



(12) 发明专利申请

(10) 申请公布号 CN 104361113 A

(43) 申请公布日 2015. 02. 18

(21) 申请号 201410717830. 6

G06F 3/06(2006. 01)

(22) 申请日 2014. 12. 01

(71) 申请人 中国人民大学

地址 100872 北京市海淀区中关村大街 59
号中国人民大学明德商学楼 110 室

(72) 发明人 张延松 张宇 王珊

(74) 专利代理机构 北京纪凯知识产权代理有限
公司 11245

代理人 徐宁 孙楠

(51) Int. Cl.

G06F 17/30(2006. 01)

权利要求书2页 说明书7页 附图2页

(54) 发明名称

一种内存-闪存混合存储模式下的OLAP查询
优化方法

(57) 摘要

本发明涉及一种内存-闪存混合存储模式下的OLAP查询优化方法,其包括:OLAP存储采用flash-aware的存储模型,在相对较小的DRAM和相对较大flash存储之间按数据访问的局部性进行划分,在异构的两级内存上进行存储优化;内存OLAP采用数组存储,每个属性列存储于连续的数组单元中,将传统的连接操作简化为数组下标访问,进行AIR算法的OLAP查询处理;其中AIR为数组下标访问;将基于数组存储和AIR算法的OLAP查询处理分解为三个顺序的数据访问过程;通过选择向量访问flash存储度量列指定度量值;采用在基于关键字的位图连接索引的基础上在DRAM-flash两级存储中优化存储K个关键字连接位图,形成二级连接位图索引结构。本发明能提高内存存储性价比、内存和CPU使用效率以及数据存储效率,能广泛应用于通用OLAP应用场景中。

1. 一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其包括以下步骤:

1) OLAP 存储采用 flash-aware 的存储模型,即根据 OLAP 星形或雪花形模型中维表较小、谓词操作较多的特点和事实表由外键和度量属性组成的特点,在相对较小的 DRAM 和相对较大 flash 存储之间按数据访问的局部性进行划分,在异构的两级内存上进行存储优化;

2) 内存 OLAP 采用数组存储,每个属性列存储于连续的数组单元中,表由相同长度的各属性数组组成,维表使用数组下标作为主键,事实表外键为维表中相应记录的数据下标,事实表记录能够根据外键值直接定位维表中对应的数组单元,将传统的连接操作简化为数组下标访问,进行 AIR 算法的 OLAP 查询处理;其中 AIR 为数组下标访问;

3) 将基于数组存储和 AIR 算法的 OLAP 查询处理分解为三个顺序的数据访问过程:维表访问、事实表外键访问和事实表度量属性访问,三个阶段产生的中间数据结构包括:维表过滤分组向量、选择向量和分组向量、分组多维数组;维表过滤分组向量、选择向量和分组向量为各个查询共用的数据结构,不同的查询只需要更新向量的内容,分组多维数组根据查询的不同而动态生成;

4) 采用在基于关键字的位图连接索引的基础上在 DRAM-flash 两级存储中优化存储 K 个关键字连接位图,从中选择 n 个高频访问关系字并将相应的位图存储于 DRAM,其余的位图存储于 flash,形成二级连接位图索引结构。

2. 如权利要求 1 所述的一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 1) 中,采用内存存储引擎,维表驻留 DRAM;事实表中的外键是多维分析中星形连接时访问频度较高的列,同样驻留在 DRAM 中;事实表的度量属性存储于 flash 中,并提供度量列上按位置访问 API 支持度量列的随机访问。

3. 如权利要求 1 所述的一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 3) 中,维表过滤分组向量、选择向量和分组向量、分组多维数组存储于 DRAM 中;事实表度量属性存储于 flash 中。

4. 如权利要求 2 所述的一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 3) 中,维表过滤分组向量、选择向量和分组向量、分组多维数组存储于 DRAM 中;事实表度量属性存储于 flash 中。

5. 如权利要求 1 或 2 或 3 或 4 所述的一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 2) 中,所述 AIR 算法的 OLAP 查询处理包括以下步骤:

①将 OLAP 查询分解为维表上的分组过滤操作:将查询中的选择和分组操作按维表进行划分,将查询分解为在各个维表上的子查询;

②生成维表过滤分组向量:每个维表根据各自的 where 子句对记录进行过滤并投影出满足选择条件的分组属性,满足选择条件的分组属性进行字典表压缩,字典表存储于数组中,字典压缩码为字典数组下标,在过滤分组属性中选择条件为 false 的位置置为 -1,否则置为分组压缩编码,并记录在与维表等长的过滤分组向量中;

③事实表外键多趟扫描创建选择和分组向量:按维表过滤分组向量选择率由低到高的顺序依次扫描事实表对应的外键列,每列扫描时按外键值映射到相应维表过滤分组向量指定的位置,向量值非负时将当前事实表记录位置插入选择向量;下一个外键列扫描时按照当前选择向量中的记录位置采用随机访问方式,并通过相应维表过滤分组向量更新选择向

