点 205 被找到，等等）计算评分。应进一步注意，各种评分算法可用于对计算机节点 205 执行未决的工作任务的适合性评分。仍然进一步注意，评分过程可包含和 / 或使用启发法和 / 或近似法。此外或可选地，评分可基于部分和 / 或不是最新的信息。

[0324] 在一些情况下，对于分数将被计算的每个计算机节点 205，基于目标的路由模块 395 可配置成对每个未决的工作任务检查计算机节点 205 是否可执行未决的工作任务。在工作与逻辑存储实体相关的情况下，基于目标的路由模块 395 也可检查计算机节点 205 是否可执行未决的工作任务，同时满足由相应的 SLS 定义的要求。在计算机节点 205 不能执行未决的工作任务（或当相关的时不能满足由 SLS 定义的要求）的情况下，该节点的分数将低于能够执行未决的工作任务的计算机节点 205 的分数（当相关的时满足由 SLS 定义的要求）。在一些情况下，也基于与连接到相应的计算机节点 205 的一个或多个存储相关资源有关的参数（例如，与存在和 / 或负荷和 / 或可用性和 / 或故障和 / 或能力和 / 或响应时间和 / 或连接性和 / 或与存储相关资源相关的成本有关的参数的数据）和这样的存储相关资源执行未决的工作任务的能力（同时当相关的时满足由 SLS 定义的要求）来计算分数。

[0325] 以示例性的方式且为了理解的容易，不能执行未决的工作任务（同时当相关的时满足 SLS 所定义的要求）的计算机节点 205 的分数为零，而能够执行未决的工作任务（同时当相关的时满足 SLS 所定义的要求）的计算机节点 205 的分数大于零。

[0326] 应注意，在一些情况下，所计算的分数可由非标量值例如由多维值表示。应进一步注意，所计算的分数可以不属于有序集合。仍然应进一步注意，适当的节点和 / 或最适当的节点的决定（例如，哪个分数“较高”的决定）可以是任意的（例如，当分数不属于有序集合时，等等）。

[0327] 在一些情况下，如果本地计算机节点 205 执行任务的适合性将与一个或多个远程计算机节点 205 的适合性相同，如果它们都具有将工作传递到其的相同传递成本，则由于与将工作传递到任何远程计算机节点 205 相关的成本，本地计算机节点 205 的分数将较高。

[0328] 在一些情况下，对于分数将被计算的每个计算机节点 205，基于目标的路由模块 395 可配置成基于对每个未决的工作任务计算的分数来计算总分数（块 730）。这样的总分数可例如是计算机节点 205 的任务分数的总和、计算机节点 205 的任务分数的平均数或基于所计算的计算机节点 205 的任务分数的任何其它计算。

[0329] 应注意，参考图 9，一些块可整合成合并的块或可分成几个块，和 / 或其它块可被

添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0330] 转到图 10,其示出根据当前公开的主题的某些例子的为了在计算机节点上执行未决的工作任务而执行的一序列操作的流程图。

[0331] 如本文详述的,工作任务模块 335 可配置成利用用于执行任务执行过程 800 的执行模块 350,任务执行过程 800 用于执行一个或多个未决的工作任务。在这样的情况下,执行模块 350 可配置成执行一个或多个未决的工作任务(块 810)。

[0332] 如本文所示的,应注意,在一些情况下,不是本地计算机节点 205 能够执行的所有未决的工作任务都由它执行,而是仅仅它被选择的未决的工作任务。此外,应进一步注意,与本地计算机节点 205 相关的 UDSP 代理 220 在一些情况下可利用用于一个或多个任务的并行和/或同时执行的多于一个处理资源(如果这样的资源存在)。在一些情况下,对于多于一个任务的这样的并行和/或同时执行,本地计算机节点 205 可利用远程处理资源(例如,与一个或多个远程计算机节点 205 相关的处理资源)。

[0333] 在一个或多个未决的工作任务的执行之后,执行模块 335 可配置成更新所执行的任务的状态以指示任务已被执行(块 820)。

[0334] 在一些情况下,任务可部分地被执行或其执行可能失败。在这样的情况下,执行模块 335 可配置成用相关的指示更新任务状态。在一些情况下,状态也可包含执行结果的数据。

[0335] 在一些情况下,执行模块 335 可配置成检查是否有检查当前 DSS 200 的配置(特别包括资源可用性和分配)的需要(块 830)。例如在与逻辑存储实体相关的一个或多个被执行的任务的执行不满足(或接近于不满足,例如根据预定的阈值等)相应的 SLS 要求的情况下,和/或如果一个或多个任务执行失败和/或如果任务的执行导致与计算机节点 205 和/或连接到其的资源有关的参数的数据的变化超过预定或计算的阈值(例如存储空间或任何其它资源的不足,等等)和/或由于任何其它原因,这样的需要可存在。

[0336] 在需要检查 DSS 200 的当前配置的情况下,执行模块 335 可配置成推荐与一个或多个计算机节点 205 相关的 UDSP 代理 220 以检查重新配置是否是需要的(块 840)。应注意,在一些情况下,推荐可由与计算机节点 205(其上一个或多个任务被执行)相关的 UDSP 代理 220 的基于目标的配置模块 390 处理。在其它情况下,推荐可被发送到与可负责执行重新配置过程的一个或多个计算机节点 205(例如专用计算机节点)相关的 UDSP 代理 220。特别关于图 11 在本文提供了关于重新配置检查的进一步的解释。

[0337] 在不需要检查 DSS 200 的当前配置的情况下或在检查是否需要重新配置的推荐之后,执行模块 335 可配置成检查一个或多个未决的工作任务(例如,与已被执行的工作相关的所有任务)的下面执行是否完成(块 850)。

[0338] 在工作未完成的情况下,过程结束(块 860)。如果工作完成,则执行模块 335 可配置成检查指示工作完成的任何通知(例如,对工作发起者的通知等)是否是需要的(块 870)。如果没有通知是需要的,则过程结束(块 860)。如果通知是需要的,则执行模块 335 可配置成根据需要发出工作执行的通知(块 880),且过程结束(块 860)。

[0339] 根据当前公开的主题的一些例子,对于每个所需的通知,发送所需通知的专用任

务可在例如本文所述的工作创建过程期间创建。在这样的情况下,可选地,块 850-880 可被忽略。

[0340] 应注意,参考图 10,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0341] 现在注意图 11,其示出根据当前公开的主题的某些例子的为了管理 DSS 的重新配置而执行的一序列操作的流程图。

[0342] 根据当前公开的主题的一些例子,在一些情况下,可执行检查 DSS200 的重新配置是否是需要的重新配置过程 900。在一些情况下,这样的检查可被周期性地(例如,根据预定的时间间隔,例如每分钟、每五分钟、每小时或任何其它预定的时间间隔)、连续地(例如,在重复的循环中等)、在触发事件(例如,所监测的参数超过预定的或计算的阈值,来自与计算机节点 205 相关的 UDSP 代理 220 的推荐的接收,如果特别关于图 10 详述的,等等)之后等被执行。

[0343] 如果本文所示的,在一些情况下,例如当利用基于目标的配置模块 390 时,与计算机节点 205 相关的每个 UDSP 代理 220 可配置成执行推荐过程 900。在一些情况下,例如当利用基于目标的配置模块 390 时,与一个或多个计算机节点 205(例如,专用计算机节点)相关的 UDSP 代理 220 可负责执行重新配置过程 900。

[0344] 在一些情况下,基于目标的配置模块 390 可配置成接收与 DSS 200 中的一个或多个逻辑存储实体相关的 SLS、指示 DSS 200 及其资源和环境的动态行为的数据、指示 DSS 200 的当前配置的数据、与 DSS 200 有关的统计数据 and 历史数据等中的任一个或任何组合(块 910)。应注意,在一些情况下,数据的全部或部分可此外或可选地从与计算机节点 205(其上重新配置过程 900 被执行)相关的 UDSP 数据存储库 330 取回。

[0345] 在一些情况下,基于目标的配置模块 390 可配置成利用用于检查任何 SLS 是否被违反(或接近于被违反,例如根据预定的阈值等)和 / 或是否有用于执行 DSS 200 的重新配置的任何其它原因(例如,未能执行一个或多个任务,而不考虑 SLS 等)的所接收的数据(块 920)。

[0346] 应注意,虽然在一些情况中,每当 SLS 被违反时(应注意,SLS 的违反有时可包括接近这样的违反,例如根据预定的阈值等),DSS 200 的重新配置可被发起,在其它情况下 DSS 200 的这样的重新配置可根据满足某个预定的标准而被发起。这样的标准可例如是预定数量的探测到的所需 SLS 违反被满足,在预定的时间帧内或不考虑时间,等等。因此例如,示例性标准可以是三个 SLS 违反的探测或在一天内三个 SLS 违反的探测,等等。在一些情况下,违反的重要性可此外或可选地被考虑为标准。为了这个目的,基于目标的配置模块 390 可配置成利用与 DSS 200 有关的统计数据 and 历史数据。

[0347] 在需要重新配置 DSS 200 的情况下,基于目标的配置模块 390 可配置成启动用于执行 DSS 200 的配置过程的基于目标的管理系统(OBMS) 100,如上面特别关于图 2-4 详述的(块 930)。应注意,如本文所示的,在 DSS 200 的重新配置的情况下,OBMS 100 可接收 DSS 200 的当前配置作为配置过程的输入的部分并在重新配置 DSS 200 时考虑它。在一些情况下,在这样的重新配置期间,OBMS 100 可配置成保留和 / 或分配和 / 或重新分配和 / 或释

放全部或部分资源。

[0348] 如果没有 SLS 被违反(或接近于违反)且没有用于执行重新配置的其它原因,或在 DSS 200 的重新配置的发起之后,重新配置过程 900 结束(块 940)。

[0349] 应注意,参考图 11,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0350] 现在注意图 12,其示出根据当前公开的主题的某些例子的为了监测计算机节点和连接到其的资源的本地参数而执行的一序列操作的流程图。

[0351] 在一些情况下,本地参数监测模块 360 可配置成监测计算机节点 205 和 / 或连接到其的存储相关资源的各种参数(块 1010)。如本文所示的,所监测的参数可以是指示存在和 / 或负荷和 / 或可用性和 / 或故障和 / 或能力和 / 或响应时间和 / 或连接性和 / 或成本(例如,网络链接的成本、不同类型的数据存储资源)的任何参数和 / 或指示计算机节点 205 和 / 或连接到其的任何存储相关资源的动态行为的任何其它参数和 / 或与计算机节点 205 和 / 或连接到其的一个或多个存储相关资源有关的任何其它数据。在一些情况下,本地参数监测模块 360 可配置成监测客户端服务器 218 和 / 或网关资源 216 的各种参数,加以必要的变更。

[0352] 应注意,这样的监测可被周期性地(例如,根据预定的时间间隔,例如每分钟、每五分钟、每小时或任何其它预定的时间间隔)、连续地(例如,在重复的循环中,等等)、在触发事件(例如,新资源连接到计算机节点 205,等等)之后等被执行。

[0353] 在一些情况下,本地参数监测模块 360 可配置成检查新参数或任何所监测的参数的值中的变化是否被探测到(块 1020)。如果否,本地参数监测模块 360 可配置成继续监测参数。然而如果新参数或任何所监测的参数的值中的变化被探测到,本地参数监测模块 360 可配置成传播(例如,当利用多播模块 340 时)指示对一个或多个本地参数的改变的通知。在一些情况下,这样的通知可被发送到一个或多个计算机节点 205 和 / 或客户端服务器 218 和 / 或网关资源 216 (例如,通过单播 / 多播 / 重播传输)(块 1030)。

[0354] 应注意,在一些情况下,本地参数监测模块 360 可配置成以不同的预定时间周期或响应于各种触发事件而发送可包括各种指示(例如,不同组的一个或多个本地参数的指示等)的各种类型的通知。应进一步注意,一些通知可选择性地被发送到例如被注册来接收这样的通知的一个或多个计算机节点 205。

[0355] 在一些情况下,本地参数监测模块 360 可配置成在 UDSP 数据存储库 330 中更新参数值且在一些情况下此外或可选地更新其导数(例如,与参数有关的各种统计数据)(块 1040)。

[0356] 在一些情况下,本地参数监测模块 360 可配置成检查是否需要检查当前 DSS 200 的配置。在所监测的参数之一超过与其相关的预定的或计算的阈值的情况下和 / 或由于任何其它原因,这样的需要可例如存在。

[0357] 在需要检查 DSS 200 的当前配置的情况中,本地参数监测模块 360 可配置成推荐与一个或多个计算机节点 205 相关的 UDSP 代理 220 以检查重新配置是否是需要的。应注意,在一些情况下,推荐可由与本地计算机节点 205 (在其上本地参数监测模块 360 正运行)

相关的 UDSP 代理 220 的基于目标的配置模块 390 处理。在其它情况下,推荐可被发送到与可负责执行重新配置过程的一个或多个计算机节点 205 (例如,专用计算机节点)相关的 UDSP 代理 220。特别关于图 11 在本文提供了关于重新配置检查的进一步的解释。

[0358] 应注意,参考图 12,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0359] 现在注意图 13,其示出根据当前公开的主题的某些例子的为了探测和管理连接到计算机节点的资源而执行的一序列操作的流程图。

[0360] 在一些情况下,资源探测和管理模块 385 可配置成执行探测和管理过程 1200。在一些情况下,资源探测和管理模块 385 可配置成扫描连接到一个或多个计算机节点 205 的存储相关资源(块 1210)。在一些情况下,资源探测和管理模块 385 可配置成连续和 / 或周期性地(例如,每预先确定的时间段,例如每分钟、每五分钟、每小时等)等执行扫描。在一些情况下,扫描可由用户(例如,系统管理员等)发起。

[0361] 资源探测和管理模块 385 可配置成检查任何新的存储相关资源是否被找到(块 1220)。如果没有新的存储相关资源被找到,则资源探测和管理模块 385 可配置成继续扫描存储相关资源。如果一个或多个新的存储相关资源被找到,则存储相关资源探测和管理模块 385 可配置成检查在一个或多个插件中是否有对使用存储相关资源的需要,且如果是,则检查是否插件在本地存在(例如,在新资源附接 / 连接到的计算机节点 205 上)(块 1230)。

[0362] 如果没有对一个或多个插件的需要且它们都在本地存在,则资源探测和管理模块 385 可配置成使插件与新存储相关资源相关,且存储相关资源可被添加到本地资源池(块 1240)。

[0363] 如果有对在本地不存在的一个或多个插件的需要,则资源探测和管理模块 385 可配置成检查一个或多个遗漏的插件是否存在于例如一个或多个计算机节点 205 和 / 或客户端服务器 218 和 / 或网关资源 216 上(例如当利用多播模块 340 时)和 / 或在如本文详述的共享虚拟软件扩展库中(块 1250)中和 / 或在 DSS 200 的任何其它位置上和 / 或在任何辅助实体上。

[0364] 如果资源探测和管理模块 385 找到所需的插件,则资源探测和管理模块 385 可配置成使插件与新存储相关资源相关,且存储相关资源可被添加到本地资源池(块 1240)。

