

多盘广播的偏斜索引算法

马小琴, 余建桥

(西南大学计算机与信息科学学院, 重庆 400715)

摘要: 在移动环境中, 多盘广播是被广泛使用的数据广播调度算法之一。为广播数据建立索引可以使移动客户机选择性地监听信道, 以减少电源消耗, 但过多地插入索引会增加数据访问时间。针对该问题, 为多盘广播建立偏斜索引, 设计索引树构造算法和广播数据的索引树算法。与可变扇出算法相比, 多盘广播的偏斜索引算法需要的平均访问时间和平均调谐时间较短。

关键词: 多盘广播; 移动环境; 调度算法; 偏斜索引; 调谐时间

Skew Index Algorithm for Multi-disk Broadcast

MA Xiao-qin, YU Jian-qiao

(College of Computer and Information Science, Southwest University, Chongqing 400715)

【Abstract】 In mobile environment, multi-disk broadcast is one of the widely used schedule algorithm of data broadcast. Building index for the broadcast data can make the mobile client monitor the channel selectively, and reduce energy consumption. But inserting too many indexes will increase data access time. Aiming at this problem, this paper builds skew index for multi-disk broadcast, designs construction algorithm for index tree and the index tree algorithm of broadcast data. Compared with Variant Fanout(VF) algorithm, skew index algorithm for multi-disk schedule needs less access time and tuning time in average.

【Key words】 multi-disk broadcast; mobile environment; schedule algorithm; skew index; tuning time

1 概述

在移动环境中, 数据广播是一种有效的数据访问方式, 它能以很小的代价支持大量移动客户机(Mobile Client, MC)同时访问数据库^[1]。MC通过侦听信道的方式获取所需数据, 在侦听过程中, MC需要不断地匹配数据, 其电源消耗很大。

目前, 树索引是数据广播中普遍采用的索引技术之一, 主要分为偏斜访问模式的索引技术和均匀访问模式的索引技术。文献[2]的索引算法是基于均匀访问模式的有效索引算法, 但在实际应用中, 其数据访问概率经常呈现倾斜性^[3]。文献[4]讨论了偏斜访问模式下的索引算法, 该算法减少了热数据的访问时间, 但增加了冷数据的访问时间, 且没有考虑索引节点的复制问题^[5]。多盘广播是一种广泛使用的调度算法^[6], 其盘内各数据块访问概率几乎相同, 盘间访问概率具有明显的偏斜性。针对多盘广播的特点, 本文提出一种偏斜索引算法, 该算法考虑了各盘数据项的访问概率和索引节点的复制问题, 减少了热数据的目录丢失率, 且不会明显增加冷数据的平均访问时间。

2 多盘广播和可变扇出算法

2.1 多盘广播调度算法

文献[6]提出的多盘广播调度是一种被广泛使用的调度算法, 它将待广播的数据组织成多个具有不同容量和转速的虚拟磁盘。其基本思想是将数据库的所有数据项按访问概率分为 K 组, 即 B_1, B_2, \dots, B_K , 其中, B_1 中的数据项访问概率最高, 而 B_K 最低。在一个广播调度中, 同一组内的数据项以相同的频率广播, 而不同组的广播频率是不同的。假设 B_i 的广播频率为 f_i , 则 $f_1 > f_2 > \dots > f_K$ 。通过将访问概率高的数据项放在相对快速的磁盘中, 可以降低所有请求的平均访问时间。

2.2 VF算法

文献[4]提出为广播数据建立基于访问概率的索引算法, 即可变扇出(Variant Fanout, VF)算法。VF算法用从上至下的

方法构造索引树, 它先将所有记录依访问概率递减的次序放在根节点之下, 然后通过一定的衡量标准, 将一些不经常访问的记录移到下一层, 以降低平均访问代价。对下一层按相同方法递归进行, 直到不满足衡量的标准时结束。

2.3 多盘广播与VF算法的联系

多盘广播调度是依照各数据项的访问概率进行分盘, 各盘之间的访问概率具有明显倾斜性。VF算法是一种基于访问概率的索引算法, 访问概率的倾斜性越大, VF算法优势越强。因此, 为减少多盘广播的调谐时间, 将VF算法应用在多盘广播中构成一种多盘广播下的偏斜索引, 以减少多盘广播的平均调谐时间, 并使其平均访问时间保持在一定范围内。

3 多盘广播的偏斜索引算法

3.1 算法基本思想

多盘广播偏斜索引(Skew Index, SI)算法的主要思想是分别为多盘广播的每个盘依据访问概率建立一个度为 d 的局部索引树, 并对各盘的根节点用VF算法构造一个偏斜的索引树。