量,删除不满足后续外键映射条件的事实表记录位置;当各外键列扫描结束后,选择向量记录了事实表中满足全部连接条件的记录位置;当总选择率较高时,在更新选择向量的同时更新分组向量,用维表过滤分组向量中当前分组的下标增量地计算分组多维数组下标;当总选择率很低时,首先生成选择向量,最后按选择向量的位置随机访问各外键列,一次性生成分组向量;

④通过选择向量访问 flash 存储度量列指定度量值:根据选择向量中的位置访问 flash 存储上的度量属性值;根据选择向量在 flash 上的随机访问采用多核并行执行;

⑤通过分组向量将度量值在分组多维数组中聚集:按照分组向量所记录的分组下标,将从 flash 存储上返回的度量属性值“推”到分组向量中下标值指示的分组多维数组单元中进行聚集计算,完成全部的 OLAP 查询处理;

⑥将分组多维数组的各维反向映射到分组字典表中获得分组属性值并输出查询结果。

6. 如权利要求 5 所述的一种内存-闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤②中,所述过滤分组向量是一个动态生成的维表附加列,代替维表与事实表进行连接操作并提供连接记录在当前维上分组属性的编码;当维表上没有分组属性时,过滤分组向量简化为一个位图,用于事实表外键的连接过滤。

7. 如权利要求 1 或 2 或 3 或 4 或 6 所述的一种内存-闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 4) 中,在关键字位图连接索引的基础上采用 DRAM-flash 两级位图连接索引方法,根据内存存储空间配额精确地在 K 个关键字连接位图中选择最频繁使用的 n 个作为 DRAM 驻留位图连接索引,其余 K-n 个关键字位图作为二级索引存储在 flash 中。

8. 如权利要求 5 所述的一种内存-闪存混合存储模式下的 OLAP 查询优化方法,其特征在于:所述步骤 4) 中,在关键字位图连接索引的基础上采用 DRAM-flash 两级位图连接索引方法,根据内存存储空间配额精确地在 K 个关键字连接位图中选择最频繁使用的 n 个作为 DRAM 驻留位图连接索引,其余 K-n 个关键字位图作为二级索引存储在 flash 中。

一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法

技术领域

[0001] 本发明涉及一种数据库领域中存储优化和 OLAP(联机分析处理)查询优化方法,特别是关于一种适用于内存数据库一体机平台的以 DRAM(动态随机存取存储器)和 Flash(闪存)两级存储为基础的内存 - 闪存混合存储模式下的 OLAP 查询优化方法。

背景技术

[0002] 内存分析处理(内存 OLAP)是大数据实时分析处理的重要技术,在大内存和多核处理器并行处理能力的支持下,内存 OLAP 具有优异的实时分析处理能力,但相对于其他存储设备,如 flash、磁盘等,内存仍然是非常昂贵的存储介质,而且在存储能耗上比 flash 高一个数量级(DRAM:~100mW/GB,NAND flash:1-10mW/GB),内存 OLAP 分析处理需要以大数据为基础,内存分析处理的硬件成本很高。PCIe Flash Card 作为一种大容量(几百 GB 至 TB 级存储容量)高速存储技术,已经被大量应用于高性能数据库领域,如 Oracle Exadata X3 内存数据库一体机配置了大容量的高速闪存卡,并提供 Smart Flash Cache 缓存热数据,通过数据库逻辑优化缓存算法并能够基于表指定缓存优化策略。高速闪存卡的应用一方面为昂贵的内存存储提供了廉价的二级存储扩展能力,但另一方面高速闪存卡主要用作数据库在 flash 上的扩展缓存,扩展了内存缓存(buffer)的容量,但并没有与内存 OLAP 的存储优化及查询处理优化技术结合起来,并没有在 OLAP 算法层面实现 flash-aware 的 OLAP 优化技术。

[0003] 当前的分析型内存数据库主要以 DRAM 为主存储设备,flash 用作替代磁盘的后备存储或磁盘缓存,还没有将 flash 纳入内存的 OLAP 算法设计中。如何将内存和大容量 flash 所形成的二级存储模式应用到内存分析处理领域,使之成为高性能、高性价比的主流平台,并且内存数据库不仅要支持完全内存模式的分析处理,也需要支持 DRAM-flash 两级存储下的内存分析处理成为亟待解决的技术问题。

发明内容

[0004] 针对上述问题,本发明的目的是提供一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,该方法基于 DRAM-flash 两级存储,通过大容量廉价的 flash 提高内存存储性价比。同时,该方法能有效提高数据存储效率。进一步,本发明提供的方法可以有效提高内存和 CPU 使用效率。

