

# 性能比拼: MySQL vs PostgreSQL

旅途散记 2025年04月07日 20:18 上海

本内容是对知名性能评测博主 Anton Putra [MySQL vs PostgreSQL Performance Benchmark \(Latency – Throughput – Saturation\)](#)<sup>[1]</sup> 内容的翻译与整理, 有适当删减, 相关指标和结论以原作为准

## MySQL vs PostgreSQL 数据库性能对比

在本内容中, 我们将对比 [MySQL](#) 和 [PostgreSQL](#) 关系型数据库的性能。我们将运行一系列测试, 其中 [第一项测试](#) 重点关注 [数据写入 \(ingestion\) 效率](#)。

首先, 我们会测量:

- [数据插入请求的延迟](#), 使用 [p99 百分位数](#) 进行评估
- [数据库的吞吐量](#), 即每秒可处理的 [请求数和插入数](#)
- [数据库的饱和度 \(saturation\)](#), 即数据库的负载情况, 在这里我们主要关注:
  - [CPU 使用率](#)
  - [内存使用率](#)

此外, 我们还会测量:

- [磁盘写入操作数](#)
- [数据库在磁盘上的存储效率](#)

值得一提的是, 这两种数据库在这些方面的差异 [非常大](#)。

最后, 我们将评估 [数据库的连接池 \(connection pool\)](#), 观察它们如何管理插入数据的连接。

在 [第二项测试](#) 中, 我们将测量 [数据读取 \(retrieval\)](#) 的效率。

我使用的是当前最新版本:

- [PostgreSQL 17.0](#)
- [MySQL 9.0](#)

为了进行测试，我在 **两个数据库** 中分别创建了 **两张表**，具体的 SQL 语句如下。

假设我们有一个 **分析系统 (analytics backend)**，用于记录 **用户在网站上的行为**，例如：

- **登录**
- **登出**
- **搜索**
- **将商品加入购物车**

数据库中有两张表：

#### 1. **Customer 表**

- 存储用户的 **姓名、地址**
- 设有 **主键 ID**

#### 2. **Event 表**

- 设有 **主键 ID**
- **外键** 指向 **Customer 表的主键**
- 记录 **用户执行的操作类型 (action type)**

#### **第一项测试：**

我们将使用以下 SQL 语句插入数据，**随机生成客户 ID 和操作类型**。

#### **第二项测试：**

我们将在包含 **约 7000 万条记录** 的数据库上执行 **简单的 JOIN 查询**：

- **随机生成一个 event ID**
- **查询并连接 customer 表**
- **返回用户的姓名、地址和操作类型**

如果你有任何关于 **改进测试设计** 的建议，请告诉我！

-----

## 代码概览

在客户端编写方面，我选择使用 [Golang](#)，因为：

- [Golang 是一门流行的编程语言](#)
- [它拥有成熟的数据库驱动库](#)

为了让 [MySQL](#) 和 [PostgreSQL](#) 的测试尽可能公平，我使用 [database/sql 接口](#) 进行数据库操作，而不是直接使用 [pgx 驱动](#)（尽管 [pgx](#) 可能会降低查询延迟）。但本次测试的重点是 [对比数据库本身的性能](#)，而不是不同的驱动实现。

此外，我确保：

- [执行相同的 SQL 查询](#)
- [使用相同的数据库连接配置](#)