3.2 算法的具体实现

3.2.1 算法实现的适用范围

为使本算法更好地应用在数据广播环境中, 作如下假设:

- (1)数据访问是偏斜的。
- (2)单信道可靠地广播数据项。
- (3)数据项广播是非均匀的, 即在一个广播周期中, 热数据比冷数据广播的次数多。
- (4)客户端每次只请求一个数据项, 如单项请求。
- (5)在广播信道中, 桶是一个逻辑传送单元, 一个索引节点只放入一个桶, 称为索引桶, 数据节点放入的桶称为数据桶。

3.2.2 算法过程及实例

算法过程如下: (1)为每个盘以自底向上的方式建立一个

作者简介: 马小琴(1982 -), 女, 硕士研究生, 主研方向: 移动数据库, 数据库技术; 余建桥, 教授、博士

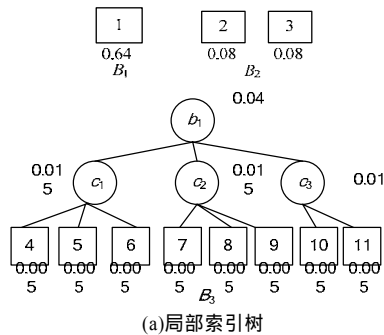
收稿日期: 2009-07-27 **E-mail:** xqma@swu.edu.cn

度为 d 的局部索引树, d 是手动设置的。(2)将各盘索引树的根节点用 VF 算法构造索引树。(3)用 $\text{allocateindexnode}(L)$ 为广播盘分配索引节点, L 是一个多盘广播周期里所有数据项的线性表。(4)检查每个子周期的第 1 个节点是否是根节点, 如果不是, 就把根节点插入到子周期之前。

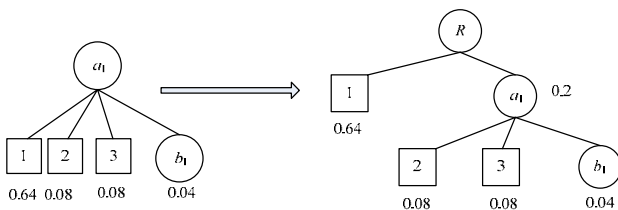
索引节点分配过程 $\text{allocateindexnode}(L)$ 描述如下: (1)取线性链表 L 的第 1 个数据项送给中间变量 C 。(2)如果 C 不为空, 则将 C 值送给中间变量 a , 否则转第(7)步。(3)当 a 不为空时, 判断 a 是否为根节点, 如果是根节点则置其父节点为空, 否则取 a 的父节点。(4)判断 a 是否是其父节点的第 1 个孩子, 如果是, 则将父节点入栈 indexstack , 并将父节点送给 a , 转第(3)步, 否则执行第(5)步。(5)将 indexstack 的内容依次出栈并输出, 直到栈为空时, 输出 C 值。(6)取 L 中的下一个数据项送给 C , 转第(2)步。(7)返回。

以图 1 为例, 其中, $B_1 \sim B_3$ 中的数据项作为一个输入。

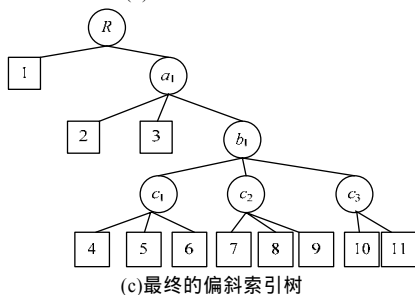
(1)为每个盘以自底向上的方式建立一个局部索引树, 最下层的数据项自左向右以度为 3 的方式构成各个索引节点, 允许最右边节点的度数小于 3。将这些索引节点以同样的方式进行处理, 直到只有一个根节点为止, 结果如图 1(a)所示。图 1 中的圆表示索引节点, 正方形表示数据节点。



(a)局部索引树



(b)VF 索引树的构造



(c)最终的偏斜索引树

图 1 偏斜索引树构造过程

(2)对各盘形成的索引树的根节点用 VF 算法构造根节点索引树, 根据表 1, 构造过程如图 1(b)所示, 得到的索引树如图 1(c)所示。

表 1 节点划分的集合

i	1	2
$(4-i-1)\sum_{1 \leq j \leq i} Pr(h_j)$	2×0.64	0.64
$\sum_{1 \leq j \leq m} Pr(h_j)$	0.20	0.11
$Y(i)$	1.08	0.53

