

# 数据库查询优化器的艺术



[数据库查询优化器的艺术 下载链接1](#)

著者:李海翔

出版者:机械工业出版社

出版时间:2014-1-1

装帧:平装

isbn:9787111447467

《数据库技术丛书·数据库查询优化器的艺术：原理解析与SQL性能优化》是数据库查询优化领域的里程碑之作，由Oracle公司MySQL全球开发团队、资深专家撰写，作者有10余年数据库内核和查询优化器研究经验。数据库领域泰斗王珊教授亲自作序推荐，PostgreSQL中国社区和中国用户会发起人以及来自Oracle、新浪、网易、华为等企业的数位资深数据库专家联袂推荐。从原理角度深度解读和展示数据库查询优化器的技术细节和全貌；从源码实现角度全方位深入分析MySQL和PostgreSQL两大主流开源数据库查询优化器的实现原理；从工程实践的角度对比了两大数据库的查询优化器的功能异同和实现异同。它是所有数据开发工程师、内核工程师、DBA以及其他数据库相关工作人员值得反复研读的一本书。

《数据库技术丛书·数据库查询优化器的艺术：原理解析与SQL性能优化》共19章，分为四个部分：第一篇（第1~4章）对数据库查询优化技术的范围、逻辑查询优化、物理查询优化，以及查询优化器与其他模块的关系做了非常细致、深入的讲解；第二篇（第5~10章）首先从源码角度对PostgreSQL查询优化器的架构、层次、设计思想、相关数据结构和实现原理进行了深入、系统的分析，然后从功能角度对PostgreSQL的逻辑查询优化、物理查询优化、查询优化器的关键算法，以及PostgreSQL查询优化器与其他模块的关系做了深入的讲解；第三篇（第11~16章）首先从源码角度对MySQL查询优化器的架构、层次、设计思想、相关数据结构和实现原理进行了深入、系统的分析，然后从功能角度对MySQL的逻辑查询优化、物理查询优化、查询优化器的关键算法，以及MySQL查询优化器与其他模块的关系做了深入的讲解；第四篇（第17~19章）对PostgreSQL与MySQL的逻辑查询优化技术、物理查询优化技术、设计思想和编码规范等各方面进行了深度的比较。

作者介绍:

李海翔，网名“那海蓝蓝”，资深数据库专家，从事数据库研发、数据库测试与技术管理等工作10余年，对数据库的内核有深入的研究，长于PostgreSQL和MySQL等开源数据库的内核与架构。现任职于Oracle公司MySQL全球开发团队，从事查询优化技术的研究和MySQL查询优化器的开发工作。曾参与了863、核高基、工信部、科技部、发改委、北京市科委等多个重大科技项目。2005年获得北京市科学技术进步奖一等奖，2006年获高级工程师（系统分析师）。

