第一图书网, tushu007.com

<<基于成本的Oracle优化法则>>

图书基本信息

书名：<<基于成本的Oracle优化法则>>

13位ISBN编号：9787302153894

10位ISBN编号：7302153892

出版时间：2007年7月

出版时间：清华大学出版社

作者：Jonathan Lewis

页数：435

译者：赵恒,李政仪

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：http://www.tushu007.com

内容概要

《基于成本的Oracle优化法则》数据库领域的超级畅销书，世界级0racle大师 Jonathan Lewis的最新力作，也是近几年来Oracle领域最重要的著作之一，荣获2006年度Oracle杂志编辑选择大奖：自该书出版以来，受到无数读者、网友和专家的好评，被翻译成多种语言，成为众多中高级Oracle用户、DBA和开发人员的必看图书。

作者简介

作者：(美国)刘易斯

第一图书网, tushu007.com

<<基于成本的Oracle优化法则>>

书籍目录

目 录 第1章 成本的含义 11.1 优化器选项 21.2 成本的定义 31.3 变换和成本计算 51.4 所见未必即所得 81.5 本章小结 81.6 测试用例 8第2章 表扫描 92.1 入门 102.2 提高 142.2.1 块大小的影响 142.2.2 计算 162.2.3 CPU成本计算的作用 222.3 BCHR 242.4 并行执行 272.5 索引快速全扫描 302.6 分区 322 小结 372.8 测试用例 37第3章 单表选择率 393.1 入门 403.2 空值 423.3 使用列表 433.4 区间谓词 48 词 523.6 多谓词的相关问题 543.7 本章小结 563.8 测试用例 57第4章 简单B树访问 594.1 索引成本计算 础知识 604.2 入门 614.2.1 有效索引选择率 634.2.2 有效表选择率 644.2.3 clustering_factor 654.2.4 综合 算 674.2.5 扩展算法 684.2.6 3个选择率 744.3 CPU成本计算 784.4 待处理的零碎问题 804.5 本章小结 测试用例 81第5章 群集因子 835.1 基本示例 845.1.1 减少表争用(多个自由列表) 865.1.2 减少叶块的争用 反转键索引，ReverseKey Index) 895.1.3 减少表的争用(ASSM) 925.1.4 减少RAC中的争用(自由列表 群) 955.2 列顺序 965.3 额外的列 995.4 校正统计信息 1015.4.1 sys_op_countchg()技术 1015.4.2 非正式 略 1055.5 待处理的零碎问题 1065.6 本章小结 1075.7 测试用例 107第6章 选择率的相关问题 1096.1 不 数据类型 1106.1.1 日期类型 1106.1.2 字符类型 1106.1.3 愚蠢的数据类型 1126.2 前导零 1166.3 致命的 值 1176.4 离散数据的风险 1196.5 令人惊奇的sysdate 1236.6 函数表示 1256.7 相互关联的列 1266.7.1 动 样 1296.7.2 优化器配置文件 1326.8 传递闭包 1336.9 产生约束的谓词 1366.10 本章小结 1396.11 测试用 例 139第7章 直方图 1417.1 入门 1427.2 普通直方图 1477.2.1 直方图和绑定变量 1477.2.2 Oracle何时忽 方图 1497.3 频率直方图 1527.3.1 伪造频率直方图 1557.3.2 注意事项 1567.4 “ 高度均衡 ” 直方图 1577 新审视数据问题 1637.5.1 愚蠢的数据类型 1637.5.2 危险的默认值 1667.6 本章小结 1677.7 测试用例 16 章 位图索引 1698.1 入门 1708.1.1 索引组件 1748.1.2 表组件 1758.2 位图合并 1778.2.1 较低的基数 17 空值列 1828.3 CPU成本计算 1858.4 一些有趣的示例 1868.4.1 多列索引 1878.4.2 位图连接索引 1878.4. 