[0365] 在一些情况下,如果资源探测和管理模块 385 未找到所需的插件,资源探测和管理模块 385 可配置成发出一个或多个插件请求。这样的插件请求在一些情况下可被发送给用户(块 1270),因而使这样的用户能够将相关的插件添加到 DSS 200 (例如,在购买它、从互联网下载它之后,等等)。在发送这样的请求之后,资源探测和管理模块 385 可配置成继续扫描存储相关资源(块 1210)。

[0366] 应注意,在一些情况下,直到所需的插件被找到、取回(如果需要)和安装为止,新存储相关资源可被标记为每当对存储相关资源的扫描被执行时识别的新存储相关资源,且因而在本文详述的过程重复,直到所需的插件被找到为止。

[0367] 在一些情况下,资源探测和管理模块 385 可此外或可选地配置成检查存储相关资源移除是否在对存储相关资源的扫描之后被探测到(块 1280)。在这样的情况下,如果存储

相关资源移除被探测到,则资源探测和管理模块 385 可配置成从本地资源池移除存储相关资源,且可选地,清理不再被需要的任何插件(例如,根据利用这样的插件的资源被移除的事实)(块 1290)。

[0368] 应注意,在一些情况下,资源探测和管理模块 385 可此外或可选地配置成对连接到一个或多个客户端服务器 218 和 / 或网关资源 216/ 与一个或多个客户端服务器 218 和 / 或网关资源 216 断开的存储相关资源执行探测和管理过程 1200,加以必要的变更。应进一步注意,资源探测和管理模块 385 的利用可实现存储相关资源到计算机节点 205 和 / 或客户端服务器 218 和 / 或网关资源 216 的无缝添加和 / 或移除和 / 或附接和 / 或分离(例如,“即插即用”),包括在 DSS 200 的操作期间,且在一些情况下不由用户执行任何管理行动(特别包括任何预备管理行动)。

[0369] 应进一步注意,在一些情况下,存储相关资源到本地资源池的添加 / 从本地资源池的移除可导致计算机节点 205 的所监测的本地参数的变化(例如,各种本地参数的添加和 / 或移除和 / 或更新和 / 或任何其它变化)。如本文所示的,当新的参数被探测到时,在一些情况下,适当的通知可被本地参数监测模块 360 发送,如在本文特别关于图 12 描述的。应注意,在一些情况下,这样的通知可触发重新配置。

[0370] 应注意,参考图 13,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0371] 现在注意图 14,其示出根据当前公开的主题的某些例子的为了将新计算机节点连接到分布式存储系统(DSS)而执行的一序列操作的流程图。

[0372] 在一些情况下,当包括 UDSP 代理 220 的新计算机节点 205 连接到网络时,新计算机节点 205 的云即插即用模块 380 可配置成探测新网络连接和 / 或对现有的网络连接的变化(例如,计算机节点 205,在其上云即插即用模块 380 连接到新网络或不同的网络)(块 1305)。在新网络连接的探测之后,云即插即用模块 380 可配置成例如通过利用多播模块 340 发送(例如,通过单播 / 多播 / 重播传输)发现消息(块 1310)。这样的发现消息可触发任何接收计算机节点 205 以例如通过发送包括至少 DSS 200 的标识符的答复来作出响应(每个 DSS 200 可具有实现其识别的唯一标识符)。

[0373] 云即插即用模块 380 可配置成听在预定的时间间隔(例如,可使接收计算机节点 205 能够对发现消息作出响应的的时间间隔)内接收的任何响应并检查任何响应是否被接收(块 1315)。如果没有响应被接收到且计算机节点 205 未连接 DSS 200,则云即插即用模块 380 可配置成重复块 1310 并重新发送发现消息。

[0374] 如果响应被接收到,则云即插即用模块 380 可配置成检查响应是否涉及单个 DSS 200(例如,根据所接收的 DSS 200 标识符)(块 1320)。如果是,云即插即用模块 380 可配置成将计算机节点 205 连接到所探测的 DSS 200(块 1325)。应注意,作为连接 DSS 200 的结果,计算机节点 205 可自动开始发送和接收各种通知,如本文所详述的。

[0375] 如果多于一个 DSS 200 被探测到(例如,多于一个 DSS 200 的标识符被接收到,作为对发现消息的响应),则云即插即用模块 380 可配置成检查默认 DSS 200 是否存在(块 1330)。为了这个目的,在一些情况下,默认 DSS 200 的指示可从本地注册库(例如,在本地

网络上可访问的数据存储库)、从域名系统(例如,在预定的 DNS 记录下等)等取回。在一些情况下,默认 DSS 200 的指示可由响应计算机节点 205 之一发送,响应计算机节点 205 的响应可包括默认 DSS 200 的指示。应注意,也可使用用于识别默认 DSS 200 的其它方法和技术。

[0376] 如果这样的默认 DSS 200 存在,则云即插即用模块 380 可配置成将计算机节点 205 连接到默认 DSS 200 (块 1325)。如果没有默认 DSS 200 被检测到,则新计算机节点 205 的指示可被提供给用户用于选择新计算机节点 205 将连接到的 DSS 200,且云即插即用模块 380 可配置成等待这样的选择(块 1335)。一旦选择被作出,云即插即用模块 380 就可配置成将计算机节点 205 连接到选定的 DSS 200 (块 1325)。

[0377] 在一些情况下,当检测到新网络连接(块 1305)时,云即插即用模块 380 可此外或可选地配置成例如在预定的网络地址上和 / 或在目录服务(例如,DNS、活动目录等)上查找本地注册库(例如,在本地网络上可访问的数据注册库)和 / 或全局注册库(例如,在互联网上可访问的数据存储库)注册服务(块 1340)。这样的注册服务可特别实现可用的 DSS 200 和 / 或默认 DSS 200 的识别。

[0378] 云即插即用模块 380 可配置成检查本地注册库是否被找到(块 1345),且如果是,它可配置成在本地注册库上注册(如果它不是已经被注册)(块 1355)。这样的注册可包括存储与注册库中的本地计算机节点 205 有关的各种配置参数。云即插即用模块 380 还可配置成检查本地注册库所定义的政策是否允许全局注册(块 1355)。如果是,或在没有本地注册库被找到的情况下,云即插即用模块 380 可配置成检查全局登记库是否被找到(块 1360)。如果是,云即插即用模块 380 可配置成在全局注册库上注册(如果它不是已经被注册)(块 1365)。这样的注册可包括存储与注册库中的本地计算机节点 205 有关的各种配置参数。

[0379] 在全局注册库上注册之后或在本地注册库所定义的政策不允许全局注册的情况下,云即插即用模块 380 可配置成跳到块 1320 并从那里继续。

[0380] 应注意,可使用其它方法,以便自动和手动地将新计算机节点 205 连接到 DSS 200,且本文提供的方法仅仅是例子。

[0381] 应注意,云即插即用模块 380 可使计算机节点 205 能够在任何时间——包括在 DSS 200 的操作期间——从网络无缝地添加和 / 或移除和 / 或附接和 / 或分离,且在一些情况下不由用户执行任何管理行动(特别包括任何预备管理行动),假定 UDSP 代理 220 安装在计算机节点 205 上(UDSP 代理 220 的详细描述在本文被提供)。应进一步注意,可选地,在一个或多个计算机节点 205 从网络的添加和 / 或移除和 / 或附接和 / 或分离之后,没有用户需要实现 DSS 200 的连续操作。

[0382] 应注意,参考图 14,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0383] 现在注意图 15,其示出根据当前公开的主题的某些例子的为了从远程计算机节点接收通知并相应地更新统一分布式存储平台(UDSP)数据库而执行的一序列操作的流程图。

[0384] 在一些情况下,计算机节点 205 的 UDSP 代理 220 的远程节点参数监测模块 370 可配置成接收起源于其它计算机节点 205 和 / 或客户端服务器 218 和 / 或网关资源 216 和 /

或用户等的各种通知(一般通知和 / 或起源于计算机节点 205 注册一般从其接收消息的源的通知)(块 1410)。

[0385] 在一些情况下,远程节点参数监测模块 370 可配置成相应地更新 UDSP 数据存储库 330 (块 1420)。

[0386] 应注意,存储在 UDSP 数据存储库 330 中的这样的数据可被使用,以便在本地维持 DSS 200 或其部分的状态的知识(例如,其动态行为等),其对于计算机节点 205 所执行的过程是相关的,如本文详述的。

[0387] 应注意,参考图 15,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0388] 在描述了 DSS 200 之后,接下来是用于管理 DSS 200 中的高速缓存资源的系统和方法的描述。如本文所示的,基础设施层 201 可包括一个或多个且在一些情况下两个或多个计算机节点 205。基础设施层 201 还可包括一个或多个高速缓存资源 212 和 / 或可用作高速缓存资源的资源(例如, RAM、DRAM、SSD 213 等)。高速缓存资源 212 和 / 或可用作高速缓存资源的资源可连接到一个或多个计算机节点 205 (例如,直接地、通过网络等)。如上面进一步指示的,每个计算机节点 205 可具有安装在其上的 UDSP 代理 220 (或以另外方式与其相关)。

[0389] 如本文所示的, UDSP 代理 220 可包括高速缓存管理模块 397。高速缓存管理模块 397 可配置成处理各种高速缓存相关操作,特别是对一个或多个高速缓存资源,在其上标准和 / 或专用高速缓存算法、方法和技术正操作。高速缓存管理模块 397 可配置成管理存储在连接到一个或多个计算机节点 205 (其连接到 DSS 200)的一个或多个高速缓存资源上的对象空间的高速缓存映射。

[0390] 图 16 是示意性示出根据当前公开的主题的某些例子的高速缓存管理模块的方框图。

[0391] 根据当前公开的主题的一些例子,高速缓存管理模块 397 可包括下列模块中的一个或多个:本地高速缓存资源管理模块 2510、远程高速缓存监测模块 2520、高速缓存切换模块 2530 和对象请求管理模块 2540。

[0392] 本地高速缓存资源管理模块 2510 可配置成管理计算机节点 205 的本地高速缓存资源,如在本文特别关于图 17 进一步详述的。

[0393] 远程高速缓存资源模块 2520 可配置成监测远程计算机节点 205 的远程高速缓存资源并相应地发出切换推荐和 / 或指令,如在本文特别关于图 18 进一步详述的。

[0394] 高速缓存切换模块 2530 可配置成执行各种高速缓存切换相关过程,如在本文特别关于图 20 和 21 进一步详述的。

[0395] 对象请求管理模块 2540 可配置成管理在切换期间接收的对象相关请求,如在本文特别关于图 22 和 23 进一步详述的。

[0396] 应注意,根据当前公开的主题的一些例子,高速缓存管理模块 397 中的一些或全部可组合并作为单个模块被提供,或作为例子,它们中的至少一个可以以两个或多个模块的形式实现。

[0397] 注意图 17。图 17 是示出根据当前公开的主题的某些例子的为了管理计算机节点的本地高速缓存资源而执行的一序列操作的流程图。

[0398] 根据当前公开的主题的一些例子，本地高速缓存资源管理模块 2510 可配置成监测各种高速缓存相关参数，包括连接到计算机节点 205 的高速缓存资源的参数(块 1510)。

[0399] 高速缓存相关参数 205 可包括相应于计算机节点 205 和 / 或连接到其的资源(包括高速缓存资源)的节点级高速缓存相关参数(例如，负荷参数、性能参数、存在参数、可用性参数、故障参数、能力参数、响应时间参数、连接性参数、成本参数、位置参数等)。高速缓存相关参数可此外或可选地包括关于高速缓存对象(例如，高速缓存在高速缓存资源上的对象)的对象级高速缓存相关参数(例如，高速缓存对象的位置、对象被高速缓存于的介质的类型等)。高速缓存相关参数可仍然此外或可选地包括外部高速缓存相关参数，例如使用高速缓存器的各种实体(例如，客户端服务器 218 等)的参数。

[0400] 本地高速缓存资源管理模块 2510 可仍然进一步配置成接收第三方推荐(例如，来自远程计算机节点 205 的推荐)以执行高速缓存切换(如在下文特别关于图 18 进一步描述的)。应注意，这样的第三方推荐可包括一个或多个高速缓存相关参数的数据，这样的数据在一些情况下可包括指示第三方推荐执行高速缓存切换所根据的原因的高速缓存相关参数。在接收到第三方推荐的情况下，本地高速缓存资源管理模块 2510 可配置成确定这样接收的参数是否相对于本地已知的参数(例如，存储在与计算机节点 205 相关的 UDSP 数据存储库 330 上的参数)是优选的(例如，新的和 / 更更新的参数被接收到)，且如果是，则在下面的块中考虑它们，除了和 / 或代替本地已知的参数。

[0401] 应注意，这样的监测可被周期性地(例如，根据预定的时间间隔，例如每分钟、每五分钟、每小时或任何其它预定的时间间隔)、连续地(例如，在重复的循环中，等等)、在触发事件(例如，存储相关资源——包括高速缓存资源——连接到计算机节点 205/ 与计算机节点 205 断开，等等)之后等被执行。

[0402] 应注意，在一些情况下，高速缓存相关参数的监测可由本地参数监测模块 360 执行，如果特别关于图 12 进一步详述的。

[0403] 在一个或多个所监测的参数的值的变化被探测到和 / 或任何新参数被探测到和 / 或任何参数不再被探测到(例如某些高速缓存资源被移除，等等)和 / 或周期性地(例如，根据预定的时间间隔，例如每分钟、每五分钟、每小时或任何其它预定的时间间隔)被探测的情况下，本地高速缓存资源管理模块 2510 可配置成检查所监测的参数是否指示一个或多个 SLS 所定义的一个或多个高速缓存相关要求的违反。可例如通过评估所监测的参数根据与当前高速缓存在高速缓存资源上的对象和 / 或被映射到高速缓存资源(使用本文详述的任何映射)的对象相关的 SLS 例如根据与这样的对象相关的逻辑存储实体相关的 SLS 来执行这样的检查。

[0404] 高速缓存资源管理模块 2510 还可配置成确定所监测的参数是否满足与这样的高速缓存相关 SLS 相关的一个或多个第一 SLS 标准，例如一个或多个高水印(例如，预定的最大阈值、所计算的最大阈值等)，指示一个或多个这样的 SLS 的接近违反(或违反)。

[0405] 本地高速缓存资源管理模块 2510 可仍然进一步配置成确定所监测的参数是否指示一个或多个阈值和 / 或涉及 DSS 200 或其部分(例如，最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 的违反或接近这样的违反(例如，根据一个或多

个第一 SLS 标准例如一个或多个高水印) (块 1520)。

[0406] 应注意,为了这些目的,本地高速缓存资源管理模块 2510 可配置成从 UDSP 数据存储库 330 取回相关的 SLS。

[0407] 如果存在 SLS 违反或 SLS 接近违反或存在涉及 DSS 200 或其部分的一个或多个参数的违反,则本地高速缓存资源管理模块 2510 可配置成搜索第一本地计算机节点 205 (切换发起者) 能够将一个或多个高速缓存对象空间的所有权(例如,用于处理的责任)转移到的一个或多个切换目标(例如,具有一个或多个高速缓存资源 212 的另一个远程计算机节点 205 和 / 或连接到其的可用作高速缓存资源的一个或多个资源),使得所有高速缓存相关 SLS 和 / 或阈值和 / 或涉及 DSS 200 或其部分(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 将满足后转移(块 1530)。高速缓存相关 SLS 是包含涉及和 / 或影响一个或多个高速缓存对象的要求的任何 SLS 和 / 或包含这样的对象和 / 或与其相关的高速缓存资源的一个或多个计算机节点 205 和 / 或与这样的计算机节点 205 相关的任何实体。