目录: 推荐序一

推荐序二

前言

第一篇 查询优化技术

第1章 数据管理系统的查询优化

1.1 数据库调优

1.2 查询优化技术

1.2.1 查询重用

1.2.2 查询重写规则

1.2.3 查询算法优化

1.2.4 并行查询优化

1.2.5 分布式查询优化

1.2.6 其他优化

1.3 本章小结

第2章 逻辑查询优化

2.1 查询优化技术的理论基础

2.1.1 关系代数

2.1.2 关系代数等价变换规则对优化的指导意义

2.2 查询重写规则

2.2.1 子查询的优化

2.2.2 视图重写

2.2.3 等价谓词重写

2.2.4 条件化简

2.2.5 外连接消除

2.2.6 嵌套连接消除

2.2.7 连接消除

2.2.8 语义优化

2.2.9 针对非SPJ的优化

2.3 启发式规则在逻辑优化阶段的应用

2.4 本章小结

## 第3章 物理查询优化

### 3.1 查询代价估算

#### 3.1.1 代价模型

#### 3.1.2 选择率计算的常用方法

### 3.2 单表扫描算法

#### 3.2.1 常用的单表扫描算法

#### 3.2.2 单表扫描代价计算

### 3.3 索引

#### 3.3.1 如何利用索引

#### 3.3.2 索引列的位置对使用索引的影响

#### 3.3.3 联合索引对索引使用的影响

#### 3.3.4 多个索引对索引使用的影响

### 3.4 两表连接算法

#### 3.4.1 基本的两表连接算法

#### 3.4.2 进一步认识两表连接算法

#### 3.4.3 连接操作代价计算

### 3.5 多表连接算法

#### 3.5.1 多表连接顺序

#### 3.5.2 常用的多表连接算法

#### 3.5.3 多表连接算法的比较

### 3.6 本章小结

## 第4章 查询优化器与其他模块的关系

### 4.1 查询优化器整体介绍

### 4.2 查询优化器与其他模块的关系

### 4.3 本章小结

## 第二篇 PostgreSQL查询优化器原理解析

## 第5章 PostgreSQL查询优化器概述

### 5.1 PostgreSQL查询执行过程

### 5.2 PostgreSQL查询优化器的架构和设计思想

#### 5.2.1 PostgreSQL查询优化器架构

#### 5.2.2 PostgreSQL查询优化器的层次

#### 5.2.3 PostgreSQL查询优化器设计思想

### 5.3 主要概念

### 5.4 代码层次结构

### 5.5 本章小结

## 第6章 PostgreSQL查询优化器相关数据结构

### 6.1 主要数据结构

#### 6.1.1 基本数据结构

#### 6.1.2 查询树

#### 6.1.3 各种对象的结构

#### 6.1.4 连接操作相关的结构

#### 6.1.5 查询执行计划相关的结构

### 6.2 各个结构之间的关系

### 6.3 各个阶段间和主要结构体间的关系

### 6.4 本章小结

## 第7章 PostgreSQL查询优化器实现原理解析

### 7.1 查询优化整体流程

### 7.2 查询优化器实现原理解析

#### 7.2.1 planner--主入口函数

#### 7.2.2 standard\_planner--标准的查询优化器函数

#### 7.2.3 subquery\_planner--生成（子）查询执行计划函数

#### 7.2.4 grouping\_planner--生成查询执行计划并对非SPJ优化

#### 7.2.5 build\_minmax\_path--聚集函数MIN/MAX的优化函数

#### 7.2.6 query\_planner--生成最优的查询路径函数

- 7.2.7 make\_one\_rel--构造多表连接路径并选出最优路径函数
- 7.2.8 make\_rel\_from\_joinlist--生成多表连接路径函数
- 7.2.9 optimize\_minmax\_aggregates--聚集操作MIN/MAX优化函数
- 7.2.10 create\_plan--创建查询执行计划函数
- 7.2.11 非SPJ处理--grouping\_planner的各个子模块
- 7.2.12 其他重要的函数与操作
- 7.3 代价估算实现原理解析
  - 7.3.1 查询代价估算
  - 7.3.2 单表扫描方式的代价估算
  - 7.3.3 两表连接的代价估算
  - 7.3.4 其他代价估算函数
  - 7.3.5 选择率的计算
- 7.4 从目录结构和文件功能角度看查询优化器
  - 7.4.1 查询优化子模块与主要文件的关系
  - 7.4.2 查询优化器代码结构
- 7.5 本章小结

第8章 从功能的角度看PostgreSQL查询优化

- 8.1 优化器之逻辑查询优化
  - 8.1.1 视图重写
  - 8.1.2 子查询优化
  - 8.1.3 等价谓词重写
  - 8.1.4 条件化简
  - 8.1.5 外连接消除
  - 8.1.6 嵌套连接消除
  - 8.1.7 连接的消除
  - 8.1.8 语义优化
  - 8.1.9 选择操作下推
  - 8.1.10 非SPJ优化
- 8.2 优化器之物理查询优化
  - 8.2.1 PostgreSQL的物理优化主要完成的工作
  - 8.2.2 启发式规则在物理查询优化阶段的使用
  - 8.2.3 两表连接
  - 8.2.4 代价估算
  - 8.2.5 PostgreSQL的索引与查询优化
- 8.3 其他
  - 8.3.1 grouping\_planner函数主干再分析
  - 8.3.2 用户指定的连接语义与PostgreSQL实现两表连接的函数及算法的关系
  - 8.3.3 集合操作优化
- 8.4 本章小结