[0005] 为实现上述目的,本发明采取以下技术方案:一种内存 - 闪存混合存储模式下的 OLAP 查询优化方法,其包括以下步骤:1)OLAP 存储采用 flash-aware 的存储模型,即根据 OLAP 星形或雪花形模型中维表较小、谓词操作较多的特点和事实表由外键和度量属性组成的特点,在相对较小的 DRAM 和相对较大 flash 存储之间按数据访问的局部性进行划分,在异构的两级内存上进行存储优化;2)内存 OLAP 采用数组存储,每个属性列存储于连续的数组单元中,表由相同长度的各属性数组组成,维表使用数组下标作为主键,事实表外键为维表中相应记录的数据下标,事实表记录能够根据外键值直接定位维表中对应的数组单

元,将传统的连接操作简化为数组下标访问,进行AIR算法的OLAP查询处理;其中AIR为数组下标访问;3)将基于数组存储和AIR算法的OLAP查询处理分解为三个顺序的数据访问过程:维表访问、事实表外键访问和事实表度量属性访问,三个阶段产生的中间数据结构包括:维表过滤分组向量、选择向量和分组向量、分组多维数组;维表过滤分组向量、选择向量和分组向量为各个查询共用的数据结构,不同的查询只需要更新向量的内容,分组多维数组根据查询的不同而动态生成;4)采用在基于关键字的位图连接索引的基础上在DRAM-flash两级存储中优化存储K个关键字连接位图,从中选择n个高频访问关系字并将相应的位图存储于DRAM,其余的位图存储于flash,形成二级连接位图索引结构。

[0006] 所述步骤1)中,采用内存存储引擎,维表驻留DRAM;事实表中的外键是多维分析中星形连接时访问频度较高的列,同样驻留在DRAM中;事实表的度量属性存储于flash中,并提供度量列上按位置访问API支持度量列的随机访问。

[0007] 所述步骤3)中,维表过滤分组向量、选择向量和分组向量、分组多维数组存储于DRAM中;事实表度量属性存储于flash中。

[0008] 所述步骤2)中,所述AIR算法的OLAP查询处理包括以下步骤:①将OLAP查询分解为维表上的分组过滤操作:将查询中的选择和分组操作按维表进行划分,将查询分解为在各个维表上的子查询;②生成维表过滤分组向量:每个维表根据各自的where子句对记录进行过滤并投影出满足选择条件的分组属性,满足选择条件的分组属性进行字典表压缩,字典表存储于数组中,字典压缩码为字典数组下标,在过滤分组属性中选择条件为false的位置置为-1,否则置为分组压缩编码,并记录在与维表等长的过滤分组向量中;③事实表外键多趟扫描创建选择和分组向量:按维表过滤分组向量选择率由低到高的顺序依次扫描事实表对应的外键列,每列扫描时按外键值映射到相应维表过滤分组向量指定的位置,向量值非负时将当前事实表记录位置插入选择向量;下一个外键列扫描时按照当前选择向量中的记录位置采用随机访问方式,并通过相应维表过滤分组向量更新选择向量,删除不满足后续外键映射条件的事实表记录位置;当各外键列扫描结束后,选择向量记录了事实表中满足全部连接条件的记录位置;当总选择率较高时,在更新选择向量的同时更新分组向量,用维表过滤分组向量中当前分组的下标增量地计算分组多维数组下标;当总选择率很低时,首先生成选择向量,最后按选择向量的位置随机访问各外键列,一次性生成分组向量;④通过选择向量访问flash存储度量列指定度量值:根据选择向量中的位置访问flash存储上的度量属性值;根据选择向量在flash上的随机访问采用多核并行执行;⑤通过分组向量将度量值在分组多维数组中聚集:按照分组向量所记录的分组下标,将从flash存储上返回的度量属性值“推”到分组向量中下标值指示的分组多维数组单元中进行聚集计算,完成全部的OLAP查询处理;⑥将分组多维数组的各维反向映射到分组字典表中获得分组属性值并输出查询结果。

[0009] 所述步骤②中,所述过滤分组向量是一个动态生成的维表附加列,代替维表与事实表进行连接操作并提供连接记录在当前维上分组属性的编码;当维表上没有分组属性时,过滤分组向量简化为一个位图,用于事实表外键的连接过滤。

[0010] 所述步骤4)中,在关键字位图连接索引的基础上采用DRAM-flash两级位图连接索引方法,根据内存存储空间配额精确地在K个关键字连接位图中选择最频繁使用的n个作为DRAM驻留位图连接索引,其余K-n个关键字位图作为二级索引存储在flash中。

