redislabs
home of redis

whitepaper

Redis on Flash with Intel NVMe SSDs: A High Performance Benchmark

>3 Million operations/sec, <1 ms latency while generating > 1GB disk throughput with a single x86 server

Benchmark performed by:

Keren Ouaknine: Redis Labs
Frank Ober: Intel

Table of Contents

Executive Summary

The next generation of Intel NVMe SSDs technology delivers leading performance, low latency and high quality of service. The extreme throughput and low latency of this new technology makes it an ideal choice for big data use cases, where DRAM cost is sometimes a limiting factor in how fast insights can be extracted.

Redis, the world's most popular in-memory NoSQL database is used by developers worldwide for its powerfully efficient operational and analytical processing of data. Redis typically runs in RAM and delivers extremely high performance—millions of operations per second with sub-millisecond latency on a single standard server. Redis on Flash (delivered as a service through Redis Cloud Private or as downloadable software in the form of Redis Labs Enterprise Cluster -RLEC) can run on a combination of RAM and "slower memory," such as SATA-based SSD or next generation NVMe-based SSD, and even allows the user to adjust the ratio of the two types of memory. With this increased flexibility, customers can now better optimize performance, making cost tradeoffs based on the importance of individual workloads.

The combination of Redis on Flash and Intel NVMe SSDs has set a new industry performance record: delivering over 3 million ops/second and generating over 1GB disk throughput on a single bare metal server while keeping latency under 1 millisecond. This new generation of NVMe based SSDs delivers a 15x increase in throughput with only incremental cost as compared to SATA-based SSDs. This is a monumental advance in the economics of in-memory computing.

This report outlines the details of the benchmark.

Introduction

Large Datasets in Memory: The Business Driver

Redis is the most popular in-memory database of choice for real-time operational or transactional processing of application data, often classified as "hot" data. Hot data typically requires throughput in the millions of operations/second and latencies at sub-millisecond levels. As Redis deployments increase in size and as Redis is used for an increasing number of use cases, the cost of memory could limit how much data can be handled at lightning fast speeds. Alternative memory technologies like Intel NVMe deliver memory-like performance, at one-tenth the cost of DRAM. With Redis on Flash, the versatility of Redis is now accessible for a range of use cases, that require hundreds of thousands of operations/sec, and latency levels in the milliseconds, with greater cost effectiveness than using DRAM.

A similar shift to in-memory technologies can be observed in data analytics use cases, where analysis tools must meet user responsiveness SLAs. Big data analytics is rapidly moving towards the real-time automation of decisions and instantaneous customization of responses to user profiles. In such scenarios, using DRAM for increasingly large datasets can prove to be expensive. Emerging memory technologies such as the Intel NVMe eliminate cost barriers while delivering groundbreaking performance, enabling blazing fast analysis of extremely large datasets on standard x86 servers.

Intel® Solid-State Drive DC P3700 Series Specifications

The Intel SSD DC P3700 Series is a PCIe* Gen3 SSD architected with a new high performance controller interface, Non-Volatile Memory Express*, which delivers leading performance, low latency and quality of service. Matching the performance with world-class reliability and endurance, Intel SSD DC P3700 Series offers a range of capacities: 400 GB, 800 GB, 1.6 TB, and 2 TB in both add-In card and 2.5-inch form factor.

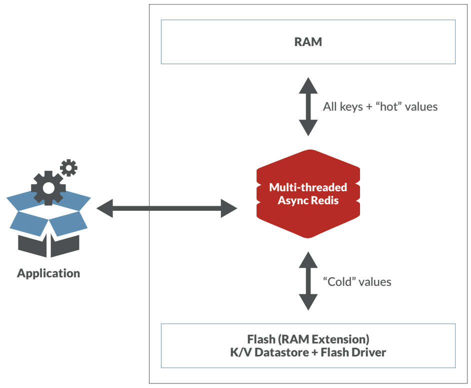
With PCIe Gen3 support and the NVMe queuing interface, the Intel SSD DC P3700 Series delivers excellent sequential read performance of up to 2.8 GB/s and sequential write speeds of up to 1.9 GB/s. Intel SSD DC P3700 Series delivers very high random read IOPS of 460,000 and random write IOPS of 180,000 for 4 KB operations. Taking advantage of the direct path from the storage to the CPU by means of NVMe, Intel SSD DC P3700 Series exhibits low latency of less than 20 μs for sequential access to the SSD.

