

# MySQL 数据库开发研究

## ——基础研究视角下的探索与应用

张岩 金刚

哈尔滨信息工程学院

**摘要:** 文章以我校教学中使用 MySQL 数据库开发为研究对象,从 MySQL 数据库基础研究角度进行了系统的研究和探索。通过剖析 MySQL 数据库核心技术、开发规范等基础内容,结合学校计算机、经管、工科等多学科方向,呈现不同领域在 MySQL 应用中的研究成果、理论突破、实验数据及技术革新。并对学术交流、知识普及于基础研究方面所具有的重要作用做了解释、分析典型案例。研究表明,MySQL 数据库开发基础研究为各学科提供了高效数据支撑,推动了学科交叉融合,基础研究的结果对学校教学和科研活动、对应用人才的培养都有很大的参考价值。

**关键词:** MySQL 数据库开发; 基础研究; 学科应用; 学术交流; 技术革新

**DOI:** 10.65976/3105-4838.2026.03.017

如今,数字化浪潮正在席卷全球,数据是社会发展的驱动力,数据库作为数据存储、管理与分析的核心,其重要性不言而喻。MySQL 数据库开源免费、性能突出、跨平台、低成本等优势在科研、教育、企业级应用等领域得到广泛应用。据相关行业报告显示,全球约 40% 的中小型企业及 60% 的高校教学科研系统采用 MySQL 数据库。

哈尔滨信息工程学院是一所以培养应用型人才为主的民办大学,十分重视学科建设和科学研究。学校计算机科学与技术、软件工程、数据科学与大数据技术专业都有相当多的数据库开发需求,MySQL 数据库开发相关基础研究的深度与广度影响各学科的教学和科研水平。如计算机专业的软件开发课程要求有稳定的数据库支撑项目开发,而数据科学与大数据技术专业需要可视化展示层需要数据库支撑。

研究 MySQL 数据库开发基础对学校具有重要意义。一是能够提高学校在数据库方面的实力,为学校的某些专业提供技术储备;二是促进多学科交叉融合,通过数据纽带连接不同学科,形成协同创新格局;三是有利于教学与实践的衔接,让学生在掌握理论知识的同时,又能积累实际开发经验,为以后就业打下基础;四是地方经济服务,为企业发展提供技术支持,助力本地中小企业快速完成数字化转型升级。

学术交流、知识传播是基础研究的重要部分,在 MySQL 数据库开发研究中占有重要地位,建立校内的讨论、校企研讨、校际交流等各种机制有助于集众智,解决发展瓶颈。本文将围绕 MySQL 数据库开发基础研

究谈学校学科特点,综述 MySQL 数据库开发技术特点、学科应用、学术交流、发展预期等。

### 一、MySQL 数据库开发基础研究概述

#### (一) 核心技术基础

MySQL 数据库最核心的核心是稳定良好的运行、高效应用的基础,如数据存储结构、查询优化机制、事务处理等各部分的技术。MySQL 灵活性主要体现在数据库存储引擎。MySQL 提供了很多种存储引擎实现模式,如 InnoDB 是 MySQL 默认存储引擎,其数据和索引结构具有聚簇索引的特点,将数据与索引紧密结合,InnoDB 存储引擎提供的事务 ACID 特性,行级锁,适合用于金融业务、电商网站订购业务这种对数据库数据一致性能要求非常高的场景。MyISAM 数据索引结构没有聚簇索引的聚簇索引形式,索引和数据内容分开存放,读取速度高效,但是 MyISAM 不支持事务,适合用来做新闻发布类、图书目录等查询频繁、数据变化不常的网站。Memory 引擎中将数据放置在内存中,处理临时数据,Archive 引擎主要处理归档数据,每个引擎都有不同的特点,不同的特点适用于不同的场景。

查询优化是提升 MySQL 性能的重要手段,本质上是用 MySQL 的优化器生成高效执行计划。基础研究中首先应掌握执行计划分析;用通过 EXPLAIN 语句查看查询语句的执行方式,如表的扫描方式、索引的使用情况、连接方式等。在学校计算机专业搞数据库性能优化时,学生对用户行为数据表中的行为数据共 50 万条的查询进行优化。开始对原查询语句实行全表扫描查询耗时 4.8 秒,加上适当地 B+ 树索引后查询的耗时减少到 0.3

秒；再加上调整连接的顺序和子查询优化后最终查询的耗时都维持在 0.12 秒，这样的实验生动地传达了查询优化的意义和价值，同时也是索引分析、查询的写法改进等一般查询优化方法有效性的一种证明。

