

本期目录 | 下期目录 | 过刊浏览 | 高级检索

[打印本

页] [关闭]

### 数据库技术

## 基于分簇的B+树数据库索引优化算法

胡廷波, 钟俊

四川大学 电气信息学院, 成都 610065

**摘要:** 在数据库中普遍采用的索引结构为适合随机查找的B+树结构, 当关键字之间存在顺序关系时, 该类索引方式效率较低。针对以上问题, 提出了基于分簇的B+树——CB+树(CB+ Tree)结构。该树在B+树的基础上充分考虑了记录集关键字之间的顺序关系, 通过降低索引树的高度来提高关键字的索引效率。仿真结果显示, 在记录数为100万的情况下, CB+树和B+树效率相当。当记录数达到500万时, CB+树插入用时6.7s, 比B+树插入用时7.6s减少了8%; CB+树查询用时9.9s, 比B+树查询用时11.1s减少了10%; CB+树删除用时10.1s, 比B+树删除用时11.2s减少了10%。由此说明, 在记录集关键字有序且记录数大于100万时, 提出的CB+树是更为高效的索引结构, 且其效率随记录数的增大提升更为明显。

**关键词:** 数据库 索引 效率 B树 CB树

## Database index optimization algorithm based on cluster B+ tree

HU Tingbo, ZHONG Jun

School of Electrical Engineering and Information, Sichuan University, Chengdu Sichuan 610065, China

**Abstract:** Currently, the most popular index structure used in database is B+ tree, which is easy to search randomly. But in some cases of which keywords are sequential, this structure is not efficient. In order to solve this problem, this paper proposed Cluster B+ tree (CB+ tree) on the basis

- 扩展功能
- 本文信息
- Supporting info
- PDF (604KB)
- [HTML全文]
- 参考文献
- [PDF]
- 参考文献
- 服务与反馈
- 把本文推荐给朋友
- 加入我的书架
- 加入引用管理器
- 引用本文
- Email Alert
- 文章反馈
- 浏览反馈信息
- 本文关键词相关文章
- 数据库
- 索引
- 效率
- B树
- CB树
- 本文作者相关文章
- 胡廷波
- 钟俊

of B+ tree. This tree considered the order between keywords to a large extent, and also reduced the tree height to improve the efficiency of search. Simulation shows that when the record number was one million, CB+ tree had the same efficiency as B+ tree. When the record number was five million, it took CB+ tree 6.7s to insert, which was 8% less than B+ tree's 7.6s, and it took CB+ tree 9.9s to inquire, which was 10% less than B+ tree's 11.1s. Otherwise, CB+ tree also reduced the time of deleting by 10% compared with that of B+ tree's, from 11.2s to 10.1s. According to the simulations, when the key words are in order and the record number is larger than one million, CB+ tree is a more efficient index structure, and its efficiency will promote obviously while the record number increases.

Keywords: database index efficiency B Tree  
CB Tree

收稿日期 2013-03-21 修回日期 2013-05-03 网络版  
发布日期 2013-10-18

DOI: 10.11772/j.issn.1001-9081.2013.09.2474

基金项目:

四川省科技支撑计划项目

通讯作者: 胡廷波

作者简介: 胡廷波(1987-),男,四川广元人,硕士研究生,

主要研究方向:嵌入式Linux;

钟俊(1972-),男,重庆人,副教授,博士,主要研究方向:信号  
与信息处理、嵌入式Linux、图像处理、扩频通信。

作者Email: 913441860@qq.com

---

参考文献:

本刊中的类似文章

1. 马健 张太红 陈燕红.中文搜索引擎分块倒排索引存储模式[J]. 计算机应用, 2013,33(07): 2031-2036
2. 刘军丹 赵书良 赵娇娇 郭晓波 陈敏 柳萌萌.家谱关系的元图表示[J]. 计算机应用, 2013,33(07): 2037-2040