如果你有任何改进建议，请告诉我，或者更好的是，[提交一个 Pull Request](#)！

## 第一项测试：数据写入

现在，我们开始 [第一项测试](#)。

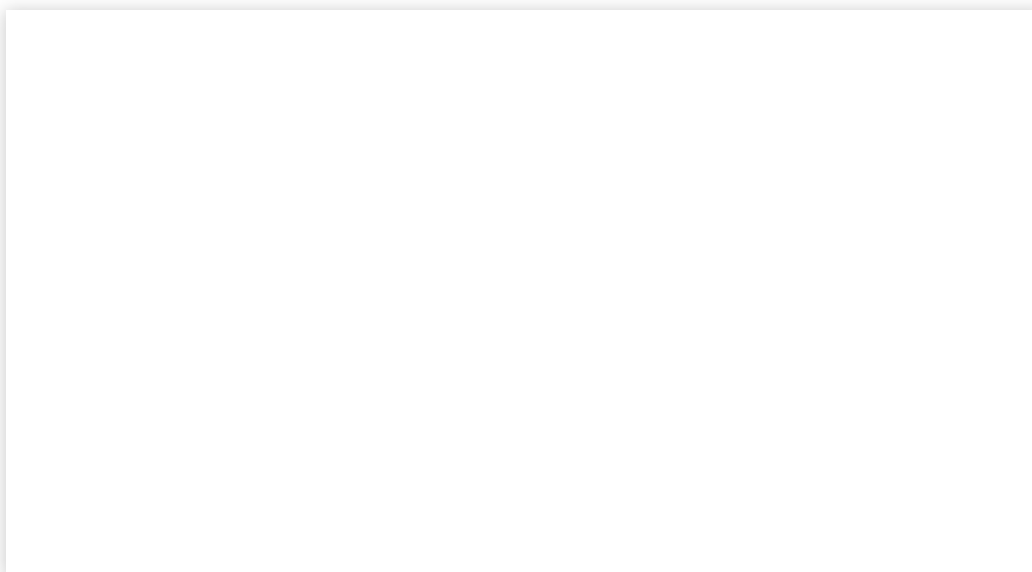
这次测试 [总共持续了近 3 小时](#)，但我会将其压缩至 [几分钟](#) 展示。测试过程中，我们会：

- [部署多个客户端](#)
- [逐步增加负载](#)
- [持续向数据库发送请求](#)

- 每 60 秒增加一个虚拟客户端

你可以在 右上方的图表 看到 每秒查询数（QPS），而 左侧图表 显示的是 从客户端测量的插入延迟。

测试结果：



- 从一开始，PostgreSQL 的插入延迟明显低于 MySQL
- MySQL 插入数据的时间几乎是 PostgreSQL 的两倍
- PostgreSQL 的 CPU 使用率更低，意味着它有更大的吞吐能力
- PostgreSQL 占用的内存也更少

最大区别：

- MySQL 需要更多的磁盘操作来存储相同数量的记录
- PostgreSQL 在磁盘上的存储效率更高
- PostgreSQL 执行的磁盘操作更少，数据占用空间也更小

连接池情况：

- 两个数据库的连接池最终都达到了 80 个连接
- 每个客户端的最大连接数设置为 40
- 每个数据库有 2 个客户端在不断生成负载

临界点：

- 当 QPS 达到 5,500 时，MySQL 开始跟不上 PostgreSQL
  - MySQL 的延迟和 CPU 使用率激增，导致整体性能下降
  - 当 QPS 达到 12,000 时，我为 PostgreSQL 额外添加了一个客户端
  - 最终，PostgreSQL 最高可处理 19,000 QPS（4 核 CPU + 直连 SSD）
  - MySQL 最高只能处理 10,000 QPS，约为 PostgreSQL 的一半
-