事务是实现数据可靠性的手段，其 ACID 特性（原子性、一致性、隔离性、持久性）是通过日志系统（Redo Log、Undo Log）与锁机制实现。高并发情况下选择什么样的事务隔离级别会严重影响系统性能与一致性。学校数据库实验室开展的并发事务实验显示：读未提交（Read Uncommitted）隔离级别下，事务吞吐量达 800TPS（每秒事务数），有脏读现象；可重复读（RepeatableRead）隔离级别（MySQL 默认）下能处理的事务吞吐量为 500TPS，无脏读、不可重复读；串行化（Serializable）隔离级别吞吐量仅 200TPS，但完全避免并发问题。实验结果为不同场景下的隔离级别选择提供了数据支撑。

## （二）开发方法与规范

建立科学规范的开发方法是实现 MySQL 数据库可维护、可扩展、安全的前提，贯穿数据库设计、开发、运维整个生命周期。数据库设计时一定要讲究合理性，按要求分表、分库避免数据冗余。如常见的传统电商订单系统的数据库设计问题，基本把用户、订单、商品信息放到一个表里，存在严重的数据库冗余、低效的数据库数据更新等问题，规范化之后将用户表、订单表、商品表、订单中的商品关联表分开设计，数据库冗余少了 60%，数据库的修改执行效率提高了 40%。

开发规范包括命名、编码、安全等维度。规范的命名是指表名定义成“模块\_功能”格式（如 user\_login），字段名采用小写字母或小写加下划线（如 user\_name），索引名标明类型（如 idx\_user\_phone 表示用户手机号索引）均体现名称直观含义。编码规范是要求 SQL 的命名关键字大写（如 SELECT），复杂 sql 语句部分添加注释、sql 查询结构嵌套不超过 3 层等等。安全规范是开发要求的深化，如：防止 SQL 注入——应用参数化查询避免 SQL 注入（项目使用字符串拼接 SQL 容易导致注入漏洞，解决方法采用 PreparedStatement，成功抵御攻击）。

## 二、学术交流与知识传播：MySQL 数据库开发基础研究的关键平台

### （一）学术交流在基础研究中的作用

学术交流是打破研究壁垒、激发创新思维的核心机制，为 MySQL 数据库开发基础研究提供了多元化的智慧碰撞平台。专业采用高年级课题方式交流数据库问题。

### （二）知识传播的途径与效果

知识传播是将研究成果转化为实际价值的关键环节，通过多元化渠道实现 MySQL 数据库开发知识的共享与应用。课程教学是知识传播的主阵地。高年级课题课程中，将最新研究成果融入教学内容。例如，将“分库分表技术”转化为实验项目，让学生在模拟高并发场景中设计分表方案；将“事务隔离级别实验数据”作为案例，帮助学生理解理论与实践的关联。

## 三、各学科方向中的 MySQL 数据库开发应用研究

计算机学科是 MySQL 数据库开发研究的主攻方向，包括性能优化、网络数据库、数据库挖掘等研究方向，是理论性和应用性结合比较紧密的一门学科。数据库性能优化研究方向大致有“速度需要提高”“占用资源需要减少”等研究课题。存储大量数据库时，“分层存储”是一个研究方向，即热点数据存储在 SSD 磁盘上（InnoDB），冷数据存储在 HDD 磁盘上（Archive），使得存储成本降低 30% 的同时满足热点数据查询速度的需求。

## 四、MySQL 数据库开发的理论探讨

### （一）数据模型理论

数据模型是数据库开发的骨架，它的合适与否决定了数据库的合理性。MySQL 是关系模型，但是在应用上需要针对实际应用情况进行改造。关系模型缺点之一是处理复杂层次数据（比如组织机构树形数据），查询效率很差。传统的递归查询方式（UNDER\_RECURSIVE）查询层次比较深（如大于 10 层）的效率非常低。本文对关系模型性能进行了“路径存储+索引”的优化，在表中增加一个路径 path 字段来存储全部路径（如“1.3.5”表示根节点为 1，子节点为 3，孙节点为 5），然后对该表建立索引，得到层次结构查询效率提高了 10 倍的性能。时空数据模型的扩展应用较普遍。MySQL 扩展时空字段，存储某个对象的地理位置和时间，时空索引（SPATIAL）便于进行距离查询。

### （二）事务处理理论

事务处理理论研究侧重于“性能与一致性的平衡”，优化隔离级别和并发控制策略提升适应性。

隔离级别变化方式的研究有新发现，提出了“基于负载的隔离级别自适应”方法。数据库 CPU 负载水平在 60% 以下时，数据库采用“可重复读”隔离级别；超过 80% 时，则自动“降低”到“读已提交”，另外通过应用层的补偿以保证一致性。长事务的解决有效处理实际问题。针对一些长事务如报表的生成等（长事务 >10s），将其拆分为“小事务+作业”的方式，

先进行少数的关键性数据的写入，其他的则做成异步处理的方式。在办理企业财务月报表时，拆分后的两个事务时间从 15s 减到 2s 左右，保证无法被处理长事务锁占用而阻塞并发访问等问题。

