

带你走入PostgreSQL数据库查询优化器
抽丝剥茧，深入剖析查询优化技术内幕

Broadview[®]
www.broadview.com.cn



PostgreSQL 技术内幕

查询优化深度探索

张树杰 / 著



中国工信出版集团



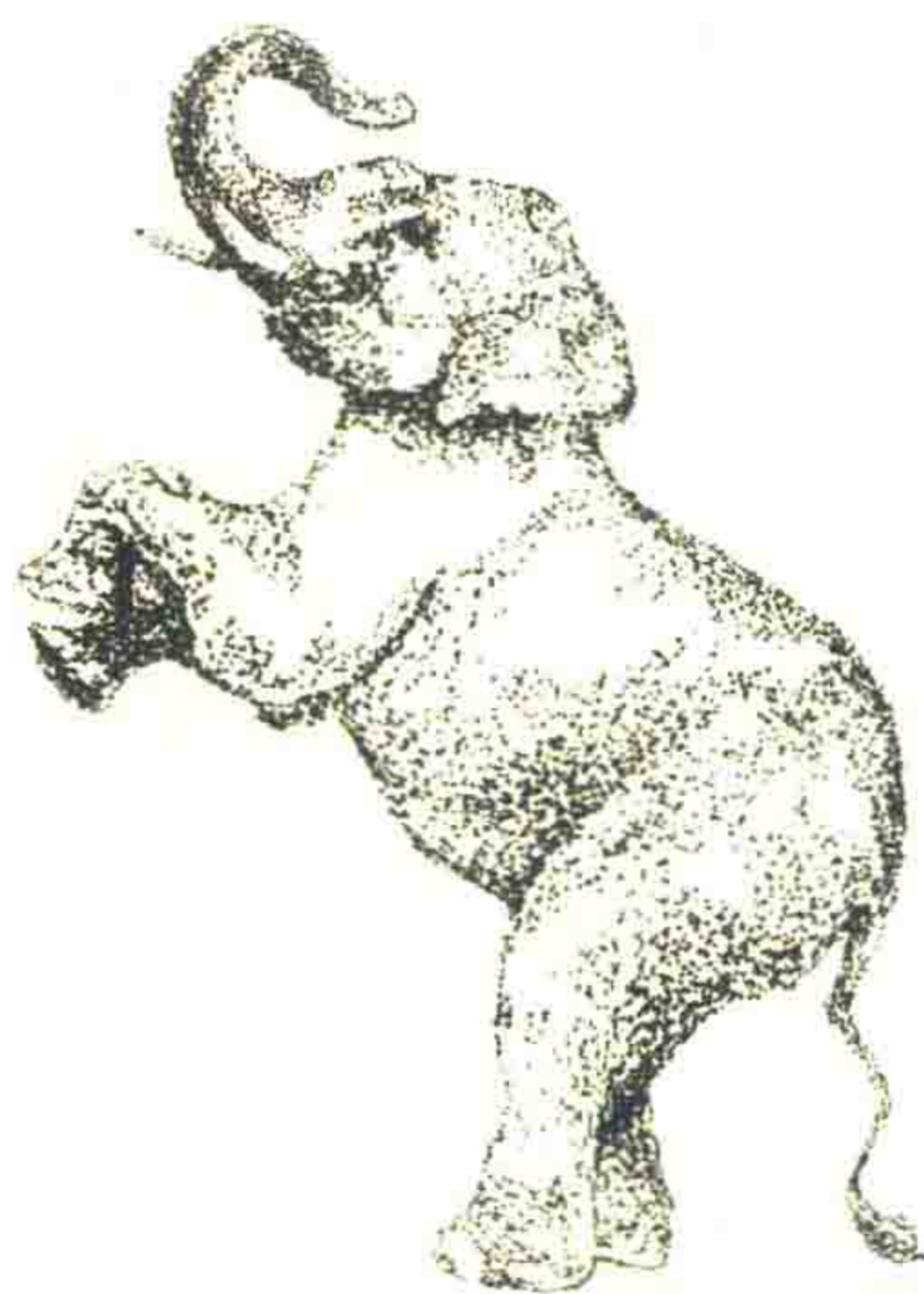
电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