For more information, visit http://www.intel.com/content/www/us/en/solid-state-drives/ssd-dc-p3700-spec.html

Redis on Flash

The enterprise-grade Redis from Redis Labs has been enhanced to run on a combination of "RAM" and Flash (which is used as "slower RAM") in the following form factors: regular SATA SSD, NVMe SSD or the new NV-DIMM technology. The enhancements are much more than a simple port of Redis, they include:

- The ability to tier memory into "fast" (RAM) and "slow" (Flash)
- The ability to configure RAM allocation with a "fast RAM" vs. "slower RAM" ratio
- The ability to store keys and "hot" values in RAM, and cold values in slow RAM
- Use of a pluggable storage engine (like RocksDB) to optimize access to Flash
- 100% compatibility with open source Redis



RocksDB, used in this benchmark as the pluggable storage engine, was further optimized and tuned for the specific characteristics of Redis on Flash.

The main advantage of RLEC on Flash is that the user can optimize tradeoffs between price and performance for various workloads.

Benchmarking Methodology

Methodology

We used memtier_benchmark, an open-source load generation tool created by Redis Labs.

We measured maximum operations/second that could be obtained from a single node at sub-millisecond latencies with different RAM:Flash hit ratios. Object sizes, read/write ratios and RAM/Flash hit ratios represent common configuration at Redis Labs customers.

Hardware and Software Setup

The benchmark was performed on industry standard hardware with

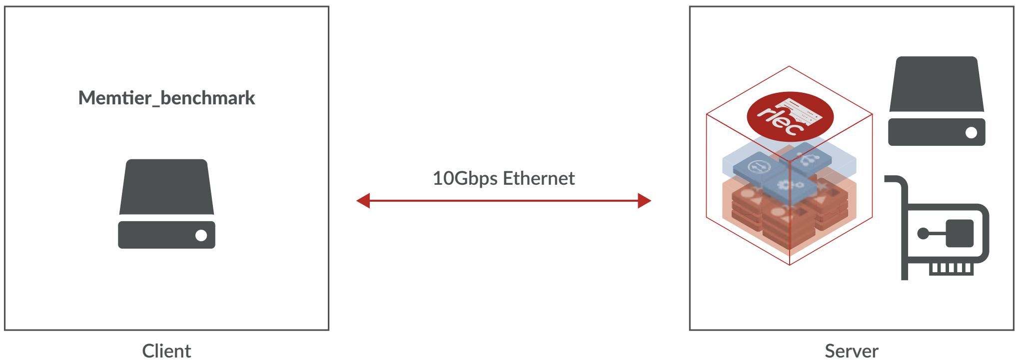
- Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz
- 2-sockets, 88 hardware threads, 44 cores (22 cores/socket)
- 128GB of memory (8GB DDR4 DIMMs at 2133 MHz)
- Four Intel® SSD DC P3700

Software versions:

Redis Labs Enterprise Cluster (RLEC) v4.3

Memtier_benchmark v1.2.6

We used the configuration in the illustration below to run the benchmark. The server used to run the memtier_benchmark was a 2-socket server with 12 cores/socket, resulting in 24 virtual cores and 128GB of memory (8GB DDR3 DIMMs at 1600 MHz).

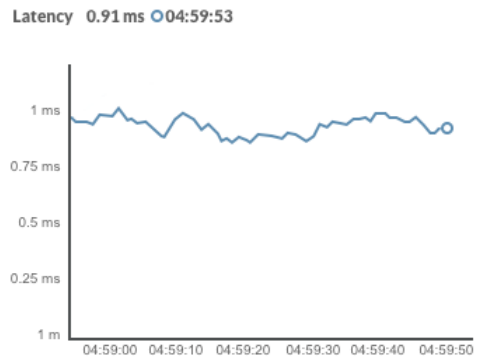


Results Observed

For object sizes of 100B with a combination of 80% reads and 20% writes, while 90% of the dataset was stored on Flash and only 10% in RAM, we achieved the following results:

An average of 3.25M ops/second, with an average of 0.92 ms latency from a single server, while reading and writing over 1GB of disk bandwidth.

Also notable was that 99.99% of requests were served with <1ms latency.



This represents a huge increase in throughput compared to regular SATA-based SSDs. Also, 3M ops/second, with <1ms latencies and >1 GB disk throughput to a single server is cause for celebration any day. This benchmark was accomplished with an 10:90 ratio of RAM:Flash (NVMe-based SSD).

While throughput is several times higher, the cost of NVMe-based SSDs is likely to be only incremental compared to SATA-based SSDs—representing a generational change in the economics of in-memory computing.

Conclusion

Intel SSD technology delivers an enormous advantage in price and performance when compared to SATA-based SSDs. Redis on Flash takes advantage of this improvement to deliver over 3M ops/second, at sub-millisecond latency with over 1GB disk bandwidth on a single server, setting a new record for highest performance with least amount of hardware.

The combination of these two technologies will herald a new age for operational in-memory databases as well as real-time big data analytics by making it both extremely fast as well as cost effective to serve and analyze very large datasets in memory