

(12) 发明专利申请

(10) 申请公布号 CN 102651007 A

(43) 申请公布日 2012.08.29

(21) 申请号 201110051574.8

(22) 申请日 2011.02.28

(71) 申请人 国际商业机器公司
地址 美国纽约阿芒克

(72) 发明人 高英明 张凯 邹献 霍嘉

(74) 专利代理机构 北京市金杜律师事务所
11256

代理人 鄞迅 李峥宇

(51) Int. Cl.

G06F 17/30(2006.01)

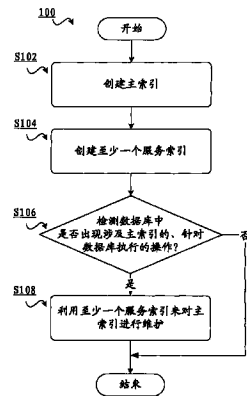
权利要求书 2 页 说明书 10 页 附图 6 页

(54) 发明名称

一种管理数据库索引的方法和装置

(57) 摘要

本发明的各实施方式涉及一种管理数据库索引的方法和装置。在一个实施方式中,提供了一种管理数据库索引的方法,包括:创建主索引;创建至少一个服务索引,其中至少一个服务索引用于记录主索引中将被更新节点的变化;检测数据库中是否出现涉及主索引的、针对数据库执行的操作;当在数据库中出现涉及主索引的、针对数据库执行的操作时,利用至少一个服务索引来对主索引进行维护,其中维护是基于至少一个服务索引中记录的、主索引中将被更新节点的变化进行的。在其他实施方式中,还提供了管理数据库索引的相应装置。



1. 一种管理数据库索引的方法,包括:
创建主索引;
创建至少一个服务索引,其中所述至少一个服务索引用于记录所述主索引中将被更新节点的变化;
检测所述数据库中是否出现涉及所述主索引的、针对所述数据库执行的操作;
响应于在所述数据库中出现涉及所述主索引的、针对所述数据库执行的操作,利用所述至少一个服务索引来对所述主索引进行维护,其中所述维护是基于所述至少一个服务索引中记录的、所述主索引中将被更新节点的变化进行的。
2. 根据权利要求1所述的方法,其中所述至少一个服务索引中的节点至少包括:与所述主索引的节点相关联的节点信息以及针对所述主索引进行的操作类型。
3. 根据权利要求1或2所述的方法,其中针对所述数据库执行的操作包括:针对所述数据库执行的数据操纵操作和数据查询操作中的至少一个。
4. 根据权利要求3所述的方法,其中:
响应于执行所述数据操纵操作,并行地更新所述主索引和所述至少一个服务索引;以及
响应于所述至少一个服务索引和所述主索引中的任一项被成功更新,标记所述数据操纵操作结束。
5. 根据权利要求4所述的方法,其中所述并行地更新所述主索引和所述至少一个服务索引包括:
当所述主索引首先被成功更新后,停止更新所述至少一个服务索引。
6. 根据权利要求4所述的方法,其中所述并行地更新所述主索引和所述至少一个服务索引包括:
当所述至少一个服务索引之一首先被成功更新后,
停止更新尚未被成功更新的其他服务索引;
继续更新所述主索引;以及
当所述主索引被成功更新时,在所述被成功更新的服务索引中删除先前被成功更新的节点。
7. 根据权利要求3所述的方法,其中当执行所述数据查询操作时,通过合并查询所述主索引和所述至少一个服务索引所得的结果集合以获得所述查询结果。
8. 根据权利要求7所述的方法,其中所述通过合并查询所述主索引和所述至少一个服务索引所得的结果集合以获得所述查询结果包括:
在所述主索引和所述至少一个服务索引中,基于操作类型而将具有相同节点信息的节点合成为一个节点。
9. 一种管理数据库索引的装置,包括:
第一创建装置,被配置用于创建主索引;
第二创建装置,被配置用于创建至少一个服务索引,其中所述至少一个服务索引用于记录所述主索引中将被更新节点的变化;
检测装置,被配置用于检测所述数据库中是否出现涉及所述主索引的、针对所述数据库执行的操作;

索引装置,被配置用于响应于在所述数据库中出现涉及所述主索引的、针对所述数据库执行的操作,利用所述至少一个服务索引来对所述主索引进行维护,其中所述维护是基于所述至少一个服务索引中记录的、所述主索引中将被更新节点的变化进行的。

10. 根据权利要求 9 所述的装置,其中所述至少一个服务索引中的节点至少包括:与所述主索引的节点相关联的节点信息以及针对所述主索引进行的操作类型。

11. 根据权利要求 9 或 10 所述的装置,其中所述索引装置包括以下至少一个:

操纵装置,被配置用于在对所述数据库执行的数据操纵操作中对所述主索引和所述服务索引中的至少一项进行维护;以及

查询装置,被配置用于在对所述数据库执行的数据查询操作中获取查询结果。

12. 根据权利要求 11 所述的装置,其中所述操纵装置包括:

用于响应于执行所述数据操纵操作,并行地更新所述主索引和所述至少一个服务索引的装置;以及

用于响应于所述至少一个服务索引和所述主索引中的任一项被成功更新,标记所述数据操纵操作结束的装置。

13. 根据权利要求 12 所述的装置,其中用于所述并行地更新所述主索引和所述至少一个服务索引的装置包括:

用于当所述主索引首先被成功更新后,停止更新所述至少一个服务索引的装置。

14. 根据权利要求 12 所述的装置,其中用于所述并行地更新所述主索引和所述至少一个服务索引的装置包括用于执行如下的装置:

当所述至少一个服务索引之一首先被成功更新后,

停止更新尚未被成功更新的其他服务索引;

继续更新所述主索引;以及

当所述主索引被成功更新时,在所述被成功更新的服务索引中删除先前被成功更新的节点。

15. 根据权利要求 11 所述的装置,所述查询装置包括:用于当执行所述数据查询操作时,通过合并查询所述主索引和所述至少一个服务索引所得的结果集合以获得所述查询结果的装置。

16. 根据权利要求 15 所述的装置,其中用于所述通过合并查询所述主索引和所述至少一个服务索引所得的结果集合以获得所述查询结果的装置包括:

用于在所述主索引和所述至少一个服务索引中,基于操作类型而将具有相同节点信息的节点合成为一个节点的装置。

一种管理数据库索引的方法和装置

技术领域

[0001] 本发明的各实施方式涉及数据库管理,并且更具体地,涉及一种管理数据库索引的方法、装置和相关计算机程序产品。

背景技术

[0002] 数据库技术的发展为各种计算机应用提供了强大的数据组织、数据管理和数据存储能力。近年来,数据库存储容量逐渐增加并且用户对数据库执行的各种操作也愈加复杂和频繁,如何确保数据库在可预期的响应时间内快速返回用户期望的结果,这已经成为数据库领域的一个研究热点。

[0003] 在现有的数据库管理系统(DBMS)中,已经广泛使用了数据库索引技术,数据库索引是对数据库表中一列或多列的值进行排序的一种结构,使用索引可快速访问数据库表中的特定信息。数据库索引可以大大加快查询操作的响应速度,然而在改进查询性能的同时,索引还可能产生一些负面影响。