图转换 1888.5 本章小结 1918.6 测试用例 192第9章 查询变换 1939.1 入门 1949.2 过滤 1979.2.1 过滤 化 2009.2.2 标量子查询 2029.2.3 子查询分解 2089.2.4 复杂视图合并 2139.2.5 推入谓词 2159.3 一般子查 询 2169.3.1 子查询参数 2189.3.2 分类 2199.3.3 半连接 2249.3.4 反连接 2269.3.5 反连接异常 2289.3.6 和Notin 2299.3.7 有序提示 2319.4 星型变换连接 2329.5 星型连接 2379.6 展望 2399.7 本章小结 2409. 用例 241第10章 连接基数 24310.1 基本的连接基数 24410.2 实际SQL的连接基数 24910.3 扩展和异常情 况 25210.3.1 使用范围的连接 25210.3.2 不等于 25310.3.3 重叠 25610.3.4 直方图 25710.3.5 传递闭包 2 三表连接 26410.5 空值 26710.6 实现问题 27010.7 困难之处 27410.8 特性 27610.9 另一观点 27810.10 结 27910.11 测试用例 279第11章 嵌套循环 28111.1 基本机制 28211.2 实际示例 28611.3 完备性检查 本章小结 29111.5 测试用例 291第12章 散列连接 29312.1 入门 29412.1.1 最优散列连接 29712.1.2 一遍 连接 29912.1.3 多遍散列连接 30412.2 追踪文件 30812.2.1 event 10104 30812.2.2 event 10053 30912.3 难 点 31112.3.1 传统成本计算 31112.3.2 现代成本计算 31212.4 比较 31312.5 多表连接 31812.6 本章小 结 32112.7 测试用例 321第13章 排序与归并连接 32313.1 入门 32413.1.1 内存的使用 32913.1.2 CPU的 用 33013.1.3 sort_area_retained_size 33313.1.4 pga_aggregate_target 33413.1.5 实际I/O 33713.2 排序的成 本 33913.3 比较 34313.4 归并连接 34613.4.1 归并机制 34713.4.2 无最初排序的归并连接 35113.4.3 笛卡 并连接 35213.5 聚集及其他 35413.5.1 索引 35813.5.2 集合运算 35913.6 最后一次提醒 36313.7 本章小 结 36513.8 测试用例 366第14章 10053 trace文件 36714.1 查询 36814.2 执行计划 36914.3 环境 37014.4 件 37114.4.1 参数设置 37214.4.2 查询块 37514.4.3 存储统计信息 37614.4.4 单表 37814.4.5 完备性检 查 37914.4.6 一般计划 38014.4.7 Join order[1] 38014.4.8 Join order[2] 38614.4.9 Join order[3] 38714.4.10 J order[4] 38814.4.11 Join order[5] 38814.4.12 Join order[6] 39214.4.13 Join order[7] 39214.4.14 Join order[8] 39514.4.15 Join order[9] 39714.4.16 Join order[10] 39814.4.17 Join order[11] 39814.4.18 Join order[12] 40114.4.19 Join order[13] 40414.4.20 Join order[14] 40514.4.21 Join order[15] 40614.4.22 Join order[16] 40714.4.23 Join order[17] 40714.4.24 Join order[18] 40914.5 连接评估小结 41014.6 测试用例 41 录A 升级问题 415A.1 dbms_stats 416A.2 频率直方图 417A.3 CPU成本计算 417A.4 舍入误差 417A.5 绑 量窥视 418A.6 连接间的空值 418A.7 B树到位图的转换 418A.8 索引跳跃扫描 419A.9 AND-Equal 419A 索引散列连接 420A.11 修正的In-List 420A.12 传递闭包 420A.13 sysdate算术修正 421A.14 对空值的索 引 422A.15 pga_aggregate_target 422A.16 排序 422A.17 分组 423A.18 完备性检查 423A.19 超出界限的情