3. 杨刚 陈越 黄惠新 于喆.改进多项式划分的XML数据库安全服务方案[J]. 计算机应用, 2013,33(06): 1637-1641
4. 吴敏 公衍超 杨楷芳.面向高效率视频编码标准的快速码率控制算法[J]. 计算机应用, 2013,33(05): 1251-1254
5. 王淼 李松阳.空间方向关系形式化描述研究进展[J]. 计算机应用, 2013,33(05): 1324-1329
6. 李成华 江小平 向文 李斌.面向管理实体对象的海量工程项目数据模型及实现[J]. 计算机应用, 2013,33(04): 1010-1014
7. 罗建平 邬群勇 朱莉.面向对象的全时域移动对象数据模型[J]. 计算机应用, 2013,33(04): 1015-1017
8. 朱媛媛 王晓京.基于GE码的HDFS优化方案[J]. 计算机应用, 2013,33(03): 730-733
9. 杨泽平 刘德强 王茜 向强铭.基于地理信息系统技术的数据采集与监视控制系统设计[J]. 计算机应用, 2013,33(02): 567-574
10. 房俊华 王翰虎 陈梅 马丹.一种具有自适应机制的闪存数据库索引结构[J]. 计算机应用, 2013,33(02): 563-566
11. 赵慧 张学 刘明 龚海刚 吴跃.实现无线传输能量效率最大化的功率控制新方法[J]. 计算机应用, 2013,33(02): 365-381
12. 柴佳晶 顾海华 包斯刚.密码芯片中二元扩域Eta双线性对安全算法[J]. 计算机应用, 2013,33(01): 163-167
13. 孙宏 张曦煌.联合能量路由和睡眠调度算法的分析与改进[J]. 计算机应用, 2013,33(01): 115-119
14. 冯汝伟 谢强 丁秋林.基于文本聚类与分布式Lucene的知识检索[J]. 计算机应用, 2013,33(01): 186-188
15. 徐翔 邹复民 廖律超 朱铨.基于GemFire的海量数据计算性能实验分析[J]. 计算机应用, 2013,33(01): 226-229
16. 杨超 张红旗.基于秘密共享方案RFID认证协议[J]. 计算机应用, 2012,32(12): 3458-3461
17. 潘冠桦 张兴忠.Sunday算法效率分析[J]. 计算机应用, 2012,32(11): 3082-3088
18. 边小勇 张晓龙 余海.基于工厂信息的实时数据流分析与全过程质量监控[J]. 计算机应用, 2012,32(10): 2935-2939
19. 杨欣欣 李培峰 朱巧明.基于查询扩展的人名消歧[J]. 计算机应用, 2012,32(09): 2488-2490
20. 李峰 罗磊.基于道路网络的时空索引方法IMon-tree[J]. 计算机应用, 2012,32(08): 2205-2222

21. 马瑞敏 陈继红.基于二维空间元素匹配的数据库水印算法[J]. 计算机应用, 2012,32(08): 2268-2270
22. 付东来 彭新光 陈够喜 杨秋翔.动态Huffman树平台配置远程证明方案[J]. 计算机应用, 2012,32(08): 2275-2282
23. 宋晓炜 杨蕾 刘忠 廖亮.基于视差矢量特点的视差估计快速算法[J]. 计算机应用, 2012,32(07): 1856-1859
24. 黄云 洪佳明 覃遵跃.基于双索引的近似子图匹配[J]. 计算机应用, 2012,32(07): 1994-1997
25. 黄纵 张晓龙 边小勇.基于实时数据流分析的钢铁产品质量监测[J]. 计算机应用, 2012,32(05): 1470-1473
26. 姚全珠 余训滨.基于最小相关实体子树的XML关键字查询算法[J]. 计算机应用, 2012,32(04): 1090-1093
27. 郭丽丽 汪洋.多用户MIMO系统下的闭环调度算法及频谱性能分析[J]. 计算机应用, 2011,31(11): 2912-2914
28. 卢秉亮 刘娜.路网中移动对象快照K近邻查询处理[J]. 计算机应用, 2011,31(11): 3078-3083
29. 姚全珠 郭祯 房美君.改进的基于小枝模式的匹配算法——cTwigStack[J]. 计算机应用, 2011,31(10): 2782-2785
30. 吴家皋 邵世伟 华正 邹志强 胡斌.混合结构P2P空间索引网络优化机制的研究与实现[J]. 计算机应用, 2011,31(09): 2301-2304
31. 赵艳梅 郑鑫福 徐立臻.内存数据库系统SwiftMMDB索引机制的设计与实现[J]. 计算机应用, 2011,31(09): 2395-2398
32. 李旻松 段琢华.基于支持向量机的隐含语意特征选择方法[J]. 计算机应用, 2011,31(09): 2429-2431
33. 公伟 刘培玉 贾娴.基于改进PrefixSpan的序列模式挖掘算法[J]. 计算机应用, 2011,31(09): 2405-2407
34. 陈旦 叶晓俊 施霖.TPC-DS性能测试工具的实现[J]. 计算机应用, 2011,31(09): 2449-2452
35. 殷小静 慕晓冬 徐义文 陈琦.海量地形数据的管理和交互策略优化[J]. 计算机应用, 2011,31(09): 2465-2467
36. 刘波 范士明 刘华.实时数据库混合索引机制的设计与实现[J]. 计算机应用, 2011,31(08): 2265-2269
37. 高健 宋奥 刘万 陈耀.基于静态串表的图像无损压缩编码方法[J]. 计算机应用, 2011,31(06): 1578-1580
38. 黄海峰 张珂珩 张鸿 季学纯 陈鹏.电力系统动态信息数据库关键技术[J]. 计算机应用, 2011,31(06): 1681-1684
39. 刘毅 林子雨.用积分法处理分布式数据库中的多副本并发控制问题[J]. 计算机应用, 2011,31(05): 1404-