### 五、实验数据与技术革新

关键实验及数据，实验是验证理论、优化技术的方法，学校在这方面基于 MySQL 性能、可靠性等做过各种实验，积累了大量丰富数据。

索引性能对比实验用到 3 种常见的索引（B+ 树、哈希、全文索引），在数据量为 100 万的用户表上进行性能测试。实验结果为 B+ 树索引在范围查询（如 `age BETWEEN 20 AND 30`）中有最高的性能（0.12 秒）；哈希索引在等值查询（如 `user_id=10086`）的性能耗时为 0.05 秒比 B+ 树索引快（0.08 秒）；全文索引在模糊查询（如 `MATCH (name) AGAINST ('张')`）的性能耗时 0.3 秒比 `LIKE '%张 %'`（1.8 秒）又快的多，实验得出的结论是对不同查询需要选择不同的索引有很有价值的依据。

存储引擎做压力测试，模拟高并发的环境。在 500 用户并发写入（用户每秒写入 500 条记录）下，InnoDB 引擎平均响应时间 0.8s，数据一致性 100%；MyISAM 引擎响应时间 0.5s，不一致（无事务支持，丢 3 条数据）；Memory 平均响应时间 0.1s，重启清数据。结论是事务场景下，InnoDB 为好，临时数据可用 Memory 引擎处理。

### 六、案例分析：MySQL 数据库开发在学校教学与科研中的应用

案例一：“MySQL 性能优化”实验课程。实验设置“劣质 SQL”优化的实验任务，给出存在全表扫描、非索引排序等问题的 SQL 语句供学生进行 EXPLAIN 分析并优化。学生用熟练掌握 explain 的方法优化一个“查询近 30 天热门图书”的慢查询语句（耗时 3.2s），建立 `create_time` 和 `borrow_count` 的联合索引、把子查询改成链接查询优化，将耗时缩小到 0.2 秒。实验反馈 85% 的同学基本能将 SQL 优化至要求，掌握了如何设计索引和 SQL 执行计划查看及分析的方法。

案例二：“模拟电商用户画像分析系统”。采集电商网站用户访问、收藏、购买等行为信息（累计 100 万条），保存在 MySQL 数据库中，使用关联规则算法从中挖掘出用户的兴趣特点。发现在“购买婴儿奶粉的用户中有 60% 会在 1 个月内购买婴儿类湿巾”，根据这一统计规律为合作电商平台制定捆绑交易方案，捆绑产品的销售额比之前提升 20%。这个系统采用“定

时执行 + 增量更新”机制，每隔一天的凌晨更新用户画像，确保数据分析时效性。

### 七、结论与展望

#### （一）结论

本文通过对我校 MySQL 数据库开发基础研究的系统梳理，得到如下结论：基础研究体系日趋完善，拥有了以核心技术部分（存储引擎、查询优化、事务处理）为重点，以开发规则为基础的研究框架，为广大学科应用开发提供了支持。

学科应用成效显著，各学科应用中，MySQL 数据库较好地解决了数据的存储、分析、管理等问题，促进了学校的教学与科研水平。基础研究与开发的技术革新实用，开发的 MySQL 索引建议工具、MySQL 数据库同步中间件等都有很大实用价值，在教学、科研、公司等方面取得较好效果，也是基础研究与开发应用化特点的体现。

#### （二）展望

云原生 MySQL 研究。研究 MySQL 与 Kubernetes 的结合，实现 MySQL 数据库的自动扩缩容、自动故障恢复等，满足云环境下的弹性需求。计划与云服务企业合作，提出针对中小企业的轻量化的云数据库方案。

AI 辅助数据库优化。利用 AI 做预测 SQL 执行性能预测，利用 AI 提供 SQL 优化建议，降低数据库运维人成本。初步设想是基于 SQL 以往执行历史记录训练模型，实现 AI 推荐索引、AI 调参数等。

同时，将进一步加强相关领域学术交流和科普平台建设，提升“数据库技术沙龙”的影响力，建立“MySQL 技术社区”，吸收接纳更多师生参与，保证讨论热度和可持续性，为学校的人才培养、学科建设与促进区域经济发展持续服务。

#### 参考文献：

- [1] 王珊, 萨师焯. 数据库系统概论 (第 6 版)[M]. 北京: 高等教育出版社, 2020.
- [2] 周志华. MySQL 性能调优与架构设计 [M]. 北京: 机械工业出版社, 2022.
- [3] 张晓明, 李红. 开源数据库技术及应用 [J]. 计算机工程与应用, 2021, 57(12): 1-10.
- [4] 刘军. 分布式 MySQL 数据库设计与实现 [J]. 数据库技术与应用, 2023, 26(03): 45-50.
- [5] 陈志强. MySQL 在电子商务中的数据处理策略 [J]. 中国管理信息化, 2022, 25(08): 16-18.
- [6] Oracle 官方文档. MySQL 8.0 Reference Manual [Z]. 2023.