

Optimizing Data Center Systems and Applications for Samsung SSDs

Enabling Data Center Applications to Leverage Samsung NVMe SSDs

Samsung Memory Solutions Lab Brochure

SAMSUNG

Optimizing Data Center Systems and Applications for Samsung SSDs

Summary

This technology overview provides a comprehensive look at the benefits that data center applications can realize when they deploy Samsung SSDs. The detailed technical papers referenced here demonstrate how to tune and configure high speed, data-rich applications and storage stacks to achieve the benefits enumerated below.

Introduction

Current data center applications present significant challenges for storage sub-systems due to rapid growth in the velocity, volume and variety of data requiring storing and processing within these applications. The latest-generation Samsung NVMe SSDs provide unprecedented throughput and latency improvements over SATA and SAS solid state storage, while simultaneously scaling up to 10's of terabytes per-device. However, applications only realize a fraction of the potential improvement when older SSD and HDDs are simply swapped for newer NVMe SSDs. By conducting extensive end-to-end performance analysis, we have identified and tuned these applications to best leverage the performance of the latest SSDs. In some cases, tuning implies firmware, operating system alterations, and application configuration changes. In other cases, it requires application-level code modifications.

NVMe over Fabrics (NVMe-OF) is a complementary technology that has matured this year. This technology enables network-attached NVMe SSDs to deliver throughput and latency that is almost as good as direct-attached NVMe SSDs. NVMe-OF enables disaggregated storage systems to deliver Samsung SSD performance in a wide variety of critical applications over most networks. Disaggregated storage enables data centers to pool, scale and manage storage much more efficiently.

The following section summarizes our recent efforts to optimize and improve the performance of some data center applications for direct attached and NVMe-OF scenarios.

Samsung SSDs for Data Center Applications

Samsung Enterprise SSDs with V-NAND and Z-NAND Memory

Applications today need advance storage devices to process large and growing amounts of data efficiently. Different application require different types of storage. Samsung has successfully established itself as an innovator in flash-based storage across the enterprise markets with cutting-edge flash memory leveraging V-NAND™ and more recently Z-NAND™ technology. Samsung's comprehensive SSD portfolio helps to meet critical speed requirements with confidence, eradicating performance and efficiency bottlenecks.



Samsung PM1725a

The Samsung PM1725a NVMe SSD, based on V-NAND technology, is optimized to excel in virtually any data center scenario. This enterprise-level, ultra-high performance SSD provides high random read performance and is particularly suitable for read-intensive data center applications. The PM1725a comes in capacity points ranging from 800GB to 6.4TB.



Samsung PM963

Samsung's PM963 NVMe SSD also is based on V-NAND technology. It provides exceptional value and is more cost effective than SATA SSDs. With approximately four times the IOPS per GB, data center operators can continue to scale workload utilizations on their servers, decreasing the cost of computing. Samsung PM963 comes in the M.2 form factor and is available in capacity points ranging from 960GB to 3.8TB.



Samsung Z-SSD™

Based on a new and highly innovative Samsung Z-NAND technology, this new class of enterprise storage SSD's achieves peak performance at low queue depths, giving better performance across all workloads. It has a unique circuit design and a controller that can maximize performance, with 5.5 times less latency than the Samsung PM1725a NVMe SSD. The Samsung Z-NAND SSD achieves peak performance at low queue depths, delivering better performance across all workloads. Customers needing a boost in high-performance data intensive, real-time analytics will greatly benefit from this disruptive, new technology.

Samsung SSDs for Data Center Applications (continued)

Performance Specifications	SZ985	PM963	PM1725A
Data Transfer Rate (128KB data size, QD=32)			
Sequential Read/Write (GB/s)	3.2 / 3.2	2.0 / 1.2	6.4 / 3.0
Data I/O Speed (4KB data size, Sustained, QD=32)			
Random Read/Write (K IOPS)	750 / 170	430 / 40	1080 / 170
Latency (Sustained random workload, QD=1)			
Read (Typical)	20us	85us	90us
Read (Best)	12us	N/A	
Write (Typical)	16 us	50us	20us
Form Factor	HHHL	2.5"/ M.2	HHHL / 2.5"
NAND technology	Z-NAND™	V-NAND™	
Capacity	800GB	960GB to 3.84TB	800GB to 6.4TB

To help simplify integration of these storage devices, Samsung Multi-Stream technology is an additional feature that ensures optimal performance and better endurance for an entire class of applications. In addition Samsung produces a wide range of SAS and SATA drives to match the need of the application.

