redislabs
home of redis

whitepaper

Redis on Samsung NVMe

Benchmarking Results

Benchmark performed by:

Keren Ouaknine: Redis Labs

Zvika Guz: Samsung

Table of Contents

Executive Summary

Samsung NVMe SSD technology delivers uncompromising performance with ultra-low latency, and is expected to bring the next revolution in storage technologies enabling big data analysis at blazing fast speeds but with substantially lower costs.

Redis, the world's most popular in-memory NoSQL database is the choice of developers worldwide for powerfully efficient operational and analytic processing of data.

Redis typically runs in RAM and delivers extremely high performance (millions of operations per second with sub-millisecond latency with a single standard server). Redis on Flash is delivered through Redis Labs Enterprise Cluster (RLEC) and runs on a combination of RAM and "slower memory" such as Flash or next generation NVMe. RLEC Flash allows configurability of the ratio between the two types of memory. With this flexibility, customers can now optimize performance-to-cost tradeoffs based on the importance of individual workloads.

The combination of RLEC Flash and Samsung NVMe SSD has set a new industry performance record: delivering over 2 million operations/second, over 1GB disk throughput and over 1Gbps network throughput on a single server while keeping latency under 1 millisecond. This represents a 40X increase in throughput with NVMe based SSDs, while the cost of the new SSDs is likely to be only incremental compared to SATA-based SSDs. This signifies a generational change in the economics of in-memory computing.

This report outlines the details of the benchmark.

Introduction

Large Datasets in Memory: The Business Driver

Big data analysis yields insights that are valuable to business decision-making. The first generation of big data tools, while revolutionary in the extreme amounts of data they could handle, were still largely oriented towards slow batch analytics. As expectations around user responsiveness and customized responses increase, processing and analyzing extremely large datasets (whether structured or unstructured) is increasingly moving to in-memory technologies. Emerging memory technologies like the Samsung NVMe are breaking down cost barriers while delivering groundbreaking performance, thereby enabling cost-effective analysis of extremely large datasets in memory.

Samsung PM1725 NVMe SSD

Samsung PM1725 NVMe SSD cards offer enterprise environments superb solid-state drives (SSDs) that deliver exceptional performance in multi-threaded applications. These high-performing SSDs also deliver outstanding reliability for continual operation in situations of unanticipated power loss.

With its revolutionary 3D vertical NAND technology, the Samsung PM1725 SSD paves the way to the terabyte era, offering a huge jump in capacity and ultra-low latencies. Samsung NVMe SSD cards are available with the following specifications:

Form Factor	2.5”U.2 devices or HHHL Cards
Host Interface	PCIe Gen3 x4
Capacities	800GB, 1.6TB, 3.2 TB
Sequential Read	3100 MB/s
Sequential Write	2000 MB/s
Random Write	120KIOPS
Random Read	750KIOPS

RLEC Flash

Redis Labs Enterprise Cluster (RLEC), the enterprise-grade Redis offering from Redis Labs, has been enhanced from open source 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 new NV-DIMM technology, like Intel’s 3DXPoint. RLEC’s 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: “slower RAM” ratios
- 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 block storage
- 100% compatibility with open source Redis

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

One main advantage of RLEC on Flash is that developers can optimize price-to-performance tradeoffs depending on the needs of various workloads, by configuring the RAM: slower RAM ratios.

Benchmarking Methodology

Methodology

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

We measured the 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 Redis Labs customers' configurations.

Benchmark performance targets - M2G2[1] on a single instance

For this benchmark, our goal was to achieve over 1M ops/sec while generating over 1GB disk throughput and over 1Gbps network throughput, and while keeping RLEC Flash latency at sub-millisecond. We use the acronym M2G2 to refer to this goal.

Hardware and Software Setup

The benchmark was performed on industry standard hardware. The specific server used to run the clients and RLEC was:

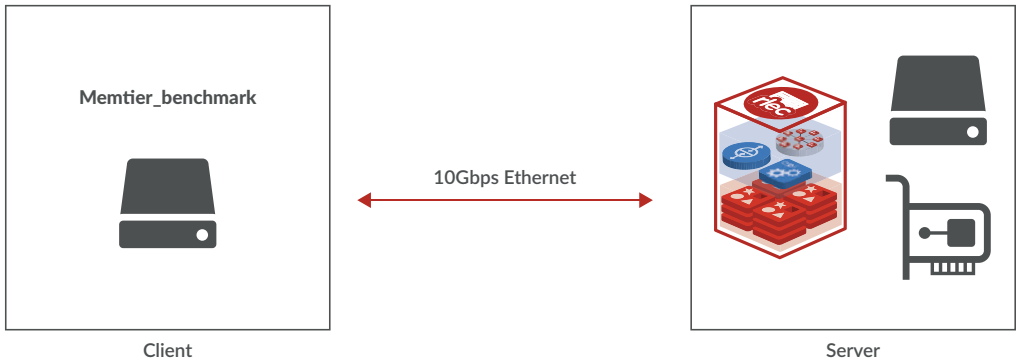
- Dell PowerEdge R730xd Rack Server with 2 X Intel(R) Xeon(R) CPU E5-2670 v3 @ 2.30GHz with a total of 24 physical cores (48 hardware threads) and 256 GB of RAM for the server and 64 GB for the client

