



(12) 发明专利

(10) 授权公告号 CN 1902704 B

(45) 授权公告日 2012.01.04

(21) 申请号 200480039846.0

(22) 申请日 2004.11.22

(30) 优先权数据

102004001207.5 2004.01.06 DE

(85) PCT申请进入国家阶段日

2006.07.04

(86) PCT申请的申请数据

PCT/EP2004/013241 2004.11.22

(87) PCT申请的公布数据

W02005/069300 EN 2005.07.28

(73) 专利权人 汤姆逊许可公司

地址 法国布洛涅

(72) 发明人 马尔科·温特 乌韦·汉森

沃尔夫刚·克劳斯伯格

斯特凡·库布施 迪特乌尔·郝佩尔

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 吕晓章

(51) Int. Cl.

G11B 27/10 (2006.01)

G06F 17/30 (2006.01)

(56) 对比文件

W0 03/079359 A1, 2003.09.25, 摘要, 权利要求 38, 说明书第 14 页第 18 行至第 15 页, 第 27 页第 11 行至第 32 页第 13 行, 第 35 页第 2-23 行、图 1, 11, 27, 32, 35, 37.

CN 1409924 A, 2003.04.09, 全文.

US 2001/0051954 A1, 2001.12.13, 全文.

审查员 段瑞玲

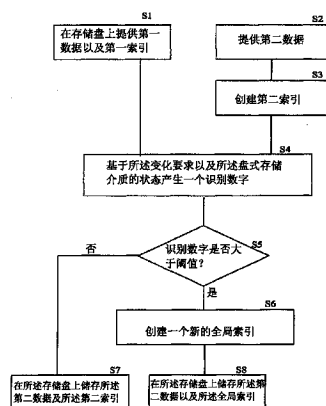
权利要求书 1 页 说明书 4 页 附图 1 页

(54) 发明名称

用于对盘式存储介质上的数据进行更新的方法和装置

(57) 摘要

为了对支持盘式存储的动态数据库的应用进行优化, 本发明提供了对盘式存储介质上的数据进行更新的方法、设备和存储介质。其中, 用于对旧数据进行更新的新数据具备第二索引, 该第二索引与所述新数据相关并为其产生 (S3), 并且所述第二索引作为与所述旧数据相关的所述第一索引 (S1) 的补充被储存在所述盘式存储介质上 (S4)。这样可以减少对所述盘式存储介质的写访问操作次数。尤其对于光介质来说具有提高使用寿命的优势。



1. 一种用于在盘式存储介质上更新第一数据的方法,所述第一数据具有相关的第一索引,

- 在所述盘式存储介质上储存第二数据 (S7, S8),所述第二数据对所述第一数据中的数据进行更新,以及

- 创建仅与所述第二数据相关的第二索引 (S3),

其中

- 确定一识别数字,所述识别数字的特征在于以下的至少一个:所述盘式存储介质的属性和所述盘式存储介质与所述第二数据的关系特征 (S4),并且将该识别数字与一个阈值相比较 (S5),以通过所述比较确定是否将所述第二索引储存在所述盘式存储介质上以作为所述第一索引的补充 (S7),或者是否创建一个与所述盘式存储介质上所存储的全部数据相关的新的全局索引 (S6) 并将所述新的全局索引储存在所述盘式存储介质上 (S8)。

2. 如权利要求 1 所述的方法,其中所述盘式存储介质是存储周期数目有限的可重写光盘。

3. 如权利要求 1 或 2 所述的方法,其中所述第二索引被储存在沿径向尽可能靠近所述第一索引的位置。

4. 如权利要求 1 所述的方法,其中所述第二索引以非分段的形式存储。

5. 如权利要求 1 所述的方法,其中对所述第二数据中的数据进行更新的第三数据被储存在所述盘式存储介质上,并且与所述第二和第三数据相关的第三索引覆盖所述第二索引。

6. 如权利要求 5 所述的方法,其中对所述第二和第三索引中的元素进行结合,从而使所述识别数字低于所述阈值。

7. 如权利要求 1 所述的方法,其中所述阈值由所述盘式存储介质的状态所决定。

8. 用于在盘式存储介质上储存数据的设备,包括:

- 写入装置,用于将第一数据和第二数据写入所述盘式存储介质 (S7, S8),所述第二数据对所述第一数据中的数据进行更新,并且所述写入装置还用于将与所述第一数据相关的第一索引写入所述盘式存储介质,