For more information:

Samsung Enterprise SSD:

<http://www.samsung.com/semiconductor/products/flash-storage/enterprise-ssd/>

Multi-Stream Technology:

http://www.samsung.com/semiconductor/global/file/insight/2016/08/Multi_stream_technology-0.pdf

Applications

Samsung Enterprise SSDs with V-NAND and Z-NAND Memory

RocksDB

Key-Value stores, like RocksDB, are the basis of many data center storage system applications. The publication referenced below shows that using Samsung NVMe SSDs provides a 3x to 4x direct benefit over Samsung's previous generation data center-class SATA SSDs. Samsung Multi-Stream SSDs can also provide additional performance improvements. These benefits can be achieved using PM963 class devices. The PM1725a, and SZ985 (Z-SSD) to a greater degree, deliver even higher throughput at lower latencies.

For more information about RocksDB:

http://www.samsung.com/us/labs/pdfs/collateral/Performance-Benefits-of-Running-RocksDB-on-SSDs_Whitepaper.pdf

Cassandra

Cassandra is very a popular NoSQL database in wide-spread use by many companies, notably Netflix. We have previously shown that NVMe SSDs can provide significant performance benefits. In a recent publication referenced below, we demonstrate Samsung Multi-Stream technology's performance benefits. Specifically, using cassandra-stress and a modified Cassandra version that leverages Multi-Stream SSDs will deliver up to a 3X performance improvement. While the publication highlights SAS SSDs, multi-stream technology is also available with PM1725 and PM983 NVMe devices.

For more information about Casandra:

http://www.samsung.com/us/labs/pdfs/collateral/Multi-stream_Cassandra_Whitepaper_Final.pdf

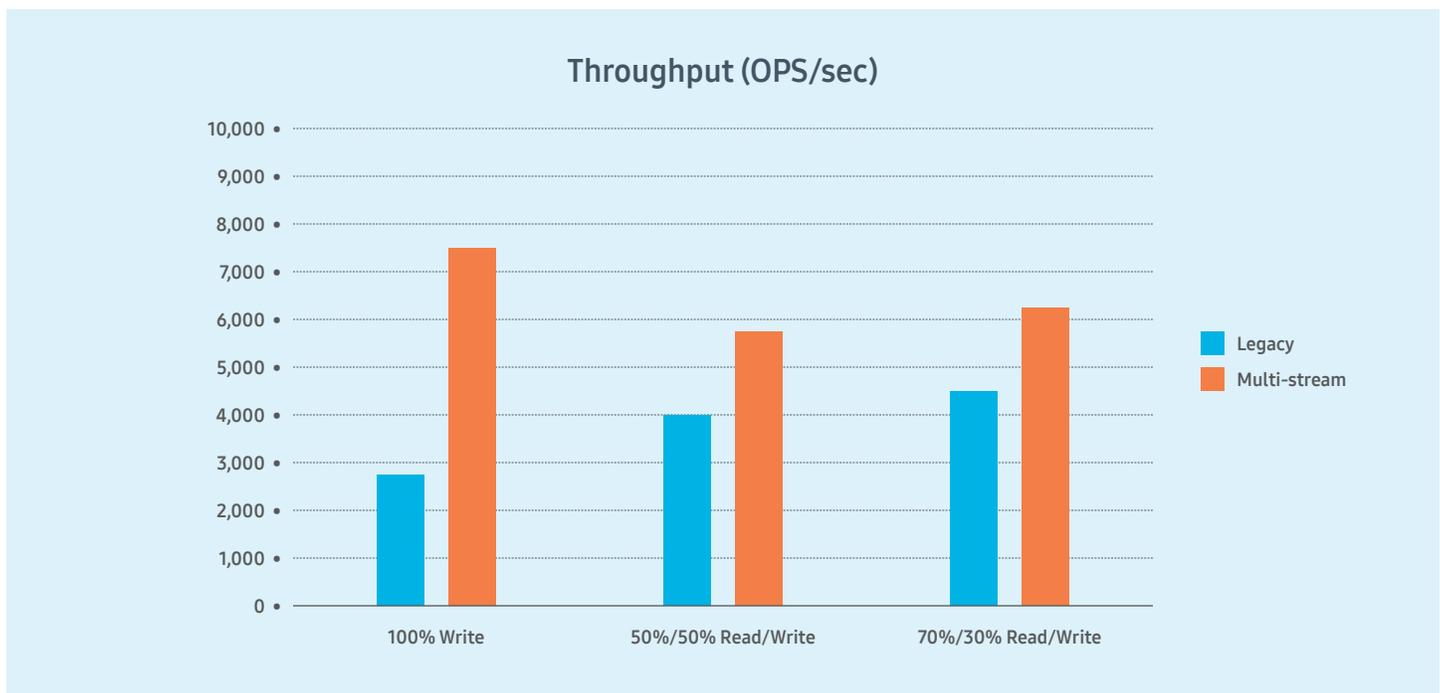


Figure 1. Samsung Multi-Stream SSD Performance with Cassandra

Applications (continued)

ScyllaDB

ScyllaDB is an open-source, high performance NoSQL data store, and can be considered a drop-in Cassandra replacement. However, ScyllaDB exhibits improved performance due to a finely crafted C++ implementation that manages low-level resources including memory, thread queues and the network stack. ScyllaDB provides a seamless migration for Cassandra applications that requires higher performance and can nicely leverage high-performance Samsung NVMe SSDs. The referenced publication outlines PM19725-class SSD performance benefits.