[0408] 应注意,切换发起者可负责处理与一个或多个逻辑存储实体有关的一个或多个对象空间。一个或多个高速缓存对象空间或其部分的切换(所有权的转移)导致切换目标接收所转移的一个或多个高速缓存对象空间或其部分的所有权。

[0409] 在一些情况下,本地高速缓存资源管理模块 2510 可进一步配置成搜索本地计算机节点 205 (切换发起者) 能够将一个或多个高速缓存对象空间的所有权转移到的切换目标,在一些情况下使得一个或多个第二 SLS 标准例如与一个或多个高速缓存相关 SLS 相关的切换目标的中间水印(例如,预定的中间阈值、所计算的中间阈值等)被满足。

[0410] 本地高速缓存资源管理模块 2510 还可配置成检查一个或多个切换目标是否被找到(块 1540)。如果否,在一些情况下,本地高速缓存资源管理模块 2510 可配置成向用户报告有不足的高速缓存资源(不足的资源错误),且可选地向用户推荐将被执行的用于允许高速缓存资源满足一个或多个高速缓存相关 SLS 的要求(例如,添加高速缓存资源,等等)的行动。

[0411] 如果一个或多个切换目标被找到,则本地高速缓存资源管理模块 2510 可配置成发起与一个或多个所探测的切换目标的切换过程(块 1560),如特别关于图 20 进一步详述的。在一些情况下,在切换过程的发起之后,本地高速缓存资源管理模块 2510 可配置返回到块 1510 并继续监测高速缓存参数。

[0412] 应注意,在一些情况下,作为非限制性的例子,可通过随机地选择(或根据任何其它规则选择,等等)在块 1530 中找到的一个或多个可能的切换目标将来执行一个或多个切换目标的选择。作为另一非限制性的例子,可通过操作任何类型的分级算法以给在块 1530 中找到的一个或多个可能的切换目标的适合性分级并选择最适合的一个切换目标来执行一个或多个切换目标的选择。

[0413] 如果在监测关于连接到本地计算机节点 205 的高速缓存资源的高速缓存相关参数之后和 / 或接收到第三方推荐以执行高速缓存切换之后没有 SLS 违反或没有第一个或多个第一 SLS 标准(例如,高水印)的交叉,则本地高速缓存资源管理模块 2510 还可配置成检查所监测的高速缓存相关参数中的任一个是否满足一个或多个第三 SLS 标准,例如与一个或多个高速缓存相关 SLS 和 / 或阈值和 / 或涉及 DSS 200 或其部分(例如,最大允许的站点级

过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 相关的低水印(例如,预定的最小阈值、所计算的最小阈值等),且因此它被认为是未充分利用的(块 1570)。如果本地计算机节点 205 的高速缓存资源未被充分利用,则本地高速缓存资源管理模块 2510 可配置成返回到块 1510 并继续监测高速缓存参数。

[0414] 如果一个或多个本地计算机节点 205 的高速缓存资源或其部分未被充分利用,则本地高速缓存资源管理模块 2510 可配置成搜索本地计算机节点 205 能够将一个或多个高速缓存对象空间的所有权转移到的一个或多个切换目标,使得所有高速缓存相关 SLS 和 / 或阈值和 / 或涉及 DSS 200 或其部分(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 将在转移后被满足(块 1580)。在一些情况下,本地高速缓存管理模块 2510 还可配置成搜索本地计算机节点 205 能够将一个或多个高速缓存对象空间的所有权转移到的切换目标,使得切换目标的第二 SLS 标准例如与一个或多个高速缓存 SLS 相关的中间水印(例如,预定的中间阈值、所计算的中间阈值等)被满足。

[0415] 应注意,在一些情况下,作为非限制性的例子,可通过随机地选择(或根据任何其它规则选择,等等)在块 1530 中找到的一个或多个可能的切换目标将来执行一个或多个切换目标的选择。作为另一非限制性的例子,可通过操作任何类型的分级算法(例如促进高速缓存对象映射的合并的算法,等等)以给在块 1530 中找到的一个或多个可能的切换目标的适合性分级并选择最适合的切换目标来执行一个或多个切换目标的选择。

[0416] 应注意,高速缓存对象的这样的转移可导致释放高速缓存资源,其可例如用于 DSS 200 或任何其它实体的其它目的(如果可能)。高速缓存对象的这样的转移也可导致能够切断高速缓存资源,如果它们被全部释放(没有更多的对象被高速缓存在其上,且没有实体正使用它),因而导致功率消耗的减小。

[0417] 本地高速缓存资源管理模块 2510 还可配置成检查一个或多个切换目标是否被找到(块 1590)。如果一个或多个切换目标被找到,则本地高速缓存资源管理模块 2510 可配置成发起与一个或多个所探测的切换目标的切换过程(块 1560),如特别关于图 20 进一步详述的。

[0418] 在一些情况下,在切换过程的发起之后,本地高速缓存资源管理模块 2510 可配置成返回到块 1510 并继续监测高速缓存参数。

[0419] 应注意,参考图 17,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0420] 现在注意图 18,其示出根据当前公开的主题的某些例子的为了管理远程计算机节点的远程高速缓存资源而执行的一序列操作的流程图。

[0421] 根据当前公开的主题的一些例子,远程高速缓存资源监测模块 2520 可配置成监测各种高速缓存相关参数,包括连接到一个或多个远程计算机节点 205 的高速缓存资源的参数(块 1510)。高速缓存相关参数 205 可包括远程计算机节点 205 和 / 或连接到其的资源(包括高速缓存资源)的节点级高速缓存相关参数(例如,负荷参数、性能参数、存在参数、可用性参数、故障参数、能力参数、响应时间参数、连接性参数、成本参数、位置参数等)。高速缓存相关参数可此外或可选地包括关于高速缓存对象的对象级高速缓存参数(例如,高速

缓存对象的位置、对象被高速缓存于的介质的类型等)。高速缓存相关参数可仍然此外或可选地包括外部高速缓存相关参数,例如使用高速缓存器(例如,客户端服务器 218 等)的各种实体的参数(块 1610)。

[0422] 为了这个目的,远程高速缓存资源监测模块 2520 可配置成从 UDSP 数据存储库 330 取回相关的高速缓存相关参数。

[0423] 远程高速缓存资源监测模块 2520 可配置成利用这样的参数以便确定一个或多个切换推荐是否应被发出(块 1620)。

[0424] 应注意,例如,如果所监测的高速缓存相关参数指示一个或多个高速缓存相关 SLS (包含涉及和 / 或影响一个或多个高速缓存对象和 / 或包含这样的对象的一个或多个计算机节点 205 和 / 或与其相关的高速缓存资源和 / 或与这样的计算机节点 205 相关的任何实体的要求的任何 SLS) 和 / 或阈值和 / 或涉及 DSS 200 或其部分(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 的违反或接近这样的违反(例如根据一个或读个第一 SLS 标准,例如一个或多个高水印),则切换推荐应被发出。

[0425] 作为另一例子,在一些情况下,如果所监测的高速缓存相关参数的任一个与一个或多个第三 SLS 标准例如低水印(例如,预定的最小阈值、所计算的最小阈值等)交叉,则切换推荐应被发出,且因此它被认为是未充分利用的,等等。

[0426] 应注意,为了这个目的,远程高速缓存资源监测模块 2520 可配置成从 UDSP 数据存储库 330 取回相关的 SLS。

[0427] 如果切换推荐应被发出,则远程高速缓存资源监测模块 2520 可配置成将这样的通知发出到一个或多个计算机节点 205,一个或多个计算机节点 205 的所监测的参数指示如本文详述的 SLS 违反、接近这样的违反或未充分利用(块 1630)。

[0428] 在一些情况下,远程高速缓存资源监测模块 2520 可配置成此外或可选地提供一个或多个切换指令。切换推荐可能被切换目标 / 发起者拒绝,而切换指令可以是执行高速缓存切换的命令。在一些情况下,这样的切换指令可被发出到一个或多个切换发起者,使它搜索切换目标并发起与其的切换。在一些情况下,这样的切换指令可被发出到一个或多个切换发起者和一个或多个相应的切换目标(与其的切换可被执行),使一个或多个切换发起者发起与相应的一个或多个切换目标的切换。

[0429] 应注意,参考图 18,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0430] 现在注意图 19,其示意性示出根据本发明的示意性实施方案的分配高速缓存资源的各种场景。

[0431] 看图 19,可认识到,具有高速缓存器的计算机节点 2010 (当提到具有高速缓存器的计算机节点时,它指具有连接到其的高速缓存资源的计算机节点 205)在一些情况下可以充当将一个或多个高速缓存对象空间或其部分的所有权转移到具有高速缓存器的计算机节点 2020 的切换发起者,并充当从具有高速缓存器的计算机节点 2030 接收一个或多个高速缓存对象空间或其部分的所有权的切换目标。应注意,在一些情况下,具有高速缓存器的计算机节点 2010 可同时充当切换发起者和切换目标。

[0432] 在一些情况下,具有高速缓存器的计算机节点例如具有高速缓存器的计算机节点 2030 在一些情况下可同时充当执行与多于一个其它切换目标例如与具有高速缓存器的计算机节点 2010 和具有高速缓存器的计算机节点 2040 的切换的切换发起者。

[0433] 在一些情况下,具有高速缓存器的计算机节点例如具有高速缓存器的计算机节点 2040 在一些情况下同时可充当执行与多于一个其它切换发起者例如与具有高速缓存器的计算机节点 2030 和具有高速缓存器的计算机节点 2050 的切换的切换发起者。

[0434] 在一些情况下,具有高速缓存器的计算机节点例如具有高速缓存器的计算机节点 2090 在一些情况下可同时充当执行与切换发起者例如与具有高速缓存器 2095 的计算机节点的切换的切换目标和充当执行与相同的切换发起者例如具有高速缓存器的计算机节点 2095 的切换的切换发起者,高速缓存器的计算机节点 2095 现在充当切换目标。因此,例如具有高速缓存器的计算机节点 2095 可发起与具有高速缓存器的计算机节点 2090 的切换,而具有高速缓存器的计算机节点 2090 可发起与具有高速缓存器的计算机节点 2095 的切换,在一些情况下同时。

[0435] 在一些情况下,某个计算机节点(有或没有连接到其的高速缓存资源)例如计算机节点 205 可配置成推荐具有高速缓存器的一个或多个计算机节点,例如具有高速缓存器的计算机节点 2080 和具有高速缓存器的计算机节点 2070,以发起与一个或多个切换目标的一个或多个切换。

[0436] 在一些情况下,某个客户端服务器例如客户端服务器 218 可配置成推荐具有高速缓存器的一个或多个计算机节点,例如具有高速缓存器的计算机节点 2070,以发起与一个或多个切换目标的一个或多个切换。

[0437] 在一些情况下,某个网关资源例如网关资源 216 可配置成推荐具有高速缓存器的一个或多个计算机节点,例如具有高速缓存器的计算机节点 2080,以发起与一个或多个切换目标的一个或多个切换(例如,具有高速缓存器的计算机节点 2080 正根据这样的推荐发起切换的这样的切换没有在附图中示出)。

[0438] 在一些情况下,具有高速缓存器的某个计算机节点例如具有高速缓存器的计算机节点 2070 和具有高速缓存器的计算机节点 2080 可配置成从一个或多个计算机节点(有或没有连接到其的高速缓存资源)和 / 或客户端服务器和 / 或网关资源例如计算机节点 205、网关资源 216、客户端服务器 218 接收一个或多个推荐,以发起与一个或多个切换目标的一个或多个切换。

[0439] 在一些情况下,具有高速缓存器的某个计算机节点例如具有高速缓存器的计算机节点 2070 可配置成从一个或多个计算机节点(有或没有连接到其的高速缓存资源)和 / 或客户端服务器和 / 或网关资源例如计算机节点 205、客户端服务器 218 接收一个或多个推荐,以发起与一个或多个切换目标的一个或多个切换,并在一些情况下同时充当将一个或多个高速缓存对象空间或其部分的所有权转移到具有高速缓存器的另一计算机节点例如具有高速缓存器的计算机节点 2080 的切换发起者。

[0440] 应注意,如本领域中的任何普通技术人员可认识到的,上面例示的场景仅仅是例子,且没有在图 19 中提供的图解中介绍的其它场景可存在。

[0441] 现在注意图 20,其示出根据当前公开的主题的某些例子的为了由切换发起者执行高速缓存切换而执行的一序列操作的流程图。

[0442] 根据当前公开的主题的一些例子,与切换发起者(具有连接到其的高速缓存资源的计算机节点 205,其将一个或多个高速缓存对象空间或其部分的所有权转移到具有连接到其的高速缓存资源的另一计算机节点 205)相关的高速缓存切换模块 2530 可配置成开始一个或多个高速缓存对象空间或其部分的所有权的切换(块 1710)。

[0443] 为了这个目的,高速缓存切换模块 2530 可配置成创建指示被切换过程影响的每个高速缓存对象空间或其部分的切换后位置的新的切换后高速缓存映射(块 1720)。

[0444] 应注意,每个客户端服务器 218 (或 DSS200 的任何其它用户)可具有指示处理高速缓存对象空间的一个或多个计算机节点 205 的一个或多个本地高速缓存映射,高速缓存对象空间与客户端服务器 218 (或 DSS 200 的任何其它用户)涉及的一个或多个逻辑存储实体有关。可认识到,高速缓存切换相应地需要这样的—个或多个高速缓存映射的更新,如特别参考块 1770 进一步详述的。

[0445] 在一些非限制性的例子中,这样的高速缓存映射可通过使用分割功能的分层结构来描述,分层结构以紧凑的方式反映从切换过程产生的高速缓存对象空间的任何分割、合并和位置改变。这样的分割功能的一些例子是散列函数、划分奇数和偶数寻址的对象等。应注意,这样的高速缓存映射描述方式和这样的分割功能仅仅是例子,且任何其它已知的方法和/或技术可此外或可选地被利用。

[0446] 高速缓存切换模块 2530 还可配置成例如通过设置指示这样的状态的局部标记而进入“进行中的切换”状态(块 1730),并将“切换开始”通知——包括切换后高速缓存映射——发送到切换被请求到的切换目标(具有连接到其的计算机节点 205,其被选择为从切换发起者接收一个或多个高速缓存对象空间或其部分的所有权),指示发起与其的切换过程的其请求(块 1740)。高速缓存切换模块 2530 还可配置成等待(例如,对于预定的或所计算的时间帧等)对来自切换目标的切换请求的响应。应注意,切换目标可接受这样的切换请求,拒绝这样的切换请求,或部分地接受这样的切换请求(例如,接受接收切换发起者试图转移到其的一个或多个高速缓存对象空间的仅仅部分或其部分的所有权,等等)。

