



(12)发明专利

(10)授权公告号 CN 103608796 B

(45)授权公告日 2016.09.21

(21)申请号 201280012607.0

(22)申请日 2012.01.10

(65)同一申请的已公布的文献号

申请公布号 CN 103608796 A

(43)申请公布日 2014.02.26

(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/IL2012/050009 2012.01.10

(87)PCT国际申请的公布数据

W02012/104847 EN 2012.08.09

(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 T·基普尼斯

G·洛埃维恩伯格

(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.

G06F 15/16(2006.01)

(56)对比文件

US 2004006589 A1,2004.01.08,

US 2006041644 A1,2006.02.23,

CN 101188569 A,2008.05.28,

石磊 等.利用ZIPF定律建立有效的WEB对象缓存机制.《计算机工程与应用》.2004,61-63.

高伟 等.一种P2P环境下分布式文件存储系统的缓存策略.《计算机工程与应用》.2004,45-48,84.

Mohsen Sardari 等.Memory allocation in distributed storage networks.

《Information Theory Proceedings (ISIT),

2010 IEEE International Symposium on》

.2010,1958-1962.

审查员 彭玉静

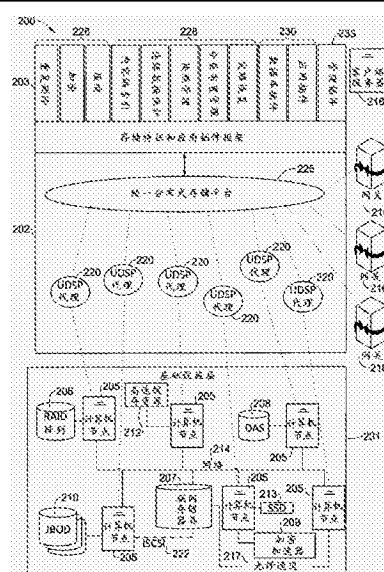
权利要求书14页 说明书48页 附图17页

(54)发明名称

大规模存储系统

(57)摘要

分布式存储系统包括: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)代理的处理资源,UDSP代理配置成:接收具有任务的工作,计算任务的分数,使得每个分数指示执行任务的适合性,同时满足至少一个服务水平规范(SLS)要求。分数基于存储相关资源参数数据来计算。工作基于所计算的分数被路由到更适当的计算机节点。响应于添加互连计算机节点并基于所计算的分数执行下面的工作的任务或将下面的工作路由到更适当的计算机节点来创建更新的基础设施层。



1. 一种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, 所述基础设施层包括互连计算机节点, 所述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, 所述计算机节点包括:

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, 其中所述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, 使得每个分数指示所述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, 所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到更适当的计算机节点, 而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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; 所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

2. 如权利要求1所述的计算机节点, 其中当计算下面的工作的所述任务的分数时, 创建所述更新的基础设施层, 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所述计算。

3. 如权利要求1所述的计算机节点, 其中当计算下面的工作的所述任务的分数时, 创建所述更新的基础设施层, 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所述计算。

4. 如权利要求1所述的计算机节点, 其中所述更新的基础设施层被动态地创建。

5. 如权利要求1所述的计算机节点, 其中仅对未决的任务计算分数, 其中未决的任务是在其执行之前没有未完成的先决条件的任务。

6. 如权利要求1所述的计算机节点, 其中所述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。

7. 如权利要求1所述的计算机节点, 其中所述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。

8. 如权利要求1所述的计算机节点, 其中所述分数由优化引擎计算。

9. 如权利要求8所述的计算机节点, 其中所述优化引擎使用下列优化技术中的一种或多种优化技术:

(a) 线性编程;

(b) 模拟退火;

(c) 遗传算法。

10. 如权利要求9所述的计算机节点, 其中所述一种或多种优化技术使用启发法或近似法。

11. 如权利要求1所述的计算机节点,其中所述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。

12. 如权利要求1所述的计算机节点,其中所述UDSP代理还配置成基于所计算的任务分数计算总分数,且其中所述执行所述任务中的一个或多个或将所述工作路由到更适当的计算机节点基于所计算的总分数。

13. 如权利要求1所述的计算机节点,其中所述工作从客户端服务器接收。

14. 如权利要求13所述的计算机节点,其中所述工作通过网关资源从所述客户端服务器接收。

15. 如权利要求1所述的计算机节点,其中所述UDSP代理还配置成:

a. 监测指示所述计算机节点的当前状态或连接到所述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,如果有一些存储相关资源连接到所述计算机节点的话;

b. 将指示对所述至少一个所监测的参数的变化的通知传播到所述互连计算机节点中的至少一个。

16. 如权利要求1所述的计算机节点,其中所述UDSP代理还配置成:

计算关于所述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;所述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;以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

17. 如权利要求1所述的计算机节点,其中所述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和与连接到所述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所述分布式存储系统的配置;以及

根据所计算的配置自动分配所述存储相关资源之一的至少部分。

18. 如权利要求17所述的计算机节点,其中所述UDSP代理还配置成:

接收与所述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、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

19. 一种操作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的方法,所述基础设施层包括互连计算机节点,所述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,所述方法包括:

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,使得每个分数指示所述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,所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到更适当的计算机节点,而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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;所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

20. 如权利要求19所述的方法,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

21. 如权利要求19所述的方法,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

22. 如权利要求19所述的方法,其中所述更新的基础设施层被动态地创建。

23. 如权利要求19所述的方法,其中仅对未决的任务计算分数,其中未决的任务是在其执行之前没有未完成的先决条件的任务。

24. 如权利要求19所述的方法,其中所述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。

25. 如权利要求19所述的方法,其中所述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。

26. 如权利要求19所述的方法,其中所述计算分数由优化引擎执行。

27. 如权利要求26所述的方法,其中所述优化引擎使用下列优化技术中的一种或多种优化技术:

(a)线性编程;

(b)模拟退火;

(c)遗传算法。

28. 如权利要求27所述的方法,其中所述一种或多种优化技术使用启发法或近似法。

29. 如权利要求19所述的方法,其中所述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。

30. 如权利要求19所述的方法,还包括:

基于所计算的任务分数计算总分数,且其中所述执行所述任务中的一个或多个或将所述工作路由到更适当的计算机节点基于所计算的总分数。

31. 如权利要求19所述的方法,其中所述工作从客户端服务器接收。

32. 如权利要求31所述的方法,其中所述工作通过网关资源从客户端服务器接收。

33. 如权利要求19所述的方法,还包括:

a. 监测指示所述计算机节点的当前状态或连接到所述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,如果有一些存储相关资源连接到所述计算机节点的

话;

b. 将指示对所述至少一个所监测的参数的变化的通知传播到所述互连计算机节点中的至少一个。

34. 如权利要求19所述的方法, 还包括:

计算关于所述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; 所述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; 以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

35. 如权利要求19所述的方法, 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和与连接到所述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所述分布式存储系统的配置; 以及

根据所计算的配置自动分配所述存储相关资源之一的至少部分。

36. 如权利要求35所述的方法, 还包括:

接收与所述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, 至少基于所述至少一个SLS、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

37. 一种分布式存储系统, 包括:

包括互连计算机节点的基础设施层, 其中:

所述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;

所述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;

所述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, 使得每个分数指示所述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, 所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到所述基础设施层的更适当的计算机节点, 而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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; 所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

38. 如权利要求37所述的分布式存储系统,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

39. 如权利要求37所述的分布式存储系统,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

40. 如权利要求37所述的分布式存储系统,其中所述更新的基础设施层被动态地创建。

41. 如权利要求37所述的分布式存储系统,其中仅对未决的任务计算分数,其中未决的任务是在其执行之前没有未完成的先决条件的任务。

42. 如权利要求37所述的分布式存储系统,其中所述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。

43. 如权利要求37所述的分布式存储系统,其中所述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。

44. 如权利要求37所述的分布式存储系统,其中所述分数由优化引擎计算。

45. 如权利要求44所述的分布式存储系统,其中所述优化引擎使用下列优化技术中的一种或多种优化技术个:

(d)线性编程;

(e)模拟退火;

(f)遗传算法。

46. 如权利要求45所述的分布式存储系统,其中所述一种或多种优化技术使用启发法或近似法。

47. 如权利要求37所述的分布式存储系统,其中所述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。

48. 如权利要求37所述的分布式存储系统,其中所述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。

49. 如权利要求37所述的分布式存储系统,其中所述工作从客户端服务器接收。

50. 如权利要求49所述的分布式存储系统,其中所述工作通过网关资源从所述客户端服务器接收。

51. 如权利要求37所述的分布式存储系统,其中所述UDSP代理还配置成:

a. 监测指示所述计算机节点的当前状态或连接到所述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,如果有一些存储相关资源连接到所述计算机节点的话;

b. 将指示对所述至少一个所监测的参数的变化的通知传播到所述互连计算机节点中的至少一个。

52. 如权利要求37所述的分布式存储系统,其中所述UDSP代理还配置成:

计算关于所述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;所述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;以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

53. 如权利要求37所述的分布式存储系统,其中所述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;

至少基于所述至少一个SLS和所述存储相关资源参数数据计算所述分布式存储系统的配置;以及

根据所计算的配置自动分配所述存储相关资源之一的至少部分。

54. 如权利要求53所述的分布式存储系统,其中所述UDSP代理还配置成:

接收与所述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,至少基于所述至少一个SLS、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

55. 一种分布式存储系统,包括:

包括互连计算机节点的基础设施层,其中:

所述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;

所述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,使得每个分数指示所述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,所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到所述基础设施层的更适当的计算机节点,而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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;所述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;以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

56. 如权利要求55所述的分布式存储系统,其中所述UDSP代理还配置成:

接收与所述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,至少基于所述至少一个SLS要求、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

57.如权利要求55所述的分布式存储系统,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

58.如权利要求55所述的分布式存储系统,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

59.如权利要求55所述的分布式存储系统,其中所述更新的基础设施层被动态地创建。

60.如权利要求55所述的分布式存储系统,其中所述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。

61.一种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,所述基础设施层包括互连计算机节点,所述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,所述计算机节点包括:

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,其中所述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,使得每个分数指示所述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,所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到所述基础设施层的更适当的计算机节点,而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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;所述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;以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

62.如权利要求61所述的计算机节点,其中所述UDSP代理还配置成:

接收与所述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,至少基于所述至少一个SLS要求、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

63.如权利要求61所述的计算机节点,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

64. 如权利要求61所述的计算机节点,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

65. 如权利要求61所述的计算机节点,其中所述更新的基础设施层被动态地创建。

66. 如权利要求61所述的计算机节点,其中所述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。

67. 一种操作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的方法,所述基础设施层包括互连计算机节点,所述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,所述方法包括:

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,使得每个分数指示所述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,所述分数也基于与连接到所述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所述任务中的一个或多个或将所述工作路由到所述基础设施层的更适当的计算机节点,而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力;

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;所述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;以及

基于所计算的分数执行下面的工作的所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

68. 如权利要求67所述的方法,还包括:

接收与所述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,至少基于所述至少一个SLS要求、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

69. 如权利要求67所述的方法,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

70. 如权利要求67所述的方法,其中当计算下面的工作的所述任务的分数时,创建所述更新的基础设施层,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所述计算。

71. 如权利要求67所述的方法,其中所述更新的基础设施层被动态地创建。

72. 如权利要求67所述的方法,其中所述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。

73. 一种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,所述基础设施层包括互连计算机节点,所述计算机节点包括:

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,其中所述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和与连接到所述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所述分布式存储系统的配置;

根据所计算的配置自动分配所述存储相关资源之一的至少部分;

接收与所述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,至少基于所述至少一个SLS、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

74.如权利要求73所述的计算机节点,其中所述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,使得每个分数指示所述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,同时满足所述至少一个SLS的至少一个用户定义的存储相关要求,所述分数也基于与连接到所述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;

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;所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

75.如权利要求73所述的计算机节点,其中所述配置由优化引擎计算。

76.如权利要求75所述的计算机节点,其中所述优化引擎使用下列优化技术中的一种或多种优化技术:

- a.线性编程;
- b.模拟退火;
- c.遗传算法。

77.如权利要求76所述的计算机节点,其中所述一种或多种优化技术使用启发法或近似法。

78.如权利要求73所述的计算机节点,其中所述至少一个用户定义的存储相关要求包括下列项中的至少一个:

- a.位置;
- b.本地保护水平;
- c.备份保留政策;

- d. 远程保护水平;
- e. 性能水平;
- f. 加密水平;
- g. 重复删除;
- h. 压缩;
- i. 存储方法。

79. 如权利要求73所述的计算机节点, 其中所述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。

80. 如权利要求73所述的计算机节点, 其中所述存储相关资源分布在至少两个位置当中。

81. 如权利要求73所述的计算机节点, 其中所述配置包括过度调拨所述分布式存储系统的存储相关资源供使用。

82. 如权利要求73所述的计算机节点, 其中所述UDSP代理还配置成响应于故障来确定所述配置, 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。