[0004] 当对数据库中的表进行数据操纵操作(即,DML操作,如,插入、删除和修改)时,还可能更新索引。此时,数据库管理系统不但需要修改数据库的各个表中的数据,还需要更新受到该DML操作影响的索引。也即,只有将数据库中的表和受影响的索引两者均成功更新之后,才能认为该DML操作完成。在仅成功更新数据库中的表而尚未更新受影响的索引时,此时的索引并不能反应数据库的真实情况。于是,若基于其更新尚未完成的索引对表进行查询,则可能出现返回结果与数据库中存储的最新数据不一致的情况。对于大规模数据库,其索引可能是十分复杂的数据结构,因而更新索引所花费的时间可能会对数据库的响应时间产生不利影响。

发明内容

[0005] 因而,为了提高对索引进行管理和维护的效率,期望能够提供一种以快速响应方式来管理数据库索引的方法、设备和相应的计算机程序产品,并且还期望可以在尽量不改变数据库管理系统现有设置的情况下进行操作。为此,本发明的各实施方式提供了一种采用服务索引来管理数据库索引的方法、装置和计算机程序产品。

[0006] 在一个实施方式中,提供了一种管理数据库索引的方法。该方法包括:创建主索引;创建至少一个服务索引,其中至少一个服务索引用于记录主索引中将被更新节点的变化;检测数据库中是否出现涉及主索引的、针对数据库执行的操作;响应于在数据库中出现涉及主索引的、针对数据库执行的操作,利用至少一个服务索引来对主索引进行维护,其中维护是基于至少一个服务索引中记录的、主索引中将被更新节点的变化进行的。

[0007] 在一个实施方式中,至少一个服务索引中的节点至少包括:与主索引的节点相关联的节点信息以及针对主索引进行的操作类型。

[0008] 在一个实施方式中,针对数据库执行的操作包括:针对数据库执行的数据操纵操作和数据查询操作中的至少一个。

[0009] 在一个实施方式中,提供了一种管理数据库索引的装置。该装置包括:第一创建装置,被配置用于创建主索引;第二创建装置,被配置用于创建至少一个服务索引,其中至少一个服务索引用于记录主索引中将被更新节点的变化;检测装置,被配置用于检测数据库中是否出现涉及主索引的、针对数据库执行的操作;索引装置,被配置用于响应于在数据库中出现涉及主索引的、针对数据库执行的操作,利用至少一个服务索引来对主索引进行维护,其中维护是基于至少一个服务索引中记录的、主索引中将被更新节点的变化进行的。

[0010] 在一个实施方式中,至少一个服务索引中的节点至少包括:与主索引的节点相关联的节点信息以及针对主索引进行的操作类型。

[0011] 在一个实施方式中,索引装置包括以下至少一个:操纵装置,被配置用于在对数据库执行的数据操纵操作中对主索引和服务索引中的至少一项进行维护;以及查询装置,被配置用于在对数据库执行的数据查询操作中获取查询结果。

[0012] 采用根据本发明的各实施方式,可以通过基于主索引和至少一个服务索引来管理数据库中的索引。并且,可以加快与数据操纵操作相关联的更新索引的响应速度;还可以将对查询操作的影响最小化并且确保索引的正确性。

附图说明

[0013] 结合附图并参考以下详细说明,本发明各实施方式的特征、优点及其他方面将变得更加明显,在附图中:

[0014] 图 1 是示出根据本发明一个实施方式的利用服务索引管理数据库索引的方法的流程图;

[0015] 图 2 是示出根据本发明一个实施方式的用于创建/调整索引的方法的流程图;

[0016] 图 3 是示出根据本发明一个实施方式的服务索引中节点的数据结构的示意图;

[0017] 图 4 是示出根据本发明一个实施方式的主索引和服务索引的树形结构的示意图;

[0018] 图 5 是示出根据本发明一个实施方式的执行数据操纵操作的方法的流程图;

[0019] 图 6 是示出根据本发明一个实施方式的合并查询主索引和服务索引所得的结果集合以获得查询结果的过程的示意图;以及

[0020] 图 7 是示出根据本发明一个实施方式的用于利用服务索引管理数据库的装置的框图。

具体实施方式

[0021] 下面参考附图详细描述本发明的各实施方式。

[0022] 附图中的流程图和框图,图示了按照本发明各种实施方式的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,所述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为备选的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算

机指令的组合来实现。

[0023] 在一个实施方式中,可以基于快速响应策略并利用主索引和至少一个服务索引来提高数据库访问速度。在一个实施方式中,响应于涉及主索引的、针对数据库执行的操作,利用至少一个服务索引来对主索引进行维护。

[0024] 下面首先概述本发明各实施方式中某些术语的含义。

[0025] 快速响应策略:快速响应策略是指一种快速更新主索引和至少一个服务索引中的节点的策略,其原理在于,一旦至少一个服务索引和主索引中的任一项被成功更新,则认为针对数据库进行的 DML 操作结束。因而采用快速响应策略可以改进数据操作的效率。

[0026] 主索引:本发明各实施方式中的主索引对应于常规索引,为了与服务索引相区别而将其称为主索引。

[0027] 服务索引:服务索引是用于记录主索引中将被更新节点的变化索引。与主索引数据结构的区别在于,服务索引的节点中还包括针对主索引的操作类型(如插入或删除)。服务索引是为主索引服务的,其节点数据是针对主索引节点的变化过程的数据。在整个索引结构(主索引及至少一个服务索引)中,服务索引的参数及类型的定义可以独立于主索引,以便加速索引结构的更新和维护,从而提高针对数据库的各种数据操作的性能。

[0028] 应当注意,在不同实施方式中可以基于不同的需求而设置一个或者多个服务索引。在一个实施方式中,全部服务索引(可以包括一个或者多个)中的节点所记录的变化最终将被更新到主索引中。尽管在某些时刻,主索引自身并不能单独反应与已更新数据库的表相对应的索引状态。然而至少一个服务索引和主索引可以作为一个整体运行,并反应与已更新数据库的表相对应的最新索引信息。由于至少一个服务索引和主索引中的任一项被成功更新时,则认为数据操纵操作结束,这大大加快了更新数据库索引的速度。

[0029] 已经概括描述了本发明各实施方式的基本原理和应用环境,在下文中,将参加图 1 至图 7 详细描述本发明的一些实施方式。图 1 是示出根据本发明一个实施方式的利用服务索引管理数据库索引的方法的流程图 100。如步骤 S102 所示,首先创建主索引;继而在步骤 S104 中,创建至少一个服务索引。应当注意,此处的主索引即对应于常规索引。在数据库中已经存在常规索引时,可以基于本发明的实施方式创建至少一个服务索引,并调整已有的常规索引适应于本发明的实施方式;否则,可以基于本发明的实施方式创建主索引和至少一个服务索引。

[0030] 在一个实施方式中,至少一个服务索引用于记录主索引中将被更新节点的变化。应当注意,本发明各实施方式中的服务索引并不同于常规索引。常规索引中保存的是对数据库表中一列或多列的值进行排序的一种结构,目的在于使用该常规索引可快速访问数据库表中的特定信息。而本发明各实施方式中的至少一个服务索引目的在于,记录主索引中将被更新节点的变化,以便在主索引尚未完成更新时,可以通过主索引和至少一个服务索引计算得出数据库中当前最新的数据。