况 423A.20 关于类型 423A.21 optimizer_mode 424A.22 降序索引 424A.23 复杂视图合并 424A.24 非嵌
查询 424A.25 标量和过滤子查询 425A.26 并行查询策略的两次变化 425A.27 动态采样 425A.28 临时
表 425A.29 字典统计 426附录B 优化器参数 427B.1 optimizer_features_enable 428B.2 10053 trace文件 430
v$sql_optimizer_env 435

媒体关注与评论

前言前　言我在Practical Oracle 8i一书的前言中曾经提到“ 如果我们想写一本关于Oracle的技术性书籍，那么在写完时，里面的内容肯定已经过时 ”。
Addison-Wesley公司出版这本书的时候正好是Larry Ellison声明发布Oracle 9i的官方版本之时。
这本书出版后两个星期，我接到了第一封E-mail询问是否已经计划撰写一本关于Oracle 9i的书。
在此之前，我拒绝了所有关于“ 升级 ”的请求，这是基于以下3个原因：(a)这需要太多的精力，(b)在我可以将Oracle 9i的信息融入到本书中之前，可能需要两年的时间来熟悉9i，(c)它仍然只是前面这本书的简单修改。
因此，我的计划就是：2003年9月开始撰写这本书(确实是2003年，我花了22个月的时间才写完)，正好是我决定撰写Practical Oracle 8i的4年之后。
(请记住，在1999年9月，如果一个Oracle专家不会调试已经40岁高龄的COBOL语言，那么不会有哪家公司会雇佣他)。
我花费了大量的时间来学习Oracle 9i，以保证能够学习一些全新的内容。
当然，在这4年中，出现了很多Oracle 8i的升级版本(最终版为8.1.7.4)，出现了Oracle 9i的两个主要版本，当开始撰写本书时，Oracle公司已经推出了Oracle 10g。
因此，我撰写Practical Oracle 9i这一想法可能已经过时了。
事实上，当结束了本书的撰写工作之后(2005年6月)，针对Oracle 10g R2的Linux端口已经在OTN上可用了！
因此，您阅读本书之后所做的第一件事情就应该是在10g R2下运行本书的示例，以检查究竟有哪些地方发生了变化。
本书并不是对Practical Oracle 8i的简单升级，本书中介绍了关于基于成本优化的相关知识。
这看起来是一个非常简单的任务—— 针对优化器所执行的操作和为什么执行这样的操作我可以谈上几个钟头；但是我要做的是要把这些知识成书出版。
遗憾的是，这一任务比想象的要困难得多。
把很多词语堆砌起来是比较容易的—— 但是将它们组织成为一本比较有用的结构良好的书籍却是另一回事。
给我关于CBO的一些问题，我能够解释其中的原因—— 也许是在创建和测试了一些理论之后。
但是，要想为其他人解释关于优化器的一些通用的信息并让他们也能得出同样的结果，这完全不是一码事。
最终，我成功地构建了本书的框架，并意识到必须撰写3本书：基础知识、某些增强的内容和所有的外围知识。
本书专门介绍关于基于成本的优化的基础知识。
0.1 写作本书的动机为什么需要了解CBO的工作原理呢？
这是因为当遇到优化器产生的执行计划非常差的情况时，需要理解这究竟是什么问题导致的，并给出正确的修复方法。
当然，我们可以通过为SQL添加一些提示或者进行一些巧妙的代码重写操作的方式来修复一些问题，但是如果采用这两种方式，当同样的问题出现在其他地方时，就必须一再重复执行类似的操作。
从另一方面讲，如果理解了问题的本质内涵，那么就可以一次性地修复该问题，而且在其他地方出现同样的问题时就可以知道该如何解决。
0.2 本书内容本书涵盖了关于优化的基本信息。
本书并不是为了成为优化器工作原理的完全参考手册—— 本书一共有14章，但是直到第10章才提到连接，由此您可能会意识到其中有大量的内容需要介绍。
优化过程中常见的重要术语有：● 选择率(selectivity)和基数(cardinality)：谓词将标识多少数据片断以及可以转换为多少数据行。
● 访问路径(access path)：当访问表时，查询应该使用B树索引、联合使用一对位图索引，还是完全忽略索引。