40. 王立 王跃清 王翰虎 陈梅.闪存数据库系统中一种高效的自适应存储模式[J]. 计算机应用, 2011,31(05): 1400-1403
41. 冯丽云 洪玫 杨秋辉 周洪宇 臧康.基于逆向查询处理算法的数据库系统测试数据生成方法[J]. 计算机应用, 2011,31(04): 948-951
42. 冯勇 方欣 徐红艳.带有高效索引的语义Web服务I/O匹配优化方法[J]. 计算机应用, 2011,31(03): 677-679
43. 陈湘 吴跃.基于概念格挖掘GIS中的关联规则[J]. 计算机应用, 2011,31(03): 686-689
44. 薛献鹏 彭明田 贺怀清.基于Memcached的日历搜索引擎系统优化设计与实现[J]. 计算机应用, 2011,31(03): 865-868
45. 黄大荣 黄惠敏.集成Spring和Aglets的远程数据库访问技术[J]. 计算机应用, 2011,31(02): 458-461
46. 余永红 柏文阳.基于加密技术的外包数据库服务集成安全[J]. 计算机应用, 2011,31(01): 110-114
47. 郝文宁 赵恩来 刘玉栋 黄亚 刘军涛.异构数据库加解密系统的关键技术研究及实现[J]. 计算机应用, 2010,30(9): 2339-2343
48. 宋鸿陟 傅熠 张丽霞 彭红星 梁华坤.分割方式的多线程快速排序算法[J]. 计算机应用, 2010,30(9): 2374-2378
49. 王黎明 程晓 柴玉梅.基于Bin位图索引的多维查询优化算法[J]. 计算机应用, 2010,30(8): 2013-2016
50. 段晓燕.基于CMAC网络的迭代学习初始控制策略[J]. 计算机应用, 2010,30(8): 2049-2051
51. 宋余庆 严振 梁成全 张勇.无线嵌入式医疗系统与数据库交互中间件研究[J]. 计算机应用, 2010,30(8): 2257-2260
52. 谭龙江.基于搜索引擎优化的网络宣传机模型[J]. 计算机应用, 2010,30(8): 2232-2234
53. 肖晖 李清泉.移动对象数据库索引研究综述[J]. 计算机应用, 2010,30(4): 1064-1067
54. 马绚.基于RBD-tree的空间索引结构设计[J]. 计算机应用, 2010,30(4): 1103-1106
55. 张好芝 刘海涛.P级文件系统搜索概述[J]. 计算机应用, 2010,30(3): 813-817
56. 徐峰 郑焱.基于本体的视频语义检索系统研究[J]. 计算机应用, 2010,30(3): 835-837
57. 彭帅 李冬梅 李朝晖.基于PC集群的三维图形并行渲染性能分析[J]. 计算机应用, 2010,30(2): 547-550
58. 吴钦阳.R\*-树空间索引的改进[J]. 计算机应用, 2010,30(2): 419-422