[0031] 发明人认为,针对数据库进行 DML 操作时可能会导致需要更新数据库索引。如果数据库的索引规模较小,则更新该索引仅需要很少的时间。也即,更新小的索引的效率通常高于更新大的索引的效率。在某些实施方式中,采取创建规模小于主索引(也即对应于常规索引)的至少一个服务索引来加快更新索引的速度,进而缩小针对数据库的数据操作的响应时间。然而更新小的索引的效率并不总是高于更新大的索引的效率,在某些特定情况

下还会出现大的索引首先被更新完毕的情况。因而在一个实施方式中,可以并行地更新主索引和至少一个服务索引,以便获得最优的效率。

[0032] 在步骤 S106 中,检测数据库中是否出现涉及主索引的、针对数据库执行的操作。由于本发明的实施方式旨在管理数据库的索引,当针对数据库执行的操作并不涉及主索引时,则不属于本发明的实施方式考虑的范围,因而在此设置检测步骤 S106。

[0033] 在步骤 S108 中,响应于在数据库中出现涉及所述主索引的、针对所述数据库执行操作,利用所述至少一个服务索引来对主索引进行维护。此时,这种维护是基于所述至少一个服务索引中记录的、所述主索引中将被更新节点的变化进行的,并且主索引和至少一个服务索引对外表现为一个整体,而在主索引和至少一个服务索引之间基于预定义的规则协调操作。

[0034] 在一个实施方式中,针对数据库执行操作可以包括:针对数据库执行数据操纵操作和数据查询操作中的至少一个。应当注意,在如下情况下可能涉及到对主索引的操作:(1)在对数据库中的表进行 DML 操作时,对各种表进行增删改的操作可以导致需要更新主索引,然而并非所有的 DML 操作均会导致更新索引。(2)在查询数据库时,通常针对主索引进行查询以便加快响应速度,因此数据查询操作也会涉及主索引。

[0035] 应当注意,在一个实施方式中,利用至少一个服务索引来对所述主索引进行维护是与针对数据库的其他操作同步执行的。也即,该维护操作并不排除针对数据库执行数据操纵操作和数据查询操作,并且本发明提供了用于数据操纵操作和数据查询操作的具体实施方式。此时,两种索引作为一个整体同步地关联于数据库中的数据。

[0036] 在一个实施方式中,当对数据库执行数据操纵操作和/或数据查询操作时,主索引和至少一个服务索引作为一个整体提供快速更新数据库索引的功能。在针对数据库执行其他操作的同时同步地维护主索引,一方面可以使主索引和至少一个服务索引实时地反映数据库中的数据;另一方面,相对于已经更新了数据库中的数据然而在一段时间后再维护数据库索引的异步方式而言,本发明的同步方式无需额外的时间来维护索引,同时还不影响针对数据库进行的其他操作。

[0037] 本发明实施方式的方法与常规数据库兼容,可以在尽量少修改现有数据库管理系统配置的情况下实现。在现有 DBMS 中,存在以下涉及维护索引的数据定义语言 (DDL) 声明,例如:

[0038] CREATE INDEX idx_name ON TABLE tb_name(column1,...)...

[0039] ALTER INDEX idx_name...

[0040] 在一个实施方式中,可以增加关键字 enable_fast_response 和 disable_fast_response。并且下文示出了根据一个实施方式的示例:

[0041] (1) 创建快速响应模式的索引:

[0042] CREATE INDEX idx_name on TABLE tb_name(column1,...)enable_fast_response with n...

[0043] (2) 将快速响应模式的索引调整为常规模式:

[0044] ALTER INDEX idx_name...disable_fast_response...

[0045] (3) 将常规模式的索引调整为快速响应模式:

[0046] ALTER INDEX idx_name...enable_fast_response with n...

[0047] 图 2 是示出根据本发明一个实施方式的用于创建 / 调整索引的方法的流程图 200。在步骤 S202 中,开始创建 / 调整索引的操作,并且在步骤 S204 中,基于上文定义的关键字判断该创建 / 调整索引的操作是基于本发明实施方式的快速响应模式还是基于常规模式。当判断结果为“是”,则操作前进至步骤 S206,以快速响应模式创建 / 调整索引;否则,操作前进至步骤 S208 以常规模式创建 / 调整索引。

[0048] 图 3 是示出根据本发明一个实施方式的服务索引中节点的数据结构的示意图 300,这仅仅是服务索引中节点的一个示意性数据结构。服务索引与主索引的数据结构的区别在于,服务索引的节点中还保持针对主索引数据的操作类型(如插入或删除)。服务索引的节点还可以采用其他的数据结构,只要能够反映主索引中数据的变化信息即可。

[0049] 在一个实施方式中,至少一个服务索引中的节点至少包括:与主索引的节点相关联的节点信息以及针对主索引进行的操作类型。如图 3 所示,字段 302 和字段 304 两者整体表示与主索引的节点相关联的节点信息,也即分别示出为行 ID 302 和值 304。字段 306 表示针对主索引进行的操作类型(例如,以“I”表示“插入”,以“D”表示“删除”)。

[0050] 例如,服务索引中的一个节点可以示出为 (R0, V1, I),该三元组的含义是,针对主索引进行“插入”操作,并且主索引中相关节点的行 ID 是 R0,值为 V1。又例如,服务索引中的一个节点可以示出为 (R2, V1, D),该三元组的含义是,针对主索引进行“删除”操作,并且主索引中相关节点的行 ID 是 R2,值为 V1。

[0051] 图 4 是示出根据本发明一个实施方式的主索引和服务索引的树形结构的示意图 400。如图 4 所示,在快速响应模式下,索引首部可以包括主索引 410,还可以包括至少一个服务索引(即服务索引一 412 至服务索引 N416)。服务索引的数量和类型可以基于不同需求而定制,将在下文中详述。

[0052] 常规索引仅包括标记为 410 的主索引(即,常规索引)部分,而并不包括后续的服务索引 412 至 416。应当注意,服务索引可以采用类似于主索引的树状结构,不同之处在于服务索引的节点内容不同于主索引。并且主索引是根据用户对数据库进行的 DML 操作随时更新的长期存在的索引,而服务索引仅记录针对主索引节点数据的变化过程数据。对于服务索引和主索引,相应的节点数据可以保存在图 4 的叶子节点中。在其他实施方式中,并不限于以图 4 所示的存储的方式保存主索引和至少一个服务索引,而是可以采用各种数据结构存储主索引和相对应的至少一个服务索引,只要主索引和至少一个服务索引可以协调操作加速数据库响应即可。

[0053] 随着时间的推移,服务索引中的节点相关的变化还被逐渐更新到主索引中,并且一旦服务索引中的节点相关的变化已经被更新到主索引中后,则删除该节点。也即,当全部服务索引中的每个节点的信息均被更新到主索引之后,服务索引可以为空(例如,服务索引一 412 所示,字段 412 下仅有一个空的根节点 422)。或者出于节约资源目的考虑,还可以删除该服务索引。例如,直接删除服务索引一 412 和根节点 422,此时服务索引二 414 直接跟随在主索引 410 之后。在后续的数据操纵操作导致需要更新主索引时,再相应地创建至少一个服务索引。

[0054] 当对数据库中的表进行插入操作导致需要对主索引进行插入操作时,可以在服务索引中加入标识为“I”类型的节点(该节点中记录与主索引的节点相关联的节点信息,例如,行 ID 和值)。当对数据库中的表进行删除操作导致需要对主索引进行删除操作时,可