[0011] 本发明由于采取以上技术方案,其具有以下优点:1、本发明由于采用 DRAM-flash 两级存储作为内存 OLAP 的存储平台。相对于完全内存存储模型, DRAM-flash 两级存储大幅度降低了对昂贵内存的需求,降低了硬件的整体成本,通过 flash-aware 存储模型及 AIR OLAP 算法的优化,对 flash 上存储的大量度量属性采用高效率的随机访问,缩小了 flash 存储的性能差距。2、本发明采用的 AIR OLAP 查询处理算法将一个完整的 OLAP 查询处理过程划分为三个独立的处理阶段,在维表和事实表外键处理阶段只涉及较小比例的数据,每个维表生成一个定长的过滤分组向量;事实表上无论查询涉及的列数量多少都只需要一个选择向量和一个分组向量,选择和分组向量的初始长度由维上最低的选择率决定,在事实表外键的星形连接中长度不断缩短,内存存储开销有限;数据库中存储空间占比最大的度量属性存储在大容量 flash 中,传统的行式访问(包括行存储以及列存储上的行式访问) OLAP 查询处理在事实表全表扫描过程中完成流水线式的哈希连接和哈希分组聚集操作, AIR OLAP 查询算法将连接、分组操作和聚集操作分解,不使用传统数据库中常用的全表扫描而是在数量有限的事实表外键上完成连接分组操作,然后根据选择率极低的选择向量按位置访问指定度量列的指定位置,将 flash 存储上的数据访问操作推到最后处理阶段,极大减少了 flash 上的数据访问负载,同时也能够发挥出 flash 良好的随机访问性能。3、本发明由于采用的位图连接索引是一种增量索引,因事实表记录增加而产生的索引更新操作是位图的顺序增长,而不是象 B+ 树索引那样更新节点内容和结构,能够消除 flash 上数据更新的数据重写代价,因此更适合 flash 存储。DRAM-flash 两级内存位图连接索引方法在基于关键字位图连接索引的基础上通过查询关键字访问频度以及内存存储空间对关键字位图采用两级存储模式,将 K 个关键字位图中 n 个高频使用的关键字位图存储于 DRAM,其余 K-n 个关键字位图存储于 flash,降低了内存索引的存储开销。4、本发明是通过 OLAP 模式和负载的特点以及 OLAP 查询处理算法中数据的局部性特征(即数据被频繁访问的程度)建立异构存储模型,以表、索引、表中不同数据局部性强度的列组以及查询依赖的中间数据结构为对象,根据内存容量及数据访问性能约束将其优化分布在高性能但容量相对较小的 DRAM 和大容量但性能相对较低的 flash 存储中,提高数据存储效率。5、本发明是在查询处理时根据 DRAM 和 flash 存储的特性,将 flash 上的数据访问推后执行,将 OLAP 完整的查询处理过程划分为内存处理和 flash 计算两个阶段,支持 OLAP 查询间不同存储访问阶段上的流水并行,提高内存和 CPU 使用效率。本发明可以广泛应用于通用 OLAP 应用场景中。

附图说明

[0012] 图 1 是本发明的内存 OLAP 在 DRAM-flash 两级存储上的存储示意图;

[0013] 图 2 是本发明实施例中内存 OLAP 在 DRAM-flash 两级存储上的查询处理示意图;

[0014] 图 3 是本发明的 DRAM-flash 两级存储上的关键字位图连接索引处理示意图。

具体实施方式

[0015] 现有的内存 OLAP 技术通常采用完全内存计算模式,或者将大容量 flash 用作高速缓存,前者增加了内存计算的成本,后者难以将 OLAP 优化从内存扩展到 flash 上。为此,本发明提出了一种基于 DRAM-flash 两级存储的内存-闪存混合存储模式下的 OLAP 查询优化方法,根据 OLAP 模式、负载以及 OLAP 算法的特点优化 OLAP 数据在内存和 flash 上的存储,

面向两级存储特点优化 OLAP 查询算法。本发明适用于通用 OLAP 应用场景。下面结合附图和实施例对本发明进行详细的描述。

[0016] 如图 1 所示,本发明提供一种内存-闪存混合存储模式下的 OLAP 查询优化方法,该方法基于 DRAM-flash 两级存储,面向采用 DRAM 和大容量 PCIe Flash Card 构成的 DRAM-flash 两级存储上的内存 OLAP 查询优化方法,其包括以下步骤:

[0017] 1) OLAP 存储采用 flash-aware 的存储模型,即根据 OLAP 星形或雪花形模型中维表较小、谓词操作较多的特点和事实表由外键和度量属性组成的特点,在相对较小的 DRAM 和相对较大 flash 存储之间按数据访问的局部性进行划分,在异构的两级内存上进行存储优化,以提高内存中高频使用数据的效率。

[0018] 由于维表较小且现代 OLAP 支持维表上的更新,因此采用内存存储引擎,维表驻留 DRAM。事实表中的外键是多维分析中星形连接时访问频度较高的列,同样驻留在 DRAM 中。事实表的度量属性较多,OLAP 查询通常只针对少数度量属性和极低选择率上的度量值,因此存储于廉价大容量的 flash 中,并提供度量列上按位置访问 API 支持度量列的按位置随机访问。

[0019] 2) 内存 OLAP 采用数组存储,每个属性列存储于连续的数组单元中,表由相同长度的各属性数组组成,其中较优地,维表使用数组下标作为主键,事实表外键为维表中相应记录的数据下标,事实表记录能够根据外键值直接定位维表中对应的数组单元,将传统的连接操作简化为数组下标访问 (Array Index Reference, AIR),进行 AIR 算法的 OLAP 查询处理。