(3)建立偏斜索引树后, 用 allocateindexnode 算法为广播周期中的数据项分配索引节点。在此算法中, 广播周期的数据项顺序存储在线性链表 L 中, 从数据节点到根节点扫描索引树, 以决定哪个索引节点应该放在数据项之前, 规则如下: 如果被访问的节点是其父节点的第 1 个孩子, 那么它的父亲应该放在这个节点之前, 该节点访问其父节点, 否则, 停止扫描索引树。相应索引节点放在其第 1 个孩子节点之前的原因是在同一个广播周期中, 索引节点能够覆盖其所有数据项。如果仅把这些索引节点放在其他孩子节点之前, 那么被这些索引节点覆盖的一些数据项可能还没有广播, 目录丢失将导致在本周期内不能收到这些数据项, 从而增加了访问时间。如果在每个数据项前面插入这些索引节点, 将增大广播周期, 进而增加访问时间。

在 allocateindexnode 中, 对于数据项 1, 由于它不是广播周期的最后一个节点, 因此将其值送给 a 并开始从它到根节点扫描索引树。由于它不是根节点, 且数据项 1 是其父节点的第 1 个孩子节点, 因此 R 进 indexstack , 如图 2(a)所示, 在第 2 层内层循环, 由于当前节点是根节点, 因此父节点置空。这样就不满足 a 是其父节点的第 1 个孩子, 跳过第(4)步, 执行第(5)步, 将所有索引节点出栈并依次放在数据项 1 之前。对于第 2 个数据项, 由于数据项 2 是索引节点 a_1 的第 1 个孩子, 因此 a_1 进栈, 如图 2(b)所示, 当 a_1 被访问时, 它不再是 R 的第 1 个孩子, 索引节点 a_1 出栈, 放在数据项 2 的前面。对于第 3 个数据项 4, 由于数据项 4 和索引节点 c_1 是它们父节点的第 1 个孩子, 因此 c_1 与 b_1 进栈, 如图 2(c)所示, 当 b_1 被访问时, 由于它不是 a_1 的第 1 个孩子, 因此 c_1 和 b_1 依次出栈, 放在数据项 4 之前。对于第 4 个数据项 5, 因为它不是父节点的第 1 个孩子, 所以其前不用插入索引节点, 这样形成了一个子周期 m_1 的广播内容, 其他子循环按相同方式, 用 allocateindexnode 处理所有剩下的数据项。一个广播周期的索引节点分配如图 3(a)所示。图 3(b)是一个带索引节点的广播周期的广播结果。

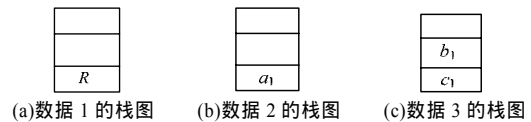
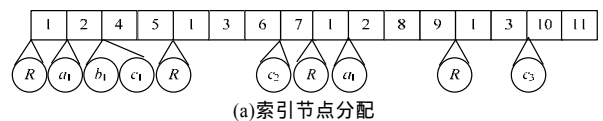
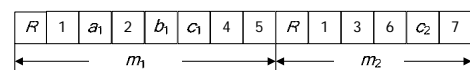


图 2 子周期 m_1 内各数据项的栈内容



(a)索引节点分配



(b)最终的带索引广播周期

图 3 索引分配

(4)由于客户端需要从根节点开始扫描索引树, 如果子周期的第 1 个节点不是根节点, 就把根节点插入到子周期之前, 这样有助于 MC 快速找到相应的索引节点。在图 3(b)中, 由于每个子周期的第 1 个节点是根节点, 因此第(4)步处理后的结果没有改变。

3.3 访问协议

算法的访问协议描述如下: (1)调谐广播信道接收当前

桶。(2)读当前桶,获取最近根节点的偏移量,进入睡眠状态。(3)调谐广播信道接收最近的根节点,并扫描索引树直到获取到所需数据项或确认没找到。

4 实验

4.1 实验模型

实验中使用 Zipf 分布来描述数据访问的偏斜性。给出 Zipf 分布的 N 个访问概率是 $PR(h)$ 的数据项 ($1 \leq h \leq N$)。Zipf 分布表示成 $PR(h) = \frac{(1/h)^\theta}{\sum_{i=1}^N (1/i)^\theta}$ ($1 \leq h \leq N$), 当 $\theta=0$ 时, 为均匀分布, 随着 θ 值的增大, 访问呈倾斜性。实验参数设置见表 2。

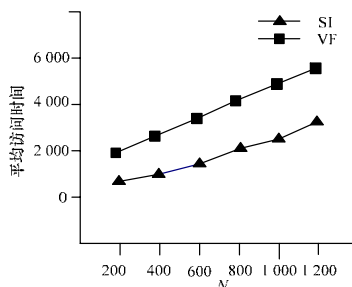
参数	含义
N	数据项数目
S	广播盘的总数
K_i	盘 j 的数据项个数满足 $1 \leq j \leq S$
θ	Zipf 函数分布因子
η	数据节点的大小
d	索引节点度数