以在服务索引中加入标识为“D”类型的节点。尽管仅示出了进行插入（表示为“I”）和删除（表示为“D”），本领域技术人员应当理解，修改操作可以等效为顺序执行的删除操作和插入操作，相应的节点中操作类型可以分别记录为“D”和“I”。也即，可以向服务节点的树状结构中分别加入删除节点和插入节点两者表示修改操作。

[0055] 在一个实施方式中，响应于执行数据操纵操作，并行地更新主索引和至少一个服务索引；以及响应于至少一个服务索引和主索引中的任一项被成功更新，标记数据操纵操作结束。

[0056] 尽管数据库索引可以加快数据库对于数据操作的响应速度，然而当数据库索引庞大到一定程度时，对数据库中的数据进行操作导致的更新数据库索引的速度成为了影响数据库响应时间的瓶颈之一。本发明实施方式的方法通过创建用于记录主索引中节点的数据变化过程的至少一个服务索引，并且将该至少一个服务索引与主索引作为整体来提供数据库索引的功能。

[0057] 在一个实施方式中，由于只要至少一个服务索引和主索引中的任一项被成功更新，则标记数据操纵操作结束，这可以在一定程度上缩短更新数据库索引的时间。尽管有可能存在主索引尚未更新完毕的情况，然而此时在至少一个服务索引的辅助下，不必等到主索引更新完毕才能标记数据操纵操作结束。

[0058] 在一个实施方式中，并行地更新主索引和至少一个服务索引包括：当主索引首先被成功更新后，停止更新至少一个服务索引。在一个实施方式中，并行地更新主索引和至少一个服务索引包括：当至少一个服务索引之一首先被成功更新后，停止更新尚未被成功更新的其他服务索引；继续更新主索引；以及当主索引被成功更新时，在被成功更新的服务索引中删除先前被成功更新的节点。

[0059] 现在将参见图 5 详细说明并行地更新主索引和至少一个服务索引的操作流程。图 5 是示出根据本发明一个实施方式的执行数据操纵操作的方法的流程图 500。在步骤 S502 中，响应于执行数据操纵操作，并行地更新主索引和至少一个服务索引。这里的“并行地”是指，当针对数据库执行的数据操作（例如，向数据库的表中插入数据、从中删除数据以及更新数据库的表中的数据）导致需要更新主索引时，在更新主索引的同时还向至少一个服务索引中插入节点，并且该节点记录主索引中节点的将被更新的变化。在存在多个服务索引的情况下，需要并行地向该多个服务索引中插入节点。

[0060] 在步骤 S504 中，响应于至少一个服务索引和主索引中的任一项被成功更新，标记数据操纵操作结束。继而，在步骤 S506 中，判断主索引是否首先被成功更新。这里的更新，指任何可以按现有方式对主索引进行的更新。当判断结果为“是”，则操作进入步骤 S508，停止更新至少一个服务索引。本发明实施方式的方法设立服务索引的目的之一在于，在更新主索引需要更长时间时，利用至少一个服务索引来记录主索引中节点的将被更新的变化。如果步骤 S506 的判断结果为“是”，则表示对主索引的更新快于对全部服务索引的更新，此时则无需服务索引参与。继而，操作流程结束。

[0061] 如果步骤 S506 的判断结果为“否”，表示服务索引之一首先被更新，则操作流程前进至步骤 S510。图 5 中右侧分支的含义是，需要利用至少一个服务索引中的节点记录所述主索引中将被更新节点的变化。进入步骤 S510 的先决条件是，至少一个服务索引中的一个首先被成功更新。由于此时首先被成功更新的服务索引已经记录了主索引中将被更新节点

的变化,因而无需在其他服务索引中重复记录这一变化。在步骤 S510 中,停止尚未被成功更新的其他服务索引。

[0062] 接着,在步骤 S512 中,继续更新主索引,并且在步骤 S514 中,当主索引被成功更新时,在被成功更新的服务索引中删除先前被成功更新的节点。设置服务索引的目的在于记录将对主索引执行的更新,由于在步骤 S514 中主索引已经更新完毕,则此时无需再保留服务索引中与该更新相关的节点,因而应当在被成功更新的服务索引中删除先前被成功更新的节点。

[0063] 在一个实施方式中,并不仅限于以图 5 所示的顺序来实现数据操纵操作。而是还可以在发布操作结束(步骤 S504)的同时,由其他线程判断是主索引还是服务索引首先被更新,并相应地执行如步骤 S508 至 S514 的操作。

[0064] 在不同实施方式中,可以采用不同机制来定义服务索引中的节点与针对主索引的更新之间的对应关系。例如,当一个服务索引首先被成功更新时,更新该服务索引的线程可以向更新主索引的线程发送消息,并告知该服务索引的标识以及被更新的节点的标识(或者位置等);当主索引更新完毕时,则基于被告知的服务索引的标识以及节点的标识(或者位置),将相应节点删除或者标记为无效。本领域技术人员还可以想到其他具体实施方式,只要能够在主索引中的被更新节点和服务索引中的被更新节点之间建立对应关系即可。

[0065] 应当注意,每个服务索引中的节点的数量是动态变化的。随着对数据库中的表执行 DML 操作而导致需要更新主索引,向至少一个服务索引中加入记录相应更新的新的节点;随着如步骤 S514 所示的继续更新主索引的成功完成,在被成功更新的服务索引中删除先前被成功更新的节点。可以理解,当对数据库中的表频繁进行 DML 操作时,可能会需要频繁地更新主索引,进而使得至少一个服务索引中的节点数量增加;而当持续一段时间无需更新主索引时,由于与至少一个服务索引中的节点相关联的变化已经被更新到主索引中,至少一个服务索引中的节点数量会相应地减少。

[0066] 在经过特定时间后,还可能出现服务索引之一可能为空的情况。应当注意,尽管此时服务索引之一为空,并不意味着本发明实施方式的方法退化成为使用常规索引(对应于本发明中的主索引)的情况。而是,当再次出现需要更新主索引时,可以继续依据上文所述的方法向服务索引中增加新的节点。本发明实施方式并不考虑那些不涉及主索引(对应于常规索引)操作的数据库操作。本发明的实施方式关注于如何更高效地管理数据库的索引,进而通过加速数据库索引的响应来达到提高针对数据库的数据操作的响应速度。

[0067] 在一个实施方式中,当执行数据查询操作时,通过合并查询主索引和至少一个服务索引所得的结果集合以获得查询结果。在一个实施方式中,当进行数据库查询时,主索引可能尚未完成更新,因而主索引自身并不能正确反映数据库中的各个表的实际内容。此时,需要合并查询主索引和至少一个服务索引所得的结果集合以获得查询结果。至少一个服务索引记录主索引中将被更新节点的变化,基于主索引和服务索引中的这些变化,可以计算出当前的真实数据库内容。

[0068] 在一个实施方式中,通过合并查询主索引和至少一个服务索引所得的结果集合以获得查询结果包括:在主索引和至少一个服务索引中,基于操作类型而将具有相同节点信息的节点合成为一个节点。现在将参考图 6 详细描述,图 6 是示出根据本发明一个实施方式的合并查询主索引和服务索引所得的结果集合以获得查询结果的过程的示意图 600。