● 连接顺序(join order)：查询应该首先访问哪一个表，为了获得指定的结果，从何处开始工作量最少。
虽然有时候我也会对需要考虑的一些微妙特性作些注释，但是，本书实际上只着重于少数核心概念的介绍。
优化器如何计算谓词将产生多少数据？
它如何用一个数值来表示表扫描的工作，如何将其与使用索引所需的工作进行对比？
用哪些数据来估计排序或散列连接所用到的资源？
我可以分析一个查询、查询中的对象以及10053 trace文件，并解释某条路径为什么比另一条路径优先使用。
遗憾的是，我无法保证读者每次都能获得与书中trace文件同样的结果，因为示例中无法调试所有的选项(我从来也没有使用过全部的选项)，那样的话本书将变成一本非常冗长乏味的书。
尽管本书无法解决所有的问题，但是，我相信本书已经给出了所有的基本方法，利用它们可以解释绝大部分情况下的问题。
0.3 本书中没有包含的内容本书不可避免地会忽略一些内容。
有些内容被忽略是因为它们不属于优化器的核心行为，还有一些被忽略是因为不是很常用，出于篇幅的原因必须排除在外。
本书根本没有提到基于规则的优化器(Rule Based Optimizer，RBO)，这是因为每个人都应尽量避免使用它。
本书也没有介绍关于可扩展优化器的任何信息(包括上下文和空间索引)，这是因为它们不是主流主题。
本书也没有提到分析函数、模型子句(10g)和OLAP (联机分析处理)，这是因为它们在进行自身类型的数据处理之前必须获取相关的数据—— 而数据获取可能是时间要求最严格的工作。
本书没有提到对象—— 这是因为就优化器而言，它们并不存在。
当您创建一个对象类型并创建对象类型的数据片断时，Oracle会将它们转换为简单的表和索引—— 优化器根本不关心对象。
最后，本书几乎没有介绍并行查询、分区表、分布式查询和Oracle中一些比较微妙的物理选项，比如群集和IOT。
这出于以下两个原因：第一，篇幅不够；第二，避免影响对主题的介绍。
对于优化器来说，存在多种不同的知识信息，要想集中介绍某一种知识将非常困难，因此，最好的办法就是在某个时间段内只介绍一部分知识。
0.4 本系列其他两本书中将出现的内容本书是这一系列中的第一本。
接下来的两本将介绍Oracle中的一些重要特性，尤其是分区表、并行执行、索引组织表、动态采样和查询重写。
对本书中已经提到的一些知识也将进一步进行介绍，比如B树索引的更多访问路径、群集访问和索引访问之间的比较以及其关于直方图的更为详细的信息。
关于基于成本优化的信息的最后一部分是支持优化的基础结构，有助于理解基于成本的优化。
此处主要的问题是理解并解释执行计划，理解提示的含义并应用提示，最有效地使用dbms_stats包。
本书是基于Oracle 9.2撰写的，其中注意了与8i的一些区别，并介绍了Oracle 10g中对应内容的变化。
接下来的两本书将基本上不介绍8i，更多地介绍Oracle 10g。
0.5 本书组织结构本书一共14章，按照顺序涵盖了如下主题：● 表扫描：比较简单，作为入门，介绍了一些关于CPU成本计算的知识。
● 简单选择率：只有一个表，但包含了大量的关于算法的知识。
● 简单B树索引：单块读取和多块读取的区别。
● 群集因子：这也许是索引中最关键的特性。
● 更微妙的选择率：介绍一个基本主题上的大量微小变化。
● 直方图：为什么需要直方图，并介绍了OLTP和DSS/DW之间的区别。
● 位图索引：并不是所有的索引都是相同的。

- 变换：看到的不一定是得到的。
- 连接：用4整章的内容来介绍两个表之间的连接。
- 10053 trace：演示性示例。
- 升级问题：从本书其余部分整理的警告和注意信息的集合。