PostgreSQL

技术内幕

查询优化深度探索

张树杰 / 著



电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

查询优化器是数据库中很重要的模块之一，只有掌握好查询优化的方法且了解查询优化的细节，在对数据库调优的过程中才能有的放矢，否则调优的过程就如无本之木、无源之水，虽上下求索而不得其法。

本书揭示了 PostgreSQL 数据库中查询优化的实现技术细节，首先对子查询提升、外连接消除、表达式预处理、谓词下推、连接顺序交换、等价类推理等逻辑优化方法进行了详细描述，然后结合统计信息、选择率、代价对扫描路径创建、路径搜索方法、连接路径建立、Non-SPJ 路径建立、执行计划简化与生成等进行了深度探索，使读者对 PostgreSQL 数据库的查询优化器有深层次的了解。

本书适合数据库内核开发人员及相关领域的研究人员、数据库 DBA、高等院校相关专业的本科生或者研究生阅读。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

PostgreSQL 技术内幕：查询优化深度探索 / 张树杰著. —北京：电子工业出版社，2018.6
ISBN 978-7-121-34148-9

I. ①P… II. ①张… III. ①关系数据库系统 IV.①TP311.138

中国版本图书馆 CIP 数据核字 (2018) 第 088240 号

责任编辑：董 英

印 刷：三河市良远印务有限公司

装 订：三河市良远印务有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱

邮编：100036

开 本：787×980 1/16 印张：25.75 字数：558 千字

版 次：2018 年 6 月第 1 版

印 次：2018 年 6 月第 1 次印刷

定 价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819, faq@phei.com.cn。

张树杰



拥有13年IT从业经验，从事国产数据库内核开发工作多年，对数据库内核各个方面均有涉猎，近些年专注于研究如何改进分布式数据库的查询优化、查询执行，目前在Pivotal公司从事Apache HAWQ数据库的内核开发工作。

邮箱: zhangshujie@pku.edu.cn

序一

查询可以说是数据库管理系统中最关键、最吸引人的功能之一，每一个生产数据库系统每天都需要处理大量的各类查询，为了让这些查询运行得更快、更好，数据库管理系统的查询优化器中包含了大量的优化技术，这些优化技术是很多研究者和技术人员数十年钻研和探索总结出来的精华。不论是数据库管理系统的开发者还是数据库应用的开发者，学习理解查询优化技术都大有裨益。

作为最先进的开源对象关系型数据库管理系统，PostgreSQL 及其源代码无疑是学习和体会查询优化技术的最佳平台。除此之外，高质量技术书籍也是研究查询优化技术必不可少的武器。本书结合 PostgreSQL 的查询优化器源代码，深入分析了一个查询进入 PostgreSQL 之后一步步被查询优化器转换成一个可执行的、优化后的执行计划的全过程。为了让读者更容易理解，本书还配备了大量的实例来讲解，确实是一部值得一读的好书。

身为一名 PostgreSQL 爱好者和数据库研究人员，我感到无比幸福和自豪——据我有限的知识，全球仅有几本分析 PostgreSQL 内核的书籍，而它们全都出自中国作者之手。希望今后有更多、更好的此类书籍面世，也祝愿中国的数据库技术和产品有朝一日能够走向世界。

彭煜玮

2018.4.25 于珞珈山

序二

中国有句古话，“巧妇难为无米之炊”，说的是再好的主妇，在没有给任何食材的情况下也做不出可口的饭菜。反过来，什么样的主妇算得上“巧妇”呢？如果给你准备好了烹调所需的所有食材，你能做出可口的饭菜吗？