[0069] 例如图 6 左上部所示,已经创建了一个主索引 610 和多个服务索引 620(包括服务索引一至服务索引 N)。在查询时,可以按值-键搜索主索引 610 和服务索引 620(分别如箭头 A 和箭头 B 所示)。在上文所示以三元组(行 ID,值,操作类型)的具体示例中,可以分别获得主集合 612 和服务集合 622。例如,主集合 612 为 $\{(R0, V1, I), (R2, V1, I)\}$,而服务集合 622 为 $\{(R1, V1, I), (R2, V1, D), (R3, V1, I), (R1, V1, D), (R1, V1, I)\}$ 。应当注意,与前文所述相同,这里“I”表示插入操作,而“D”表示删除操作。在经过并集运算后(如箭头 C 所示),所得的结果集合 630 为 $\{(R0, V1, I), (R2, V1, I), (R1, V1, I), (R2, V1, D), (R3, V1, I), (R1, V1, D), (R1, V1, I)\}$ 。

[0070] 由于此时服务索引中的变化信息尚未被更新至主索引中,因而需要在主索引基础上结合服务索引中的节点信息才能得知数据库的当前数据。应当注意,在各个索引中,具有相同节点信息的节点是针对数据库中表的相同数据项的操作。在此实施方式中,具有相同的行 ID 和相同值的三元组对应于对相同数据项的操作。例如在结果集合 630 中,第 3 个三元组 $(R1, V1, I)$ 和第 6 个三元组 $(R1, V1, D)$ 两者都是针对行 ID 为 R1、且值为 V1 的数据项的操作。不同之处在于前者是插入操作而后者是删除操作,这两个操作可以合并。

[0071] 接着如箭头 D 所示,将结果集合 630 按行 ID 和值排序,以便合并具有相同行 ID 和相同值的节点,排序后的结果集合'示出为 632。此时,如结果集合'632 左侧括号所示,第 1 个三元组是针对第一数据项的操作,第 2 至 4 个三元组是针对第二数据项的操作,第 5-6 个三元组是针对第三数据项的操作,而最后一个三元组是针对第四数据项的操作。最后,如箭头 E 所示,合成后的最终查询结果 640 包括三个三元组,也即对应于三个节点 $\{(R0, V1, I), (R1, V1, I), (R3, V1, I)\}$ 。

[0072] 应当注意在合并操作中,针对相同节点,“I”操作和“D”操作可以相互抵消,例如在结果集合'632 中所示的第 5-6 个三元组所示的操作可以相互抵消。

[0073] 应当注意,本发明实施方式的方法在服务索引 620 为空时同样适用,此时主索引 610 即可正确反映数据库中的数据。当搜索服务索引 620 时,则返回的服务集合 622 也将为空,并且在主集合 612 和服务集合 622 的并集(即,结果集合 630)中的数据即为查询主索引 610 所得结果。可见,本发明实施方式的方法此时也是正确有效的。

[0074] 基于数据库的各个表中数据的特征,以及基于针对数据库中的表进行的不同类型的操作,各种类型的索引可以具有不同的操作性能。在一个实施方式中,至少一个服务索引中可以是相同类型或者不同类型。在一个实施方式中,还可以以不同参数定义至少一个服务索引中的任一项,以便满足不同操作需要。应当注意,以不同参数定义的服务索引可以是相同类型或者不同类型的服务索引。

[0075] 例如,可以采用普通索引,也可以采用逆向索引。例如逆向索引适用于连续数据,诸如在 1、2、3、4、... 这种非离散的集中式插入数据的情况下可以使用逆向索引。又例如,如果在一段时间内插入的数据是离散的,那么可以使用普通索引;如果在一定的时间段内,产生的数据是连续的,并需要在一定时间段内集中地插入连续的数据,则可以使用逆向索引。

[0076] 具体而言,对于一个银行的数据库系统,在白天的工作时间,当银行的众多客户频繁执行存取现金时,需要对数据库系统进行大量操作。那么在短时间内产生连续插入顺序数据的可能性很小(这是由于多个客户同时进行同一类型操作的可能性较小,并且每个客

户进行操作的具体时间也不同),此时可以使用普通索引。而数据库管理员在夜间可能会集中进行清算操作,并且该操作在短时间内完成,此时可能会产生在短时间内连续插入顺序数据的情况,因而可以使用逆向索引。

[0077] 关于如何采用不同参数定义服务索引,本领域技术人员可以参考创建索引时涉及的各种参数的定义,根据具体情况实现,在此不再赘述。

[0078] 上文已经参见图 1 至图 6 详细示出了根据本发明的方法的某些实施方式,现在将参考图 7 说明如何实现相应的装置。图 7 是示出根据本发明一个实施方式的用于利用服务索引管理数据库的装置的框图 700。在一个实施方式中,提供了一种管理数据库索引的装置,包括:第一创建装置 710,被配置用于创建主索引;第二创建装置 720,被配置用于创建至少一个服务索引,其中至少一个服务索引用于记录主索引中将被更新节点的变化;检测装置 730,被配置用于检测数据库中是否出现涉及主索引的、针对数据库执行的操作;索引装置 740,被配置用于响应于在数据库中出现涉及主索引的、针对数据库执行的操作,利用至少一个服务索引来对主索引进行维护,其中维护是基于至少一个服务索引中记录的、主索引中将被更新节点的变化进行的。

[0079] 在一个实施方式中,至少一个服务索引中的节点至少包括:与主索引的节点相关联的节点信息以及针对主索引进行的操作类型。

[0080] 在一个实施方式中,索引装置 740 还可以包括:操纵装置 742,被配置用于在对数据库执行的数据操纵操作中对主索引和服务索引中的至少一项进行维护;以及查询装置 744,被配置用于在对数据库执行的数据查询操作中获取查询结果。

[0081] 在一个实施方式中,操纵装置还包括:用于响应于执行数据操纵操作,并行地更新主索引和至少一个服务索引的装置;以及用于响应于至少一个服务索引和主索引中的任一项被成功更新,标记数据操纵操作结束的装置。

[0082] 在一个实施方式中,用于并行地更新主索引和至少一个服务索引的装置包括:用于当主索引首先被成功更新后,停止更新至少一个服务索引的装置。

[0083] 在一个实施方式中,用于并行地更新主索引和至少一个服务索引的装置包括用于执行如下的装置:当至少一个服务索引之一首先被成功更新后,停止更新尚未被成功更新的其他服务索引;继续更新主索引;以及当主索引被成功更新时,在被成功更新的服务索引中删除先前被成功更新的节点。

[0084] 在一个实施方式中,查询装置还包括:用于当执行数据查询操作时,通过合并查询主索引和至少一个服务索引所得的结果集合以获得查询结果的装置。

[0085] 在一个实施方式中,用于通过合并查询主索引和至少一个服务索引所得的结果集合以获得查询结果的装置包括:用于在主索引和至少一个服务索引中,基于操作类型而将具有相同节点信息的节点合成为一个节点的装置。

[0086] 在一个实施方式中,至少一个服务索引是相同类型或者不同类型。

[0087] 应当注意,根据本发明各种实施方式的方法和装置特别适用于在数据库中存在大规模索引、并且针对数据库的频繁操作导致需要频繁更新该大规模索引的情况。采用常规索引,当针对数据库进行 DML 操作时,需要等到索引被完全更新完毕时才能标识该 DML 操作结束,这在时间和计算资源的利用方面都是巨大浪费。而使用根据本发明各实施方式的方法和装置,采用创建多个服务索引的方式提供了一种快速响应的管理数据库索引的方法。