[0447] 高速缓存切换模块 2530 也可配置成检查切换请求是否被切换目标接受(块 1750)。如果切换请求被拒绝(或在一些情况下,如果没有响应在预定的时间帧内被接收到,等等),则高速缓存切换模块 2530 可配置成退出“进行中的切换”状态(块 1760)。

[0448] 然而如果请求被接受,则高速缓存切换模块 2530 可配置成将切换后高速缓存映射发送到一个或多个客户端服务器 218 (或 DSS 200 的任何其它相关用户),例如与待转移的一个或多个高速缓存对象空间(例如,其具有对与其相关的逻辑存储实体的访问权)或其部分相关的客户端服务器 218 (块 1770)。应注意,在一些情况下,新的高速缓存映射可被发送到连接到 DSS 200 的每个客户端服务器 218 (或 DSS 200 的任何其它相关用户)。

[0449] 此外,高速缓存切换模块 2530 可配置成将在切换后高速缓存映射中已被映射到切换目标的不脏的高速缓存对象(自从上次保存到与相关的高速缓存对象空间相关的持久性存储器以来未修改的高速缓存对象)的所有权转移到切换目标,例如通过将这样的高速缓存对象发送(例如,通过将其数据传输)到切换目标和/或通过“忘记”它们(例如,通过将它们标记为被删除或从切换发起者的高速缓存资源物理地删除它们)(块 1780)。

[0450] 仍然此外,高速缓存切换模块 2530 可配置成转移在切换后高速缓存映射中已被映射到切换目标的脏高速缓存对象(自从上次保存到与相关的高速缓存对象空间相关的持

久性存储器以来修改的高速缓存对象)的所有权转移到切换目标,通过冲洗它们(将它们保存到与相关的高速缓存对象空间相关的持久性存储器)和/或通过将这样的高速缓存对象发送(例如,通过将其数据传输)到切换目标并接着“忘记”它们(块 1790)。

[0451] 在一些情况下,高速缓存切换模块 2530 还可配置成将“切换完成”通知发送到切换目标,指示切换被执行(块 1795),并退出“进行中的切换”状态(块 1760)。在一些情况下,块 1795 可只在块 1780 和 1790 的执行完成之后被执行。

[0452] 应注意,参考图 20,一些块可整合成合并的块或可分成几个块,和/或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件以外的元件执行。

[0453] 现在注意图 21,其示出根据当前公开的主题的某些例子的为了由切换目标执行高速缓存切换而执行的一序列操作的流程图。

[0454] 根据当前公开的主题的一些例子,与切换目标相关的高速缓存切换模块 2530 可配置成从切换发起者接收“切换开始”通知(包括切换后高速缓存映射),保存本地切换前高速缓存映射的备份拷贝,根据从切换发起者接收的切换后高速缓存映射更新高速缓存映射,并进入“进行中的切换”状态(块 1810)。

[0455] 高速缓存切换模块 2530 在一些情况下可进一步配置成例如根据高速缓存相关 SLS 和/或阈值和/或涉及 DSS 200 或其部分(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 的切换目标知识和/或存储在与它相关的 UDSP 数据存储库 330 中的高速缓存相关参数来检查从切换发起者接收的切换请求是否是可接受的(块 1820)。

[0456] 可认识到,在一些情况下,切换目标可访问关于高速缓存相关 SLS 和/或阈值和/或涉及 DSS 200 或其部分(例如,最大允许的站点级过度调拨、最大允许的总过度调拨、各种安全参数等)的 SLS 和/或与其相关和/或与连接到其的高速缓存资源相关的高速缓存相关参数的信息,其不同于切换发起者访问的这样的信息(且在一些情况下,新的和/或更更新的信息)。例如,在一些情况下,到切换请求由切换目标接收时为止,与连接到其的高速缓存资源有关的一个或多个参数已经改变。

[0457] 如果切换请求不是可接受的(例如,基于切换目标知识),高速缓存切换模块 2530 可配置成将拒绝通知发送到切换发起者(指示切换目标不接受切换发起者所发送的切换请求),恢复本地切换前高速缓存映射(在块 1810 中被保存用于备份),并退出“进行中的切换”状态(块 1830)。

[0458] 如果切换请求是可接受的,高速缓存切换模块 2530 可配置成将可接受的通知发送到切换发起者(指示切换目标接受切换发起者所发送的切换请求)(块 1840)。在这样的情况下,高速缓存切换模块 2530 可配置成等待“切换完成”通知,指示一个或多个高速缓存空间或其部分转移到切换目标的责任,且一旦这样的通知被接收到,就退出“进行中的切换”状态(块 1850)。

[0459] 应注意,参考图 21,一些块可整合成合并的块或可分成几个块,和/或其它块可被添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件

以外的元件执行。

[0460] 现在注意图 22, 其示出根据当前公开的主题的某些例子的为了处理在切换期间处理切换发起者所接收的对象相关请求而执行的一序列操作的流程图。

[0461] 根据当前公开的主题的一些例子, 与切换发起者相关的对象请求管理模块 2540 可配置成在切换期间例如从客户端(例如, 客户端服务器 218、网关资源 216 或任何其它资源)接收与切换前和 / 或切换后高速缓存映射内的对象有关的对象相关请求(例如, 读 / 写请求)(块 1910)。

[0462] 应注意, 在一些情况下, 请求客户端可根据切换后高速缓存映射在更新其局部高速缓存映射之前发送请求(因为在一些情况下例如由于网络上的拥挤的业务或由于任何其它原因, 切换后高速缓存映射不被客户端立即接收到)。因此, 这样的对象相关请求可被发送到切换发起者, 同时所请求的对象不再被它拥有(例如, 处理)。因此, 当接收到这样的对象相关请求时, 对象请求管理模块 2540 可配置成根据切换后高速缓存映射来检查所请求的对象是否在切换发起者的所有权下(块 1920), 且如果是, 对象请求管理模块 2540 可配置成处理请求(块 1930)。

[0463] 然而, 如果根据切换后高速缓存映射, 切换发起者不是所请求的对象的所有者, 则对象请求管理模块 2540 可配置成检查所请求的对象是否仍在切换发起者的所有权下(例如, 因为切换发起者仍然不在所请求的对象上转移所有权)(块 1940)。

[0464] 如果这样的所请求的对象仍然被切换发起者拥有, 则对象请求管理模块 2540 可配置成处理请求(块 1930)。然而, 如果这样的所请求的对象不再被切换发起者拥有, 则对象请求管理模块 2540 可配置成将对象相关请求转播到切换目标(块 1950)。

[0465] 应注意, 参考图 22, 一些块可整合成合并的块或可分成几个块, 和 / 或其它块可被添加。此外, 在一些情况下, 可以用与本文所述的不同的顺序执行块。也应注意, 虽然也参考实现它们的系统元件描述了流程图, 这决不是限制性的, 且块可由除了本文所述的元件以外的元件执行。

[0466] 现在注意图 23, 其示出根据当前公开的主题的某些例子的为了处理在切换期间处理切换目标所接收的、与包括在切换后高速缓存映射中的对象有关的对象相关请求而执行的一序列操作的流程图。

[0467] 根据当前公开的主题的一些例子, 与切换目标相关的对象请求管理模块 2540 可配置成例如从客户端(例如, 客户端服务器 218、网关资源 216 或任何其它资源)或从切换发起者接收与切换后高速缓存映射内的对象有关的对象相关请求(例如, 读 / 写请求)(块 2110)。当接收到这样的请求时, 对象请求管理模块 2540 可配置成检查请求是否起源于切换发起者(块 2120)。

[0468] 如本文所示的, 在根据切换后高速缓存映射确定切换发起者不是所请求的对象的所有者以及切换发起者不是所请求的对象当前所有者(切换发起者仍然是所有者的指示可例如是对象在切换发起者的高速缓存器中仍然是脏的, 等等)之后, 切换发起者可将对象相关请求转播到切换目标。因此, 每当请求起源于切换发起者时, 对象请求管理模块 2540 可配置成处理请求(块 2130), 因为这指示切换发起者已经将所请求的对象的所有权转移到切换目标。

[0469] 如果请求不起源于切换发起者, 则对象请求管理模块 2540 可配置成检查所请求

的对象是否被切换目标拥有(因为例如切换目标已经从切换发起者接收到所请求的对象的
所有权,等等)(块 2140)。如果它被切换目标拥有,则对象请求管理模块 2540 可配置成处
理请求(块 2130)。

[0470] 如果所请求的对象不被切换目标拥有(因为例如所请求的对象所有权还没有转
移,且切换过程没有结束),则对象请求管理模块 2540 可配置成将请求转播到切换发起者
(块 2150)。应注意,这样的场景可例如从下列事实产生:某个用户可具有更新的切换后高速
缓存映射,指示所请求的对象被映射到切换目标,而当它请求所请求的对象时,切换过程仍
然正在进行,且因此所请求的对象所有权还没有转移到切换目标。

[0471] 应注意,参考图 23,一些块可整合成合并的块或可分成几个块,和 / 或其它块可被
添加。此外,在一些情况下,可以用与本文所述的不同的顺序执行块。也应注意,虽然也参
考实现它们的系统元件描述了流程图,这决不是限制性的,且块可由除了本文所述的元件
以外的元件执行。