[0020] 3) 将基于数组存储和 AIR 算法的 OLAP 查询处理分解为三个顺序的数据访问过程:维表访问、事实表外键访问和事实表度量属性访问,三个阶段产生的中间数据结构包括:维表过滤分组向量、选择向量和分组向量、分组多维数组。维表过滤分组向量、选择向量和分组向量为各个查询共用的数据结构,不同的查询只需要更新向量的内容;分组多维数组根据查询的不同而动态生成。这三类中间数据结构在查询处理过程中被重复使用,属于强局部性数据集,需要存储于 DRAM 中;事实表度量属性所占存储空间较大,但查询中通常只访问较少的度量属性,而且通过选择向量只需要随机访问度量属性列中极低比例的记录,因此度量属性列可以存储于 flash 中以减少内存 OLAP 对 DRAM 的需求,降低系统的硬件成本。

[0021] 本发明所采用的 AIR OLAP 算法需要选择向量、分组向量、维表过滤分组向量、分组多维数组等数据用于 OLAP 查询处理过程,这些数据结构在 OLAP 查询间可以重复使用,所占内存空间固定,因此常驻于 DRAM。

[0022] 4) 数据仓库通常使用位图连接索引来优化或消除维表和事实表之间的连接代价。本发明采用在基于关键字的位图连接索引(即按关键字而不是整个属性成员来建立位图连接索引)的基础上在 DRAM-flash 两级存储中优化存储 K 个关键字连接位图(即每个关键字对应一个与事实表等长的位图,记录该关键字所对应的事实表记录位置),从中选择 n 个高频访问关系字并将相应的位图存储于 DRAM,其余的位图存储于 flash,形成二级连接位图索引结构;

[0023] 其中,本发明在关键字位图连接索引的基础上采用 DRAM-flash 两级位图连接索引方法,根据内存存储空间配额精确地在 K 个关键字连接位图中选择最频繁使用的 n 个作为 DRAM 驻留位图连接索引,其余 K-n 个关键字位图作为二级索引。具体为:索引项为某个

维表属性值所对应的事实表位置位图,位图索引的大小为定长的位图大小与关键字数量的乘积,位图索引可以通过数据压缩进一步缩减内存存储空间。通过查询执行日志和查询关键字分析,可以确定 K 个最频繁使用的维属性关键字并为其建立位图索引,根据内存空间配额可以将其中 n 个最频繁使用的关键字位图驻留于内存,其余 K-n 个关键字位图存储于 flash,形成二级连接位图索引结构。

[0024] 上述步骤 2) 中,AIR 算法的 OLAP 查询处理包括以下步骤:

[0025] ①将 OLAP 查询分解为维表上的分组过滤操作:将查询中的选择和分组操作按维表进行划分,将查询分解为在各个维表上的子查询。

[0026] ②生成维表过滤分组向量:每个维表根据各自的 where 子句对记录进行过滤并投影出满足选择条件的分组属性,满足选择条件的分组属性进行字典表压缩,字典表存储于数组中,字典压缩码为字典数组下标,在过滤分组属性中选择条件为 false 的位置置为 -1,否则置为分组压缩编码,并记录在与维表等长的过滤分组向量中。

[0027] 过滤分组向量是一个动态生成的维表附加列,代替维表与事实表进行连接操作并提供连接记录在当前维上分组属性的编码。当维表上没有分组属性时,过滤分组向量简化为一个位图,用于事实表外键的连接过滤。

[0028] ③事实表外键多趟扫描创建选择和分组向量:按维表过滤分组向量选择率由低到高的顺序依次扫描事实表对应的外键列,每列扫描时按外键值映射到相应维表过滤分组向量指定的位置,向量值非负时将当前事实表记录位置插入选择向量;下一个外键列扫描时按照当前选择向量中的记录位置采用随机访问方式,并通过相应维表过滤分组向量更新选择向量,删除不满足后续外键映射条件的事实表记录位置;当各外键列扫描结束后,选择向量记录了事实表中满足全部连接条件的记录位置。当总选择率较高时,在更新选择向量的同时更新分组向量,用维表过滤分组向量中当前分组的下标增量地计算分组多维数组下标;当总选择率很低时,首先生成选择向量,最后按选择向量的位置随机访问各外键列,一次性生成分组向量。

[0029] ④通过选择向量访问 flash 存储度量列指定度量值:根据选择向量中的位置访问 flash 存储上的度量属性值,只需要返回极少的度量值进行分组聚集计算;flash 具有良好的并行随机访问性能,根据选择向量在 flash 上的随机访问采用多核并行执行,减少 flash 数据访问延迟。