数据库是一个比较神奇的软件，我们都知道可以用 SQL 和数据库沟通，让数据库处理 SQL 和让主妇做饭是一样的道理，数据库能不能及时响应 SQL 请求，能不能用最优的计划完成 SQL 请求，取决于数据库本身提供了哪些“料”，以及数据库打算怎么“烹调”用户提交的 SQL。例如，一个简单查询 SQL，数据库的扫描方法（我暂且把它称为数据库的“料”之一）就有全表扫描、索引扫描、位图扫描、跳跃扫描等。一个 SQL 中包括了多个函数、表达式时，数据库先处理哪个表达式或函数，又或者在什么时候处理这些表达式或函数呢？数据库“烹调”一条 SQL 时，如何“烹调”，如何分解，是靠什么来做决定的？在数据库决定了怎么做之后，又是如何按部就班地执行的？如果说数据库的“扫描方法、表达式、操作符、UDF、索引接口”等是数据库的“料”，那么数据库的优化器就是“巧妇”之手，它包括了“JOIN 算法、SQL 重写规则、多表 JOIN 的遗传算法、动态路径规划、选择性算法、各种 NODE 的成本计算算法、成本因子、并行计算成本算法”等方方面面，为数据库如何执行 SQL 提供了全套流程。

PostgreSQL 作为一个非常经典的 ORDBMS，包含了很多“料”，同时有着非常先进的优化器，为高效地执行 SQL 提供了良好的基础。

本书作者长期致力于数据库内核的研发，有非常丰富的理论与实践经验，本书以 PostgreSQL 为背景，详细介绍了 PostgreSQL 查询优化器中的核心概念，从“查询树、SQL 重写、UNION 优

化、逻辑分解”到“下推、JOIN、选择性、统计信息、扫描路径、动态规划、遗传算法”等方面面，实为作者呕心沥血之作，同时也是数据库工作者，特别是 PGer 之福。

本书是不可多得的教科书级 PostgreSQL 内核读物，同时不乏实战性。建议想了解数据库优化器工作原理的读者及 PostgreSQL 爱好者深入学习。

感谢作者为 PostgreSQL 生态的辛勤付出，期待本书大卖。

PGer, Digoal

前言

为什么写这本书

我参加过很多次查询优化的培训，也查阅过很多查询优化的资料，但总是感觉对查询优化似懂而非，我总结其原因是多数培训和资料的时长或篇幅较短，内容多是对查询优化的概述，“巧妙”地避开了查询优化的难点，难以触及查询优化的本质，导致查询优化的“大道理”人人都懂，遇到问题却难以发力。

2016 年年末，我做了一次查询优化的培训，结合之前培训的经验，我对这次查询优化的培训打了一个“持久战”，不只是拿出几个小时的时间对查询优化进行一个总体描述，而是将查询优化器拆解开来，分阶段地进行详细的解读，大约做了十几次培训，最终的效果是非常显著的。在培训的过程中我发现，目前 PostgreSQL 数据库查询优化器实现细节相关的资料市场上少之又少，和数据库从业人员对查询优化器的热情远远不成正比，本着抛砖引玉的原则，我写了这本书。

为什么阅读这本书

- 在数据库内核开发的过程中，你是否有了解查询优化器的实现细节的欲望？
- 在对数据库进行调优的过程中，你是否感觉无从下手？

- 在分析查询优化的源码时，你是否会陷入某一细节而不可自拔？
- 在学习查询优化的理论时，你是否感觉理论与实践之间无法一一对应？

如果你希望深入地了解查询优化，那么最好的办法就是了解它的理论基础，然后细致地剖析查询优化器的源代码，通过理论和实践的结合，达到真正掌握相关知识的目的。本书细致地解读了 PostgreSQL 10.0 的查询优化器的大部分源码，对其中比较重要的理论都给出了说明，足以让读者了解 PostgreSQL 数据库查询优化器的全貌。

虽然本书已经尽量尝试将复杂问题简单化，但是鉴于 PostgreSQL 数据库的查询优化器的实现本身就具有一定的复杂性，读者阅读的过程可能是“痛苦”的，但请相信“梅花香自苦寒来”，只要坚持阅读就能收获很多。