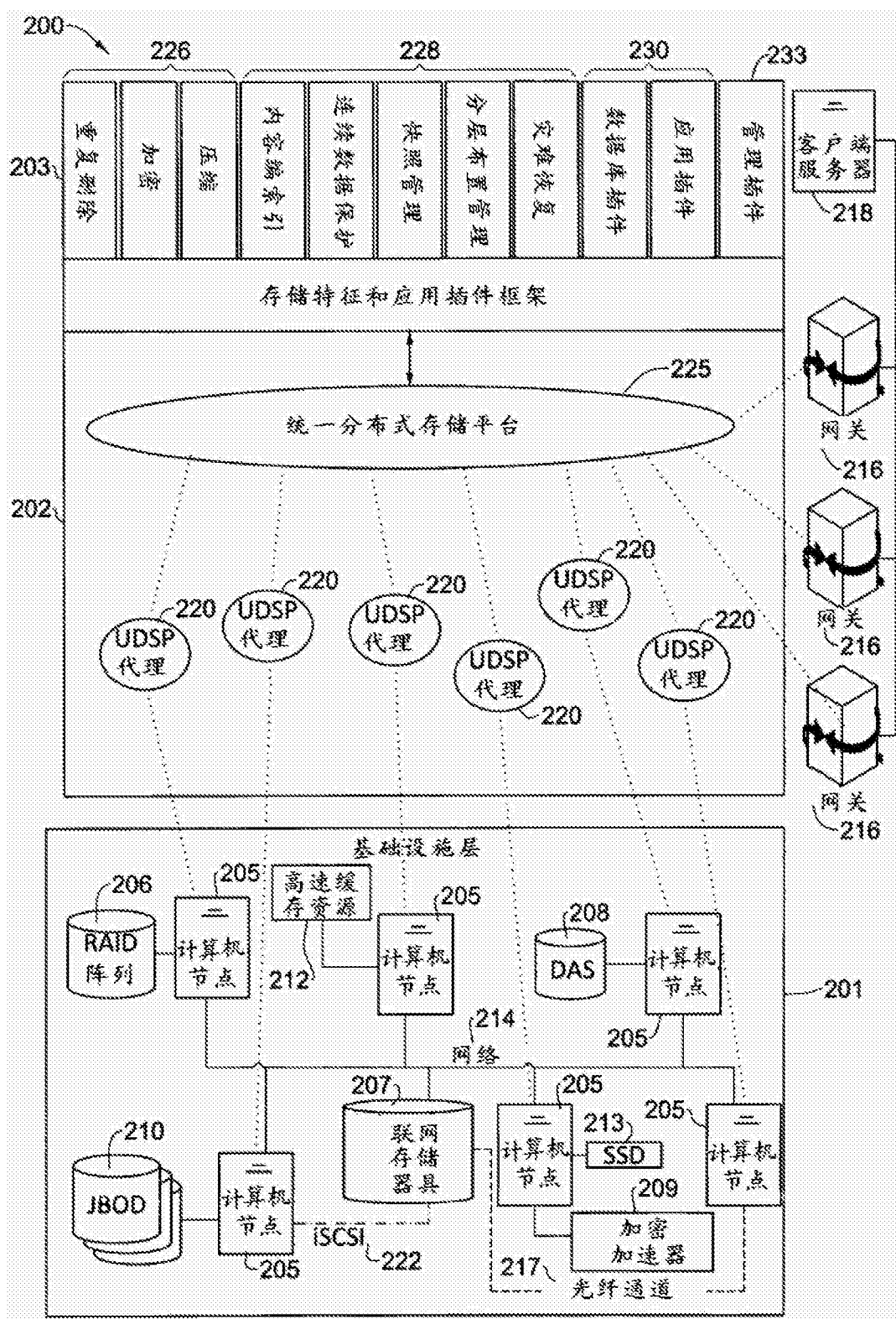


图 1

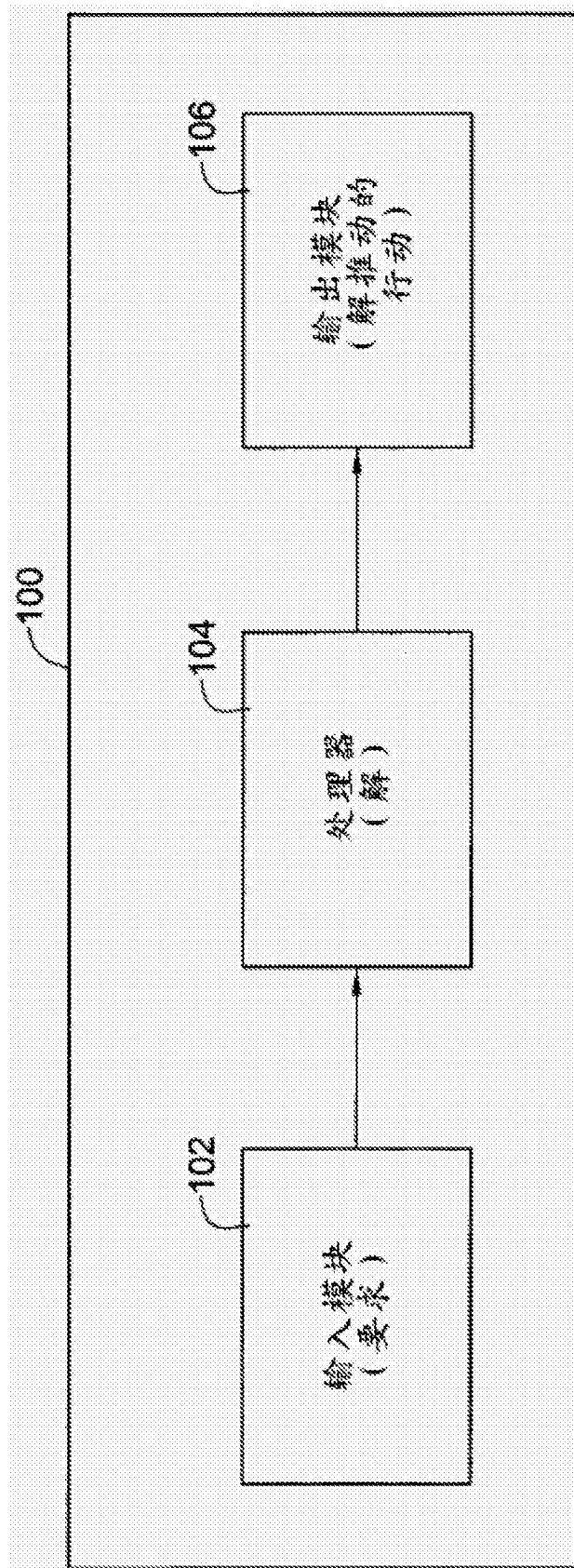


图 2

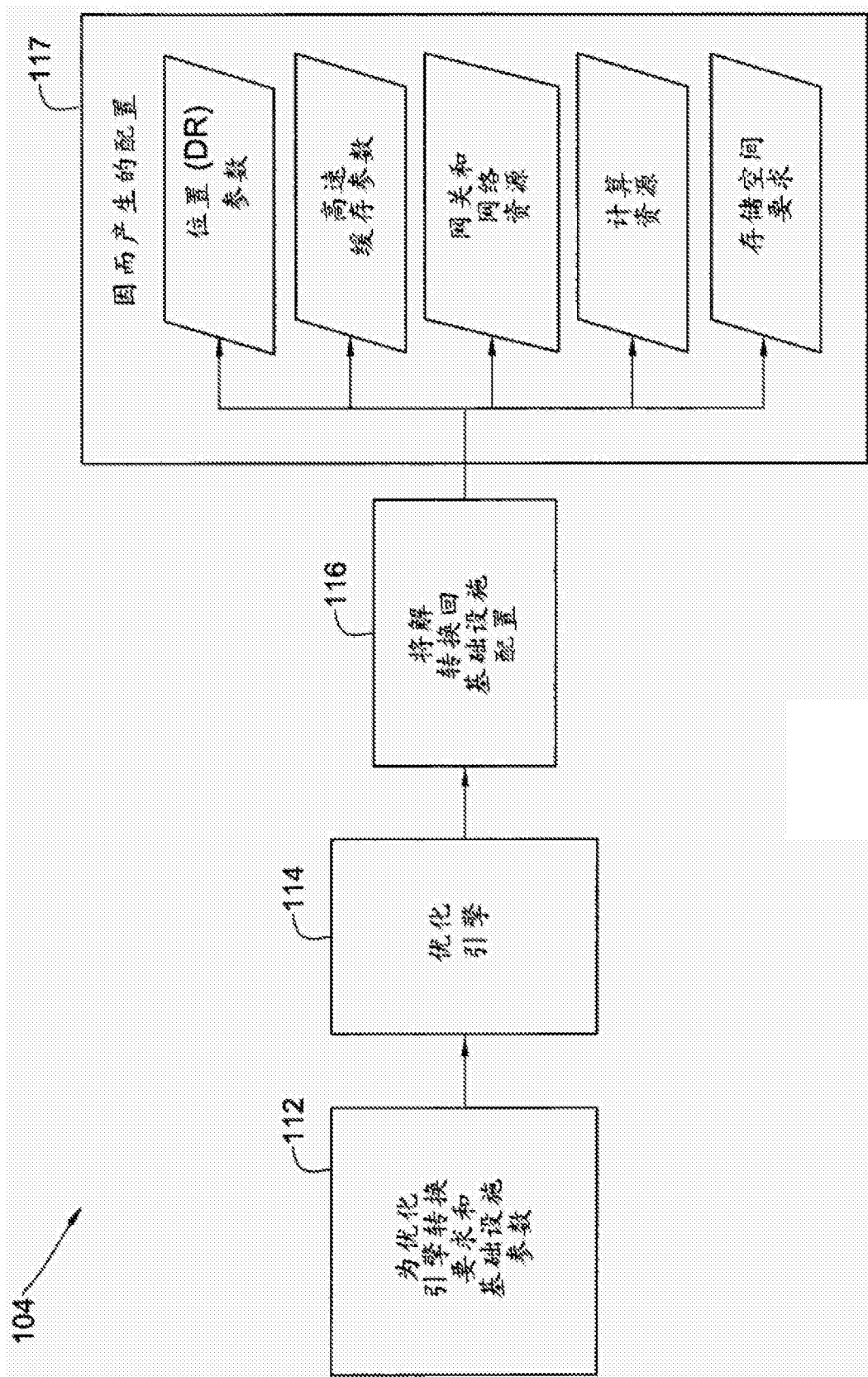


图 3

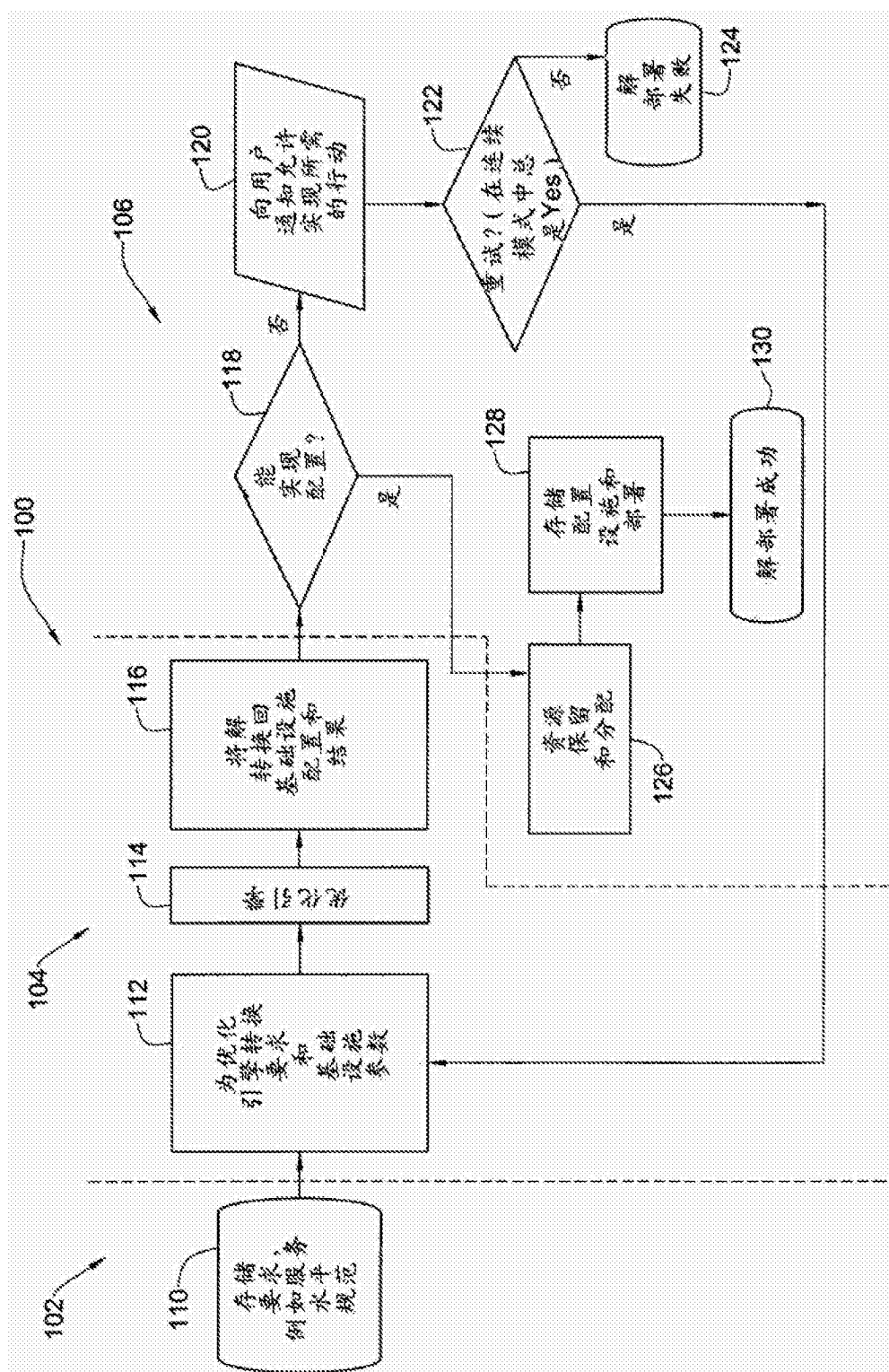


图 4

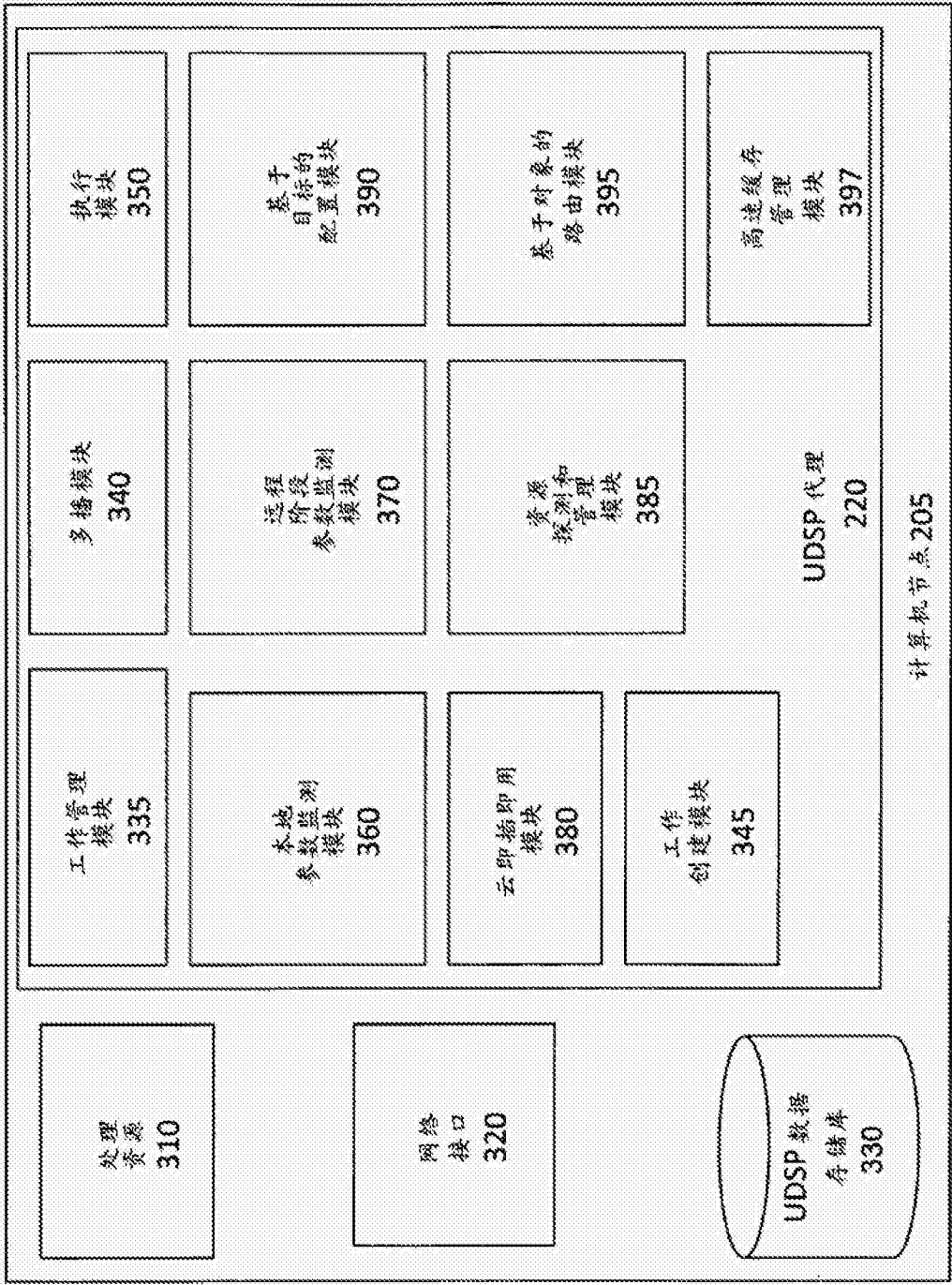


图 5

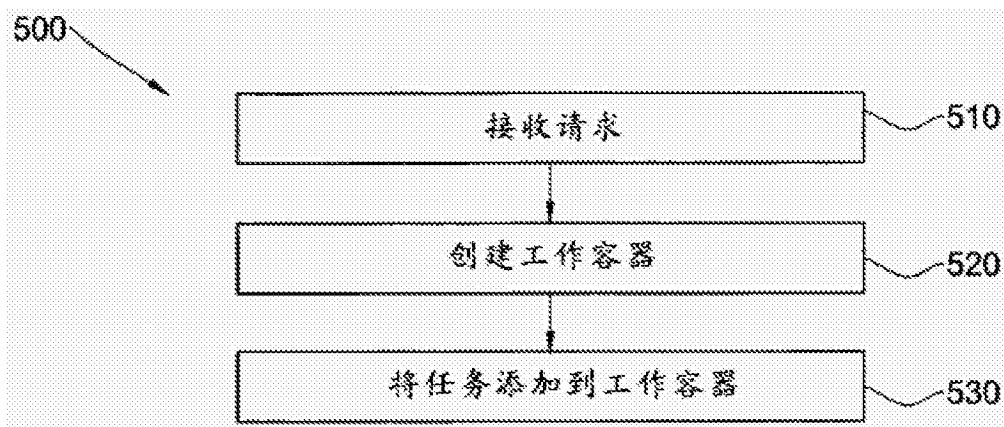