该方法可以通过提高更新数据库索引的工作效率来提高针对数据库的数据操作的效率。

[0088] 本发明可以采取硬件实施方式、软件实施方式或既包含硬件组件又包含软件组件的实施方式的形式。在优选实施方式中,本发明实现为软件,其包括但不限于固件、驻留软件、微代码等。

[0089] 而且,本发明还可以采取可从计算机可用或计算机可读介质访问的计算机程序产品的形式,这些介质提供程序代码以供计算机或任何指令执行系统使用或与其结合使用。出于描述目的,计算机可用或计算机可读机制可以是任何有形的装置,其可以包含、存储、通信、传播或传输程序以由指令执行系统、装置或设备使用或与其结合使用。

[0090] 介质可以是电的、磁的、光的、电磁的、红外线的、或半导体的系统(或装置或器件)或传播介质。计算机可读介质的例子包括半导体或固态存储器、磁带、可移动计算机磁盘、随机访问存储器(RAM)、只读存储器(ROM)、硬磁盘和光盘。目前光盘的例子包括紧凑盘-只读存储器(CD-ROM)、压缩盘-读/写(CD-R/W)和DVD。

[0091] 适合于存储/或执行程序代码的数据处理系统将包括至少一个处理器,其直接地或通过系统总线间接地耦合到存储器元件。存储器元件可以包括在程序代码的实际执行期间所利用的本地存储器、大容量存储器、以及提供至少一部分程序代码的临时存储以便减少执行期间从大容量存储器必须取回代码的次数的高速缓存存储器。

[0092] 输入/输出或I/O设备(包括但不限于键盘、显示器、指点设备等等)可以直接地或通过中间I/O控制器耦合到系统。

[0093] 网络适配器也可以耦合到系统,以使得数据处理系统能够通过中间的私有或公共网络而耦合到其他数据处理系统或远程打印机或存储设备。调制解调器、线缆调制解调器以及以太网卡仅仅是当前可用的网络适配器类型的几个例子。

[0094] 从上述描述应当理解,在不脱离本发明真实精神的情况下,可以对本发明各实施方式进行修改和变更。本说明书中的描述仅仅是用于说明性的,而不应被认为是限制性的。本发明的范围仅受所附权利要求书的限制。