每一章中都包含很多代码，它们都来自于一组SQL脚本，可以从Apress网站(www.apress.com)或www.tupwk.com.cn/downpage上下载。
您可以在自己的系统上运行这些脚本，以重新产生和分析本书中给出的结果。
必须认真地分析这些脚本，因为其中包含的一些额外注释和额外测试在书中并没有提及。
我同时还在自己的个人网站(www.jlcomp.demon.co.uk)上公布了这些脚本，并不时地增加新的内容。
这些脚本非常重要—— 如果进行修改，会对生产系统产生严重的影响。
如果采用脚本来测试基本机制，则可以在每次升级时都重新运行这些测试，以观察会出现什么样的变化。
需要注意的一点是书中出现的脚本中经常会包含一些代码行来标准化测试环境， 例如：alter session set "_optimizer_system_stats_usage" = false;不要因为本书中使用了就在生产系统中使用这样的命令。
这并不是一个很好的编程习惯；采用这样的命令的目的在于避免当某个数据库(比如)的系统统计信息和另外一个的完全不同时所导致的副作用。
联机代码包中还包括3个init.ora文件和1个用于在Oracle 9i和10g下创建测试表空间tablespace的脚本。
这4个文件必须被重新编辑，以避免出现文件路径命名导致的问题；Oracle 9i和10g下的init.ora文件也必须进行调整以适应于rollback/undo的管理选项。
我选择针对Oracle 9i和10g运行init.ora文件以避免spfile的偶然变化，但是，您也可以选择将init.ora文件中的设置组合到spfile中。
0.6 必须指出的警告每当有人要求我为Practical Oracle 8i一书签名时，我总是在签名的后面附上自己的座右铭：尽信书则不如无书(Never believe all you read)。
(如果您要求我为本书签名，座右铭将是：印在书上的并不都是真理(Just because it ’ s printed, doesn ’ t mean it ’ s true))。
因为总是会有一些特殊情况、不同的参数设置方式和bug是书中没有介绍的。
(不用说我所不知道的数千种知识，就是我所知道的也不能保证全部正确。
)参见如下的简单示例(联机代码包中的脚本in-list.sql)—— 在一个本地托管的表空间中运行，采用手动片断空间管理，8KB的块大小：create table t1 asselecttrunc((rownum-1)/100) n1,rpad('x',100) paddingfromall_objectswhererownum

编辑推荐

★数据库领域的超级畅销书，世界级Oracle大师Jonathan Lewis的最新力作！
★荣获2006年年度Oracle杂志编辑选择大奖! ★受到无数读者、网友和专家的好评，被翻译成多种语言，成为众多中高级Oracle用户、DBA和开发人员的必看图书! ［精彩试读一］ ［精彩试读二］ ［精彩评论］ 对于基于成本的优化器的执行方式，Jonathan有着深刻的见解，《基于成本的Oracle优化法则》能够帮助DBA成为更好的设计者，还能帮助开发人员开发出更卓越的代码。
—— Thomas Kyte Oracle公司全球副总裁 基于成本的优化是数据库根据统计以预测的方式选择一种可能最快的执行方式。
对执行性能高度敏感的应用使用CBO可以成倍加速SQL的执行速度。
《基于成本的Oracle优化法则》的目的在于解释优化器执行计算过程中的主要特性，以此分析计算成本，并确定SQL语句的执行计划，作者还深入剖析了优化器产生的成本与实际资源消耗之间的关系。
相信Jonathan Lewis对CBO精湛的阐述将使得Oracle DBA能够更好地管理数据库的执行，开发人员将能够更好地开发数据库应用程序。
—— 罗铁庚 国防科技大学计算机学院教授 这是一本非常优秀的介绍Oracle系统的技术图书，该书以专家的视角对Oracle系统的基于成本的优化模型进行了深入浅出的分析和介绍。
另外，这《基于成本的Oracle优化法则》的内容组织和案例选用也别具匠心、引人入胜。
—— 闪四清 著名数据库专家

第一图书网, tushu007.com

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:http://www.tushu007.com