图 6

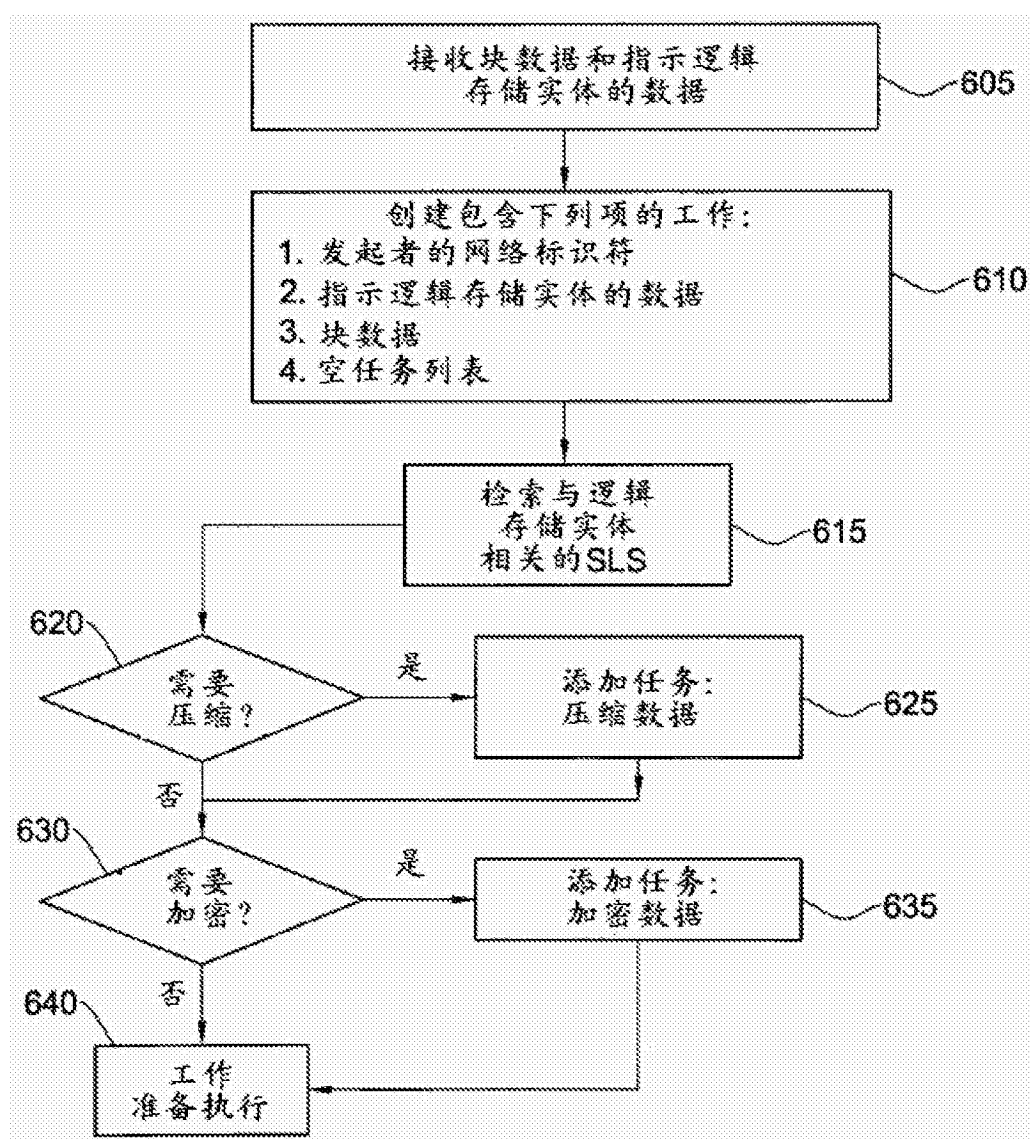


图 7

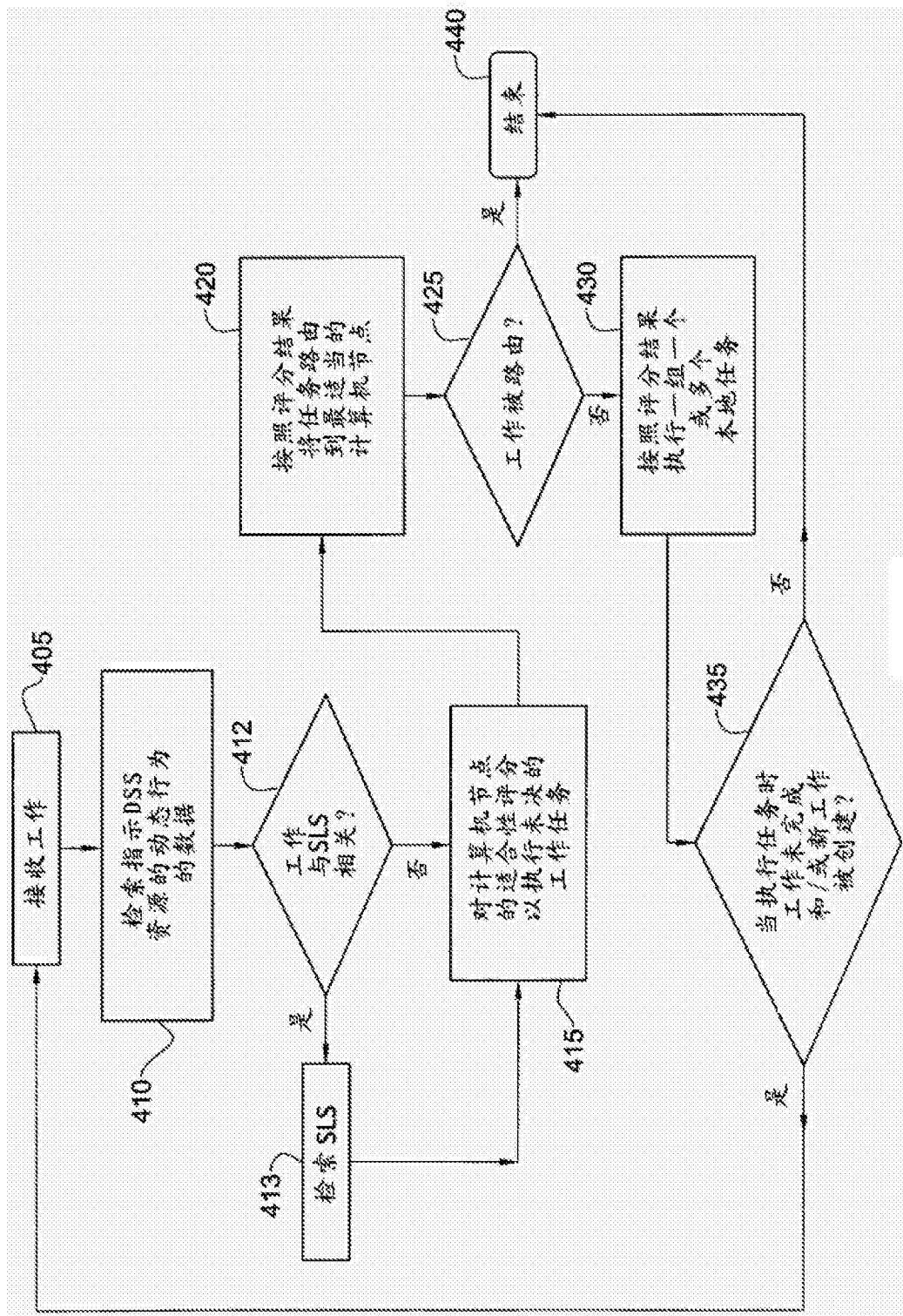


图 8

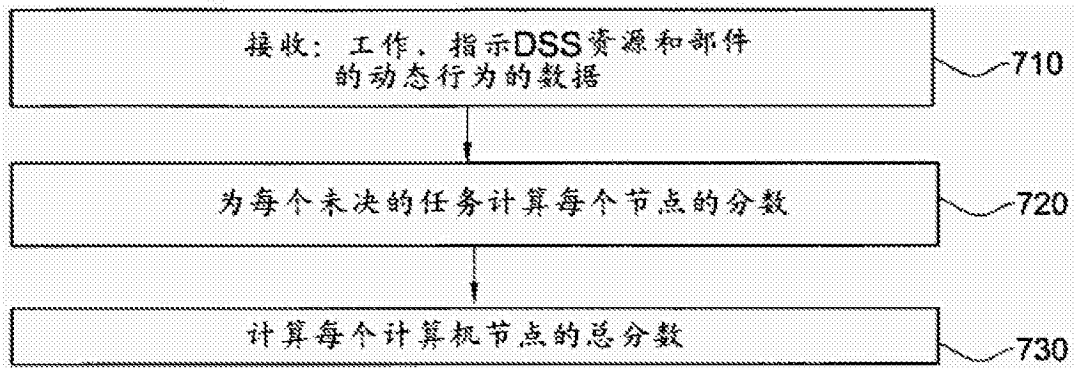


图 9

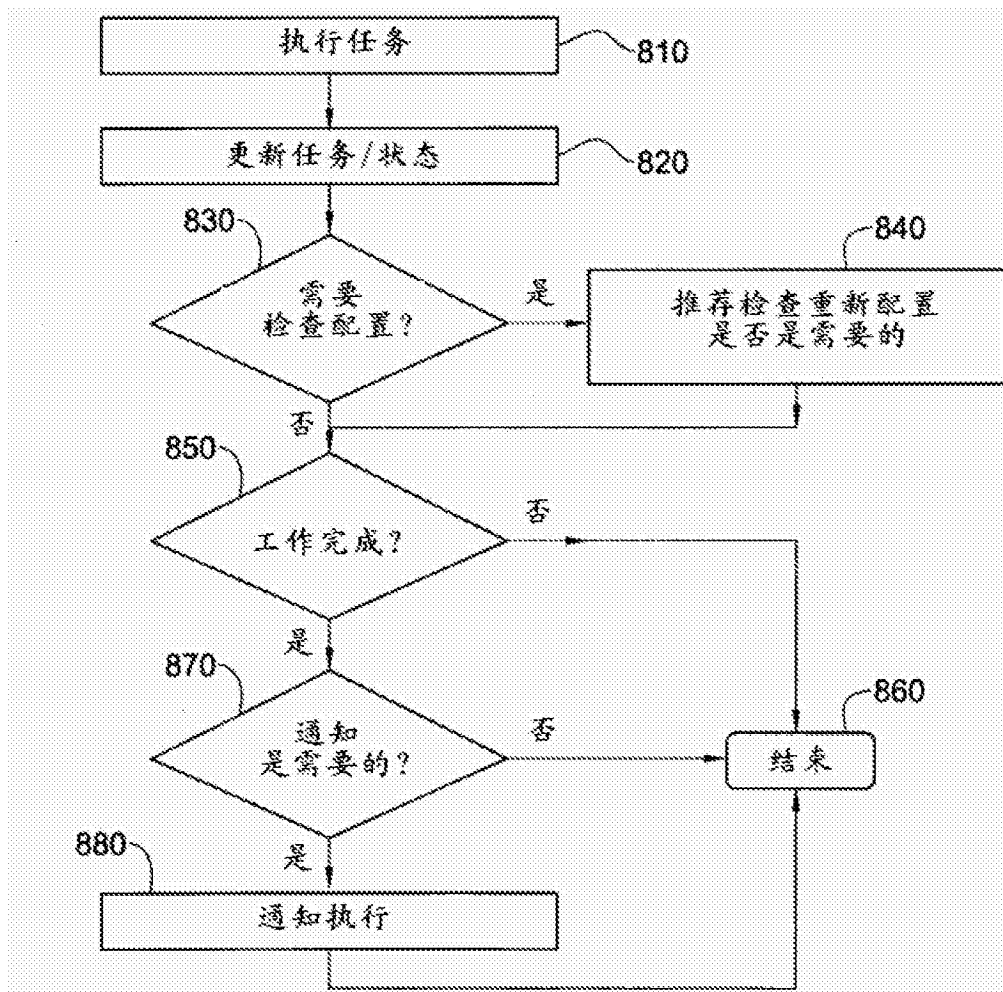


图 10

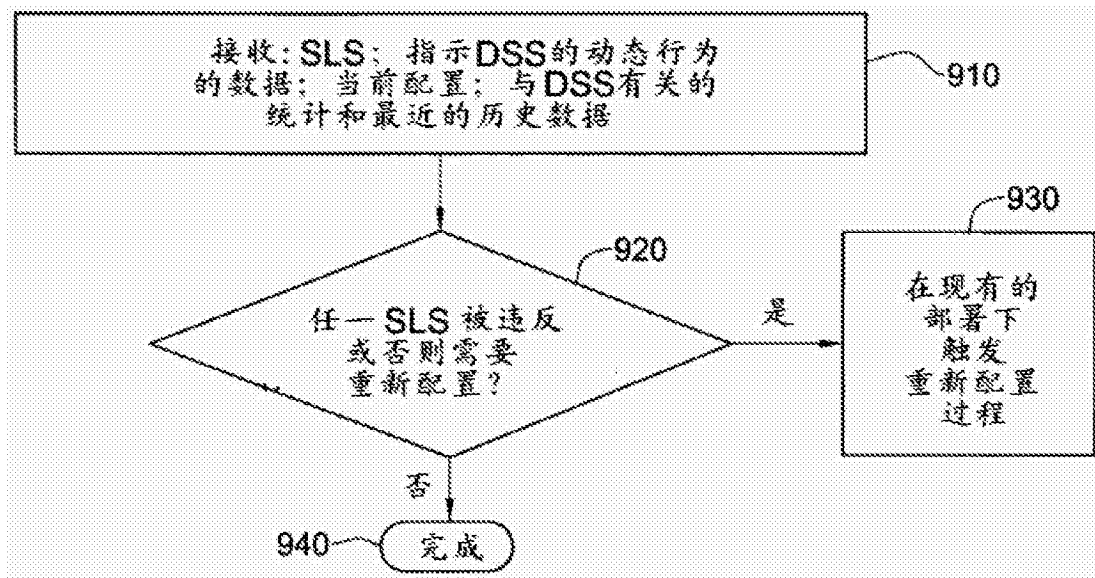


图 11

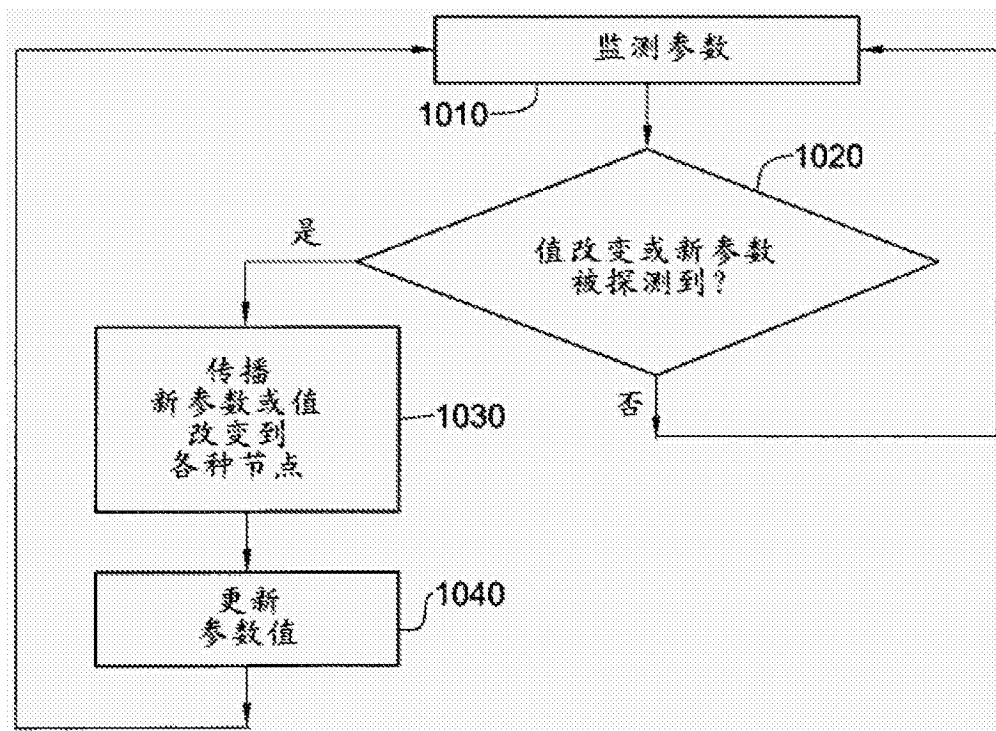