83. 一种操作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的方法, 所述基础设施层包括互连计算机节点, 所述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, 所述方法包括:

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和与连接到所述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;

至少基于所述至少一个SLS和所述存储相关资源参数数据计算所述分布式存储系统的配置;

根据所计算的配置自动分配所述存储相关资源之一的至少部分;

接收与所述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, 至少基于所述至少一个SLS、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

84. 如权利要求83所述的方法, 还包括:

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, 使得每个分数指示所述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, 所述分数也基于与连接到所述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所述任务中的一个或多个或将所述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; 所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

85. 如权利要求83所述的方法, 其中所述计算配置由优化引擎执行。

86. 如权利要求85所述的方法, 其中所述优化引擎使用下列优化技术中的一种或多种优化技术:

- a. 线性编程;
- b. 模拟退火;
- c. 遗传算法。

87. 如权利要求86所述的方法, 其中所述一种或多种优化技术使用启发法或近似法。

88. 如权利要求83所述的方法, 其中所述至少一个用户定义的存储相关要求包括下列项中的至少一个:

- a. 位置;
- b. 本地保护水平;
- c. 备份保留政策;
- d. 远程保护水平;
- e. 性能水平;
- f. 加密水平;
- g. 重复删除;
- h. 压缩;
- i. 存储方法。

89. 如权利要求83所述的方法, 其中所述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。

90. 如权利要求83所述的方法, 其中所述存储相关资源分布在至少两个位置当中。

91. 如权利要求83所述的方法, 其中所述配置包括过度调拨所述分布式存储系统的存储相关资源供使用。

92. 如权利要求83所述的方法, 还包括:

响应于故障来确定所述配置, 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资源的添加的推荐。

93. 一种分布式存储系统, 包括:

包括互连计算机节点的基础设施层, 其中:

所述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;

其中所述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)和与连接到所述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;

至少基于所述至少一个SLS和所述存储相关资源参数数据计算所述分布式存储系统的配置;

根据所计算的配置自动分配所述存储相关资源之一的至少部分;

接收与所述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，至少基于所述至少一个SLS、所述存储相关资源参数数据和所述动态行为参数数据计算对所述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所述存储相关资源之一的至少部分。

94. 如权利要求93所述的分布式存储系统，其中所述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，使得每个分数指示所述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，同时满足所述至少一个SLS的至少一个用户定义的存储要求，所述分数也基于与连接到所述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所述任务中的一个或多个或将所述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；所述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所述任务中的一个或多个或将所述下面的工作路由到所述更新的基础设施层的更适当的计算机节点。

95. 如权利要求93所述的分布式存储系统，其中所述配置由优化引擎计算。

96. 如权利要求95所述的分布式存储系统，其中所述优化引擎使用下列优化技术中的一种或多种优化技术：

- a. 线性编程；
- b. 模拟退火；
- c. 遗传算法。

97. 如权利要求96所述的分布式存储系统，其中所述一种或多种优化技术使用启发法或近似法。

98. 如权利要求93所述的分布式存储系统，其中所述至少一个用户定义的存储相关要求包括下列项中的至少一个：

- a. 位置；
- b. 本地保护水平；
- c. 备份保留政策；
- d. 远程保护水平；
- e. 性能水平；
- f. 加密水平；
- g. 重复删除；
- h. 压缩；
- i. 存储方法。

99. 如权利要求93所述的分布式存储系统，其中所述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。

100. 如权利要求93所述的分布式存储系统,其中所述存储相关资源分布在至少两个位置当中。

101. 如权利要求93所述的分布式存储系统,其中所述配置包括过度调拨所述分布式存储系统的所述存储相关资源供使用。

102. 如权利要求93所述的分布式存储系统,其中所述UDSP代理还配置成响应于故障来确定所述配置,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资源的添加的推荐。

103. 一种分布式存储系统,包括:

包括互连计算机节点的基础设施层,其中:

所述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;

其中所述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和与连接到所述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;

至少基于所述至少一个SLS和所述存储相关资源参数数据计算所述分布式存储系统的配置;以及

根据所计算的配置自动分配所述存储相关资源之一的至少部分。

104. 一种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,所述基础设施层包括互连计算机节点,所述计算机节点包括:

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,其中所述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和与连接到所述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;

至少基于所述至少一个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所述存储相关资源之一的至少部分。

105. 一种操作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的方法,所述基础设施层包括互连计算机节点,所述方法包括:

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和与连接到所述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;

至少基于所述至少一个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所述存储相关资源之一的至少部分。

106. 一种分布式存储系统,包括:

包括互连计算机节点的基础设施层,其中:

所述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；

所述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；

所述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，使得每个分数指示所述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所述任务中的一个或多个或将所述工作路由到所述基础设施层的更适当的计算机节点，而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力。

107. 一种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，所述基础设施层包括互连计算机节点，所述计算机节点包括：

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，其中所述UDSP代理配置成：

接收包括至少一个任务的工作；

计算关于连接到所述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，使得每个分数指示所述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所述任务中的一个或多个或将所述工作路由到更适当的计算机节点，而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力。

108. 一种操作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的方法，所述基础设施层包括互连计算机节点，所述方法包括：

接收包括至少一个任务的工作；

计算关于连接到所述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，使得每个分数指示所述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所述任务中的一个或多个或将所述工作路由到更适当的计算机节点，而不考虑所述计算机节点执行所述任务中的所述一个或多个的能力。

大规模存储系统

技术领域

[0001] 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。

背景技术

[0002] 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。

[0003] 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。

[0004] 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”。

[0005] 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”。

[0006] Ms.S.V.Patil等人的“Unified Virtual Storage:Virtualization of Distributed Storage in a Network”(International Journal of Computer Applications2010(0975-8887),Volume1-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”。

[0007] 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”。

发明内容

[0008] 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。

[0009] 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。

[0010] 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。

[0011] 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。

[0012] 根据当前公开的主题的某些例子,还提供了计算机节点,其中仅对未决的任务计算分数。

[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。

[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。

[0015] 根据当前公开的主题的某些例子,还提供了计算机节点,其中任务分数由优化引擎计算。

[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。

[0017] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一种或多种优化技术使用启发法或近似法。

[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。

[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。

[0020] 根据当前公开的主题的某些例子,还提供了计算机节点,其中工作从客户端服务器接收。

[0021] 根据当前公开的主题的某些例子,还提供了计算机节点,其中工作通过网关资源从客户端服务器接收。

[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: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,

[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。

[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。

[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: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。

[0026] 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。

[0027] 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。

[0028] 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。

[0029] 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。

[0030] 根据当前公开的主题的某些例子,还提供了一种方法,其中仅对未决的任务计算分数。

[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。

[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。

[0033] 根据当前公开的主题的某些例子,还提供了一种方法,其中计算分数由优化引擎执行。

[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。

[0035] 根据当前公开的主题的某些例子,还提供了一种方法,其中一种或多种优化技术使用启发法或近似法。

[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。

[0037] 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。

[0038] 根据当前公开的主题的某些例子,还提供了一种方法,其中工作从客户端服务器

接收。

[0039] 根据当前公开的主题的某些例子,还提供了一种方法,其中工作通过网关资源从客户端服务器接收。

[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: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。

[0041] 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。

[0042] 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。

[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: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。

[0044] 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。

[0045] 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。

[0046] 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。

[0047] 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。

[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,其中仅对未决的任务计算分数。

[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。

[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。

[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,其中任务分数由优化引擎计算。

[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。

[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,其中一种或多种优化技术使用启发法或近似法。

[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。

[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。

[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,其中工作从客户端服务器接收。

[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,其中工作通过网关资源从客户端服务器接收。

[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: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。

[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。

[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。

[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: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。

[0062] 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。

[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。

[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。

[0065] 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。

[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,其中更新的基础设施被动态地创建。

[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。

[0068] 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。

[0069] 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。

[0070] 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。

[0071] 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。

[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,其中更新的基础设施被动态地创建。

[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。

[0074] 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。

[0075] 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。

[0076] 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。

[0077] 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。

[0078] 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。

[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。

[0080] 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。

[0081] 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。

[0082] 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。

[0083] 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。

[0084] 根据当前公开的主题的某些例子,还提供了计算机节点,其中一种或多种优化技术使用启发法或近似法。

[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。

[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。

[0087] 根据当前公开的主题的某些例子,还提供了计算机节点,其中存储相关资源分布在至少两个位置当中。

[0088] 根据当前公开的主题的某些例子,还提供了计算机节点,其中配置包括过度调拨存储系统资源供使用。

[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。

[0090] 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。

[0091] 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。

[0092] 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。

[0093] 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。

[0094] 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。

[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。

[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。

[0097] 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。

[0098] 根据当前公开的主题的某些例子,还提供了一种方法,其中配置包括过度调拨存储系统资源供使用。

[0099] 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。

[0100] 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。

[0101] 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。

[0102] 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。

[0103] 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。

[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,其中一种或多种优化技术使用启发法或近似法。

[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。

[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。

[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,其中存储相关资源分布在至少两个位置当中。

[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,其中配置包括过度调拨存储系统资源供使用。

[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。

[0110] 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。

[0111] 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。

[0112] 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。

[0113] 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。

[0114] 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。

[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,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。

[0116] 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。

[0117] 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。

[0118] 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。

[0119] 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。

[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。

[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。

[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。

[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。

[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。

[0125] 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。

[0126] 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。

[0127] 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。

[0128] 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。

[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。

[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。

[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。

[0132] 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。

[0133] 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。

[0134] 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。

[0135] 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。

[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。

[0137] 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。

[0138] 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。

附图说明

[0139] 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:

[0140] 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；

[0141] 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；

[0142] 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；

[0143] 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；

[0144] 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；

[0145] 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；

[0146] 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；

[0147] 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；

[0148] 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；

[0149] 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；

[0150] 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；

[0151] 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；

[0152] 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；

[0153] 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；

[0154] 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；

[0155] 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；

[0156] 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；

[0157] 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；

[0158] 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；

[0159] 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；

[0160] 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；

[0161] 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

[0162] 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。

具体实施方式

[0163] 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。

[0164] 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。

[0165] 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由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。

[0166] 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。

[0167] 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。

[0168] 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。

[0169] 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。

[0170] 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)、SSD213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SSD213)等;(c)网络资源214;以及(d)向DSS200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。

[0171] 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。

[0172] 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。

[0173] 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iSCSI222等)到一个或多个计算机节点205。

[0174] 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)。

[0175] 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DSS200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DSS200利用,因而实现DSS200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DSS200利用,因而实现DSS200的可量测性。应注意,在本文特别关于图5更详细描述关于将新资源(特别包括计算机节点205)连接到DSS200的过程的更详细的解释。

[0176] 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DSS200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DSS200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)。

[0177] 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DSS200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DSS200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。

[0178] 在一些情况下,UDSP层202可包括UDSP225,其包括DSS200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DSS200(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。

[0179] 测量系统可使用户能够执行与DSS200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DSS200的动态行为,监测SLS顺应性,涉及DSS200的不同范围(例如,在整个DSS200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、趋势报告和各种参数——包括关键性能指示器——的预测报告等)的产生,管理DSS200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DSS200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。

[0180] 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。

[0181] 在一些情况下,管理系统可使用户能够给DSS200提供用户定义的存储要求,其定

义指定用户要求DSS200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,RT0——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。

[0182] 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。

[0183] 根据当前公开的主题的一些例子,UDSP225可包括用于分配资源和管理DSS200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DSS200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DSS200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,OBMS100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DSS200(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。

[0184] 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DSS200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DSS200,使DSS200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。

[0185] 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,OBMS100)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,OBMS100)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DSS200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DSS200上的数据的加密/解密的数据加密插件、启用例如存储在DSS200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DSS200上的数据的编索引的内容编索引插件、实现例如存

储在DSS200上的数据的快照的管理的快照管理插件、实现例如存储在DSS200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DSS200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DSS200的管理工作的性能和与用户、客户端服务器218和连接到DSS200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。

[0186] 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DSS200。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,OBMS100)等设置插件相关的要求。

[0187] 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。

[0188] 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DSS200的添加/从DSS200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。

[0189] 在DSS200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DSS200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,OBMS100可特别配置成自动分配和管理基础设施层201中的资源。OBMS100可包括输入模块102、一个或多个处理器104和输出模块106。

[0190] 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DSS200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DSS200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。

[0191] 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DSS200和/或一个或多个逻辑存储实体满足的各种要求。

[0192] 在一些情况下,与DSS200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DSS200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。

[0193] 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DSS200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:

[0194] 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):

[0195] 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;

[0196] 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);

[0197] 3.硬盘驱动器功率消耗;

[0198] 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)。

[0199] b.计算机节点205的参数:

[0200] 1.CPU和每CPU核心的数量;

[0201] 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;

[0202] 3.结构(例如,CPU和/或核心支持64位计算?它是小endian还是大endian?);

[0203] 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;

[0204] 5.可用的硬盘驱动器插槽的数量;

[0205] 6.可用的存储接口(SATA、SAS等);

[0206] 7.最大存储量;

[0207] 8.所支持的存储器配置;

[0208] c.高速缓存资源参数:

[0209] 1.高速缓存资源类型(例如,DRAM、SSD)、尺寸和性能;

[0210] 2.高速缓存存储空间是本地的还是远程的?