本书的组织结构

本书的组织结构基本是按照 PostgreSQL 数据库的查询优化器处理一个查询的流程来安排的，由简入繁、由易入难。

第 1 章介绍一些查询优化基础理论，这些理论是对查询优化的概述，读者在阅读第 1 章时可以参考一些经典的数据库实现理论书籍，更详细地了解数据库的基本理论，这样能给后面的阅读打好基础。

第 2 章介绍查询树，查询树是 PostgreSQL 数据库查询优化器的输入，查询优化器本身是对查询树的等价改造及等价分解。

第 3 章介绍逻辑重写优化，逻辑重写优化是逻辑优化的一部分，它主要是对查询树进行基于规则的等价重写，比较重要的有子查询提升、表达式预处理、外连接消除等。

第 4 章介绍逻辑分解优化，逻辑分解优化仍然是逻辑优化的一部分，和逻辑重写优化不同，它开始尝试分解查询树，经过谓词下推、连接顺序交换、等价类推理等对查询树进行改造。

第 5 章介绍统计信息和选择率，统计信息是代价计算的基石，因此了解统计信息的类型、了解选择率的含义对代价计算有非常重要的意义。

第 6 章介绍扫描路径的建立过程，扫描路径是为了对基表进行扫描的物理算子创建的路径，它负责将物理存储或者缓存中的数据读取上来并进行处理，通常包括顺序扫描、索引扫描、位图扫描等。

第 7 章介绍路径搜索的两个算法，PostgreSQL 数据库采用了动态规划方法和遗传算法进行

路径搜索，本书对这两种方法的实现都做了详细的介绍。

第 8 章介绍连接路径的建立过程，PostgreSQL 数据库的物理连接路径有嵌套循环连接、哈希连接、归并连接等，由于采用的扫描路径不同，导致同一种类型的物理连接路径产生的代价不同。

第 9 章介绍 Non-SPJ 的相关优化，PostgreSQL 数据库对集合操作、聚集操作、分组操作、排序操作等都做了优化处理。

第 10 章介绍执行计划的生成，在扫描路径、连接路径及 Non-SPJ 路径分别处理之后，会选择一个“最优”的连接树，PostgreSQL 数据库需要将这个连接树修正成执行计划树。

错误

限于我的能力，书中难免有错误，在写作的过程中我也尝试尽量多查阅相关的资料，尽量避免错误的出现，但是相关的资料实在是太少了，因此，欢迎广大读者对本书提出纠正、批评和意见，这也有益于我本身能力的提升。

致谢

感谢彭煜玮、周正中（德哥 Digoal）为本书作序，感谢蒋志勇、文继军、王颖泽、杨瑜、赵殿奎对本书的评价，这对我是极大的鼓励。

在写作过程中，卢栋栋、彭信东、李茂增通读了大部分书稿，给出了很多有益的意见和建议，在此表示感谢。林文、翁燕青、白洁对书稿的格式及内容提出了修改建议，在此一并表示感谢。

感谢董英编辑，在写稿及后续的审校过程中董英编辑一直在和我沟通，不厌其烦地解答我的各种问题。

感谢我的家人。我的父母和妻子在我写作的过程中给予了极大的支持，写作的过程非常枯燥，他们为我提供了最好的写作环境。另外我的两个儿子也经常在我离开电脑的间隙帮我修改书稿，虽然他们的意见一条也没有被采纳，但这里仍然对他们的“贡献”表示感谢。

读者服务

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- **提交勘误：**您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/34148>



好书分享



目 录

第 1 章	概述	1
1.1	查询优化的简介	1
1.2	逻辑优化	3
1.2.1	关系模型	3
1.2.2	逻辑优化示例	8
1.3	物理优化	10
1.3.1	物理优化的 4 个“法宝”	12
1.3.2	物理路径的生成过程	14
1.4	文件介绍	17
1.5	示例的约定	18
1.6	小结	19
第 2 章	查询树	20
2.1	Node 的结构	20