59. 袁璐 刘峰 廖洁.基于知识的土地利用数据综合[J]. 计算机应用, 2010,30(12): 3385-3387
60. 王跃清 黄烨 王翰虎 陈梅.混合结构数据库中基于页迁移的存储分层算法[J]. 计算机应用, 2010,30(11): 2962-2964
61. 周帆 李树全 肖春静 吴跃.不确定数据库中概率top-k和排序查询算法[J]. 计算机应用, 2010,30(10): 2605-2609
62. 李东艳 李绍滋 柯道.基于外部数据库的图像自动标注改善模型[J]. 计算机应用, 2010,30(10): 2610-2613
63. 余永红 柏文阳.外包数据库服务隐私保护方法[J]. 计算机应用, 2010,30(10): 2672-2676
64. 王文奇 李乔良.高效的数字指纹方案[J]. 计算机应用, 2010,30(10): 2684-2686
65. 彭智勇 黄泽谦 刘俊 李越 徐波.基于对象代理数据库的微生物信息服务系统[J]. 计算机应用, 2010,30(1): 5-9
66. 张辉 刘永山 高云.一种基于BX树的移动对象混合索引方法[J]. 计算机应用, 2010,30(1): 22-24
67. 崔斌 吕雁飞 陈学轩.闪存环境下B+树索引重访[J]. 计算机应用, 2010,30(1): 1-4
68. 张祖莲 卡米力·木衣丁 王命全.一种有效预防点击欺诈的算法[J]. 计算机应用, 2010,30(07): 1794-1796
69. 孙存群 周顺平 杨林.基于分级地名库的中文地理编码的研究[J]. 计算机应用, 2010,30(07): 1953-1955
70. 姜欣 姜怡 方淼.文本翻译索引的互文度量方法[J]. 计算机应用, 2010,30(07): 1938-1940
71. 王宝文 韩静静 陈子军 刘文远.基于NNlists的路网k路径近邻查询[J]. 计算机应用, 2010,30(07): 1947-1949
72. 王蜜宫 陈锻生 林超.基于局部形状图的三维人脸特征点自动定位[J]. 计算机应用, 2010,30(05): 1255-1258
73. 郝晓弘 段晓燕 李恒杰.基于BP神经网络的迭代学习初始控制策略研究[J]. 计算机应用, 2009,29(4): 1025-1027
74. 付国瑜 黄贤英.基于Web页面平均质量的Web搜索模型和优化算法[J]. 计算机应用, 2009,29(4): 1114-1116
75. 李宏友 汪同庆 刘青 郭春华 蒋建春.基于LOGHM的视频运动目标检测算法[J]. 计算机应用, 2009,29(3): 736-738
76. 况夯 罗军.基于遗传FCM算法的文本聚类[J]. 计算机应用, 2009,29(2): 558-560
77. 韦云凯 毛玉明 于秦.基于中继技术的B3G/4G蜂窝

网络成本优化模型[J]. 计算机应用, 2009,29(12):