[0211] 3.NUMA参数。

[0212] d.网关资源参数:

[0213] 1.CPU和每CPU核心的数量;

[0214] 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;

[0215] 3.结构(例如,CPU和/或核心支持64位计算?它是小endian还是大endian?);

[0216] 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;

[0217] 5.外壳中的可用硬盘驱动器插槽的数量;

[0218] 6.可用的存储接口(SATA、SAS等);

[0219] 7.最大存储量;

[0220] 8.所支持的存储器配置;

[0221] 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)。

[0222] e.网络资源参数:

[0223] 1.切换和路由能力;

[0224] 2.网络类型;

[0225] 3.安全参数。

[0226] 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。

[0227] 在一些情况下,与DSS200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DSS200的资源的各种参数,包括例如硬件资源,例如:

[0228] 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):

[0229] 1.硬盘驱动器自由空间。

[0230] 2.硬盘驱动器的S.M.A.R.T.参数。

[0231] 3.硬盘驱动器的功率状态(关闭、在启动阶段中、准备等)。

[0232] 4.硬盘驱动器上的最近和当前负荷。

[0233] 5.现有的分配和保留。

[0234] b.计算机节点205的参数:

[0235] 1.对每个核心的最近和当前的负荷统计。

[0236] 2.现有的分配和保留。

[0237] 3.当前存储量。

[0238] c.高速缓存资源参数:

[0239] 1.可用尺寸。

[0240] 2.高速缓存器的占用水平。

[0241] 3.最近和当前的交换/页面故障统计。

[0242] 4.现有的分配和保留。

[0243] d.网关资源参数:

[0244] 1.最近和当前的网络连接统计。

[0245] 2.最近和当前的节点负荷统计。

[0246] 3.最近和当前的时延统计。

[0247] 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)。

[0248] 5.现有的分配和保留。

[0249] e.网络资源参数:

[0250] 1.网络段的最近和当前负荷。

[0251] 2.网络段的最近和当前的可靠性和质量参数。

[0252] 3.现有的分配和保留。

[0253] 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。

[0254] 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DSS200(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OBMS100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。

[0255] 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DSS200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。

[0256] 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DSS200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。

[0257] 应注意,在一些情况下,当首次部署DSS200和/或一个或多个逻辑存储实体时和/或在施加到DSS200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DSS200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DSS200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。

[0258] 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。

[0259] 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)。

[0260] 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。

[0261] 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)。

[0262] 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变性、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)。

[0263] 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。

[0264] 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)。

[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)。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)。

[0266] 在一些情况下,输出模块可比较所需的配置与DSS200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DSS200是否可满足所需配置(块118)。应注意,在一些情况下,实际DSS200的资源可以涉及DSS200的资源的当前可用的那些部分。如果实际DSS200的资源和/或环境可满足所需的配置,则OBMS100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,OBMS100可配置成建立DSS200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。

[0267] 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DSS200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DSS200的过程成功地结束(块130)。

[0268] 如果实际DSS200的资源和/或环境不能满足所需的配置,OBMS100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,OBMS100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,OBMS100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DSS200的过程失败。在一些情况下,OBMS100可配置成报告故障。

[0269] 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。

[0270] 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。

[0271] 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的资源和/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。

[0272] 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DSS200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)。

[0273] 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DSS200或其部分的各种参数有关的数据和/或与DSS200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DSS200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DSS200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。

[0274] 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DSS200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。

[0275] 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。

[0276] 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DSS200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。

[0277] 工作创建模块345可配置成创建用于在DSS200中执行的新工作,如特别关于图8和9进一步详述的。

[0278] 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。

[0279] 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。

[0280] 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。

[0281] 云即插即用模块380可配置成实现计算机节点205到DSS200的自主和/或自动连接,如特别关于图14进一步详述的。

[0282] 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。

[0283] 基于目标的配置模块390可配置成配置和/或重新配置DSS200,如特别关于图2-4和11进一步详述的。

[0284] 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。

[0285] 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。

[0286] 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。

[0287] 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。

[0288] 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DSS200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。

[0289] 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)。

[0290] 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。

[0291] 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)。

[0292] 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。

[0293] 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。

[0294] 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DSS200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)。

[0295] 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)。

[0296] 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。

[0297] 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)。

[0298] 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)。

[0299] 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。

[0300] 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。

[0301] 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。

[0302] 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)。

[0303] 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)。

[0304] 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DSS200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)。

[0305] 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。

[0306] 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。

[0307] 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)。

[0308] 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)。

[0309] 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。

[0310] 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。

[0311] 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DSS200的操作期间。

[0312] 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。

[0313] 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)。

[0314] 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。

[0315] 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DSS200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。

[0316] 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DSS200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。

[0317] 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。

[0318] 应注意,可对连接到DSS200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。

[0319] 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。

[0320] 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。

[0321] 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)。

[0322] 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。

[0323] 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。

[0324] 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。

[0325] 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。

[0326] 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)。

[0327] 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)。

[0328] 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)。

[0329] 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。

[0330] 在一些情况下,执行模块335可配置成检查是否有检查当前DSS200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。

[0331] 在需要检查DSS200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。

[0332] 在不需要检查DSS200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)。

[0333] 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)。

[0334] 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。

[0335] 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。

[0336] 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。

[0337] 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。

[0338] 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。

[0339] 在一些情况下,基于目标的配置模块390可配置成接收与DSS200中的一个或多个逻辑存储实体相关的SLS、指示DSS200及其资源和环境的动态行为的数据、指示DSS200的当前配置的数据、与DSS200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。

[0340] 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DSS200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)。

[0341] 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),DSS200的重新配置可被发起,在其它情况下DSS200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DSS200有关的统计数据 and 历史数据。

[0342] 在需要重新配置DSS200的情况下,基于目标的配置模块390可配置成启动用于执行DSS200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DSS200的重新配置的情况下,OBMS100可接收DSS200的当前配置作为配置过程的输入的部分并在重新配置DSS200时考虑它。在一些情况下,在这样的重新配置期间,OBMS100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。

[0343] 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任何其它原因,或在DSS200的重新配置的发起之后,重新配置过程900结束(块940)。

[0344] 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。

[0345] 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。

[0346] 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。

[0347] 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。

[0348] 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)。

[0349] 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。

[0350] 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)。

[0351] 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DSS200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。

[0352] 在需要检查DSS200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。

[0353] 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。

[0354] 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。

[0355] 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。

[0356] 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)。

[0357] 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)。

[0358] 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DSS200的任何其它位置上和/或在任何辅助实体上。

[0359] 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)。

[0360] 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DSS200(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)。

[0361] 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。

[0362] 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)。

[0363] 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DSS200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)。

[0364] 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。

[0365] 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。

[0366] 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。

[0367] 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DSS200的标识符的答复来作出响应(每个DSS200可具有实现其识别的唯一标识符)。

[0368] 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DSS200,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。

[0369] 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DSS200(例如,根据所接收的DSS200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DSS200(块1325)。应注意,作为连接DSS200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。

[0370] 如果多于一个DSS200被探测到(例如,多于一个DSS200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DSS200是否存在(块1330)。为了这个目的,在一些情况下,默认DSS200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DNS200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DSS200的指示。应注意,也可使用用于识别默认DSS200的其它方法和技术。

[0371] 如果这样的默认DSS200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DSS200(块1325)。如果没有默认DSS200被探测到,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DSS200,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DSS200(块1325)。

[0372] 在一些情况下,当探测到新网络连接(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DSS200和/或默认DSS200的识别。

[0373] 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。

[0374] 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。

[0375] 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DSS200,且本文提供的方法仅仅是例子。

[0376] 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DSS200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DSS200的连续操作。

[0377] 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。

[0378] 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。

[0379] 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)。

[0380] 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)。

[0381] 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DSS200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。

[0382] 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。

[0383] 在描述了DSS200之后,接下来是用于管理DSS200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、SSD213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)。

[0384] 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DSS200)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。

[0385] 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。

[0386] 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。

[0387] 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。

[0388] 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。

[0389] 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。

[0390] 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。

[0391] 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。

[0392] 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。

[0393] 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)。

[0394] 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。

[0395] 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/或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。

[0396] 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。

[0397] 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。

[0398] 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。

[0399] 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)。

[0400] 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DSS200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)。

[0401] 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。

[0402] 如果存在SLS违反或SLS接近违反或存在涉及DSS200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DSS200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。

[0403] 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。

[0404] 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。

[0405] 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。

[0406] 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。

[0407] 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。

[0408] 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DSS200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。

[0409] 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DSS200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。

[0410] 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。

[0411] 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DSS200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。

[0412] 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。

[0413] 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。

[0414] 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。

[0415] 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。

[0416] 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)。

[0417] 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。

[0418] 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)。

[0419] 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DSS200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。

[0420] 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。

[0421] 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。

[0422] 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)。

[0423] 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。

[0424] 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。

[0425] 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。

[0426] 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。

[0427] 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。

[0428] 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。

[0429] 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。

[0430] 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。

[0431] 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。

[0432] 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)。

[0433] 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。

[0434] 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。

[0435] 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。

[0436] 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。

[0437] 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)。

[0438] 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)。

[0439] 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DSS200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。

[0440] 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。

[0441] 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)。

[0442] 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)。

[0443] 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DSS200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DSS200的每个客户端服务器218(或DSS200的任何其它相关用户)。

[0444] 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)。

[0445] 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)。

[0446] 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。

[0447] 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。

[0448] 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。

[0449] 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)。

[0450] 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DSS200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)。

[0451] 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DSS200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。

[0452] 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)。

[0453] 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)。

[0454] 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。

[0455] 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。

[0456] 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)。

[0457] 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)。

[0458] 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)。

[0459] 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)。

[0460] 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。

[0461] 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。

[0462] 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)。

[0463] 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的当前所有者(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。