- 信号处理装置,用于创建所述第一索引,

其中

- 所述信号处理装置还适用于创建仅与所述第二数据相关的第二索引 (S3) 且适用于创建一个与所述盘式存储介质上所存储的全部数据相关的新的全局索引,其中确定一识别数字,所述识别数字的特征在于以下的至少一个:所述盘式存储介质的属性和所述盘式存储介质与所述第二数据的关系特征,并且将该识别数字与一个阈值相比较,以通过所述比较确定是否创建所述新的全局索引,并且

- 所述写入装置也适用于将所述第二索引作为所述第一索引的补充写入所述盘式存储介质 (S7),且适用于将所述新的全局索引写入所述盘式存储介质,其中,所述比较确定是否储存所述第二索引或所述新的全局索引。

9. 如权利要求 8 所述的设备,其中所述盘式存储介质是可重写光盘。

10. 如权利要求 8 或 9 所述的设备,其中对所述第二数据中的数据进行更新的第三数据被储存在所述盘式存储介质上,并且与所述第二和第三数据相关的第三索引覆盖所述第二索引。

用于对盘式存储介质上的数据进行更新的方法和设备

技术领域

[0001] 本发明涉及对盘式存储介质上具有相关第一索引的第一数据进行更新的方法,通过在所述盘式存储介质上储存第二数据,由所述第二数据对所述第一数据中的数据进行更新,在所述盘式存储介质上创建并储存第二索引。本发明还涉及用于储存数据的相应设备以及相应的盘式存储介质。

背景技术

[0002] 数据库系统(DBMS 数据库管理系统)对数据进行访问。该数据通常储存在硬盘上。在很多情况下,所述数据还可以储存在 ROM 中,如用于移动无线电话的语音数据库 T9 就属于这种情况。另外,举例来说,在某情况下电话簿可以被储存在光盘例如 CD-ROM 或者 DVD-ROM 上。

[0003] 然而,目前所要避免的是在光介质上储存动态数据库。其原因是光学介质读取头的跳跃时间很长,特别是在起始点和跳跃目的点间相隔很大的径向距离时,而且与硬盘相比光介质仅允许有限数目的重写周期。

[0004] 在数据库中的数据常常具备用于优化搜索操作的索引。对索引进行访问可以有数个选择:

[0005] 1 将所述索引装载到主存储器中(例如 DRAM)并在主存储器中对该索引进行访问;

[0006] 2 将该索引留在所述存储盘上并在需要的时候将其装载到所述主存储器中并使用,

[0007] 3 在所述盘上对该索引进行搜索,

[0008] 4 将该索引缓冲储存在硬盘上;在需要的时候将相应的索引装载到所述主存储器中并对其进行搜索,或者

[0009] 5 将该索引缓冲储存在硬盘上;在需要的时候直接在硬盘上对相应的索引进行搜索。

[0010] 第一种情况需要大量昂贵的主存储器。因此这种情况不太可行。第二种情况更加实际,特别是在有大量索引列表或者索引的情况下。第三种情况节省主存储器,但是也是最缓慢的情况。

[0011] 第四种情况是一种替代方案,其中适当的索引列表被临时储存在硬盘上(如果有的话)以确保更快速的访问。类似的,第五种情况使得完全在所述硬盘上进行搜索成为可能。

[0012] 相对于硬盘来说,对于可写入光盘有一些具体的限制:

[0013] - 光盘需要非常长的时间从一个任意扇区跳转到另一任意扇区(达到 1 秒)。

[0014] - 然而,对于扇区的持续读入与目前硬盘读入的数量级相仿。

[0015] - 可能对一个扇区进行重写的次数是有限的(在 1000 到 100000 次之间,视介质类型而定)。

[0016] - 光盘如 DVD 和蓝光采用“ECC”(错误校正码)对扇区进行保护。在这种情况下,16 个(DVD)或者 32 个(蓝光)扇区被结合在一起形成一个块(ECC 块)并且具备相应的附加码。这意味着只能对整个 ECC 块进行读和写,不论是要访问一个扇区的一个字节还是要访问整个 ECC 块。

[0017] 因此本发明的目的是对盘式存储介质上的数据库的应用进行优化。

发明内容

[0018] 本发明通过在盘式存储介质上更新具备相关第一索引的第一数据的方法实现上述目标,通过在所述盘式存储介质上储存能够更新所述第一数据中的数据的第二数据,创建第二索引并将所述第二索引储存在所述盘式存储介质上,其中所述第二索引仅与所述第二数据相关并且作为所述第一索引的补充被储存在所述盘式存储介质上。