2.2	Var 结构体	21
2.3	RangeTblEntry 结构体	23
2.4	RangeTblRef 结构体	25
2.5	JoinExpr 结构体	26
2.6	FromExpr 结构体	27
2.7	Query 结构体	27
2.8	查询树的展示	31
2.9	查询树的遍历	31
2.10	执行计划的展示	32
2.11	小结	33
第 3 章	逻辑重写优化	34
3.1	通用表达式	35
3.2	子查询提升	36
3.2.1	提升子连接	37
3.2.2	提升子查询	51
3.3	UNION ALL 优化	68
3.4	展开继承表	69
3.5	预处理表达式	71
3.5.1	连接 Var 的溯源	71
3.5.2	常量化简	72
3.5.3	谓词规范	73
3.5.4	子连接处理	79
3.6	处理 HAVING 子句	80
3.7	Group By 键值消除	81
3.8	外连接消除	82
3.9	grouping_planner 的说明	91
3.10	小结	92

第 4 章 逻辑分解优化 93

4.1 创建 RelOptInfo 94

4.1.1 RelOptInfo 结构体 94

4.1.2 IndexOptInfo 结构体 97

4.1.3 创建 RelOptInfo 100

4.2 初识等价类 102

4.3 谓词下推 106

4.3.1 连接条件的下推 106

4.3.2 过滤条件的下推 112

4.3.3 连接顺序 113

4.3.4 deconstruct_recurse 函数 118

4.3.5 make_outerjoininfo 函数 124

4.3.6 distribute_qual_to_rels 函数 132

4.3.7 reconsider_outer_join_clauses 函数 151

4.3.8 generate_base_implied_equalities 函数 156

4.3.9 记录表之间的等价关系 157

4.4 PlaceholderVar 的作用 158

4.5 Lateral 语法的支持 161

4.5.1 Lateral 的语义分析 162

4.5.2 收集 Lateral 变量 164

4.5.3 收集 Lateral 信息 164

4.6 消除无用连接项 166

4.7 Semi Join 消除 171

4.8 提取新的约束条件 172

4.8.1 提取需要满足的条件 173

4.8.2 提取流程 174

4.8.3 选择率修正 176

4.9 小结 177

第 5 章 统计信息和选择率	178
5.1 统计信息	178
5.1.1 PG_STATISTIC 系统表	181
5.1.2 PG_STATISTIC_EXT 系统表	185
5.1.3 单列统计信息生成	187
5.1.4 多列统计信息生成	196
5.2 选择率	200
5.2.1 使用函数依赖计算选择率	204
5.2.2 子约束条件的选择率	208
5.2.3 基于范围的约束条件的选择率修正	211
5.3 OpExpr 的选择率	213
5.3.1 eqsel 函数	215
5.3.2 scalargtsel 函数	217
5.3.3 eqjoinsel 函数	220
5.4 小结	226
第 6 章 扫描路径	227
6.1 代价 (Cost)	228
6.1.1 代价基准单位	228
6.1.2 启动代价和整体代价	231
6.1.3 表达式代价的计算	233
6.2 路径 (Path)	236
6.2.1 Path 结构体	236
6.2.2 并行参数	237
6.2.3 参数化路径	239
6.2.4 PathKey	242
6.3 make_one_rel 函数	244
6.4 普通表的扫描路径	245

6.4.1	顺序扫描	246
6.4.2	索引扫描	248
6.4.3	位图扫描	281
6.5	小结	291
第 7 章	动态规划和遗传算法	292
7.1	动态规划	293
7.1.1	make_rel_from_joinlist 函数	297
7.1.2	standard_join_search 函数	298
7.1.3	join_search_one_level 函数	298
7.2	遗传算法	301
7.2.1	种群初始化	303
7.2.2	选择算子	308
7.2.3	交叉算子	310
7.2.4	适应度计算	311
7.3	小结	312
第 8 章	连接路径	313
8.1	检查	314
8.1.1	初步检查	314
8.1.2	精确检查	316
8.1.3	“合法”连接	318
8.2	生成新的 RelOptInfo	324
8.3	虚表	327
8.4	Semi Join 和唯一化路径	328
8.5	建立连接路径	331
8.5.1	sort_inner_and_outer 函数	334
8.5.2	match_unsorted_outer 函数	345