3207-3210

78. 张凯 钱焕延 徐延贵.混合型网络中DDoS攻击的入侵扫描研究[J]. 计算机应用, 2009,29(11): 2964-2968

79. 林镜华 雷为民 李凌南 白松.IPTV点播业务中减少特技播放交互延迟的方法[J]. 计算机应用, 2009,29

(11): 2897-2900

80. 高琪 张永平.超链接导向搜索算法中主题漂移的研究[J]. 计算机应用, 2009,29(11): 3100-3102

81. 张秋余 邓涛 袁占亭.一种基于激励机制的P2P流媒体数据调度模型[J]. 计算机应用, 2009,29(10): 2620-

2623

82. 李时文 卢建朱.快速有效的XML访问控制新方案[J]. 计算机应用, 2009,29(09): 2336-2338

83. 朱泓丞 徐志广.分布式数据库数据副本最优安置问题的研究[J]. 计算机应用, 2009,29(09): 2509-2511

84. 李方圆 何海江.由贪心策略构造Chebyshev多项式概要[J]. 计算机应用, 2009,29(08): 2253-2253

85. 汪清明.基于领域本体的Web服务动态组合模型[J]. 计算机应用, 2009,29(07): 1957-1959

86. 才藏太.班智达藏文语料切分词典的建立与算法研究[J]. 计算机应用, 2009,29(07): 2019-2021

87. 张革伏 徐琪.基于语义Web服务的分布式服装搜索引擎系统设计[J]. 计算机应用, 2009,29(06): 1601-

1604

88. 齐学玲 吕国斌 费江涛.Hypertable底层存储结构的改进[J]. 计算机应用, 2009,29(06): 1748-1750

89. 祝青 阳王东.利用关系表构建XML文档解析的树模型[J]. 计算机应用, 2009,29(06): 1719-1721

90. 贾惠芹 李周利.网络测控实验室的安全访问策略[J]. 计算机应用, 2009,29(05): 1467-1469

91. 张鹏 冯建华 房志峰.一种基于二叉树的Native XML数据库文档编码机制[J]. 计算机应用, 2008,28(9):

2331-2334

92. 吴晟 李星.分布式搜索中节点索引量大小估计算法[J]. 计算机应用, 2008,28(9): 2345-2348

93. 黄忍冬 彭舰 冯灏.并行模式下分子散射模型的求解[J]. 计算机应用, 2008,28(9): 2371-2374

94. 罗军 况夯.基于Boosting算法集成遗传模糊分类器的文本分类[J]. 计算机应用, 2008,28(9): 2386-2388

95. 潘昊 谭龙远.领域相关自适应的PageRank算法搜索策略[J]. 计算机应用, 2008,28(9): 2192-2194

96. 汤迪斌 王劲林 倪宏.一种CDN中的动态数据存储方案——UbDP[J]. 计算机应用, 2008,28(8): 1991-1993

97. 陈伟 Qiu Robin.面向大型数据库的审计数据采集方法[J]. 计算机应用, 2008,28(8): 2144-2146

98. 向军 李国徽 杨兵.基于不精确计算移动实时数据库服务质量管理[J]. 计算机应用, 2008,28(7): 1709-1712
99. 陶剑文 赵杰煜.一种优化的基于LPI的文档索引算法FLPI[J]. 计算机应用, 2008,28(6): 1566-1569
100. 李媛媛 马永强.基于潜在语义索引的文本特征词权重计算方法[J]. 计算机应用, 2008,28(6): 1460-1462
101. 黄国顺.基于数据库系统的决策表核与属性约简算法[J]. 计算机应用, 2008,28(5): 1180-1182
102. 陈冬旭 程小辉.基于线性规划的移动数据库广播调度算法研究[J]. 计算机应用, 2008,28(4): 874-876
103. 张鹏 冯建华 房志峰.Native XML数据库的结构连接算法研究[J]. 计算机应用, 2008,28(4): 869-871
104. 朱勤 陈继红 乐嘉锦.基于数字水印的外包数据库查询验证[J]. 计算机应用, 2008,28(3): 605-608
105. 吴敏琦 丁岳伟.应用IP影响因子实现Intranet搜索个性化[J]. 计算机应用, 2008,28(3): 811-812
106. 黄崇本 陶剑文 程光华.一种面向图包容搜索的图索引模型[J]. 计算机应用, 2008,28(2): 479-483
107. 黄景廉.Ad Hoc网络中基于DCF的MAC协议能量效率分析[J]. 计算机应用, 2008,28(12): 3037-3040
108. 曾倩 金敏.基于道路分布的移动对象动态组合索引方法[J]. 计算机应用, 2008,28(12): 3251-3253
109. 张孝 王珊 廉娜.ProGen:海量数据的出处数据库生成器[J]. 计算机应用, 2008,28(11): 2737-2740
110. 彭智勇 杨麇丞 任毅.可信数据库—概念、发展和挑战[J]. 计算机应用, 2008,28(11): 2741-2744
111. 王元佶 陈闾中.MSCR树:一种移动终端电子地图索引结构[J]. 计算机应用, 2008,28(10): 2561-2564
112. 王淑娥 孙劲光.基于 $\gamma$ 划分策略的高维数据索引结