[0464] 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)。

[0465] 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。

[0466] 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。

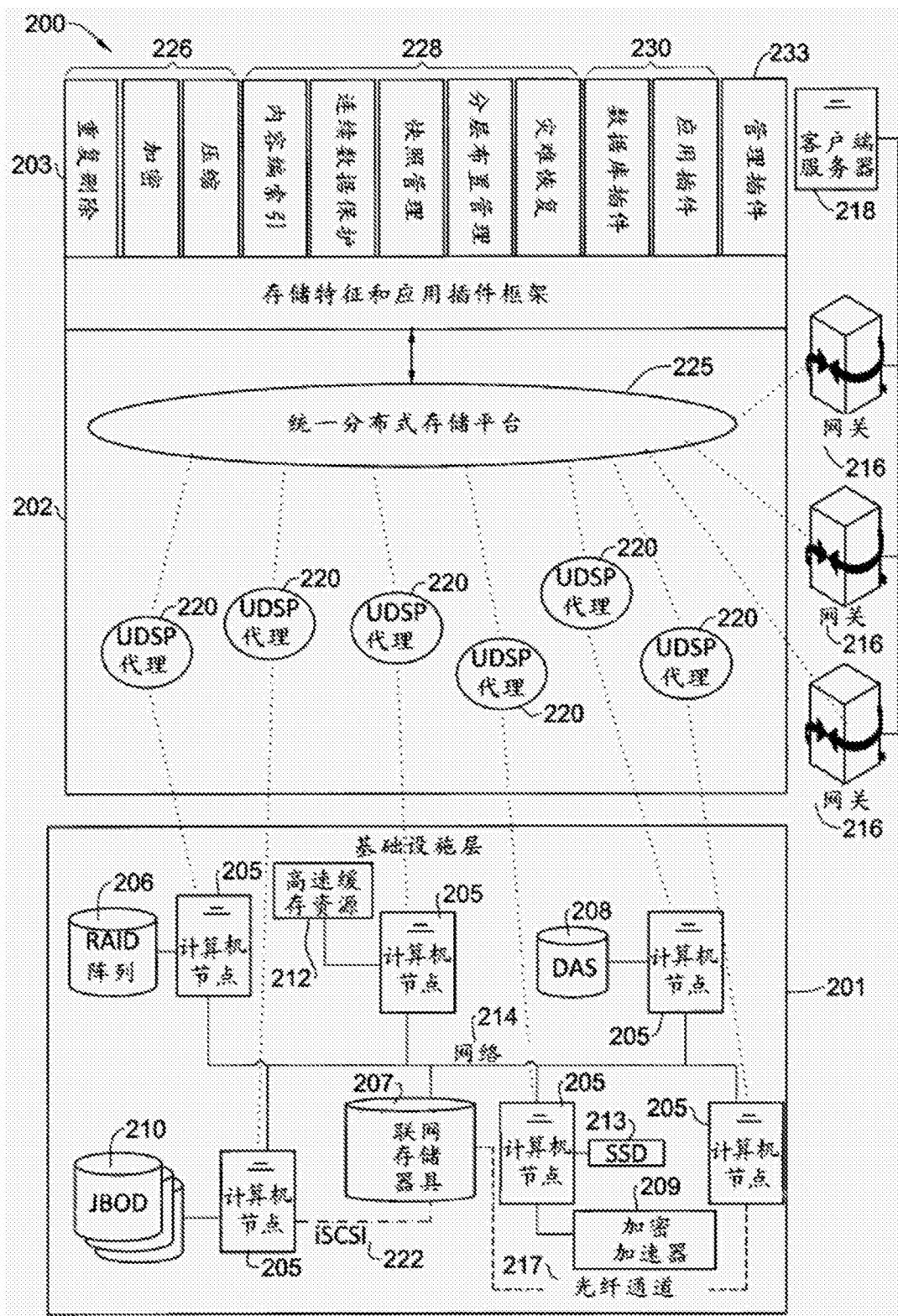


图1