因为 SI 与 VF 有相似的环境假设, 所以本文对它们进行比较。VF 算法要求数据项以访问概率递减排列, 因此, 给出 N 个基于 Zipf 分布的访问概率依次递减的数据项。VF 算法依据访问概率动态调整索引树, SI 算法分别为每个盘建立度为 d 的索引树, 并将各盘索引树的根节点用 VF 算法构成一棵索引树。

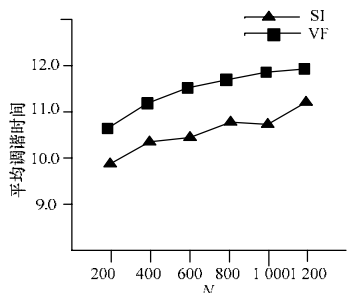
4.2 实验结果及分析

4.2.1 N 取值的影响

本实验取 $\theta = 0.8$, 数据项的数目 N 取值为 200~1 200, $d=4$, $\eta = 5$ 。从图 4(a)可以看出, 随着 N 值的增大, SI 比 VF 的访问时间增长缓慢。这是因为随 N 变大, VF 算法因为没有复制索引节点而导致目录的丢失增多, 所以增加了访问时间。而 SI 算法合理地插入了索引节点, 可以快速找到所需数据。从图 4(b)可以看到, SI 比 VF 有更短的调谐时间, 这是因为随 N 的增大, SI 探索的平均索引节点数目比 VF 算法少。



(a) N 对访问时间的影响



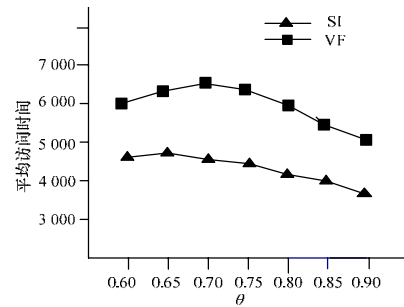
(b) N 对调谐时间的影响

图 4 固定 θ 时 N 对调谐时间和访问时间的的影响

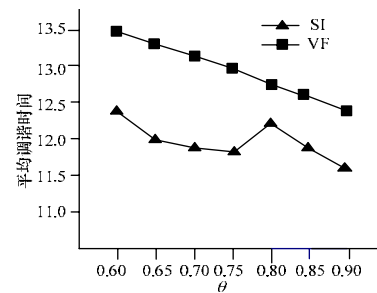
4.2.2 θ 取值的影响

本实验设定一个广播周期广播的数据项数为 1 000, 且

$d=4$, $\eta=5$, $\theta = [0.6, 0.9]$ 。随着 θ 的增大, 访问呈倾斜性, 从图 5 可以看出, SI 在一个广播周期中为热数据分配更多的索引节点次数, 减少了热数据的访问时间和调谐时间, 且由于盘内建立固定度的索引树, 因此不会明显增加冷数据的访问时间。而 VF 算法没有考虑冷数据的访问时间和调谐时间, 因此, SI 比 VF 需要更少的平均访问时间和平均调谐时间。



(a) θ 对访问时间的的影响

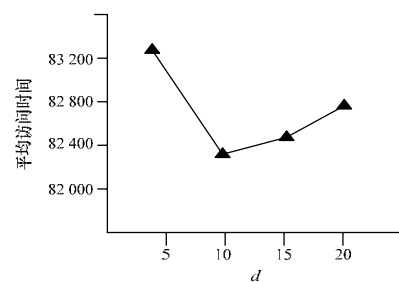


(b) θ 对调谐时间的影响

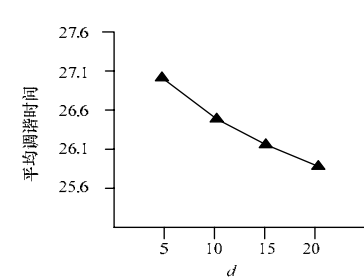
图 5 固定 N 时 θ 对调谐时间和访问时间的的影响

4.2.3 d 取值的影响

d 的取值为 5~20, 其他参数设为默认值。访问时间如图 6(a)所示, 在 $d=10$ 时, 平均访问时间最小。这是因为平均访问时间受索引节点分配的影响而与度数没有直接关系。随着 d 的增大, 在一个桶中需要检测更多索引目录, 因此, 访问时间随之增加。调谐时间如图 6(b)所示, 随着 d 的增加, 平均调谐时间减少。这是因为对于相同数目的数据项, 随着 d 的增加, 索引树的高度降低而减少了扫描的索引节点数目。



(a) d 对访问时间的的影响



(b) d 对调谐时间的影响

图 6 d 对访问时间和调谐时间的的影响

(下转第 68 页)