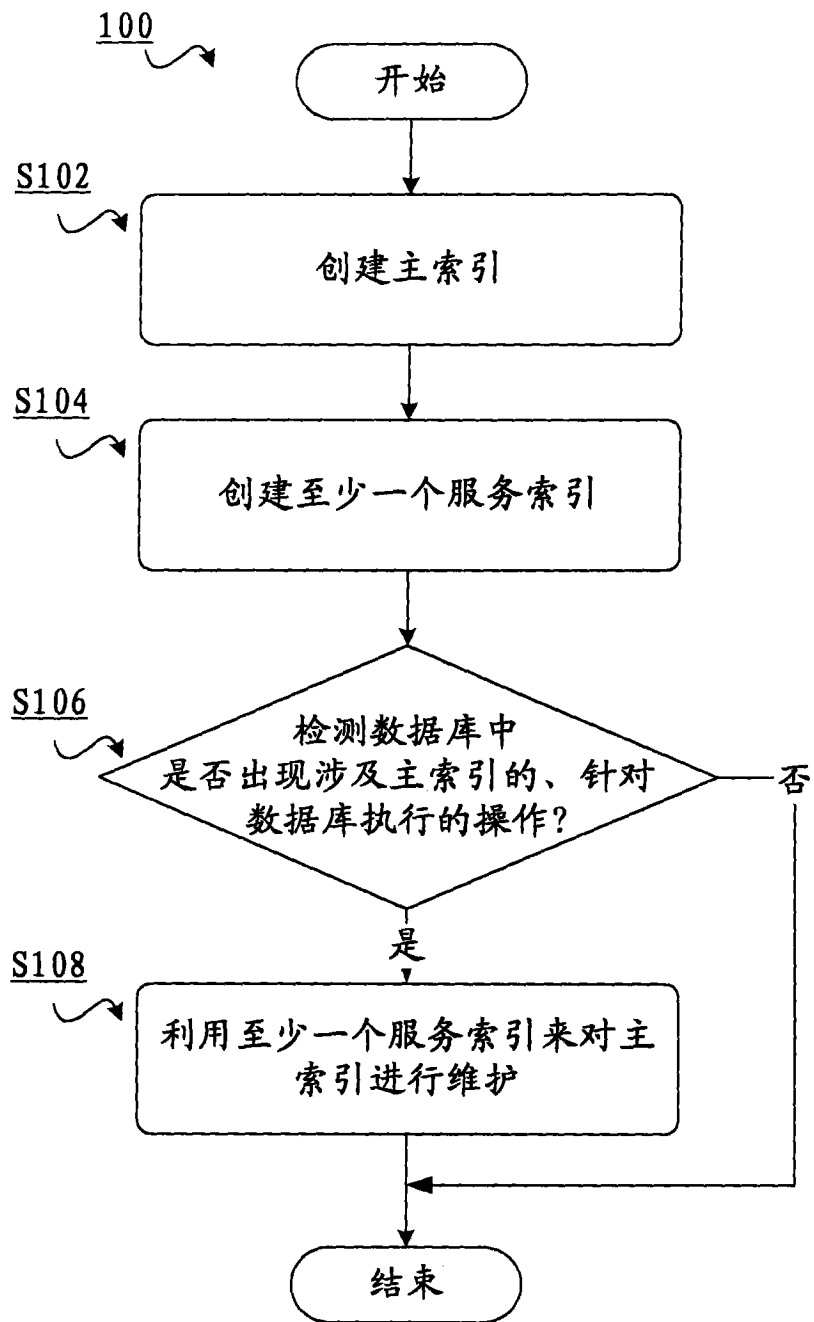


图 1

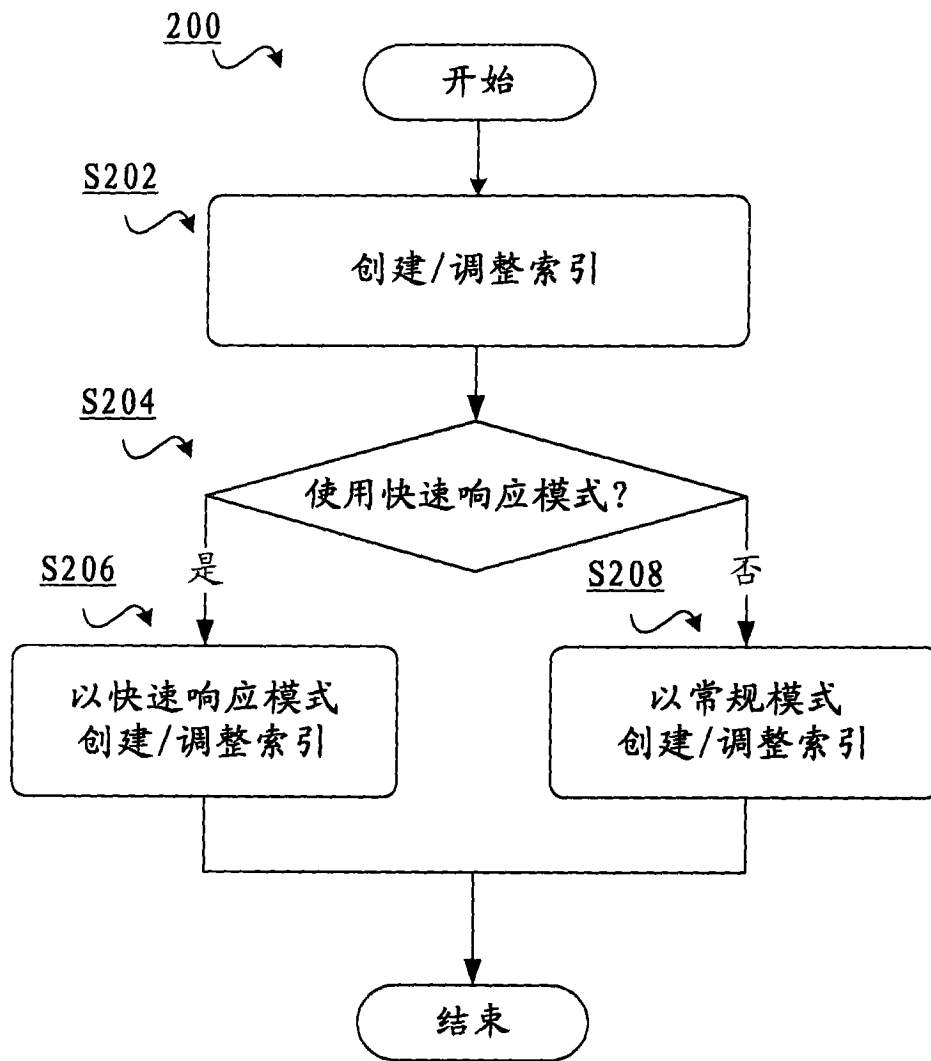


图 2

300

行ID <u>302</u>	值 <u>304</u>	针对主索引的 操作类型 <u>306</u>
-------------------	-----------------	------------------------------