[0030] ⑤通过分组向量将度量值在分组多维数组中聚集:按照分组向量所记录的分组下标,将从 flash 存储上返回的度量属性值“推”到分组向量中下标值指示的分组多维数组单元中进行聚集计算,完成全部的 OLAP 查询处理。

[0031] ⑥将分组多维数组的各维反向映射到分组字典表中获得分组属性值并输出查询结果。

[0032] 实施例:

[0033] 如图 2 所示,flash-aware 体现在通过对 OLAP 查询算法的优化将对 flash 上存储的度量属性的访问推到最后,将事实表顺序扫描转换为根据选择向量对事实表度量属性执行低选择率的随机访问,降低 flash 上度量属性访问的延迟。

[0034] 步骤 2) 以下面的查询命令为例:

[0035] SELECT c_nation,s_nation,sum(l_revenue),sum(l_price)

[0036] FROM customer, lineorder, supplier

[0037] WHERE lo_custkey = c_custkey

[0038] and lo_suppkey = s_suppkey

[0039] and c_region = 'AMERICA'

[0040] and s_region = 'ASIA'

[0041] group by c_nation, s_nation;

[0042] 查询命令需要将事实表 lineorder 和维表 customer, supplier 连接, 然后按维表的 c_region 和 s_region 属性对事实表度量列 l_revenue 和 l_price 求累加和。

[0043] 如图 3 所示, 本发明步骤 4) 中 DRAM-flash 两级存储上的关键字位图连接索引处理实施例:

[0044] 传统的数据库中, 索引是一种平面结构, 假设在相同的存储层次上使用。B+- 树索引等磁盘数据库的索引技术通过缓冲区机制实现内存缺失节点在磁盘和内存的数据交换, 但仍然是一种不透明的机制, 其索引访问效率取决于缓冲区替换算法的效率, 不能被数据库定制式优化。在 OLAP 应用中, 索引项存在于维表属性, 但索引的对象则是事实表中对应的连接记录, 普通的 B+- 树索引只能检索单表上的数据, 对 OLAP 应用而言, 较小的维表上的索引访问性能的提高对 OLAP 查询整体性能的加速作用有限, 而对事实表外键建立索引能够加速事实表与维表之间的连接操作, 但事实表外键上的索引存储开销极为庞大, 而且现代 OLAP 应用中事实表需要实时大量更新, 索引的更新代价巨大。本发明采用的是位图连接索引方法, 即通过连接操作为维表上指定的属性成员建立事实表连接位图, 每个成员建立一个位图, 指示该成员在事实表记录上的连接位置。位图连接索引是一种增量索引, 在事实表大量数据插入时只需要增量地扩展位图长度即可, 不需要对已建立的连接位图进行重构。传统数据库中使用的位图连接索引是以属性为粒度为所有成员建立连接位图, 面临着低势集属性位图连接索引空间开销小但选择率过高, 而高势集属性选择率低但连接位图数量多, 存储开销过大的矛盾。本发明采用按查询负载中的维属性关键字的 TOP K 访问频度选择位图连接索引关键字, 并为 K 个全局高频访问关键字建立全局位图连接索引, 形成一个 key/value 索引结构, key 为关键字的全局名称, 包括表及关键字信息, value 则是连接位图。不同维表、不同属性的关键字具有相同的连接位图长度, 可以存储在全局位图连接索引中。其中较优地, 本发明在关键字位图连接索引的基础上提出了 DRAM-flash 两级位图连接索引方法, 根据内存存储空间配额精确地在 K 个关键字连接位图中选择最频繁使用的 n 个作为 DRAM 驻留位图连接索引, 其余 K-n 个关键字位图作为二级索引。如图 3 所示, 显示了查询关键字位图分别存在于 DRAM 和 flash 两级存储的应用场景, 首先 DRAM 中的位图完成选择操作, 生成过滤位图, 当过滤位图的选择率在某个阈值范围 (S_{low} , S_{high}) 之间时, 按照生成的过滤位图中“1”的位置向 flash 存储中相应位图的指定位置进行随机访问, 确定该位置最终的逻辑结果。Flash 具有较好的随机访问能力, 本发明采用对过滤位图顺序访问, 对 flash 位图并行访问的策略, 提高 flash 上数据的并行访问性能, 减少位图索引计算的延迟, 如图 3 中所示的并行 flash 访问线程。假设 DRAM 中过滤位图的选择率为 S_1 , flash 中位图的选择率为 S_2 (位图的选择率可以由“1”的数量精确地给出), $T_F(S_1)$ 为在 flash 上的位图访问延迟, $T_p(S_2, S_1)$ 为在索引选择率为 S_2 和 S_1 时查询处理的时间差, 当 $T_p(S_2, S_1) - T_F(S_1) > 0$ 时, flash 上的位图访问具有查询性能收益。 S_{low} 和 S_{high} 为满足查询收

