Database Performance Comparison

Eric Lee, Andrew Loader, Bryce Uota

Project Motivation

Every modern application needs a database

Our Fundamental Question:

- Should you use a traditional relational database (PostgreSQL)?
- Or a NoSQL database (MongoDB)?

Our Goal:

- Generate real performance data
- Evaluate realistic workloads at scale (10K → 1M records)

What We're Addressing

Real applications don't do just one thing:

E-Commerce Example (Amazon):

- User browsing products (Reads)
- System running inventory (Complex queries)
- Processing new orders (Writes)

We Evaluated:

- Write Performance: Bulk data digestion
- Read Performance: Mixed query types (Lookup, count, average)
- Performance degradation as data scales

Key Question: How do databases perform under realistic conditions?

Infrastructure Setup

Hardware (Identical for both databases)

- AWS EC2 t2.Large instances
- CPU: 2 vCPUs per instance
- Memory: 8 GB RAM per instance
- Storage: 20 GB SSD per instance
- Network: AWS US-West region

Software

- PostgreSQL: Version 15.14
- MongoDB: Version 7.0
- OS: Amazon Linux 2023
- Benchmark Strategy: Custom Python script

Scale and Cost

Dataset Sizes:

- 10,000 records
- 50,000 records
- 100,000 records
- 500,000 records
- 1,000,000 records
- 10 tests (5 sizes ✕ 2 workload types)

AWS Cost Summary

- $18.19

Methodology - Data Model

E-Commerce Product Catalog

- ID
- Name
- Price
- Category (Electronics, Books, Clothing, Home)

Sample Data

- ID: 12345
- Name: "Product_847213"
- Price: 459.99
- Category: "Electronics"

Methodology - Workloads

Write Performance Evaluation

- Insert all N records as fast as possible
- PostgreSQL: Individual inserts
- MongoDB: Batches, 5,000 per batch
- Measures: Bulk data ingestion speed

Read Performance Evaluation (2,000 queries)

- 50% lookup queries
- 30% count queries
- 20% average queries

Write Performance Results

Dataset Size	PostgreSQL	MongoDB	Winner
10,000	1.45s 6,881 rec/sec	0.15s 65,256 rec/sec	MongoDB (9.5x faster)
50,000	7.16s 6,982 rec/sec	0.70s 71,391 rec/sec	MongoDB (10.2x faster)
100,000	14.25s 7,017 rec/sec	1.38s 72,261 rec/sec	MongoDB (10.3x faster)
500,000	70.40s 7,103 rec/sec	6.72s 74,354 rec/sec	MongoDB (10.5x faster)
1,000,000	140.32s 7,127 rec/sec	13.61s 73,500 rec/sec	MongoDB (10.3x faster)

Read Performance Results

Dataset Size	PostgreSQL	MongoDB	Winner
10,000	2.69s 743 q/sec	14.02s 143 q/sec	PostgreSQL (5.2x faster)
50,000	12.19s 164 q/sec	61.09s 33 q/sec	PostgreSQL (5.0x faster)
100,000	24.00s 83 q/sec	121.60s 16 q/sec	PostgreSQL (5.2x faster)
500,000	71.88s 28 q/sec	599.74s 3 q/sec	PostgreSQL (9.3x faster)
1,000,000	133.11s 15 q/sec	1,191.75s 2 q/sec	PostgreSQL (7.5x faster)

Statistical Comparison of throughput for Write PostgreSQL vs MongoDB
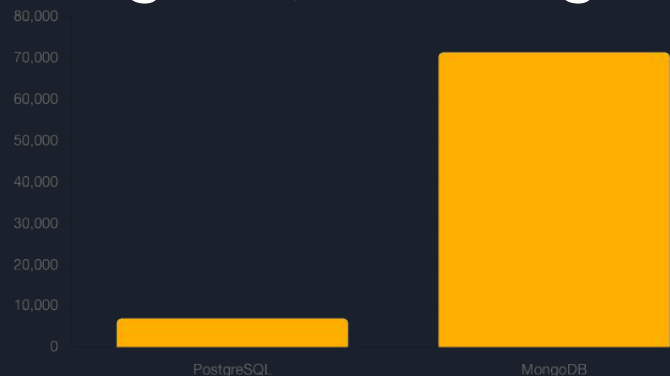
PostgreSQL
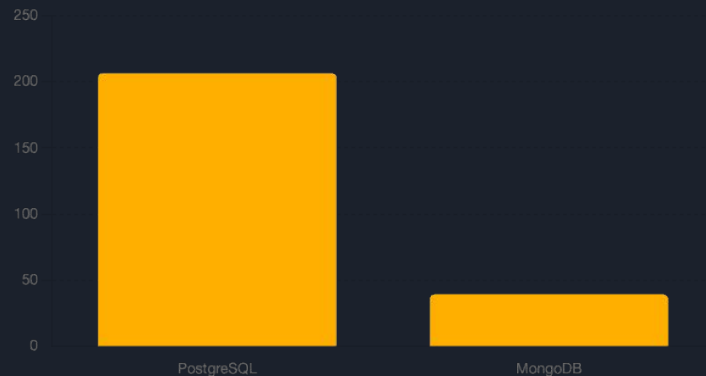
Mean: 7,022 rec/sec

MongoDB

Mean: 71,352 rec/sec

We are 95% confident that MongoDB is 59,984 to 68,676 rec/sec faster than PostgreSQL for write operations

CI does not include 0 → Performance difference is significantly different

Statistical Comparison of throughput for Read PostgreSQL vs MongoDB

PostgreSQL

Mean: 206.6 rec/sec

MongoDB

Mean: 71,352 rec/sec

We are 95% confident that PostgreSQL is 3.46x to 9.95x faster than MongoDB for read operations

CI does not include 1x → Performance difference is significantly different

Resource Utilization

Peak: 44% during 1M record write test (AWS CloudWatch)

Average: 20-25%

By Workload:

- 10K write: 5-10%
- 100K write: 25-30%
- 1M write: 40-44% (Peak)
- Read tests: 5-10%

Memory Usage:

- 6.3% (Peak)

Resource Utilization

CPU Headroom:

- 56% CPU available even at peak load
- Performance differences are architectural, not hardware-limited

Memory Efficiency

- Both databases highly efficient
- Minimal overhead for 100x data increase
- No memory bottlenecks observed

Scalability Validation

Linear Scalability

- Both databases scale predictably from 10K → 1M records
- No performance cliffs or unexpected degradation

Realistic Scaling Behavior

- Read times increase with data size (2.69s → 133s)
- Pure lookup queries would hide scaling issues

Resource Efficiency

- Performance differences are architectural
- Both databases operate efficiently

Performance Summary & Conclusion

Write Performance

- MongoDB: ~10x faster (Consistent across all scales)
- Maintains ~73,000 records/sec throughput
- Use case: High-volume data ingestion

Read Performance

- PostgreSQL: 3.46-9.95x (Advantage grows at larger scales)
- Use case: Analytics, reporting, complex queries

Performance Summary & Conclusion

Write Workload

- Why is MongoDB faster? Schema-less design, append-only architecture, minimal index overhead

Read Workload

- Why is PostgreSQL faster? 30+ years of query optimization, efficient B-tree indexing

Conclusion: For general-purpose applications with typical read/write patterns, PostgreSQL delivers better performance despite being slower at pure writes.