图 12

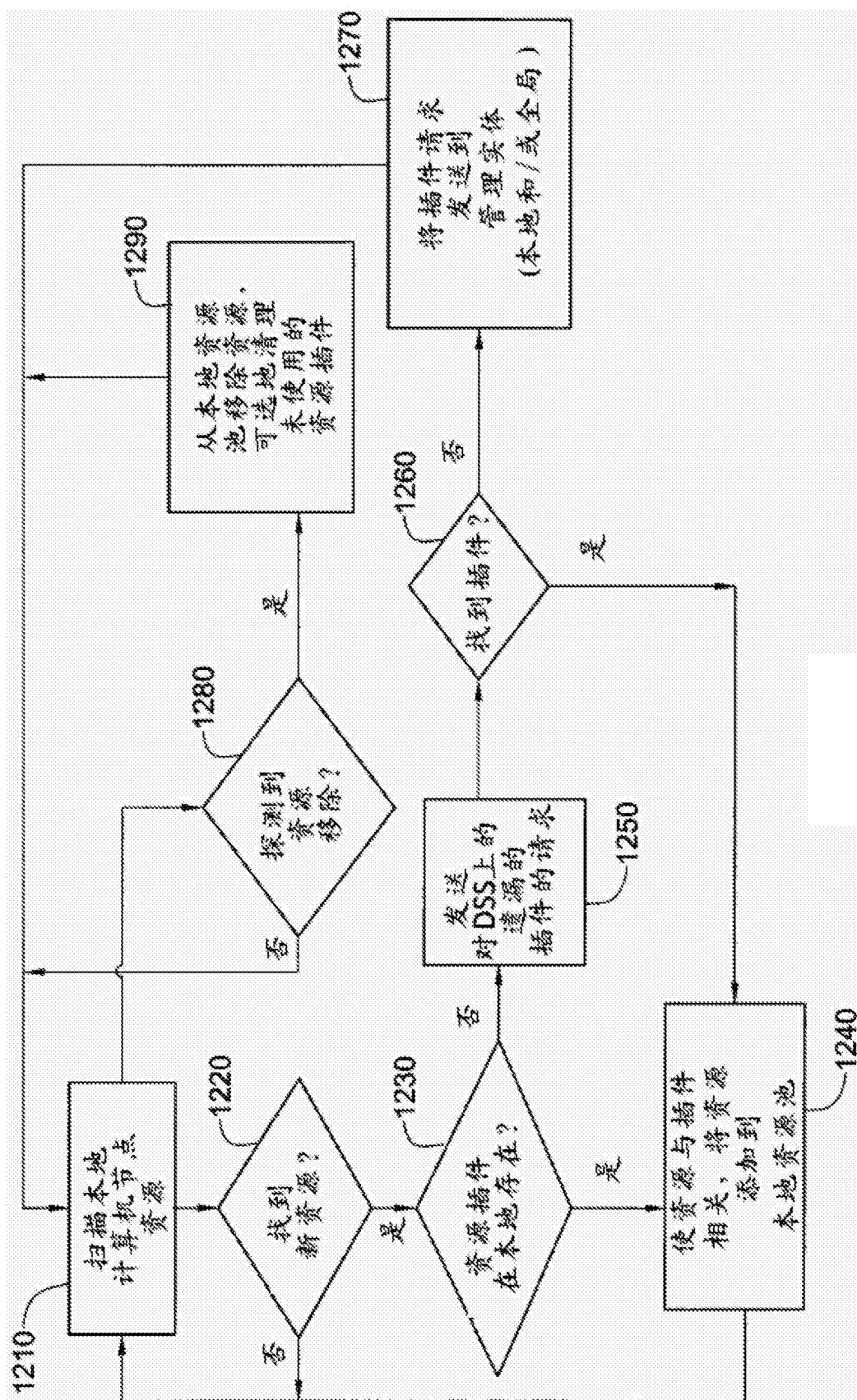


图 13

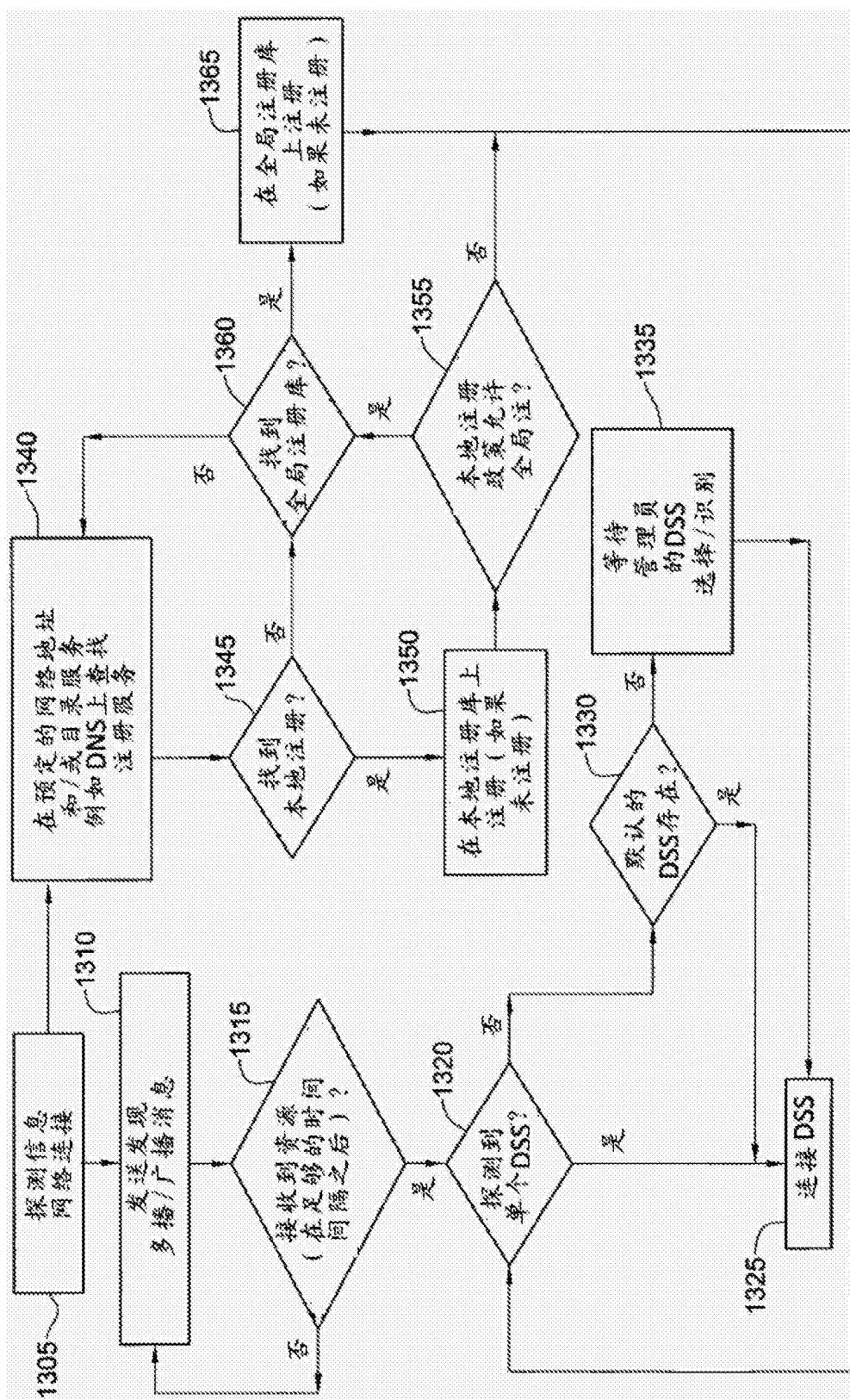


图 14

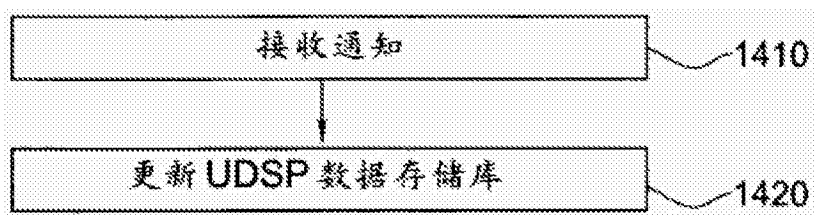


图 15

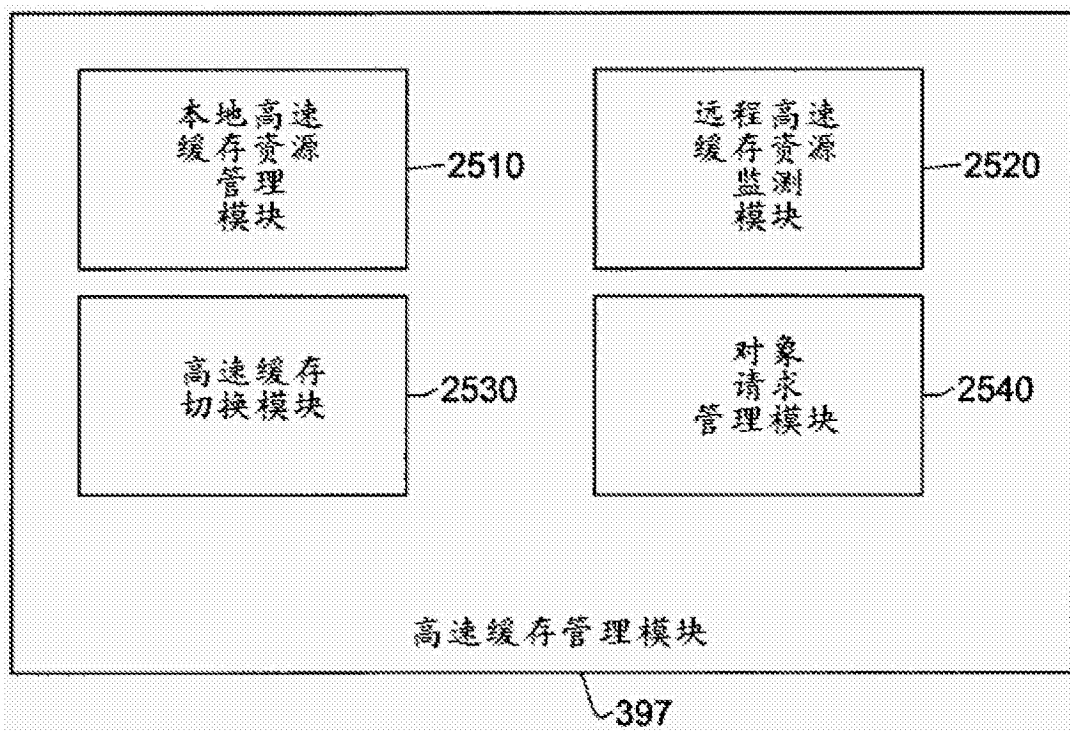


图 16

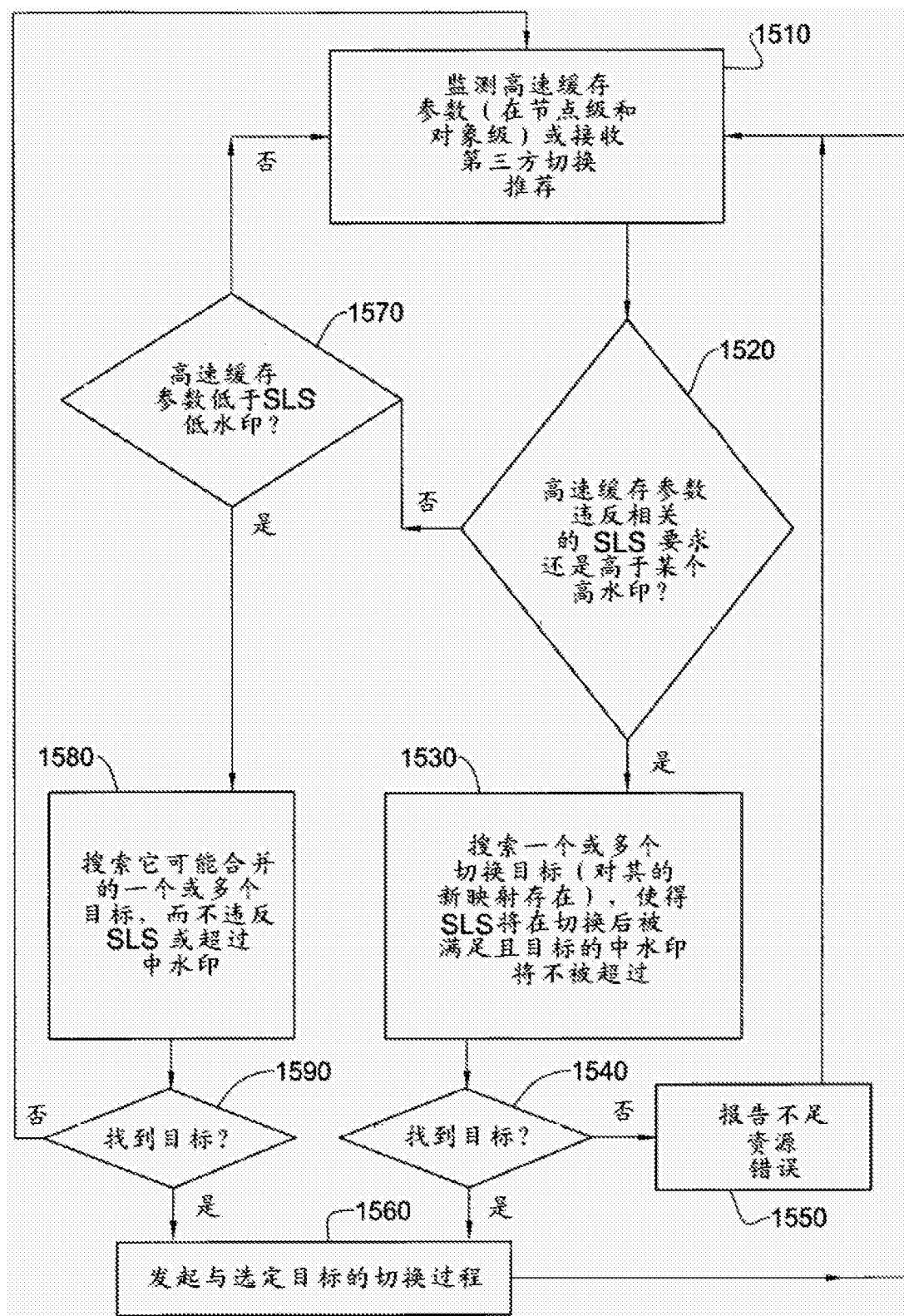


图 17