益的最低和最高选择率差值（如 S_2-S_1 ）。

[0045] 综上所述,与现有技术相比,本发明采用基于 DRAM-flash 两级存储的内存 OLAP 查询优化方法,支持 DRAM-flash 两级存储上的内存 OLAP 查询处理,降低内存 OLAP 对昂贵内存的需求,提高内存 OLAP 的性价比。通过基于 DRAM-flash 两级存储的存储优化方法、OLAP 查询优化方法和索引优化方法,透明地优化在 OLAP 查询处理过程中的数据访问和查询处理性能。本发明同时提高了内存 OLAP 大数据的访问数据管理能力和查询处理性能。

[0046] 上述各实施例仅用于说明本发明,其中各部件的结构、连接方式和制作工艺等都是可以有所变化的,凡是在本发明技术方案的基础上进行的等同变换和改进,均不应排除在本发明的保护范围之外。

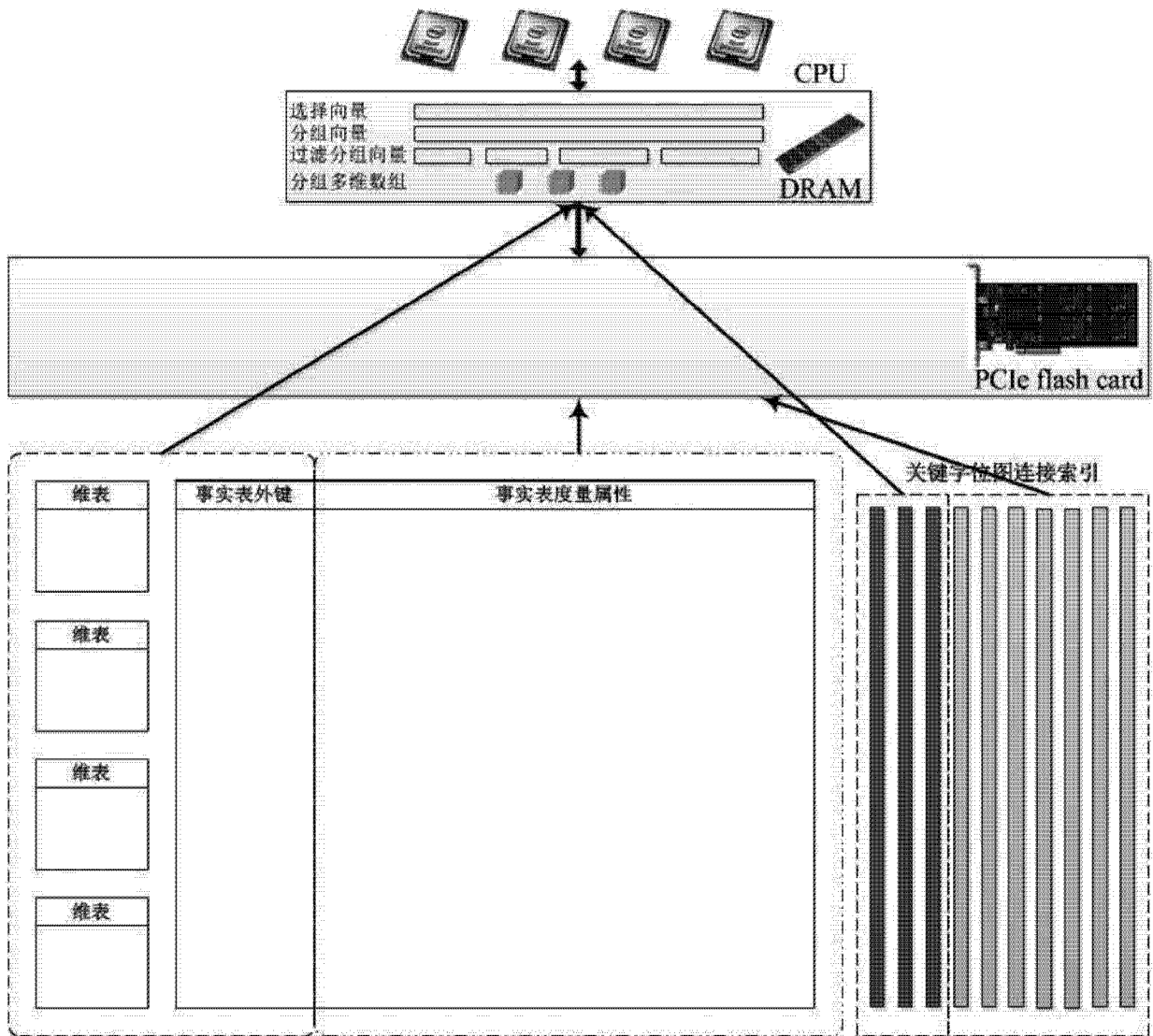


图 1

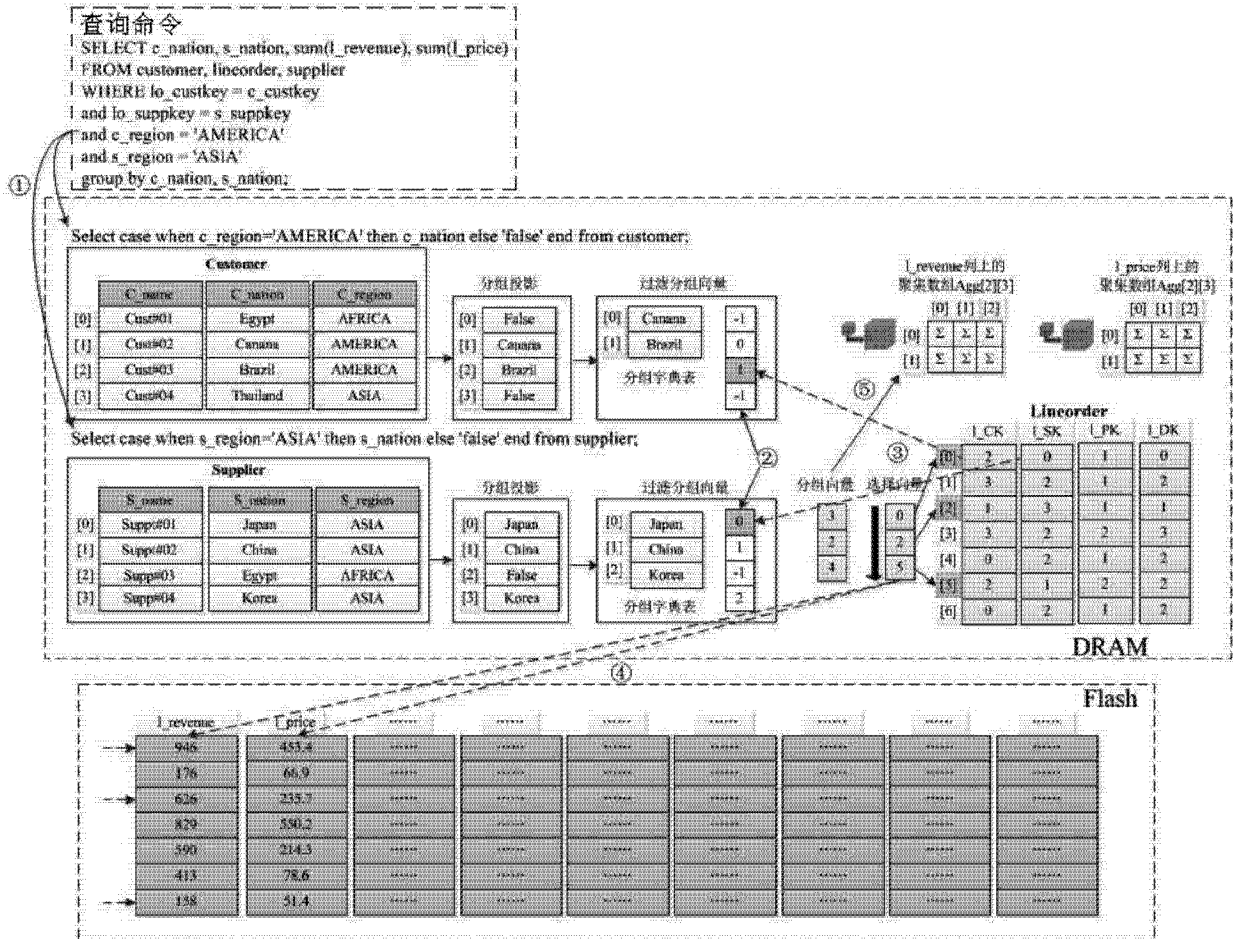


图 2

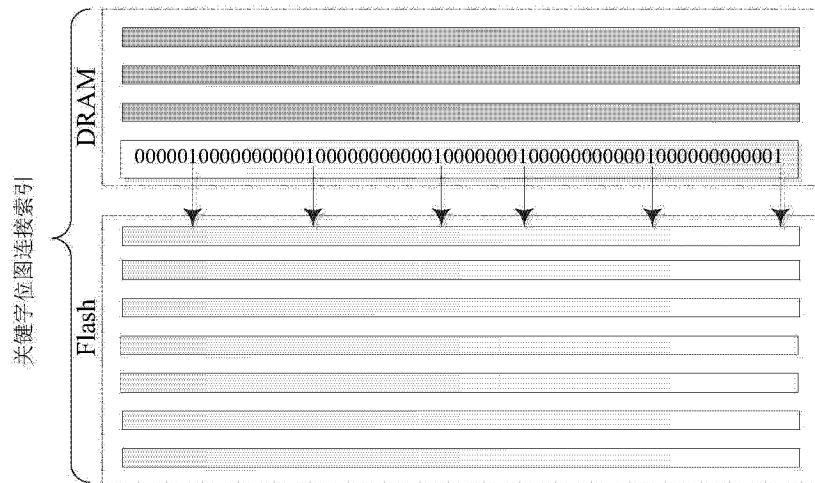


图 3