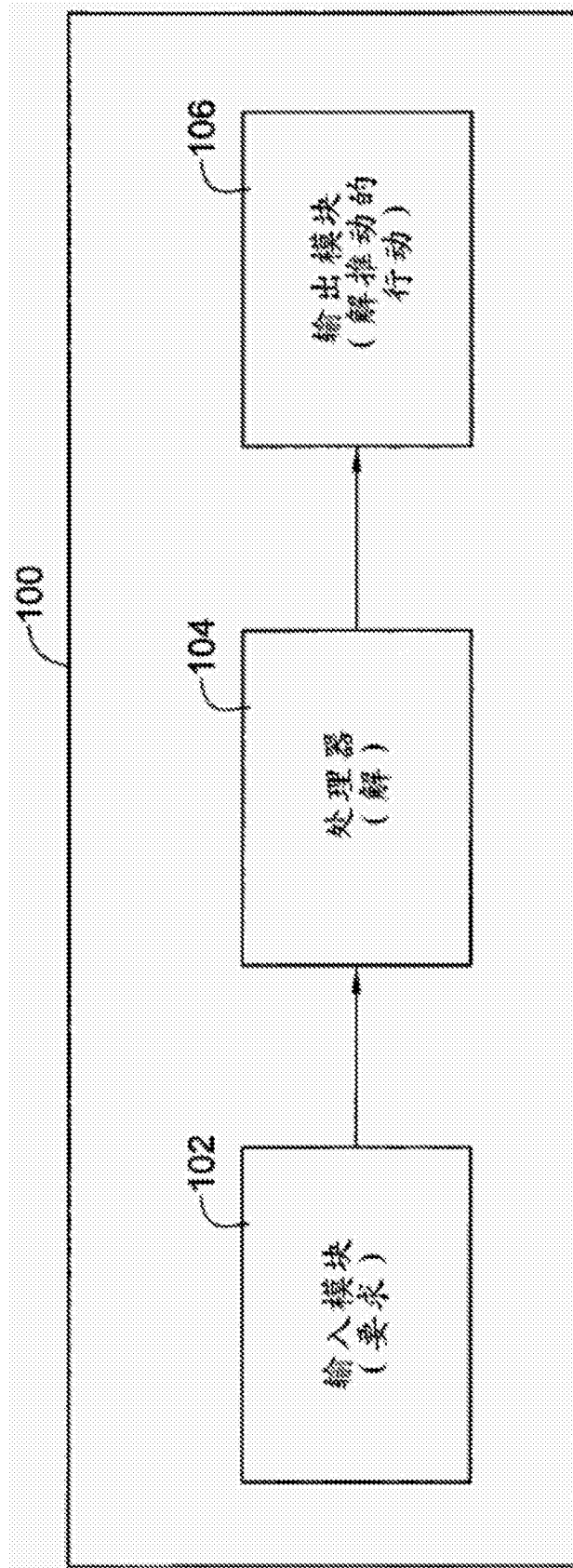


图2

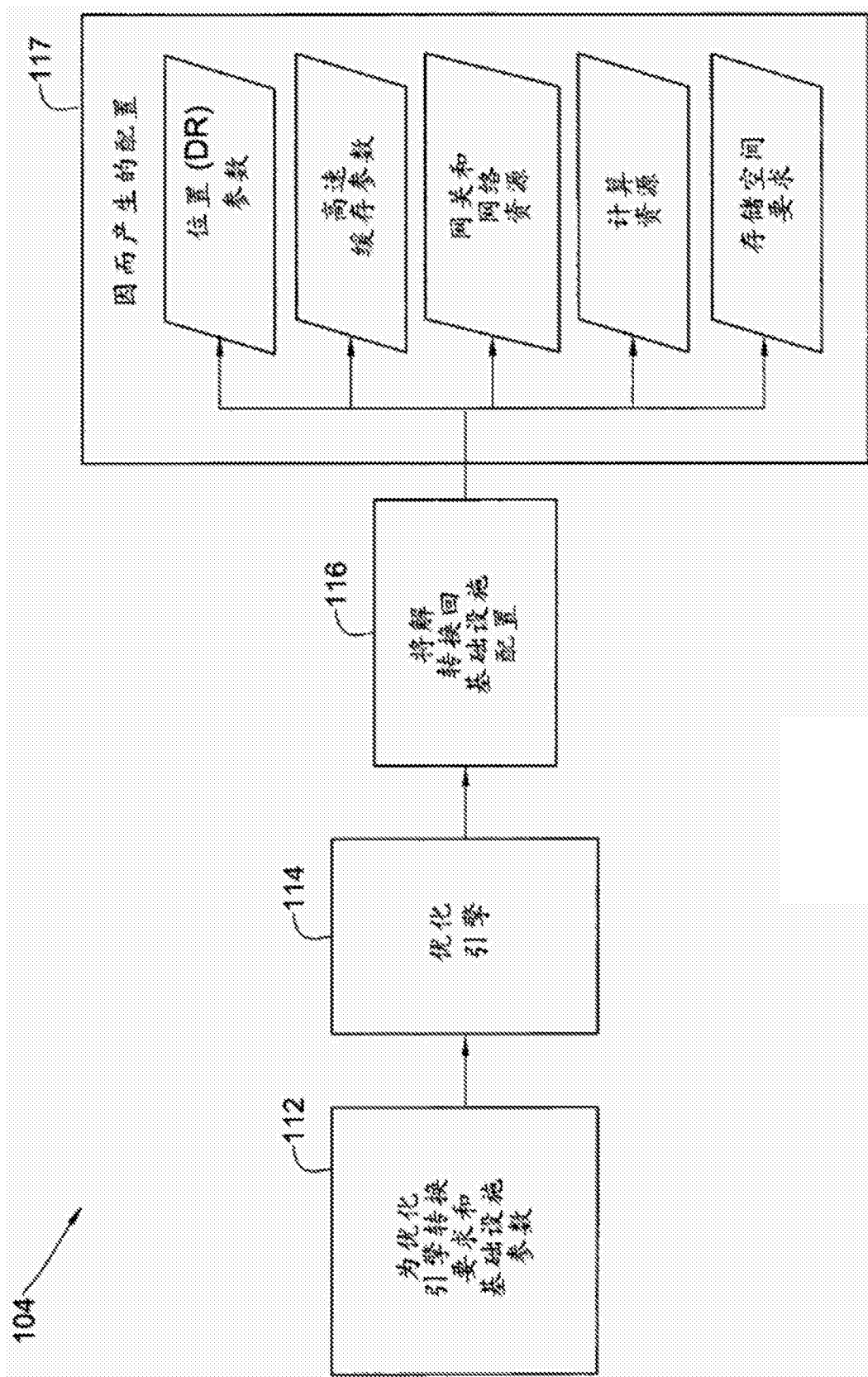


图3

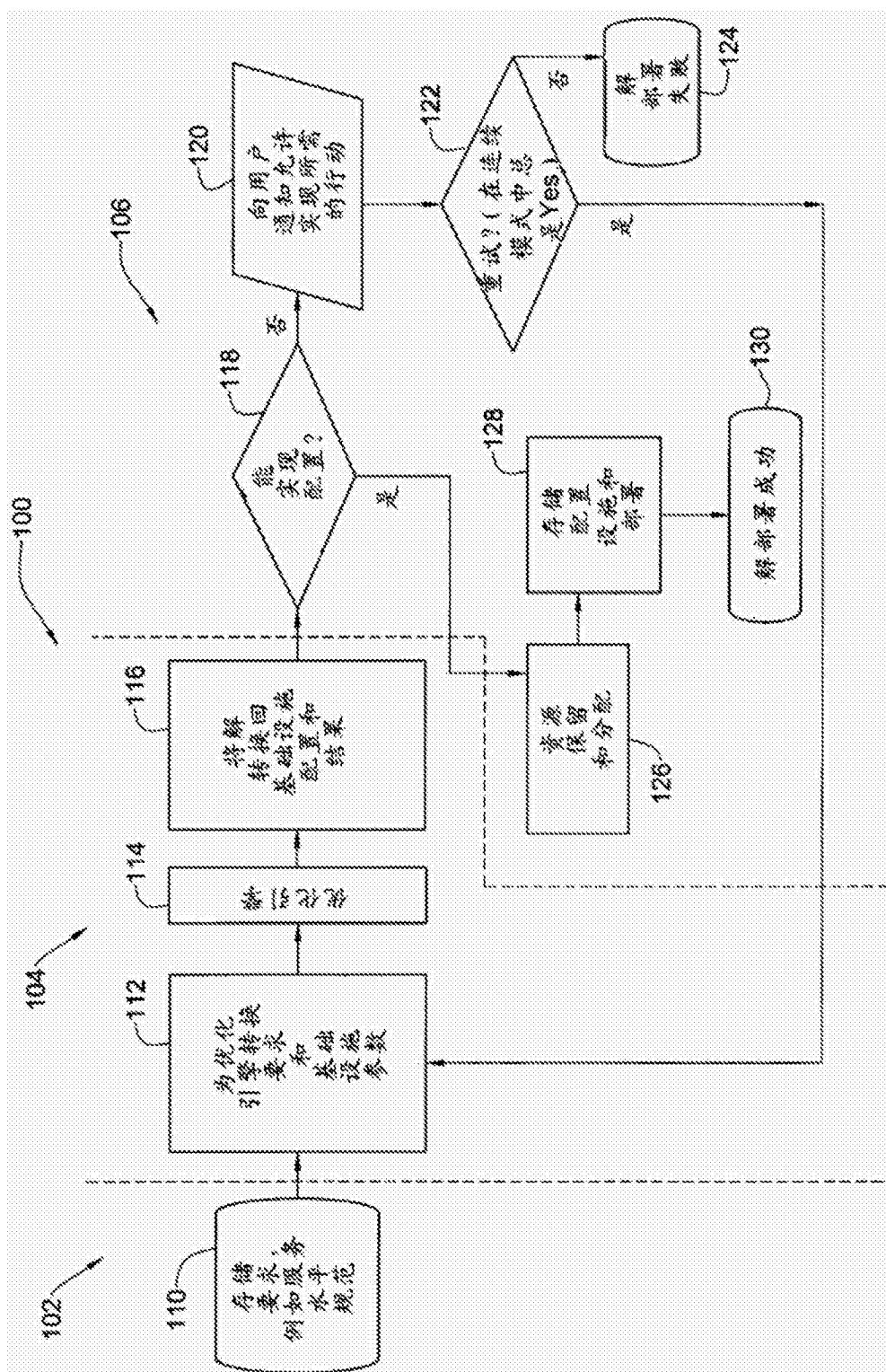


图4

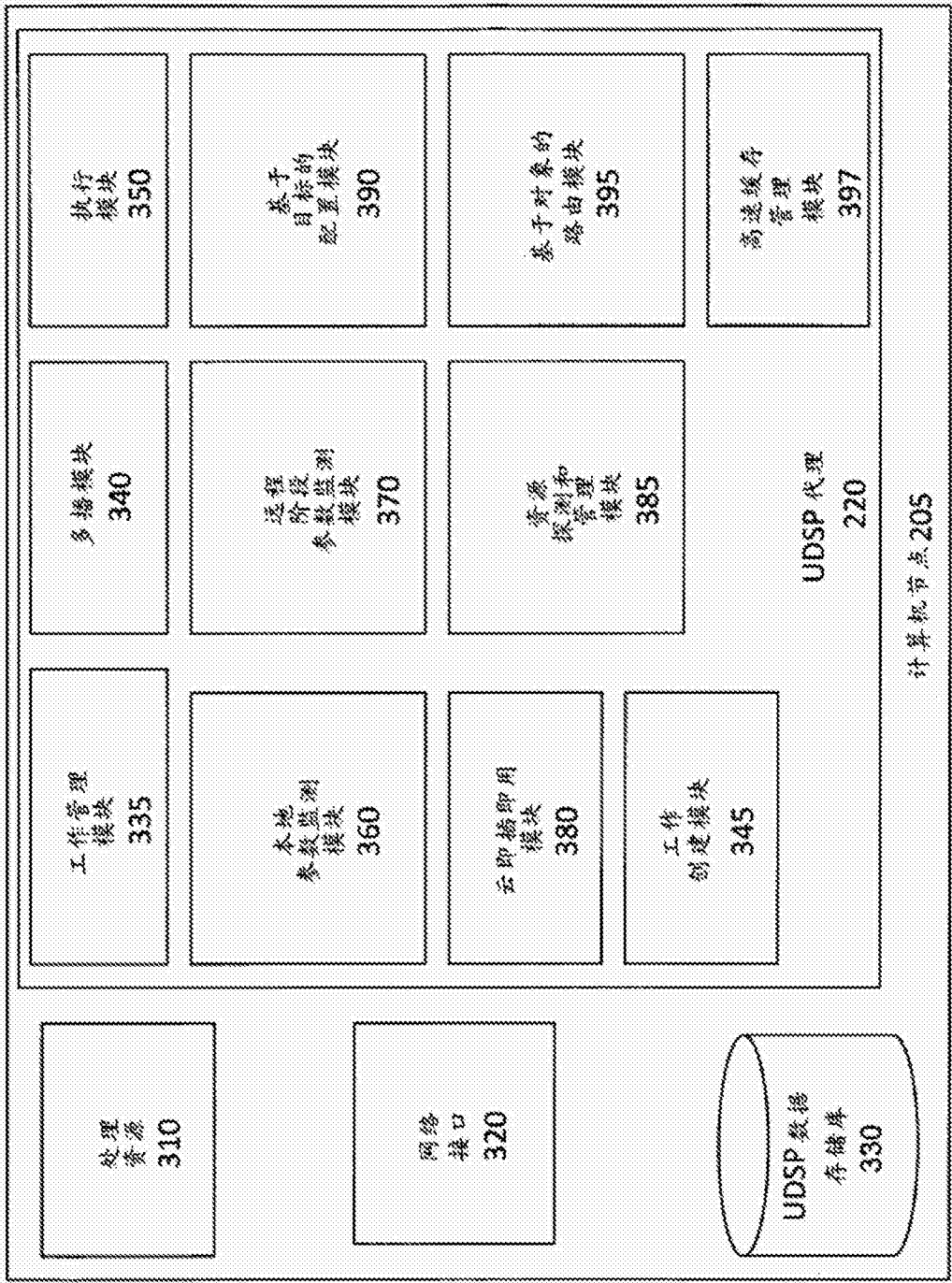


图5

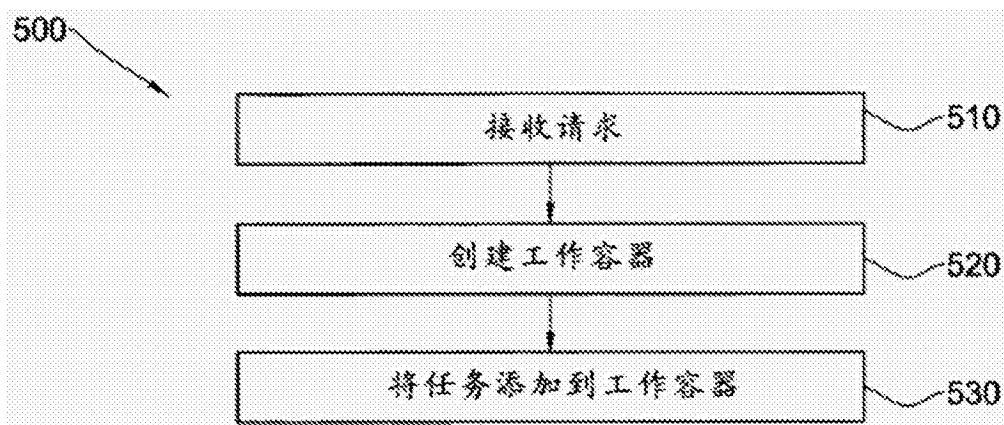


图6

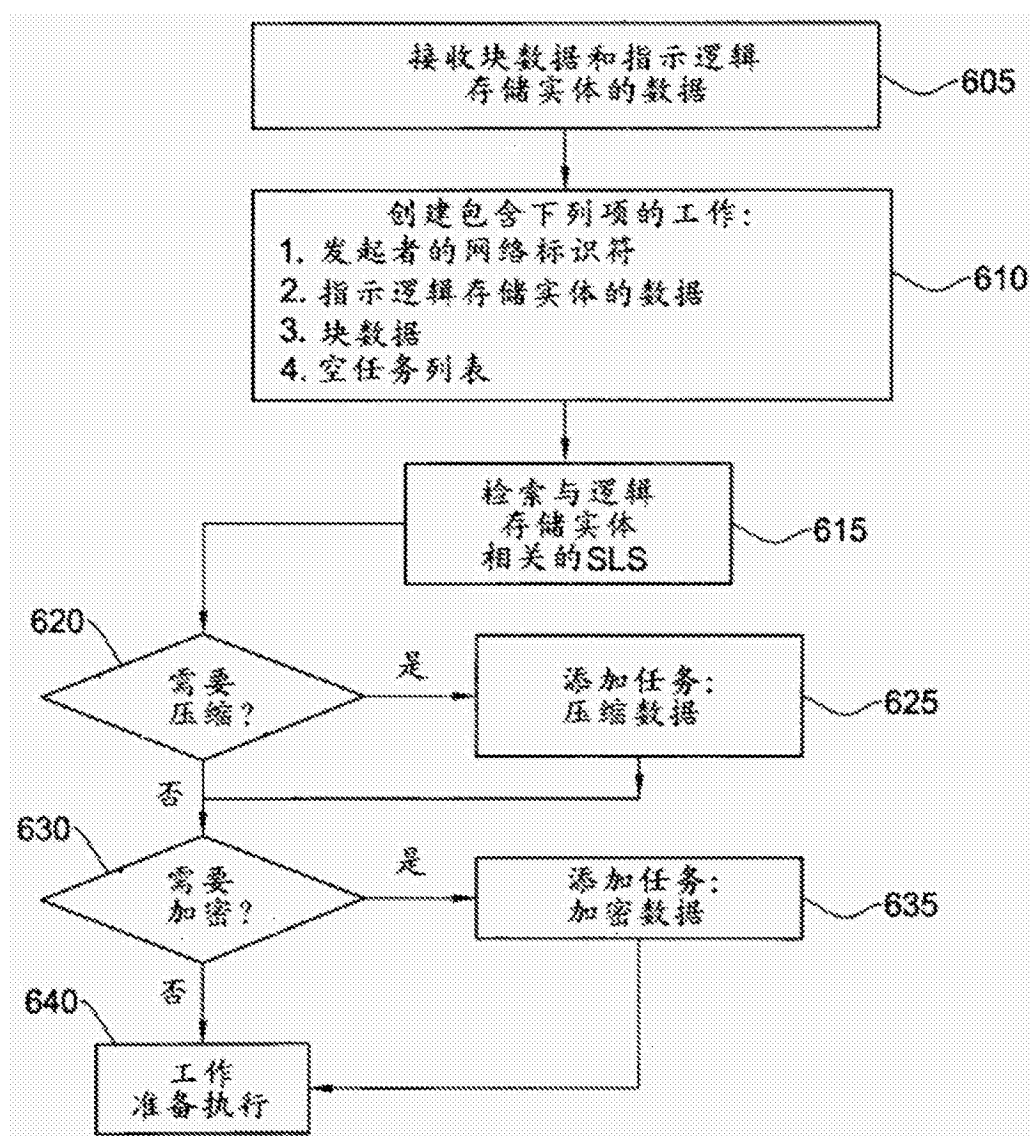