图 3

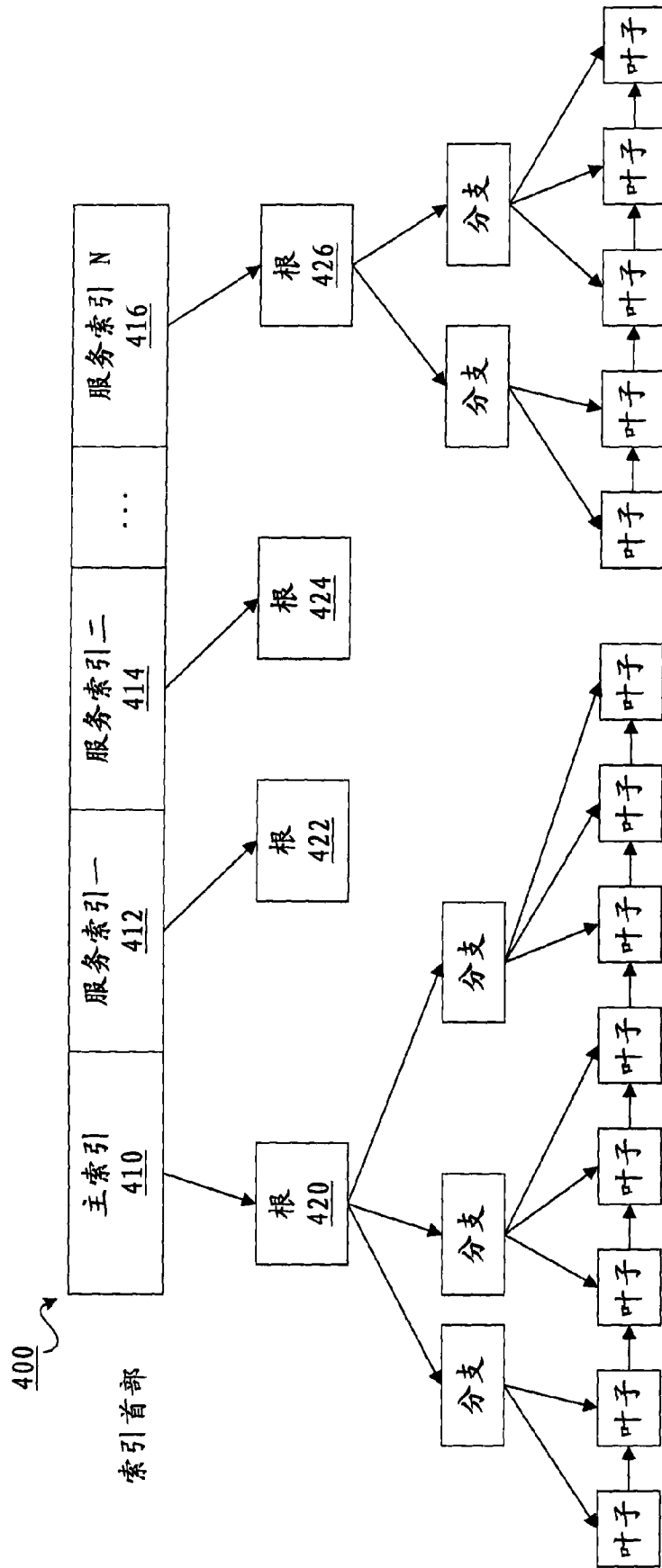


图 4

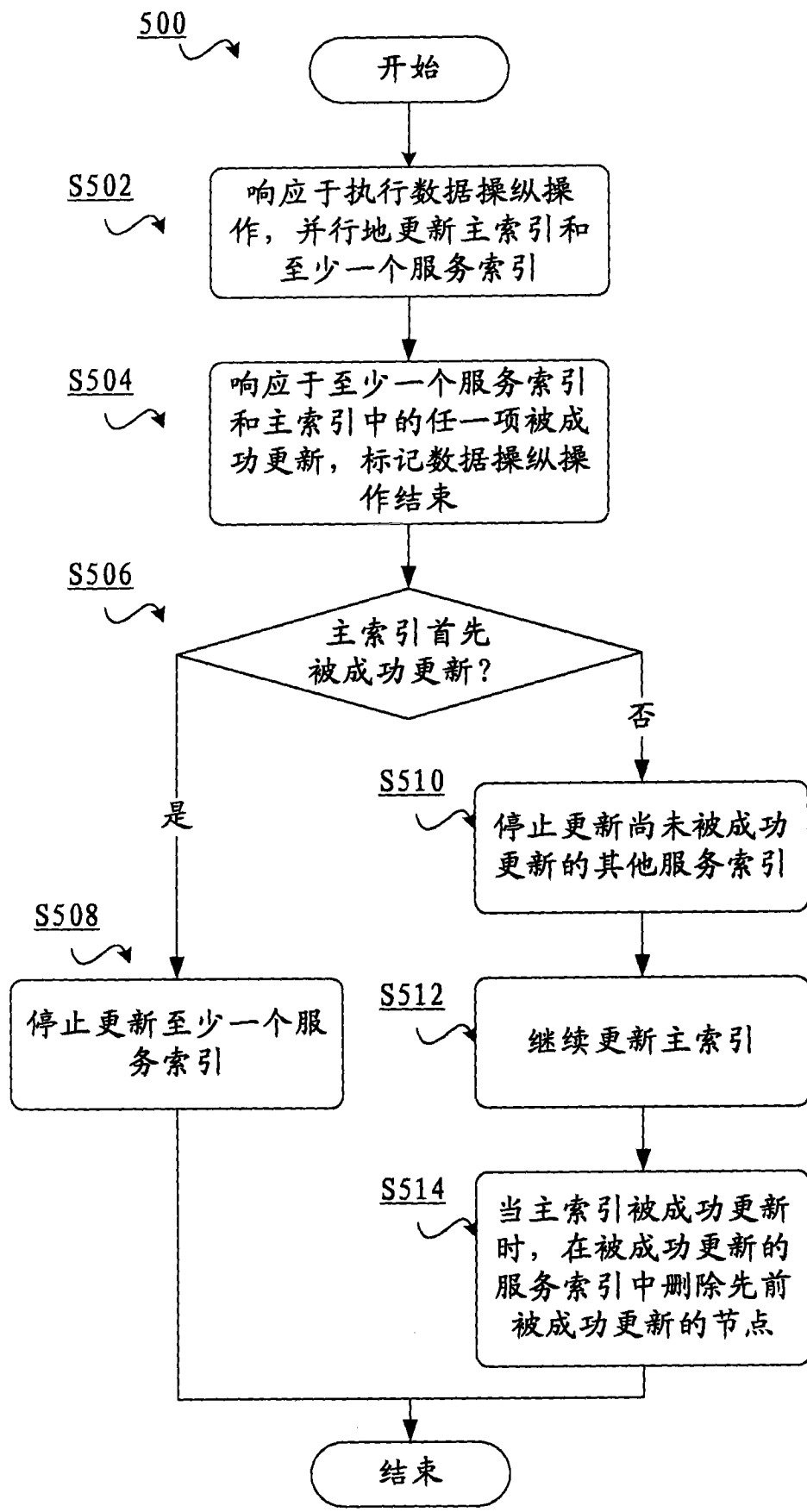


图 5

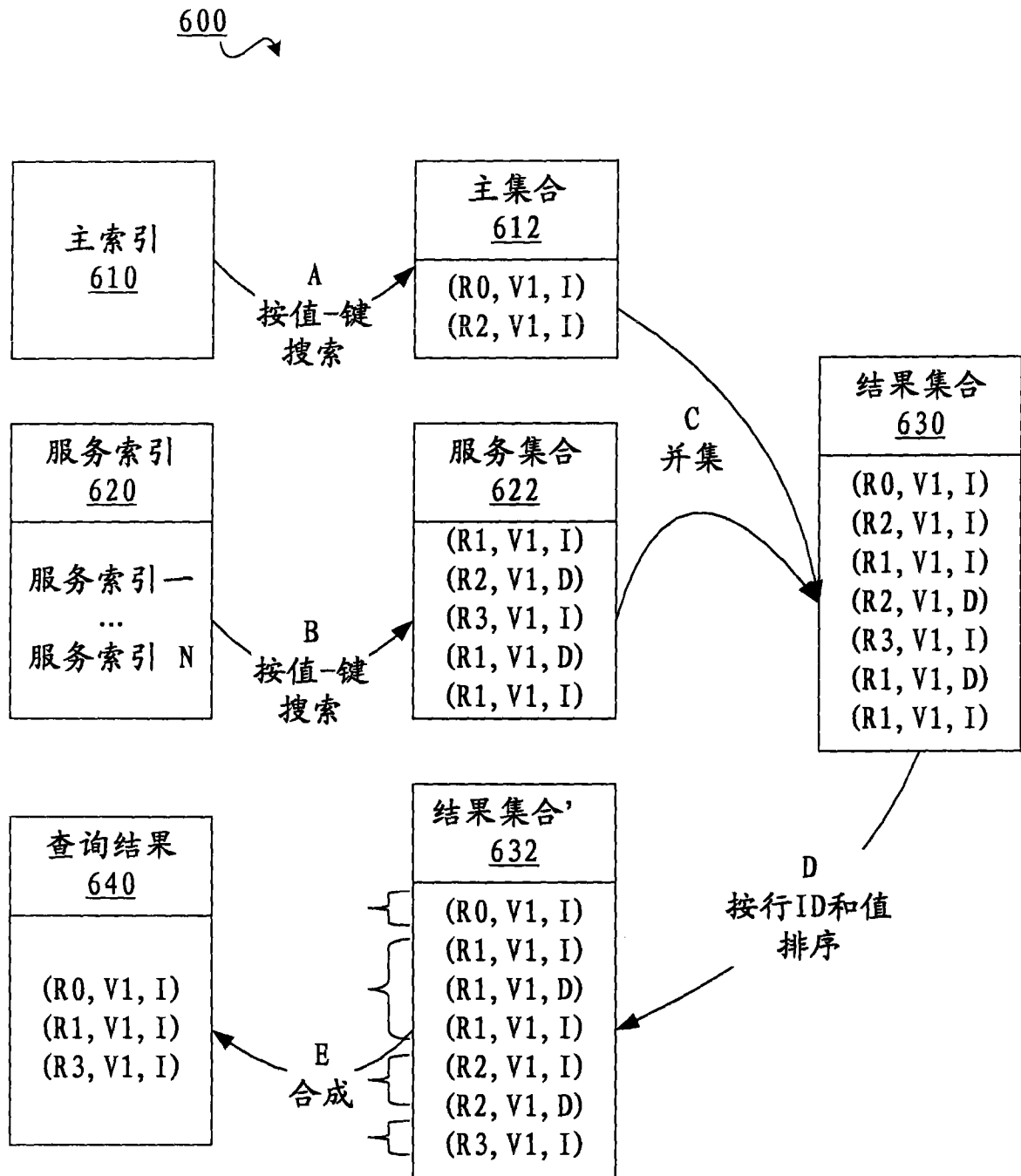


图 6

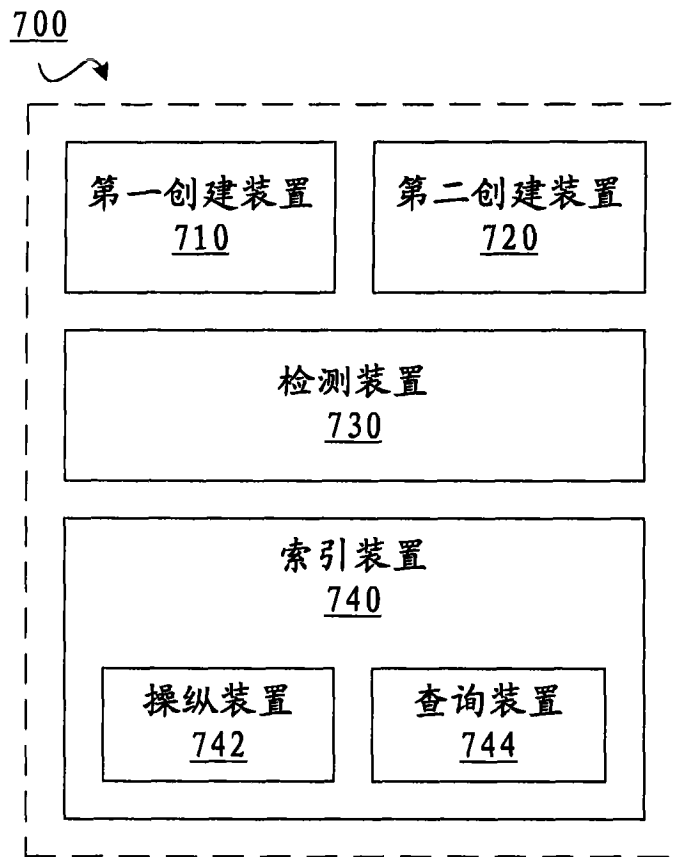


图 7