For more information about ScyllaDB:

http://www.samsung.com/us/labs/pdfs/collateral/ScyllaDB_Report_v41.pdf

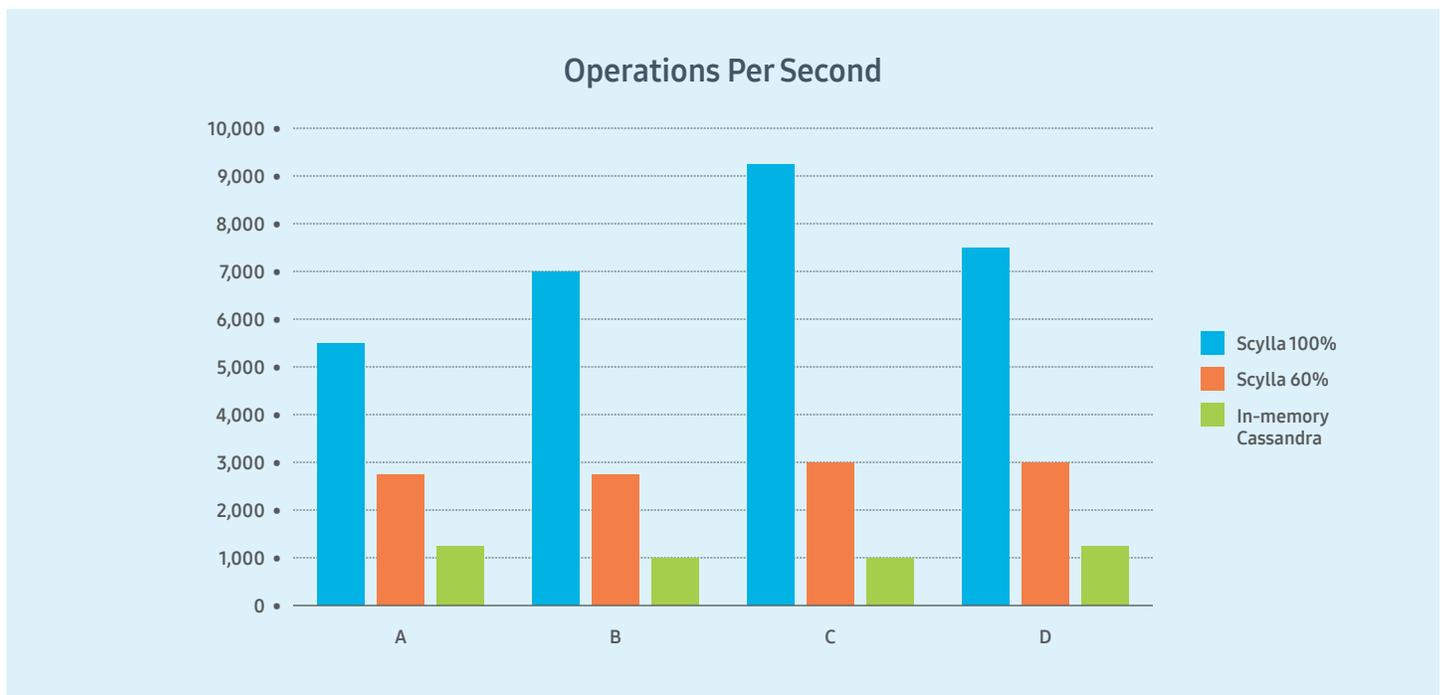


Figure 2. ScyllaDB with 2TB Data Compared to Cassandra

Applications (continued)

Redis-on-Flash

Redis-on-Flash is a RedisLabs product that leverages NVMe SSD capacity and performance to enable customers to build applications with memory requirements that exceed DRAM capacities. Samsung PM1725-class SSDs can deliver 2 million+ ops/sec at sub-millisecond latencies for today's demanding applications.