图7

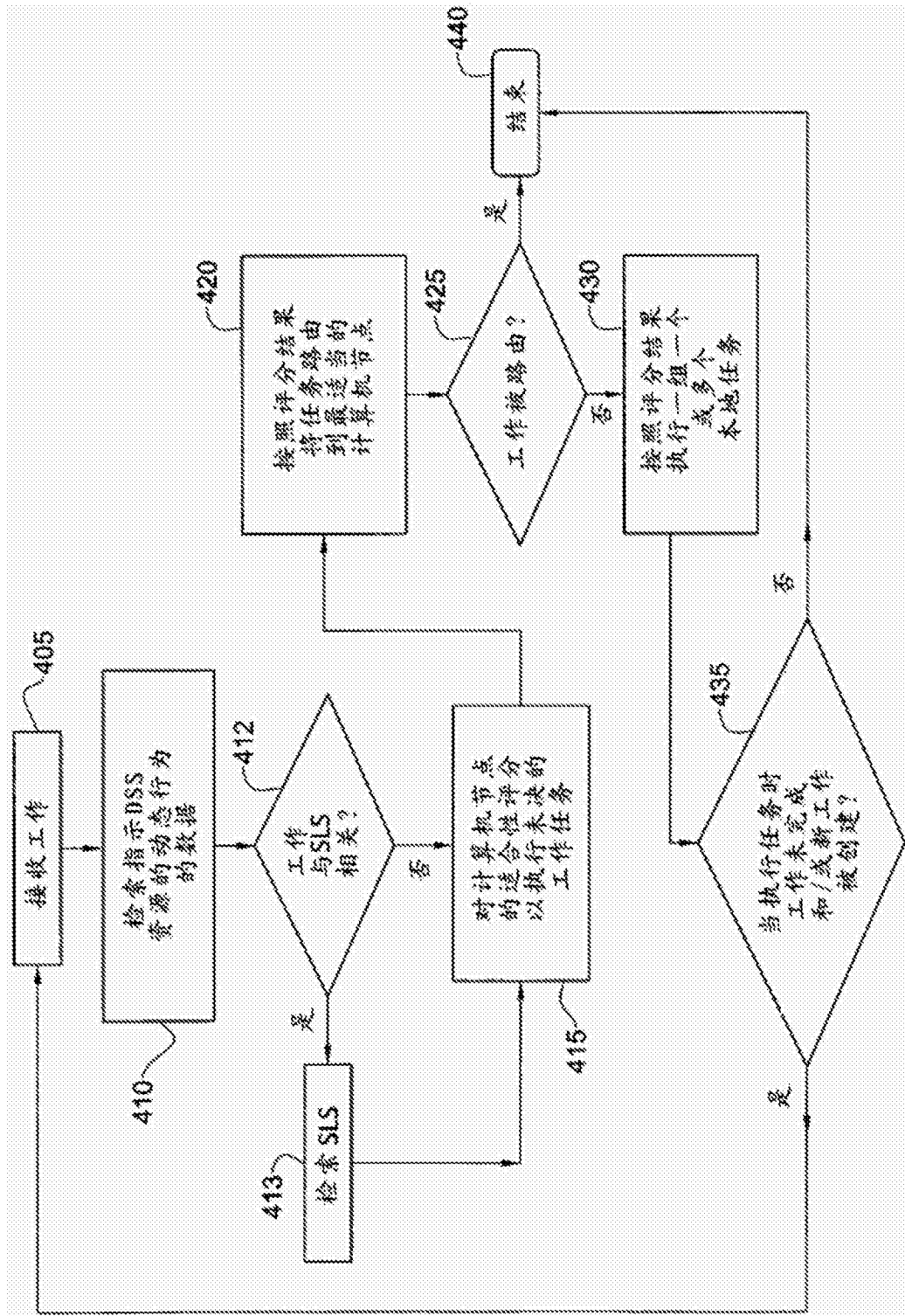


图8

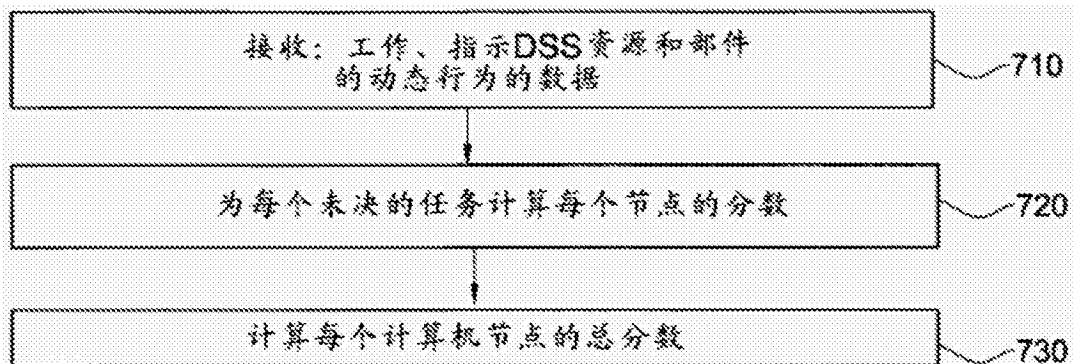


图9

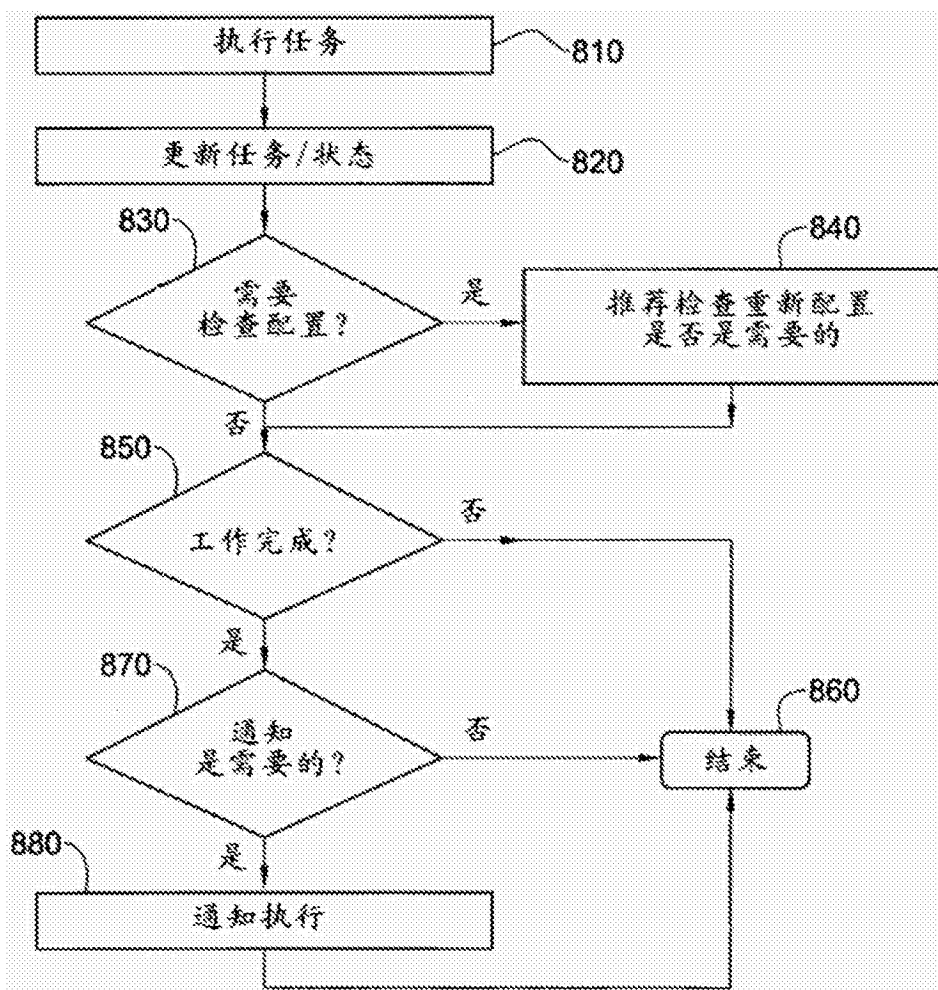


图10

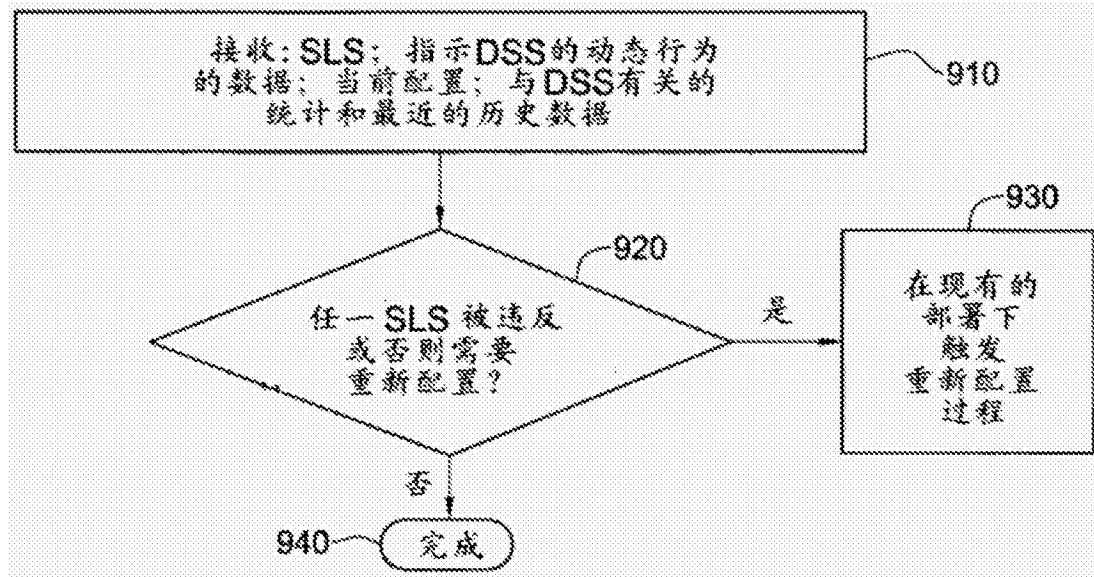


图11

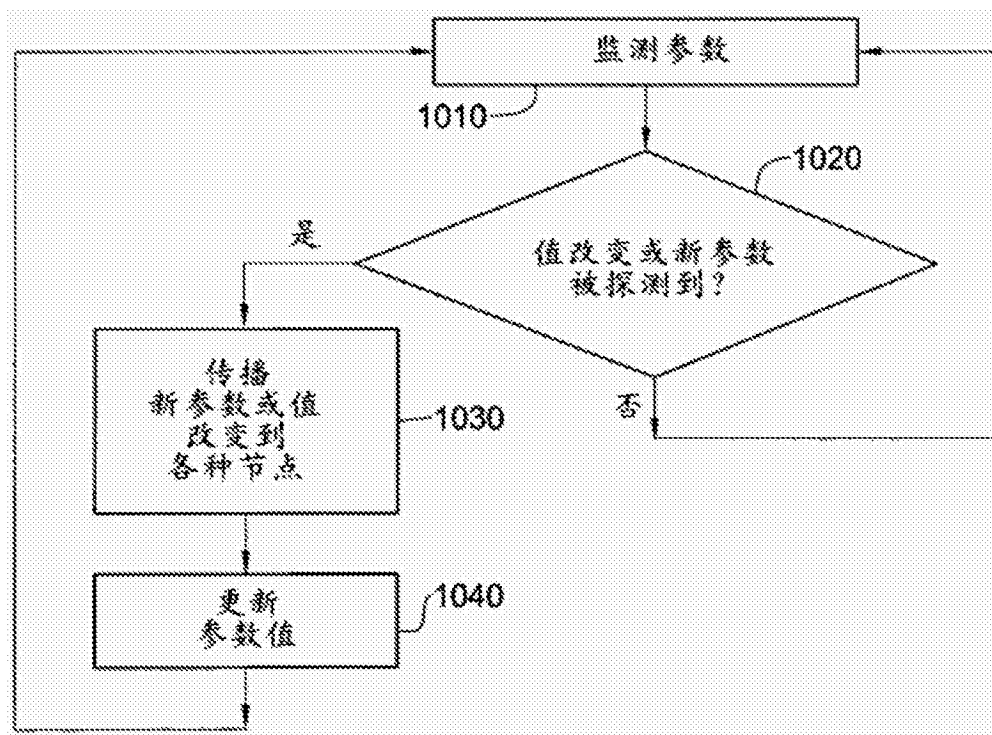


图12