[0019] 本发明还提供了一种用于在盘式存储介质上储存数据的设备,所述设备具备一写入装置,该写入装置用于将第一和第二数据写入所述盘式存储介质并且将与所述第一数据相关的第一索引写入所述盘式存储介质,其中所述第二数据对所述第一数据中的数据进行更新,所述设备还包括用于创建所述第一索引的信号处理装置,该信号处理装置也可以用于创建仅与所述第二数据相关的第二索引,并且所述写入装置可以用于将所述第二索引作为所述第一索引的补充写入所述盘式存储介质。

[0020] 上述目标还可以通过一种盘式存储介质实现,在该盘式存储介质上储存着第一数据,能够更新所述第一数据中的数据的第二数据,以及与所述第一数据相关的第一索引,还有仅与所述第二数据相关的第二索引,该第二索引作为所述第一索引的补充被储存在所述盘式存储介质上。

[0021] 这同样提供了一种为在光介质上的动态数据库有效地实现动态索引的方法。有利地,所述动态数据库的索引并不是在每次变化后都发生替换,在变化较小的情况下仅根据所发生的变化对所述索引进行扩充。只有在发生一定水平的变更时才对所述索引进行完全替换。这样减少了对于光存储介质的写访问操作,意味着光盘使用寿命的延长。

[0022] 优选地,所述盘式存储介质包括可重写光盘,如可写 CD 或 DVD。这些都是可利用的具备高存储空间的低成本存储介质。

[0023] 所述第二索引应被储存于沿径向尽可能靠近所述第一索引的位置。理想地,所述第二索引直接被写在紧邻所述第一索引的位置。然而如果在当时已经对存储盘进行了写入操作,所述第二数据或者被更新数据可能已经被置于所述第一索引之后。在这种情况下,所述第二索引应随后被写入下一个具有充足容量的空闲存储位置,这意味着从所述第一索引到所述第二索引的跳跃要尽可能小。这是因为在光盘上更短的跳跃比更长的跳跃执行的更快。由于在所述两个索引之间的小间隔或者二者以彼此紧邻的方式被写入的事实,使对所述索引的访问速度得到提高,这是因为在搜索操作中通常对两个索引都要进行搜索。不言而喻的是所述第二索引也可以被储存在所述第一索引之前,只要具备可以利用的存储空间。这种情况特别对某些仅要求对所述第二索引进行搜索以查询变化或者更新的搜索策略来说是有优势的。

[0024] 另外,所述第二索引应该以非分段形式存储。这样做同样具有在进行读操作时避免跳跃,从而提高访问速度的优势。

[0025] 通常在规律的时间间隔对数据进行更新。在具体情况中,这意味着所述第二数据被第三数据所更新。在这种情况下,对所述第二数据中的数据进行更新的所述第三数据可以储存在所述盘式存储介质上,并且与所述第二和第三数据相关的第三索引可以覆盖所述第二索引。这意味着所述第二索引被丢弃且创建并储存一个全新的变化索引。相对于将第三索引紧邻所述第二索引进行储存的策略,这样做的优势是仅需要采用两个索引,并且因此提高所述搜索速度。

[0026] 如果所述第一数据中的数据变化非常普遍,那么为全部数据创建并且储存一个包括所述变化的新的全局索引可能更为有效。为了产生一个判断标准,可以引入一个识别数字从而对所述盘式存储介质的属性和/或所述盘式存储介质与所述第二或者第三数据关系的特征进行描述,并将所述识别数字与一可指定的阈值进行比较。所述比较结果随后可以被用来确定是否要创建第二或者第三索引或者一个与所述盘式存储介质上的全部数据相关的新的全局索引。这意味着如对存储器的要求,读入时间,重写周期数目等条件都将被考虑在内。

[0027] 在某些情况下,将所述第二及第三索引的元素相结合是有益的,这样所述识别数字就不会达到所述指定的阈值。当两个索引中的一个非常小时这种情况尤为明显,并因为所述结合避免了所述读取头的跳跃。

