THE LINUX FOUNDATION

OPEN SOURCE SUMMIT INDIA

Cut Database Costs, Fund Innovation with Valkey

Roberto Luna-Rojas
Sr Developer Advocate Valkey OSS, AWS

What Is Valkey?

Open source, high-performance in-memory key/value data store

THE LINUX FOUNDATION

The Invisible Tax

Every redundant query is a tax on innovation.

A hidden cost that compounds at scale, paid on every single request.

Caching Is Architecture, Not a Band-Aid

Bolt-On Cache

A patch added after the bill arrives.

- Stapled onto a finished system
- Inconsistent and hard to reason about
- Considered only once costs spike

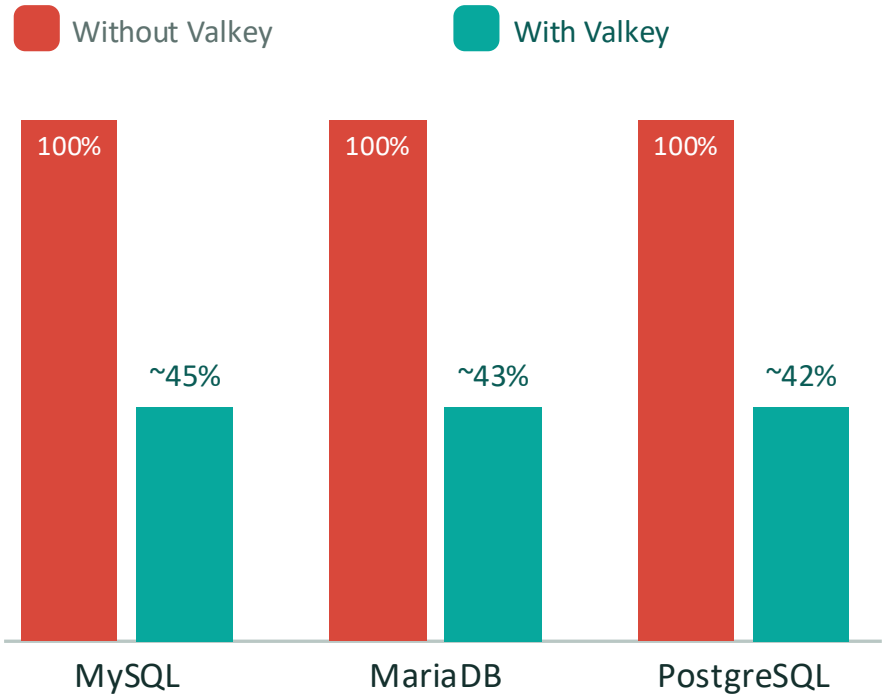
vs

Cache-First Design

Designed into the data layer from day one.

- ✓ The default path for every read
- ✓ Predictable, scalable economics
- ✓ Built in before the first deploy

Benchmarks: Traditional Databases



>55%

lower database spend

- Response time: orders of magnitude faster
- Database CPU utilization: dramatically lower

Illustrative relative DB spend. Read-heavy and mixed workloads, same hardware, Valkey as a look-aside cache.

The AI Cost Problem

Two users. Different words. Same intent.

A large share of LLM calls are semantically redundant. You pay for inference every single time.

Semantic Caching with Valkey

Up to

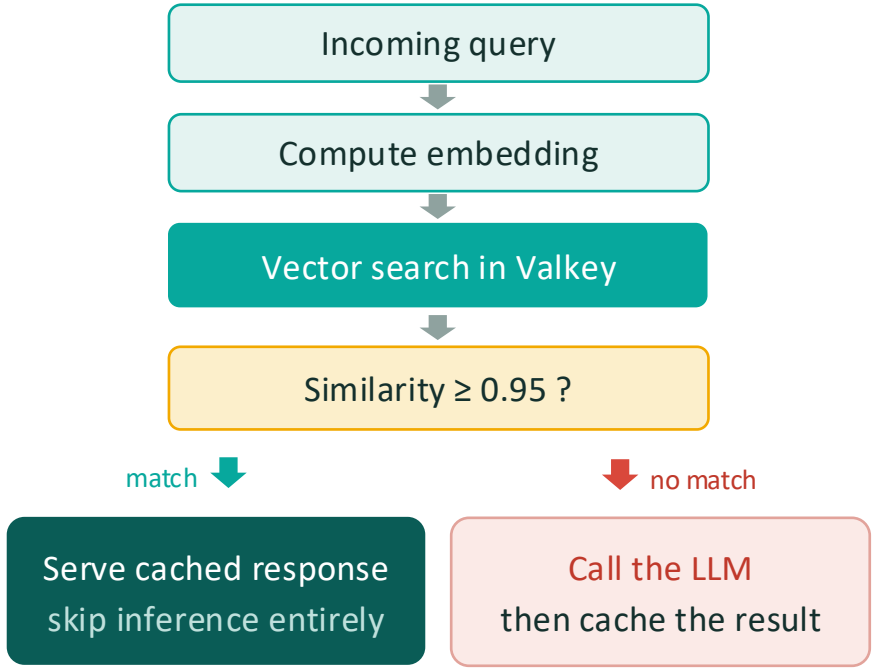
86%

lower LLM inference cost

Seconds → milliseconds

response latency on a cache hit

– Not by making the model cheaper. By not calling it.



Fund Innovation, Not Redundancy

1 Caching is architecture

Design it in from day one, not bolted on after the bill arrives.

2 Savings fund innovation

55% off database + up to 86% off LLM inference = budget to build.

Come find me at the Valkey booth

Benchmarks, architecture, mapping it to your workload.

valkey.io · github.com/valkey-io

Scan for valkey.io