- Four Samsung NVMe PM1725 cards with a total capacity of 3.2 TB

Software versions:

Redis Labs Enterprise Cluster (RLEC) v4.3
Memtier_benchmark v1.2.6

We used a client server and one server running RLEC for the benchmark, as shown below:

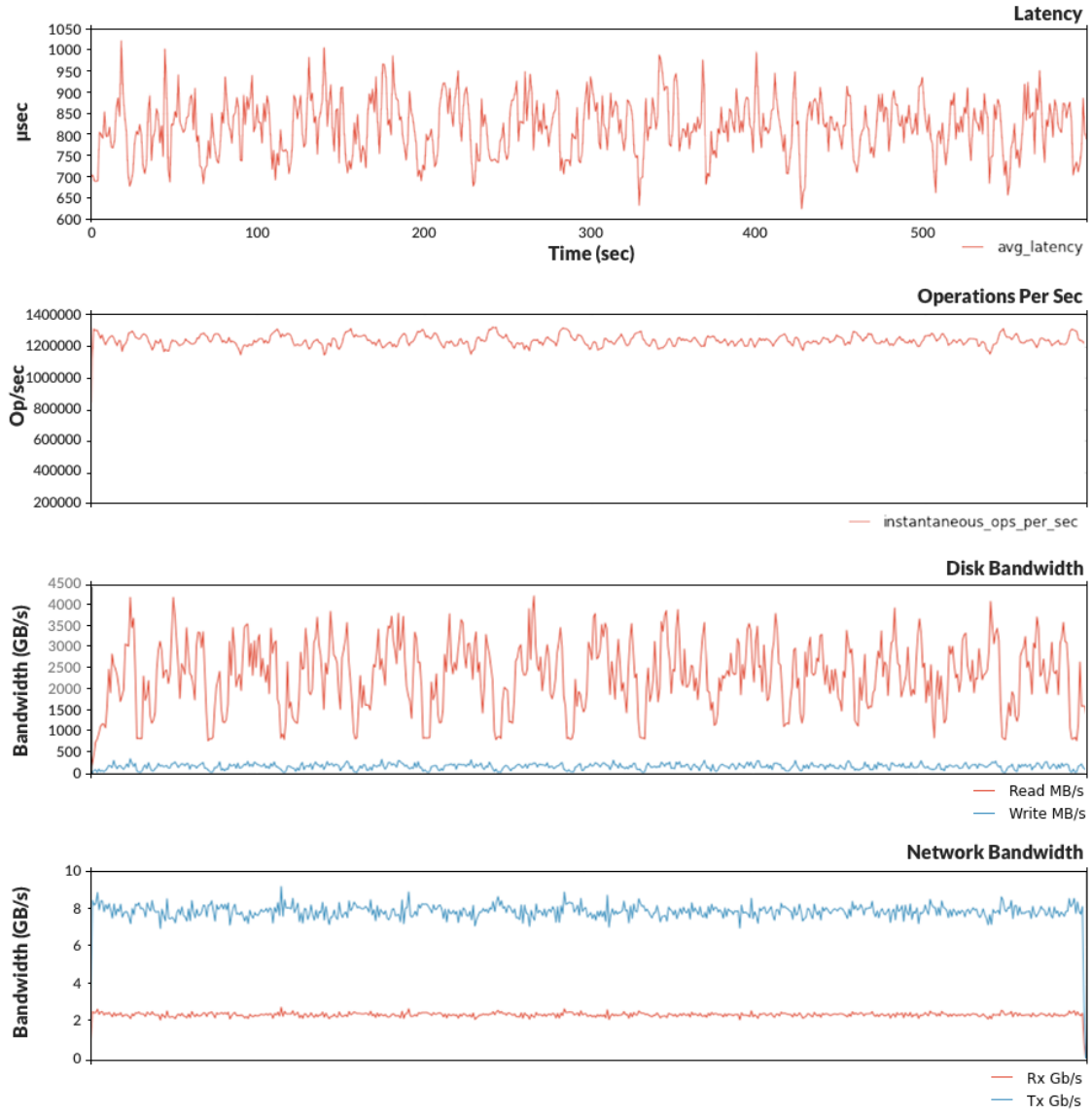


[1]M2G2 : >1 M ops/sec, <1ms latency -->M2, >1 GB disk throughput, >1Gbps network throughput-->G2