[0028] 优选地,所述阈值是可变的并且自动与所述盘式存储介质的状态相匹配。这样做可以对仍然能够进行的重写周期数目加以考虑。

附图说明

[0029] 以下将结合附图对本发明进行更加详细的描述,附图概要示出了更新数据库的方法流程图。

具体实施方式

[0030] 以下更加详细记载的实施例为本发明的优选实施例。

[0031] 如前所述的五种索引处理情况并不绝对的要求所述索引以粘着的方式储存。然而,在每种情况并且特别在第三种情况下,所述可利用的索引至少应该以基本上粘着的方式储存在所述存储盘上。这样做是为了避免不必要的减缓从所述光盘上读入或者扫描所述索引列表。该原理同样适用于本发明所述的方法。

[0032] 对数据的更新需要特殊的索引产生过程。当在所述光盘上已经有一个索引作为起始点,并且该索引以粘着方式储存在所述存储盘上(如附图中步骤 S1 所示)时,所述索引产生过程基本上如下进行。

[0033] 所述(实际)数据库现在开始变化,例如,由于用户向所述数据库中加入了新的或者第二数据(如步骤 S2 所示)。这代表着有新的数据可以被加入到所述索引列表中。这将引起所述索引列表的变更。为了不必在此重新创建并且储存整个索引列表,仅对当前已经储存的列表进行修改(如步骤 S3 和 S7 所示)。因此在对所述索引进行下一个读访问操作的时候,所述旧索引列表(第一索引)被读取,且随后所述索引列表中的变化信息(第二索引)也被读取。

[0034] 如果所述变化信息(第二索引)很少(例如< 1MB),那么仅需再占用少量时间进

行附加的读取。在此重要的是首先所述变化信息（第二索引）沿径向靠近所述旧索引列表（第一索引），并且该变化信息也以尽可能粘着的方式存储，从而降低在所述光介质上的跳跃次数。例如，对另一个索引变化进行储存（第三索引）可能仅涉及对前面一个索引变化（第二索引）进行替换，例如所述新的变化也包括了已经作出的变化。这样可以减少在读取所述索引列表及其变化时的跳跃次数。

[0035] 通过确定一个识别数字（步骤 S4）并且将其与一阈值进行比较（步骤 S5），可以确定一点，在该点之后的变化会引起所述索引列表的彻底更新（S8）并且粘着的存储（尽可能粘着）。

[0036] 当确定所述识别数字时（步骤 S4），可能会引入下列标准，例如：

[0037] - 全部变化（第二，第三索引等）对于存储器的要求，所述全部变化是对于各个索引列表来说已经作出并且是新的。

[0038] - 全部变化（第二，第三索引等）的预计读入时间：分别是跳跃时间和所述变化扇区的读取时间以及扫描变化可能的时间。

[0039] - 在所述存储盘上仍然可以利用的总的空闲存储器空间。

[0040] - 目前已空闲的扇区已经被重写过的次数；也就是说这个次数越多所述阈值应该被设置的越高。

[0041] 如果在步骤 S5 经过检测所述识别数字超过了所述阈值，则执行以下程序：

[0042] - 所述识别数字是否有可能再次低于所述阈值，例如一个或多个已经存在的变化（第二索引）被结合到了所述新变化之中（第三索引）并且所述旧变化（第二索引）所占据的存储器被释放？如果是，选择此程序（在本图中未示出）。

[0043] 如果没有找到使所述识别数字低于所述阈值的方法，则重新产生所述索引列表（步骤 S6）并且将所述旧索引列表（第一索引）所占据的存储器与全部变化（第二和第三索引）所占据的存储器一起释放。其次，尽可能粘着的储存所述新的索引列表（步骤 S7）。

[0044] 所述阈值可以是一个固定的变量或者基于所述存储盘的状态以类似于所述识别数字的方式被确定。例如，所述识别数字可以评估所述索引的读入时间，而所述阈值可以根据所述介质已经被重写的次数，在所述存储盘上仍然可以利用的所述存储器空间以及所述存储盘的分段程度计算得出。作为选择，可以对重新储存所述索引列表以及仅储存一个变化的增益（对所述介质读入 / 重写时间上的节省）进行比较。