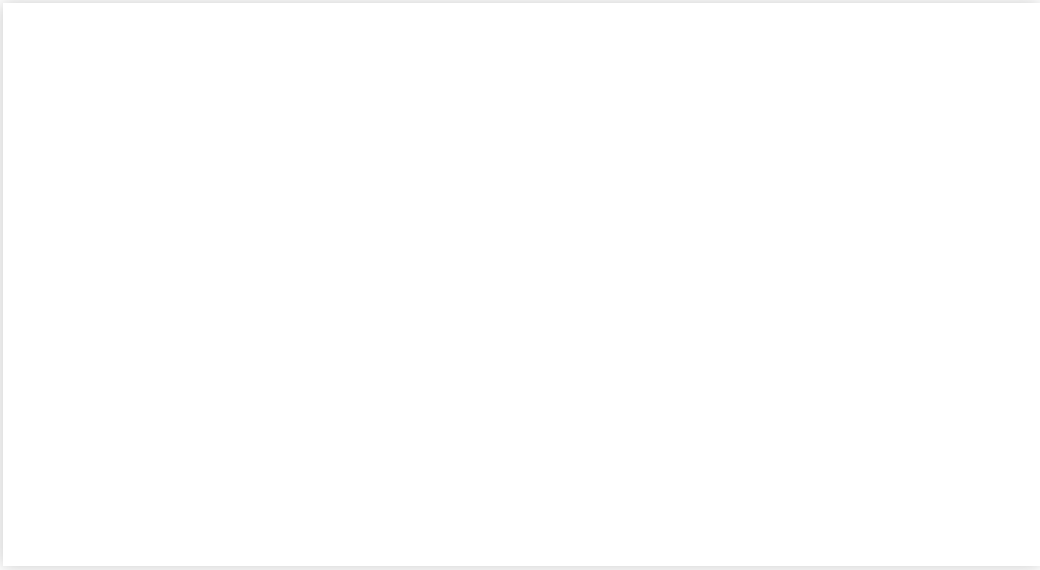
第一项测试：数据分析

现在，我们查看 [整个测试期间的数据](#)：

1. [每秒查询数（QPS）](#)



1. [插入延迟（客户端测量）](#)



1. [CPU 使用率](#)

## 1. 总插入记录数

- PostgreSQL 插入 约 6500 万条记录
- MySQL 插入 约 5000 万条记录
- 在第二项测试前，我会为 MySQL 增加一些记录，使数据量一致

## 1. 内存使用

### 1. 磁盘写入操作



### 1. 连接池大小



### 1. 数据库大小

- PostgreSQL 存储数据的 **磁盘效率更高**
- **CPU 和磁盘操作占用更少**
- PostgreSQL 在数据写入方面明显优于 MySQL



## 第二项测试：数据读取

在运行 **第二项测试** 之前，我为 **MySQL** 额外添加了一些记录，使其与 **PostgreSQL** 的数据量保持一致。

测试内容：

- 查询 **event** 表中的数据
- 通过 **event ID** 进行 JOIN
- 返回用户的姓名、地址和操作类型
- 测试前，重启两个数据库，以清除缓存

磁盘读取数据问题：

- 由于可能的配置错误，我没有获取到 **PostgreSQL** 的磁盘读取数据
- 但我仍然有 **MySQL** 的磁盘操作数据

测试结果：

- 延迟差距不像第一项测试那么大
- 但 PostgreSQL 仍然比 MySQL 低延迟
- MySQL 的 CPU 使用率 显著高于 PostgreSQL
- 这个趋势一直持续，直到 MySQL 达到 CPU 使用瓶颈，性能下降

临界点：



- 当 QPS 达到 18,000 时，MySQL 的性能下降更明显
- MySQL 的延迟出现波动，而 PostgreSQL 仍然保持稳定
- 当 QPS 达到 20,000 时，MySQL 无法再处理更多查询
- 我为 PostgreSQL 额外增加了一个客户端，加速测试
- 最终，PostgreSQL 在 32,000 QPS 时达到 90% CPU 使用率，开始下降