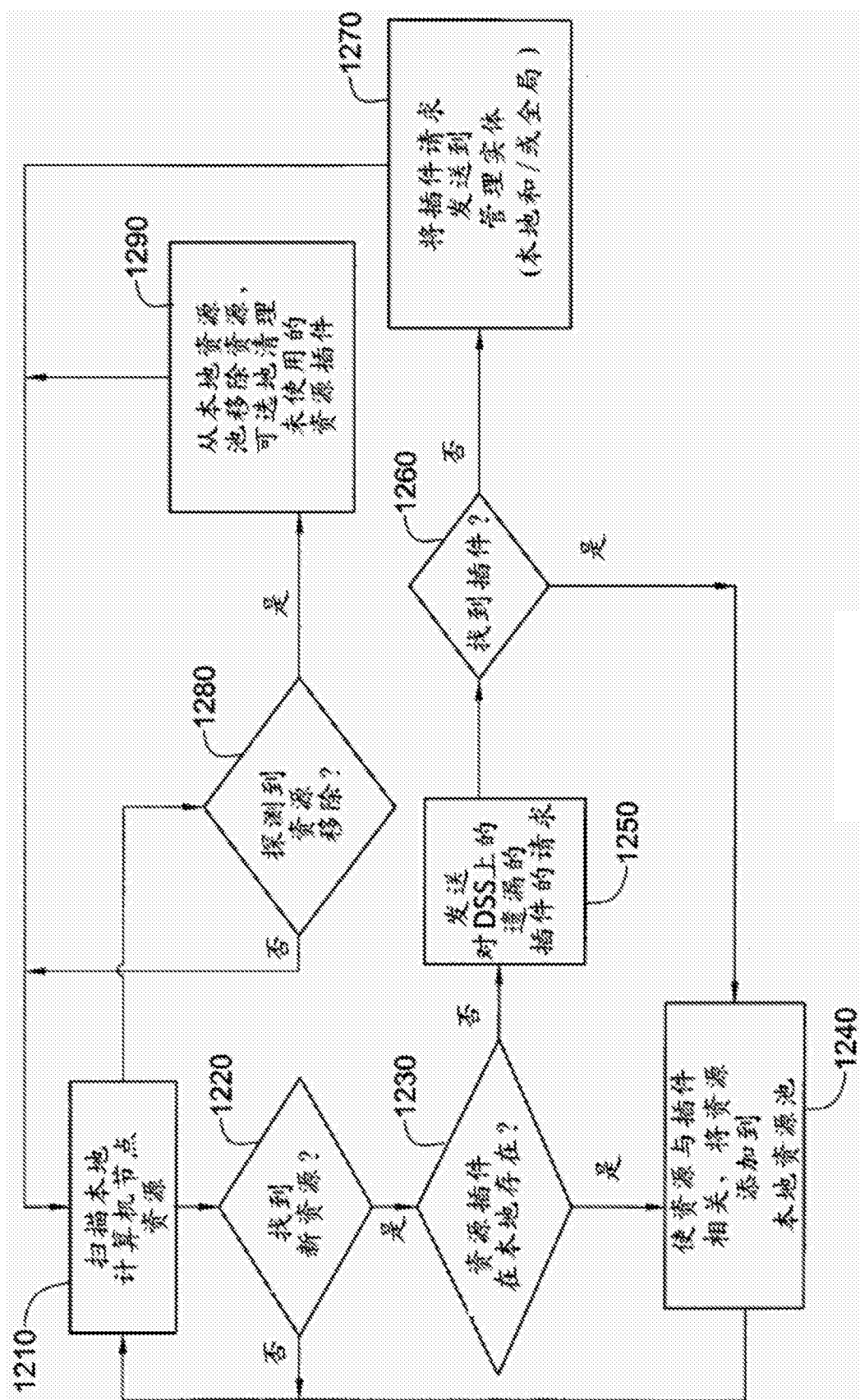


图13

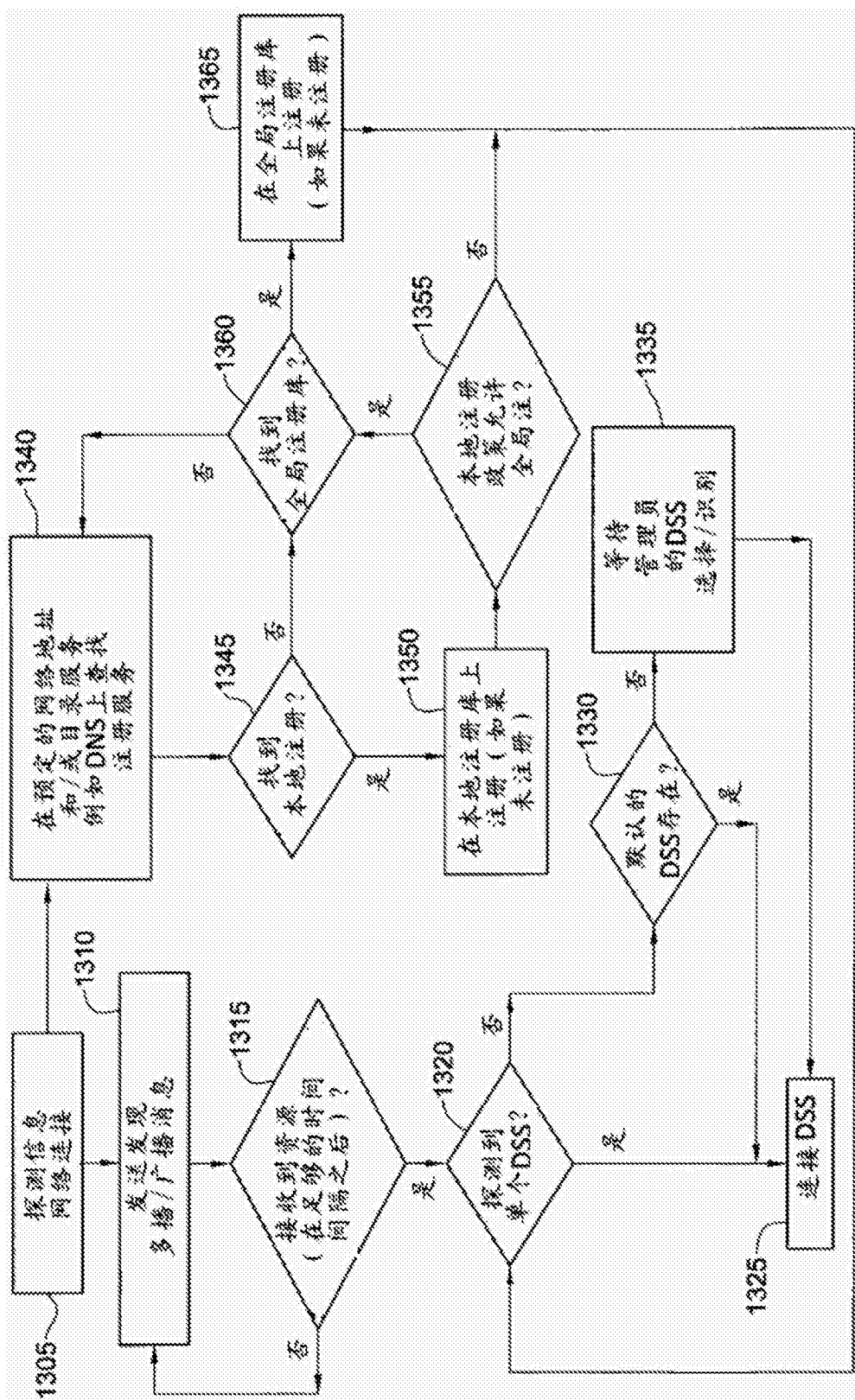


图14

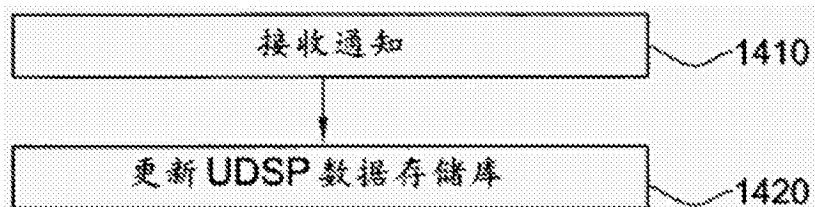


图15

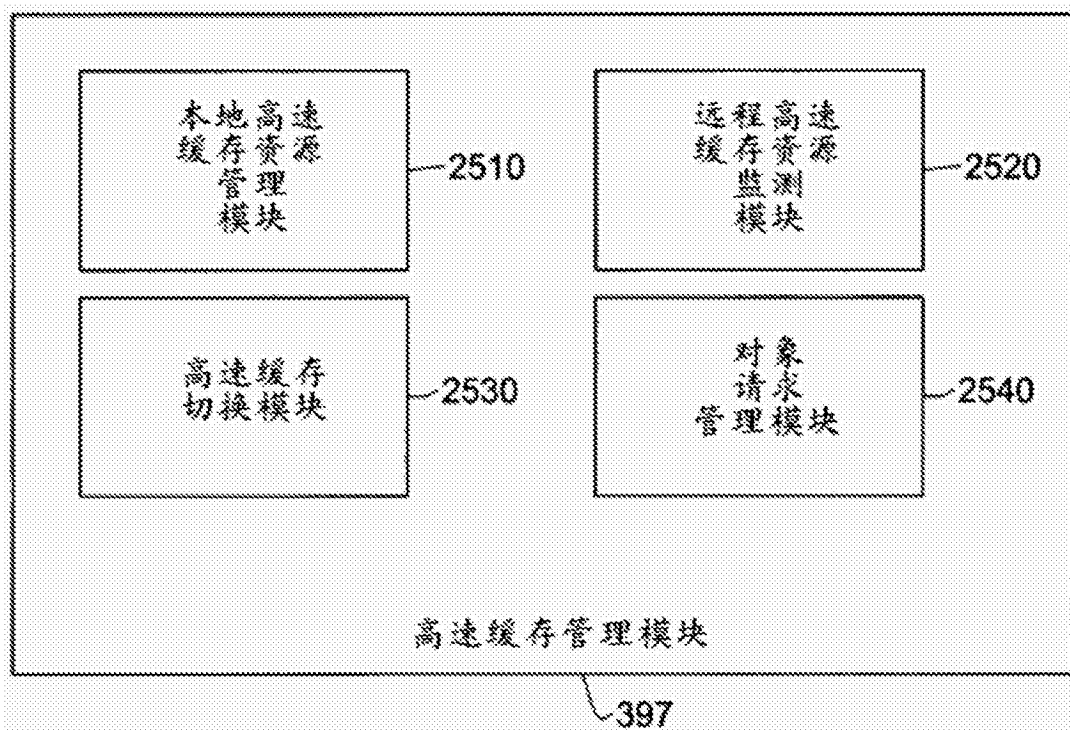


图16

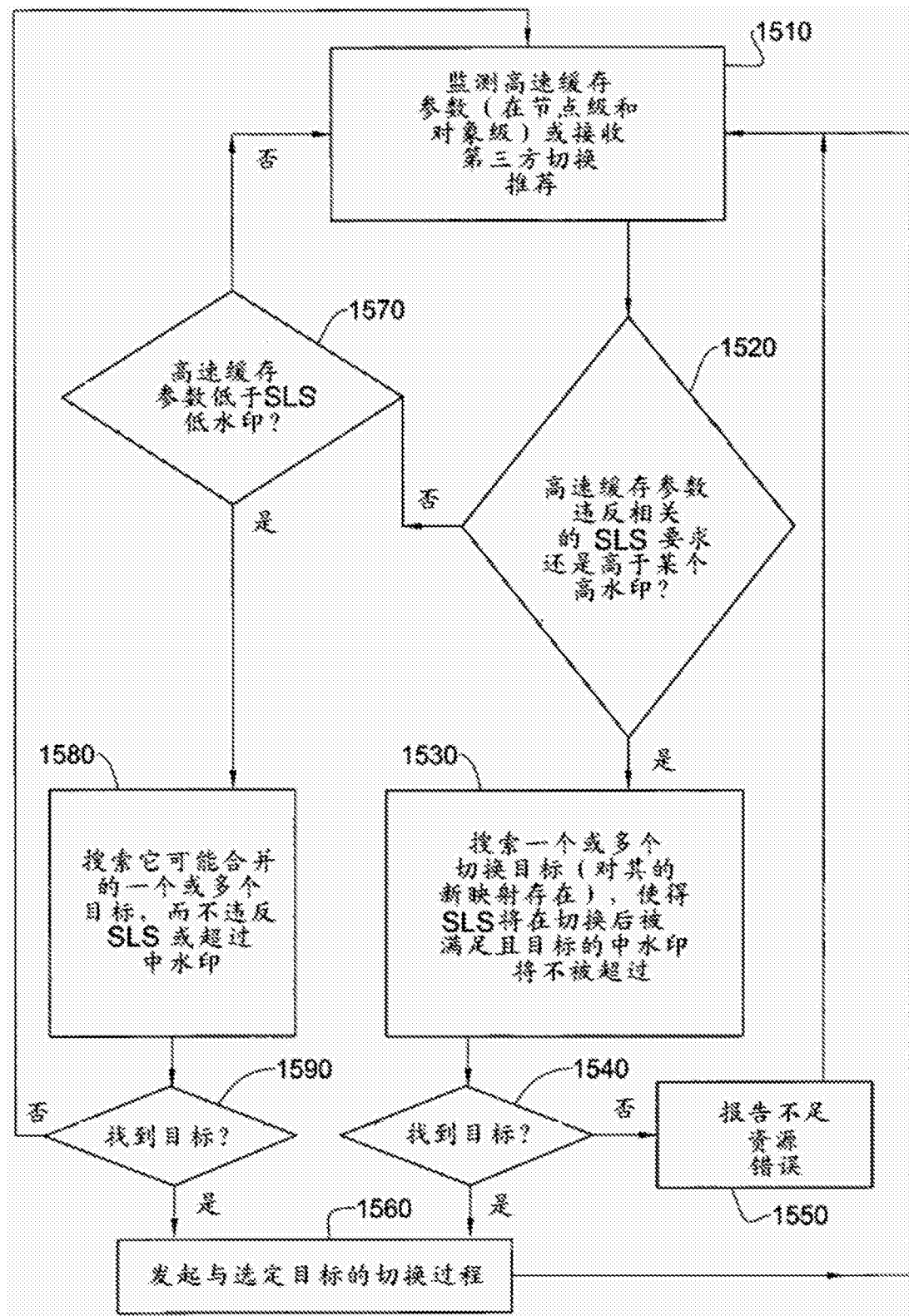


图17

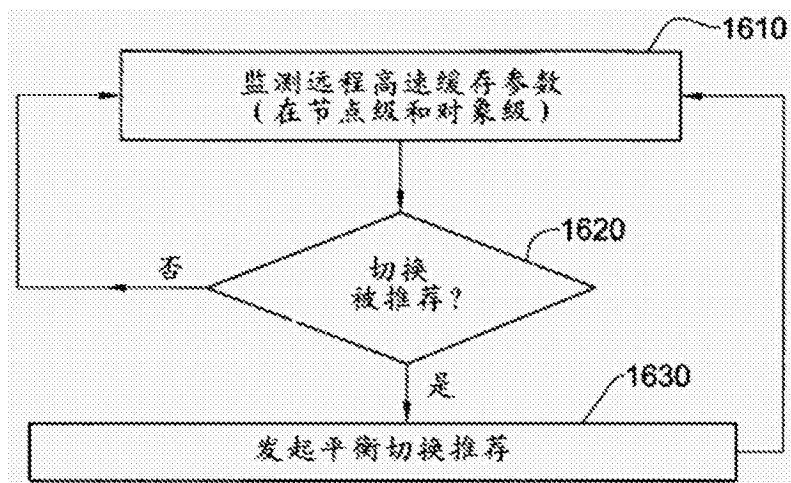