第9章 PostgreSQL查询优化的关键算法

- 9.1 动态规划算法
  - 9.1.1 动态规划算法的处理流程
  - 9.1.2 紧密树处理流程
- 9.2 遗传算法
  - 9.2.1 PostgreSQL遗传算法的处理流程
  - 9.2.2 主要的数据结构
  - 9.2.3 主要的函数和变量
  - 9.2.4 应用遗传算法实现表连接的语义
  - 9.2.5 应用遗传算法计算适应度
  - 9.2.6 进一步理解PostgreSQL的遗传算法
- 9.3 动态规划算法与遗传算法对比
- 9.4 本章小结

第10章 PostgreSQL查询优化器与其他部分的关系

- 10.1 查询优化器与语法分析器

- 10.2 查询优化器与执行器
- 10.3 查询优化器与缓冲区管理模块
- 10.4 查询优化器与对象访问模块
- 10.5 查询优化器与统计模块
- 10.6 查询优化器与索引模块
- 10.7 本章小结
- 第三篇 MySQL查询优化器原理解析
- 第11章 MySQL查询优化器概述
  - 11.1 MySQL查询执行过程
  - 11.2 MySQL查询优化器的架构和设计思想
    - 11.2.1 MySQL查询优化器架构
    - 11.2.2 MySQL查询优化器的层次
    - 11.2.3 MySQL查询优化器设计思想
  - 11.3 主要概念
    - 11.3.1 常量表
    - 11.3.2 表数据的访问方式
  - 11.4 代码层次结构
  - 11.5 本章小结
- 第12章 MySQL查询优化器相关数据结构
  - 12.1 主要的类和数据结构
    - 12.1.1 查询树
    - 12.1.2 基本对象
    - 12.1.3 连接对象与执行计划
    - 12.1.4 代价估算类
  - 12.2 各个阶段主要结构体间的关系
  - 12.3 本章小结
- 第13章 MySQL查询优化器的原理解析
  - 13.1 查询优化器整体流程
  - 13.2 优化器的代码详解
    - 13.2.1 JOIN.prepare--优化前的准备工作
    - 13.2.2 JOIN.optimize--优化器主入口方法
    - 13.2.3 make\_join\_statistics--计算最优的查询优化执行计划
    - 13.2.4 choose\_table\_order--求解多表连接最优连接路径
    - 13.2.5 make\_join\_statistics函数的其他子函数
    - 13.2.6 make\_join\_select--对条件求值、下推连接条件到表中
    - 13.2.7 test\_if\_skip\_sort\_order--排序操作的优化
    - 13.2.8 make\_join\_readinfo--为连接的每个表构造信息
    - 13.2.9 JOIN.exec--执行查询执行计划的函数
  - 13.3 代价估算
    - 13.3.1 查询代价估算模型
    - 13.3.2 查询代价估算过程
    - 13.3.3 其他的代价估算
    - 13.3.4 对存储引擎的调用接口
    - 13.3.5 统计信息
  - 13.4 本章小结
- 第14章 从功能的角度看MySQL查询优化
  - 14.1 优化器之逻辑查询优化
    - 14.1.1 视图重写
    - 14.1.2 子查询优化
    - 14.1.3 等价谓词重写
    - 14.1.4 条件化简
    - 14.1.5 外连接消除
    - 14.1.6 嵌套连接消除
    - 14.1.7 连接的消除

- 14.1.8 语义优化
- 14.1.9 非SPJ优化
- 14.2 优化器之物理查询优化
  - 14.2.1 MySQL的物理优化主要完成的工作
  - 14.2.2 启发式规则在物理查询优化阶段的使用
  - 14.2.3 MySQL的索引与查询优化
  - 14.2.4 用户指定的连接语义与MySQL实现两表连接的算法
- 14.3 本章小结
- 第15章 MySQL查询优化的关键算法
  - 15.1 深入理解MySQL的多表连接算法
  - 15.2 本章小结
- 第16章 MySQL查询优化器与其他部分的关系
  - 16.1 查询优化器与语法分析器
  - 16.2 查询优化器与执行器
  - 16.3 查询优化器与缓冲区管理模块
  - 16.4 查询优化器与索引模块
  - 16.5 本章小结
- 第四篇 PostgreSQL查询优化器VSMySQL查询优化器
- 第17章 PostgreSQL和MySQL的逻辑查询优化技术
  - 17.1 查询重写
    - 17.1.1 子查询优化
    - 17.1.2 视图重写
    - 17.1.3 等价谓词重写
    - 17.1.4 条件化简
    - 17.1.5 外连接消除
    - 17.1.6 嵌套连接消除
    - 17.1.7 连接消除
    - 17.1.8 语义优化
  - 17.2 非SPJ的优化
  - 17.3 本章小结
- 第18章 PostgreSQL和MySQL的物理查询优化技术
  - 18.1 查询代价估算模型比较
  - 18.2 单表扫描算法
  - 18.3 索引
  - 18.4 两表连接算法
  - 18.5 多表连接算法
  - 18.6 本章小结
- 第19章 PostgreSQL和MySQL的其他异同
  - 19.1 启发式规则的使用比较
  - 19.2 综合比较
    - 19.2.1 基本概念的比较
    - 19.2.2 数据结构的比较
    - 19.2.3 设计思想的比较
    - 19.2.4 编码规范的比较
  - 19.3 本章小结
- 附录A 如何掌握数据库内核
- 附录B 如何阅读本书
- 附录C 如何阅读查询执行计划
- 附录D 如何跟踪查询执行计划
- • • • • ([收起](#))