[0045] 本发明还记载了一种方法，其中当要求通过储存变化索引对所述索引列表进行特殊存储的时候，对所述光盘进行写访问操作的次数可以被最小化。这样，可以显著提高光盘的使用寿命。在这种情况下，搜索速度的轻微减缓一般不会带来任何后果。本发明还提供了一种用于在盘式存储介质上储存数据的设备，所述设备包括一写入装置，该写入装置用于将第一数据和第二数据写入所述盘式存储介质（S7，S8），所述第二数据对所述第一数据中的数据进行更新，并且所述写入装置还用于将与所述第一数据相关的第一索引写入所述盘式存储介质，所述设备还包括信号处理装置，该信号处理装置用于创建所述第一索引，其中，所述信号处理装置还适用于创建仅与所述第二数据相关的第二索引（S3），并且所述写入装置也适用于将所述第二索引作为所述第一索引的补充写入所述盘式存储介质（S7）。

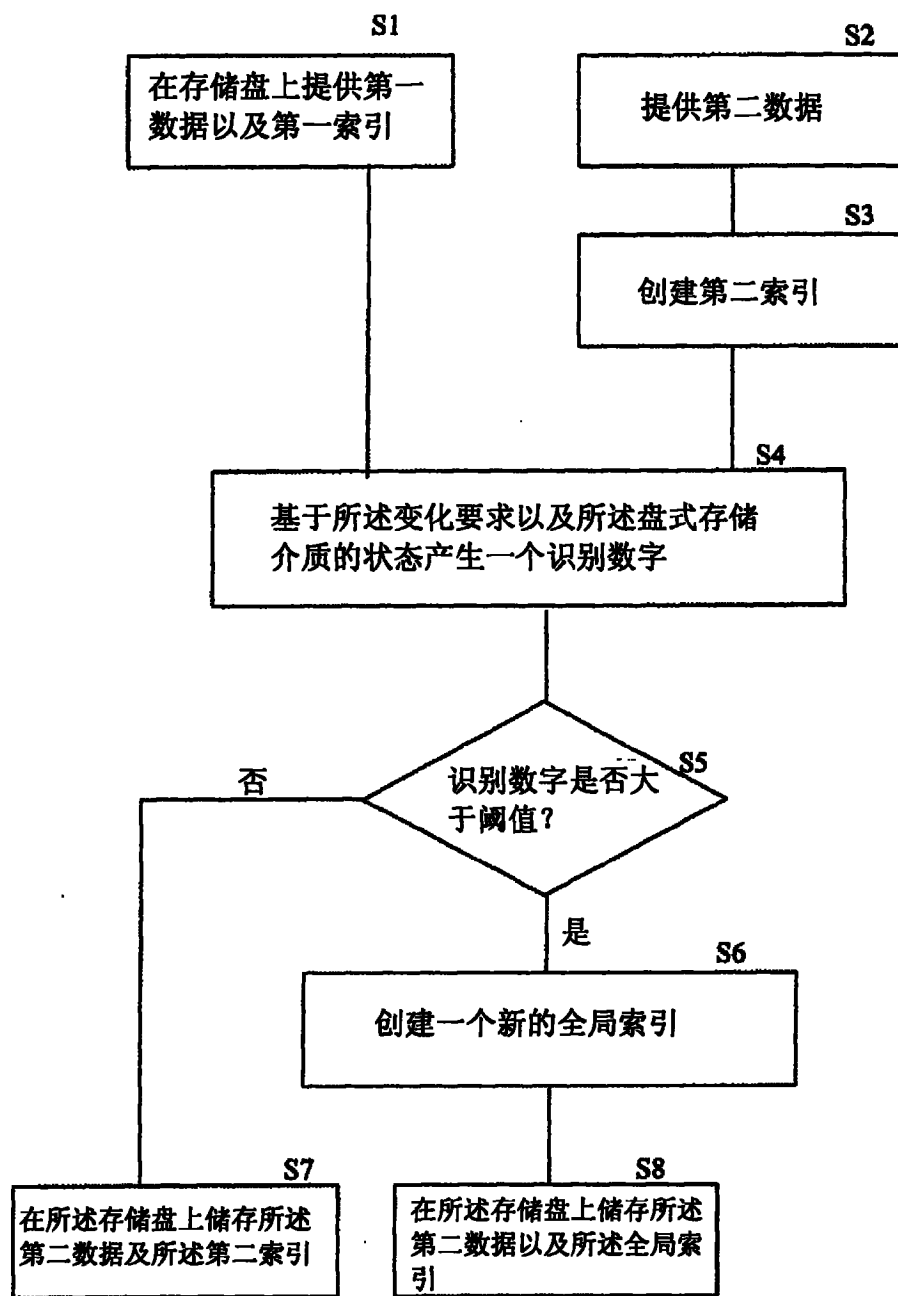


图 1