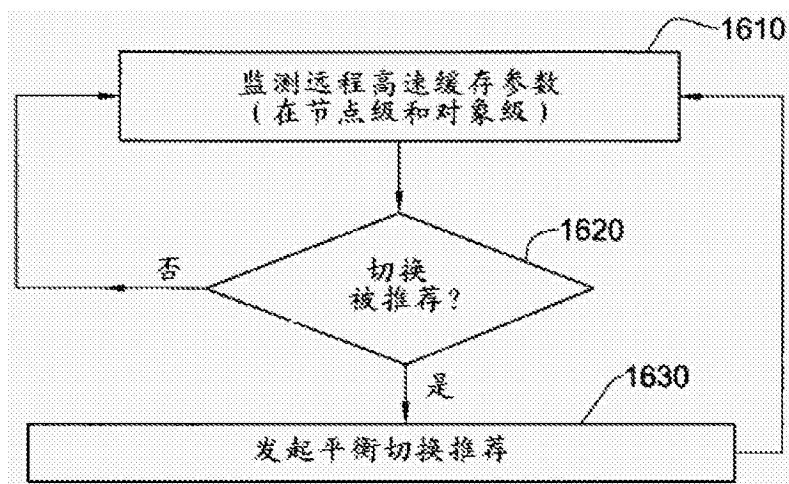


图 18

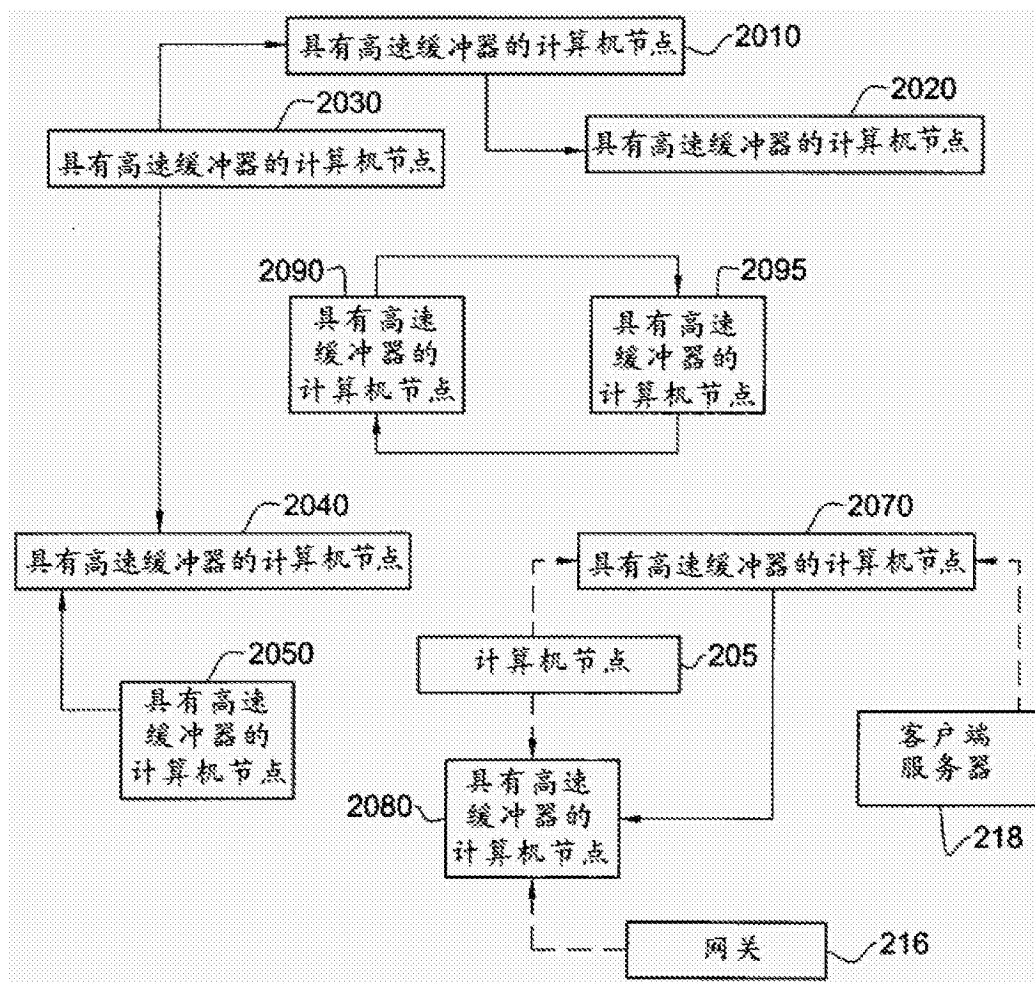


图 19

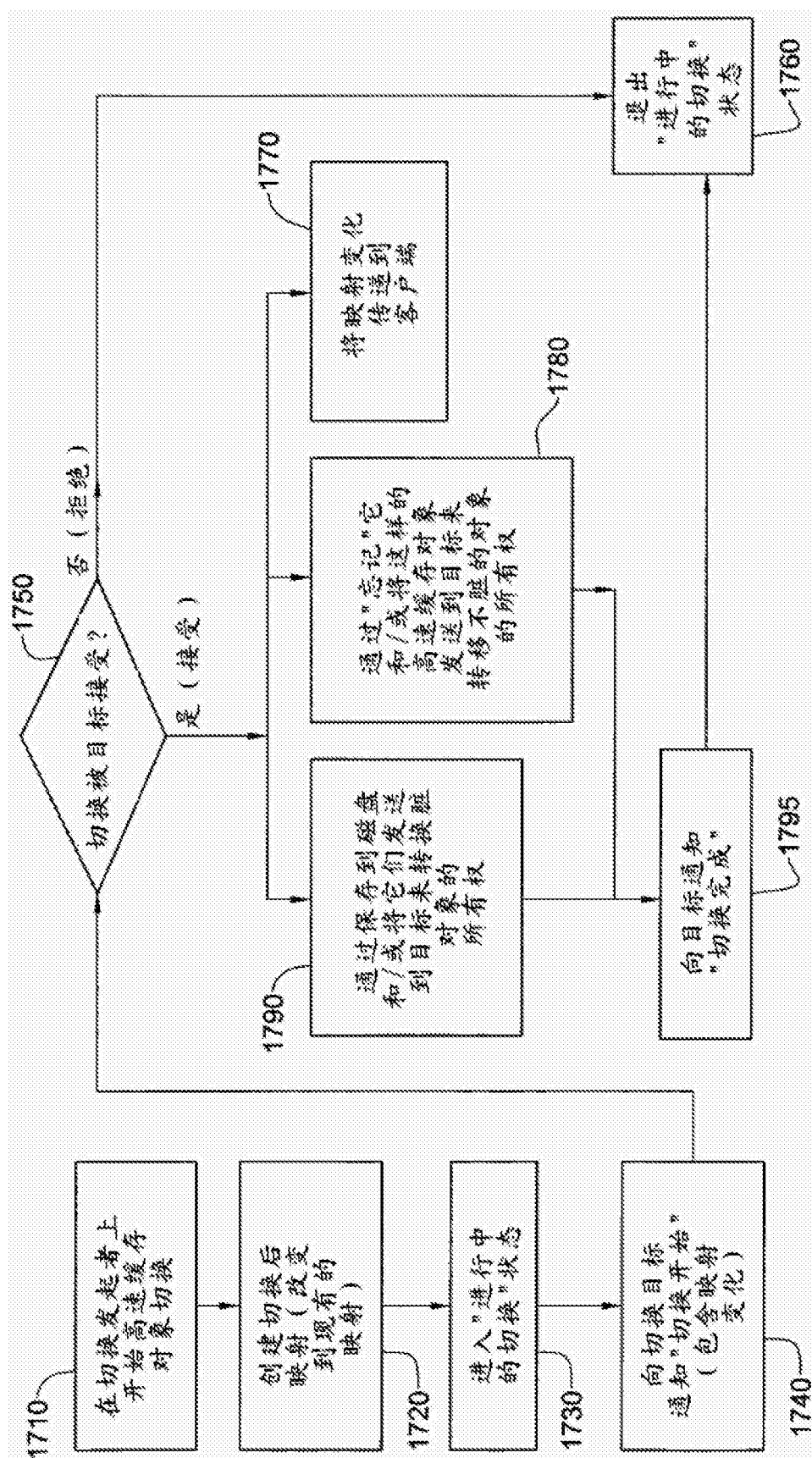


图 20

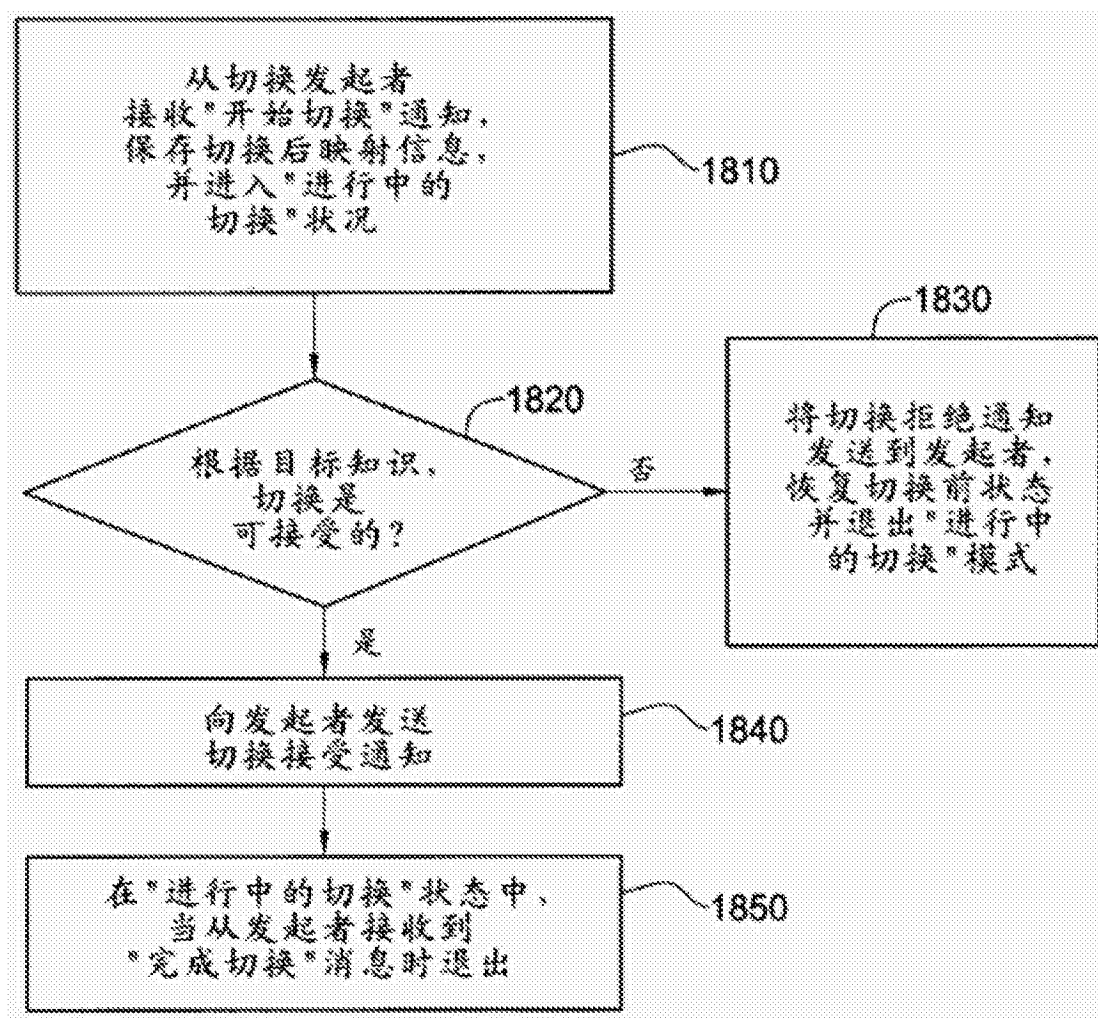


图 21

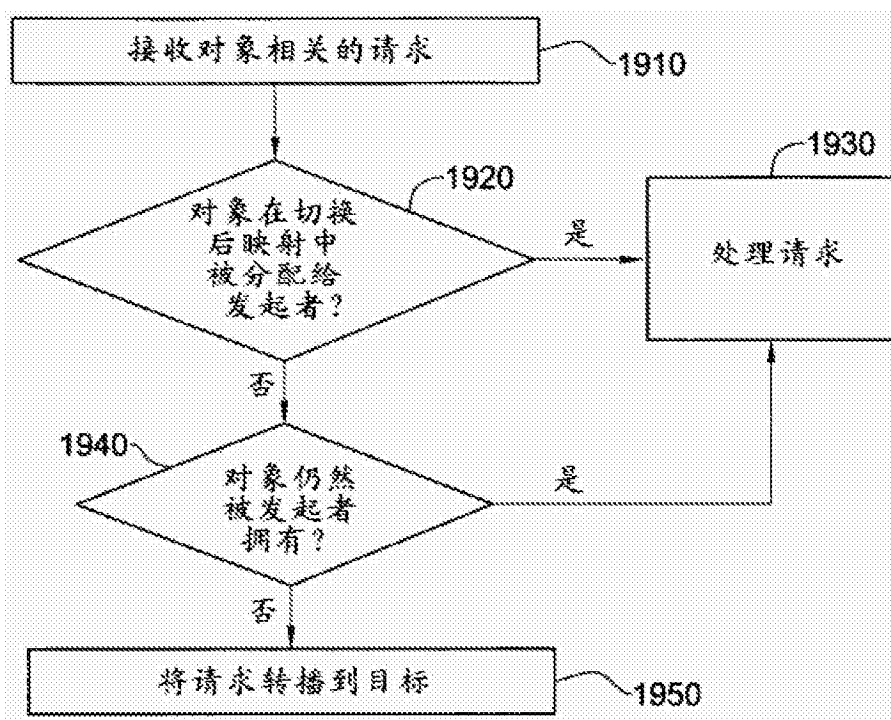


图 22

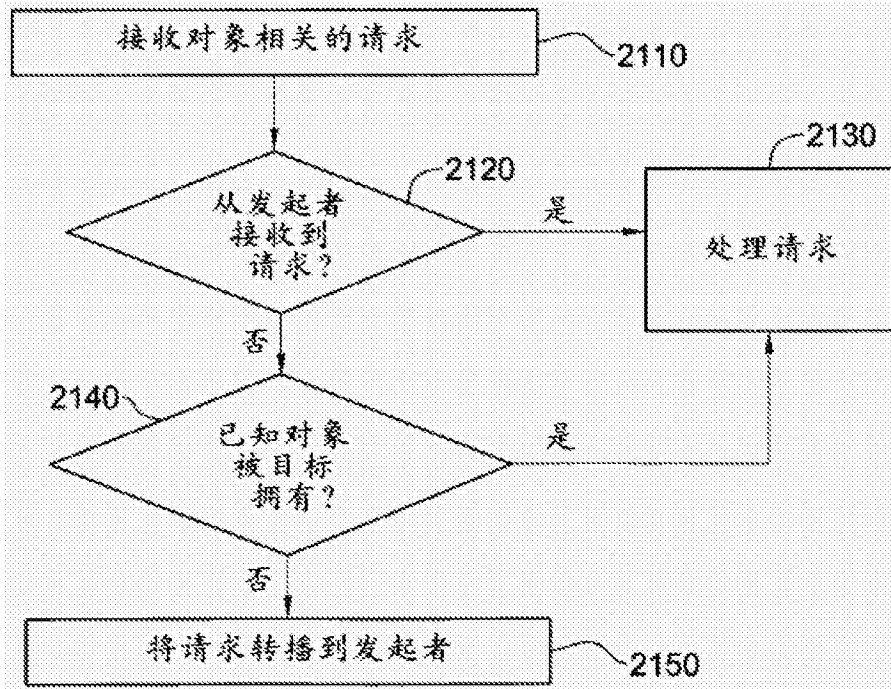


图 23