[数据库查询优化器的艺术](#) [下载链接1](#)

## 标签

数据库

MySQL

性能优化

计算机

database

源码分析

查询优化器

数据

## 评论

目前还不涉及太多查询优化的工作，粗略过一遍，以后再细看。

-----  
还是很给力的一本书的，对PostgreSQL的剖析相当到位，由顶而下介绍软件架构，然后由下而上介绍具体代码，代码分析很详细，只是一些地方稍显生硬。没有看MySQL部分。

-----  
听了他在dataguru的课程，一般般吧，实战性不强，有些太书面

-----

PG和MySQL能分开来就好了，我只需要MySQL这一半的内容

-----  
读了一大半，pg的并未研究。

-----  
知识点较多，但缺乏清晰的组织逻辑；当工具书来看可能更合适

-----  
粗略翻看了mysql的部分，非常多的篇幅在从源码上解读mysql查询优化，太硬核了

-----  
李海翔老师对优化器有深入了解，建议先找一本数据库经典教材过一遍，然后对照着PG和MySQL的源代码阅读这本书。

-----  
对我来说，看第2，3，14章就够了，其他的看不懂

-----  
简单的浏览一遍，以前没有接触过sql优化这些知识，通过这本书了解了一些sql优化的概念，逻辑优化、物理优化，优化的原理等知识。  
总体来说这本还是挺好的，对比分析了MySQL和PostgreSQL这两个主流开源数据库的优化器实现。

-----  
国内写技术书籍很多时候缺主线脉络，这本挺好，顺便通过大量的图表和实例给DBA们做sql优化提供了很好的指引。但是，code talks while code sucks，代码有点多且看到SI截出来的引用图……也就码农能忍了。

-----  
虽然有些章节大量代码，但整体上还是不错的

-----  
褚霸推荐

-----  
查询优化，从理论开始，深入浅出

-----  
刚开始讲解SQL优化器的通用理论的时候比较无聊，因为都是提纲挈领的东西，后面结合PostgreSQL和MySQL的实例来讲解，前面的统筹概括就好理解了，醍醐灌顶！建议前面几章快速过掉，进入实例，然后遇到了理论再回来前面复习一遍。Anyway，推荐！

-----  
强烈推荐给用数据库,想用好数据库的同学,我等资质愚笨的人都有诸多感悟,同学们应该体会更多  
这书讲了一件事情:SQL语句到数据库之后,数据库引擎如何根据选择更好的方案来更快的执行出来,其中包括逻辑优化部分(更有效率的SQL语句同义转换)和物理优化部分(如何更好的利用索引,如何更好的处理多表连接)  
大体理出了思路,诸多细节还没看懂,还需要重读一遍,加深体会  
书中2/3是接合源码的,我直接跳了,另外1/3是原理介绍,受益匪浅

-----  
书写得真是不怎么样，内容重复、逻辑不通，代码讲解不清晰……  
相比之下那本《PostgreSQL技术内幕》要好上一个档次。  
Anyway，只看第一篇作为一个优化器入门还是可以的，其它部分适合作为字典，有需要时再查。

-----  
[数据库查询优化器的艺术\\_下载链接1](#)

书评

-----  
[数据库查询优化器的艺术\\_下载链接1](#)