结论：

- 在数据写入和读取方面，PostgreSQL 都比 MySQL 更高效
- 如果你有任何改进建议，例如更改客户端代码或测试设计，请告诉我！

---

## 第二项测试：数据分析

现在，我们查看 整个测试期间的数据：

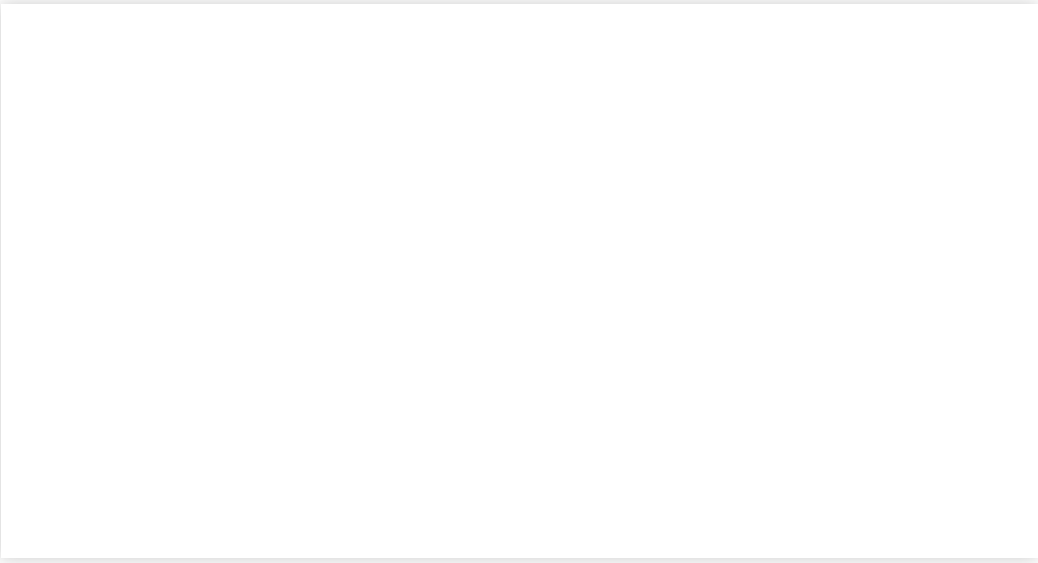
1. 每秒查询数（QPS）



1. 数据读取延迟



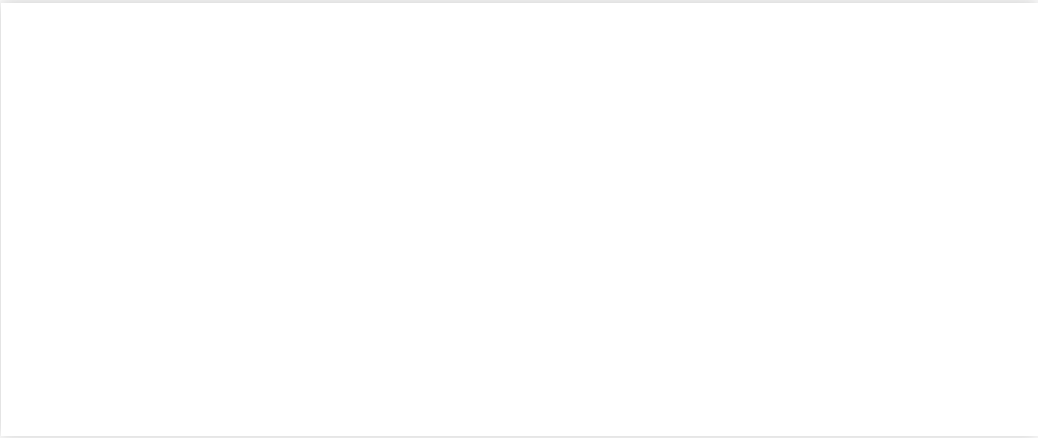
1. CPU 使用率



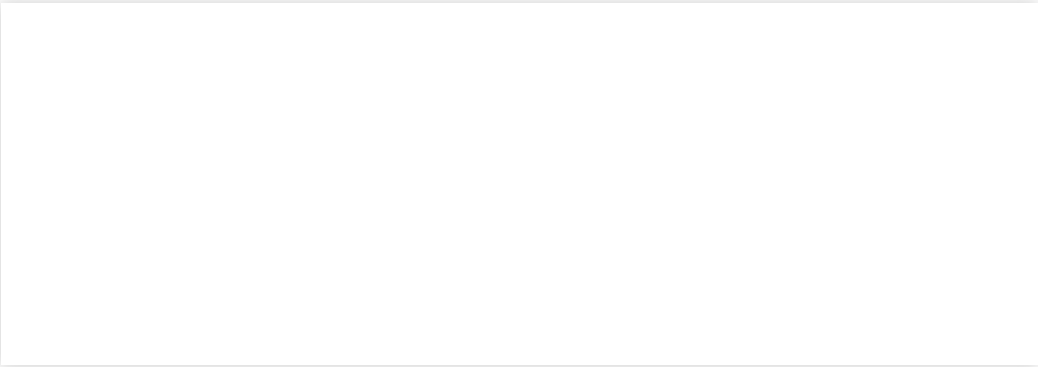
1. 内存使用



1. 数据库连接数



1. 磁盘读取操作



参考资料

[1] MySQL vs PostgreSQL Performance Benchmark (Latency – Throughput – Saturation): <https://www.youtube.com/watch?v=R7jBtnrUmYI>

性能测试 · 目录

上一篇

性能比拼: Go标准库 vs Python FastAPI(第二轮)

下一篇

性能比拼: gRPC vs REST vs GraphQL