图18

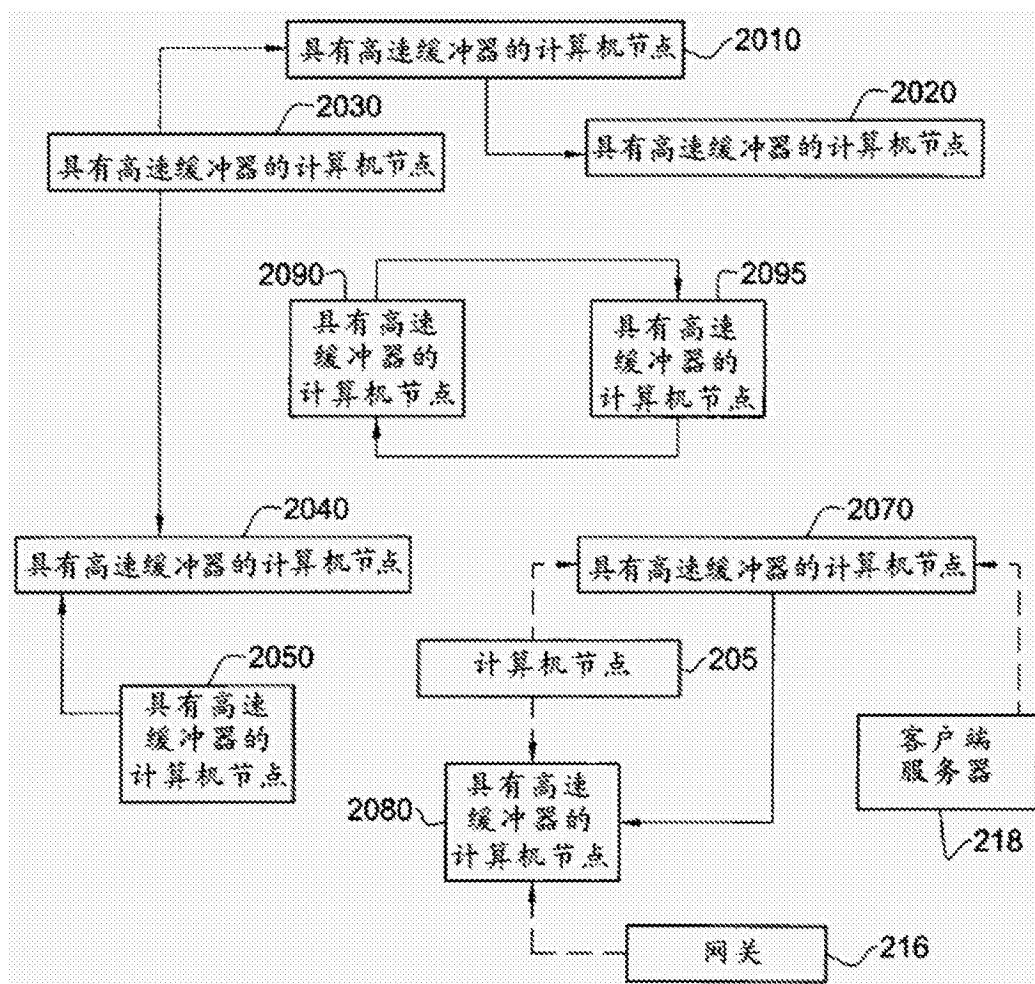


图19

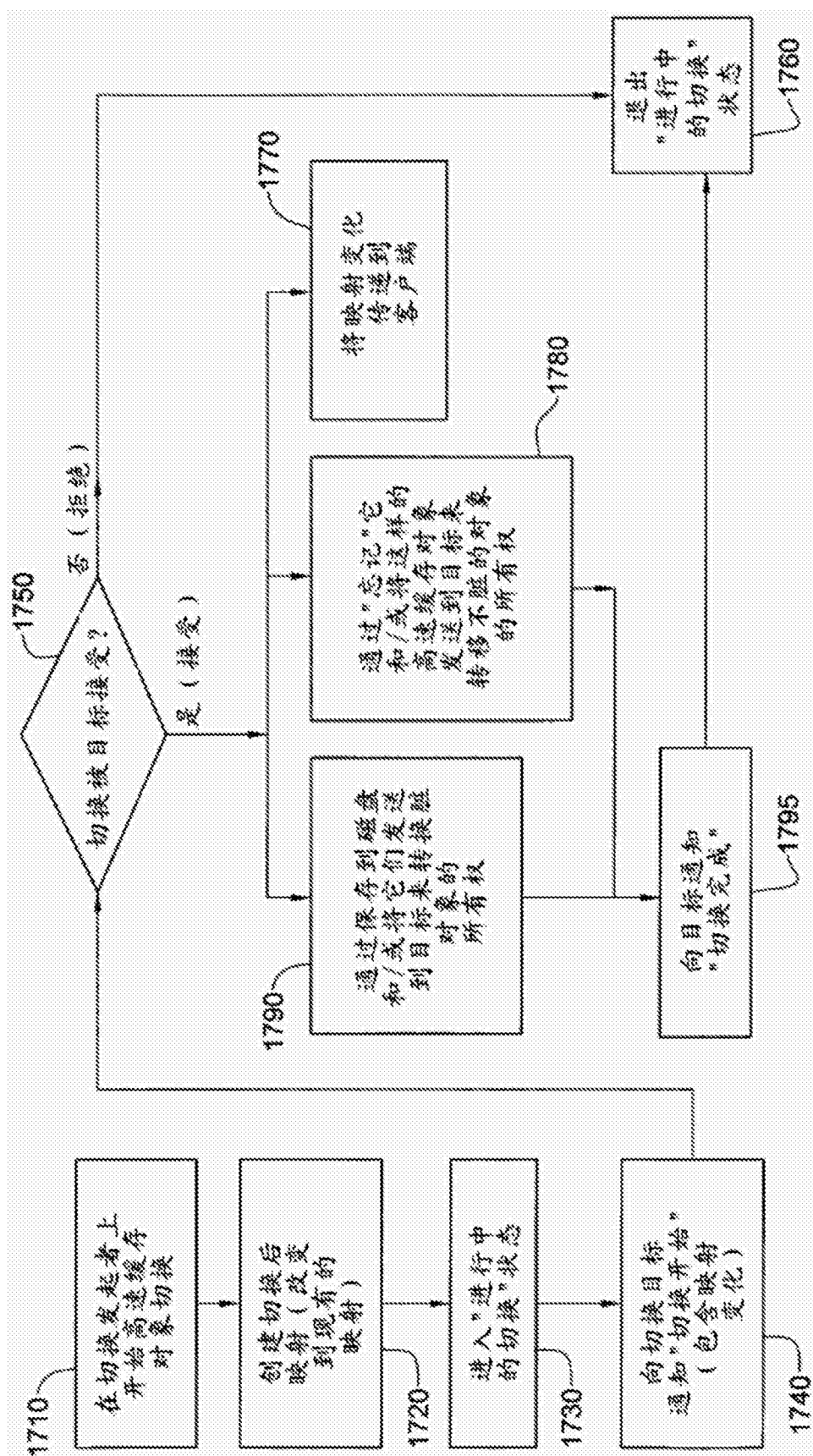


图20

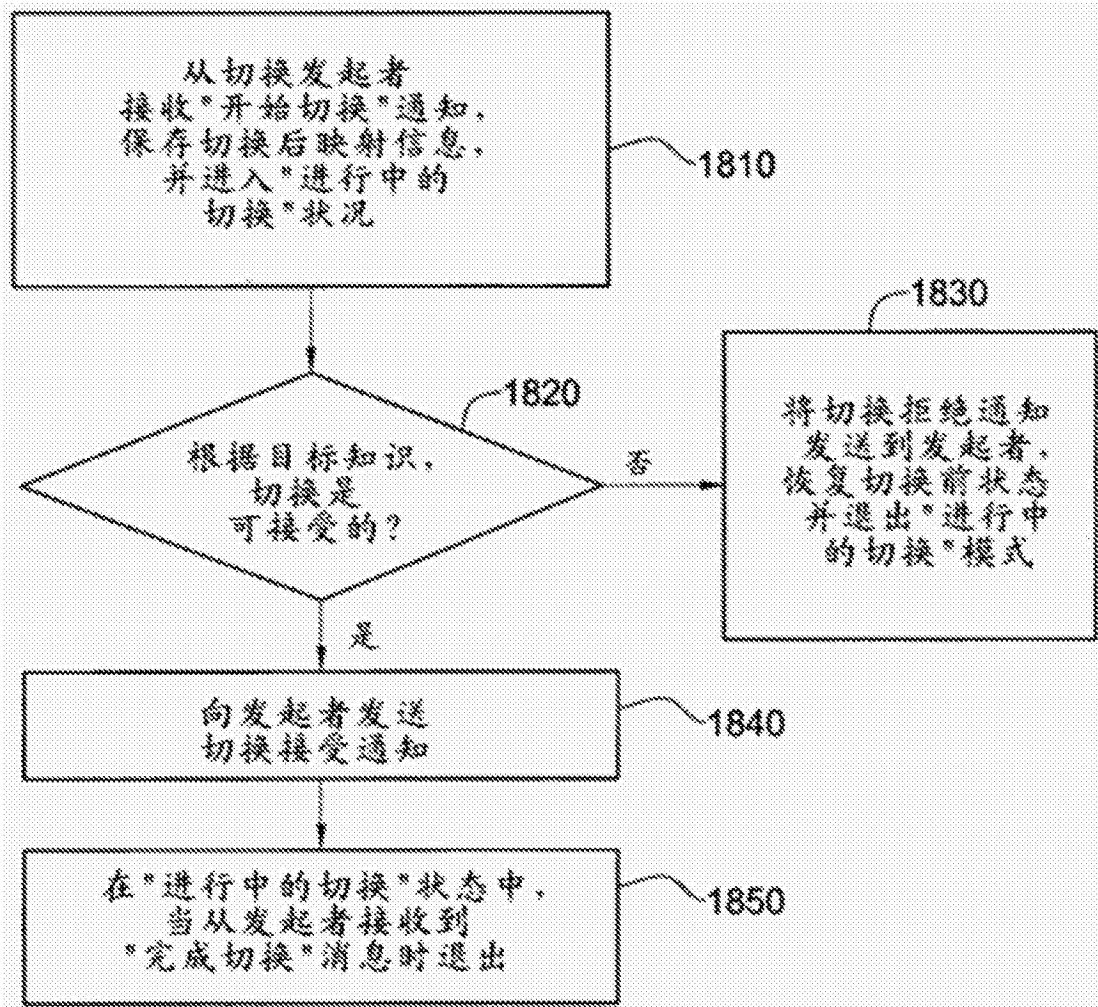


图21

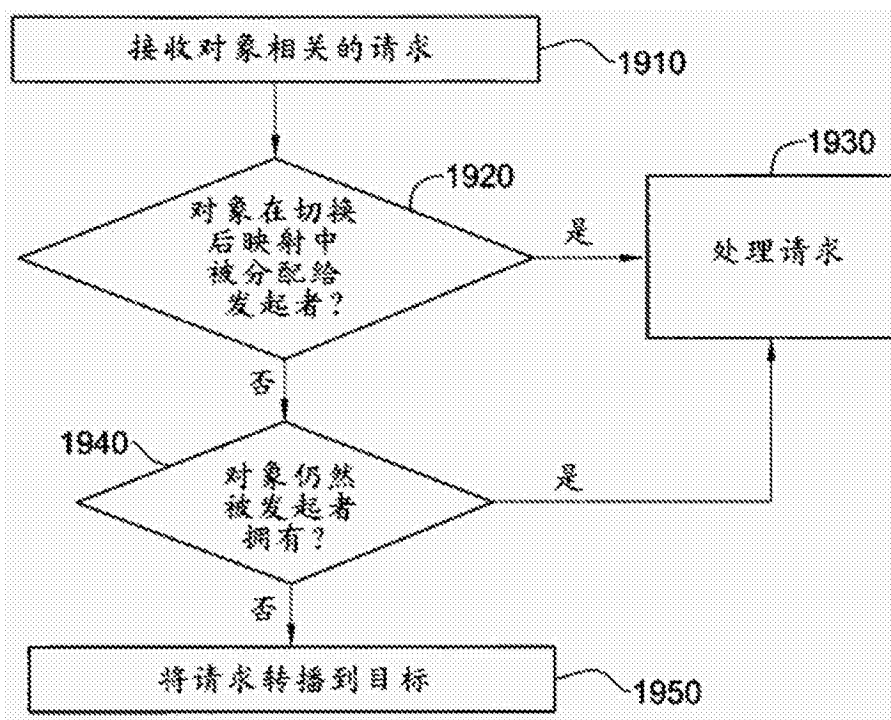


图22

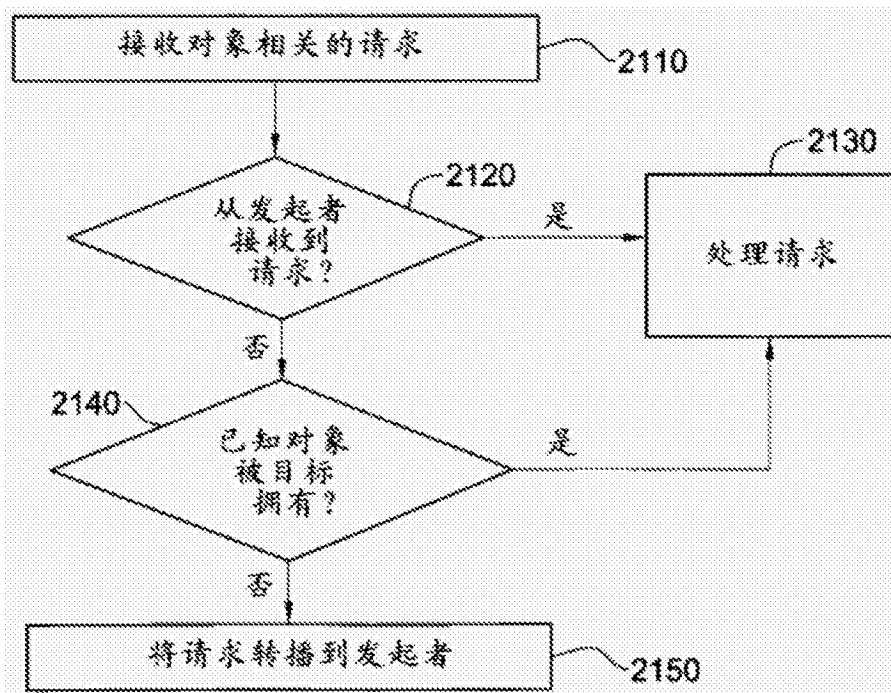


图23