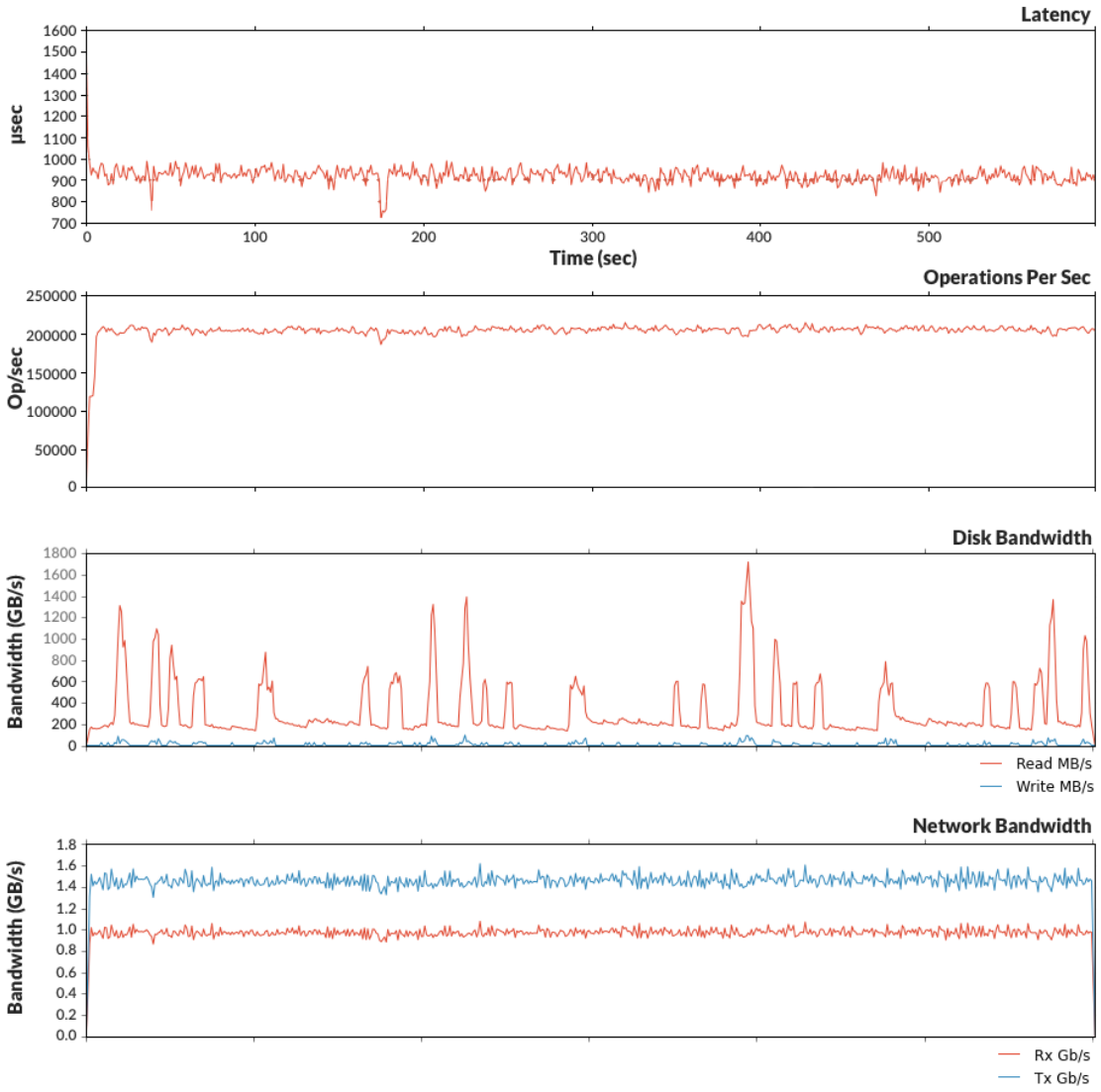
Results

1) For object sizes of 1 KB, with a combination of 80% reads and 20% writes, we observed an average of 1.24 M ops/second, with an average of 0.82 ms latency from a single server.

Also notable was that 99.99% of requests were served with <1ms latency. Since we met >1M ops/sec, at <1 ms, >1Gbps network bandwidth and >1 GB throughput, we met the M2G2 goal!



2) For object sizes of 100 B, with a combination of 80% reads and 20% writes, we observed an average of **2.04 M ops/second**, with an average of 0.91 ms latency from a single server. 100% of requests were served with < 1ms latencies.



Why is this a big deal? It represents a 40X increase in throughput compared to regular SATA-based SSDs. 2M ops/second, with <1ms latencies and >1 GB disk throughput on a single server is a cause for applause any day.

This benchmark was accomplished with a 20:80 ratio of RAM: NVMe. While throughput is 40X higher, the cost of NVMe-based SSDs is likely to be only incremental compared to SATA-based SSDs.

Conclusion

Samsung NVMe SSD technology delivers an enormous improvement in price-to-performance compared to SATA-based SSDs. Redis Labs Enterprise Cluster (RLEC Flash) takes advantage of this improvement to deliver over 2M ops/second, over 1 GB disk throughput and over 1Gbps network throughput over a single server—all while keeping latency under 1msec. This sets a new industry record for the highest performance with the least amount of hardware.

The combination of these two technologies will herald a new age for big data analysis by making it both extremely fast as well as cost-effective to analyze very large datas