For more information about Redis-on-Flash:

<https://www.slideshare.net/RedisLabs/redis-on-nvme-ssd-zvika-guz>

<https://redislabs.com/docs/redis-on-flash-with-samsung-nvme-benchmark/>



Figure 3. RedisOnFlash Throughput with PM1725 SSD

Applications (continued)

MySQL

Using TPC-C as the workload, the referenced papers show how to optimize MySQL and Linux to use Samsung PM1725 NVMe SSDs efficiently. It shows that, when selecting SSDs, careful analysis and tuning is required to leverage all of the advantages of Samsung NAND technology. We have shown that using PM1725-class of enterprise NVMe SSDs delivers more than 180X improvement over traditional HDDs, using Percona MySQL Server.

For more information about MySQL:

<http://www.samsung.com/us/labs/pdfs/collateral/samsung-tpcc-mysql-whitepaper-final-1.pdf>

https://www.snia.org/sites/default/files/SDC/2016/presentations/solid_state_storage/VeronicaLaGrange_Accelerating_OLTP_Performance_V6.pdf

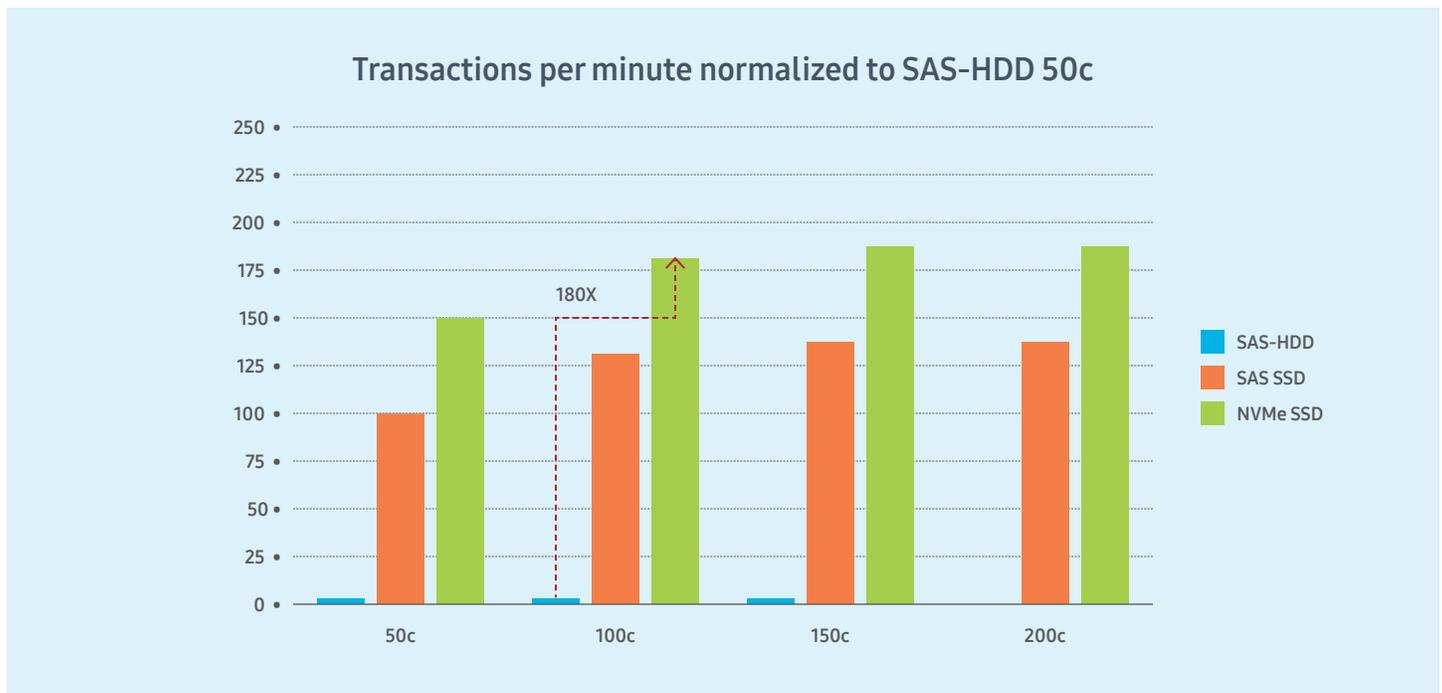


Figure 4. Percona MySQL with TPC-C (Percona Implementation)

Applications (continued)

Kafka

Kafka is a lightweight high-throughput distributed log system designed to persistently store and forward large amounts of data in the form of records, between data sources and data processing units. Its high performance and low overhead have enabled many online businesses from social media to shopping and music streaming to scale to the demands of increased usage. With respect to storage, Samsung NVMe SSDs offer high performance in the throughput range of GBs per second, while also providing extremely low latency. We can show that PM963 eliminates bottlenecks introduced by SATA devices and delivers a cost-effective solution for Kafka nodes.

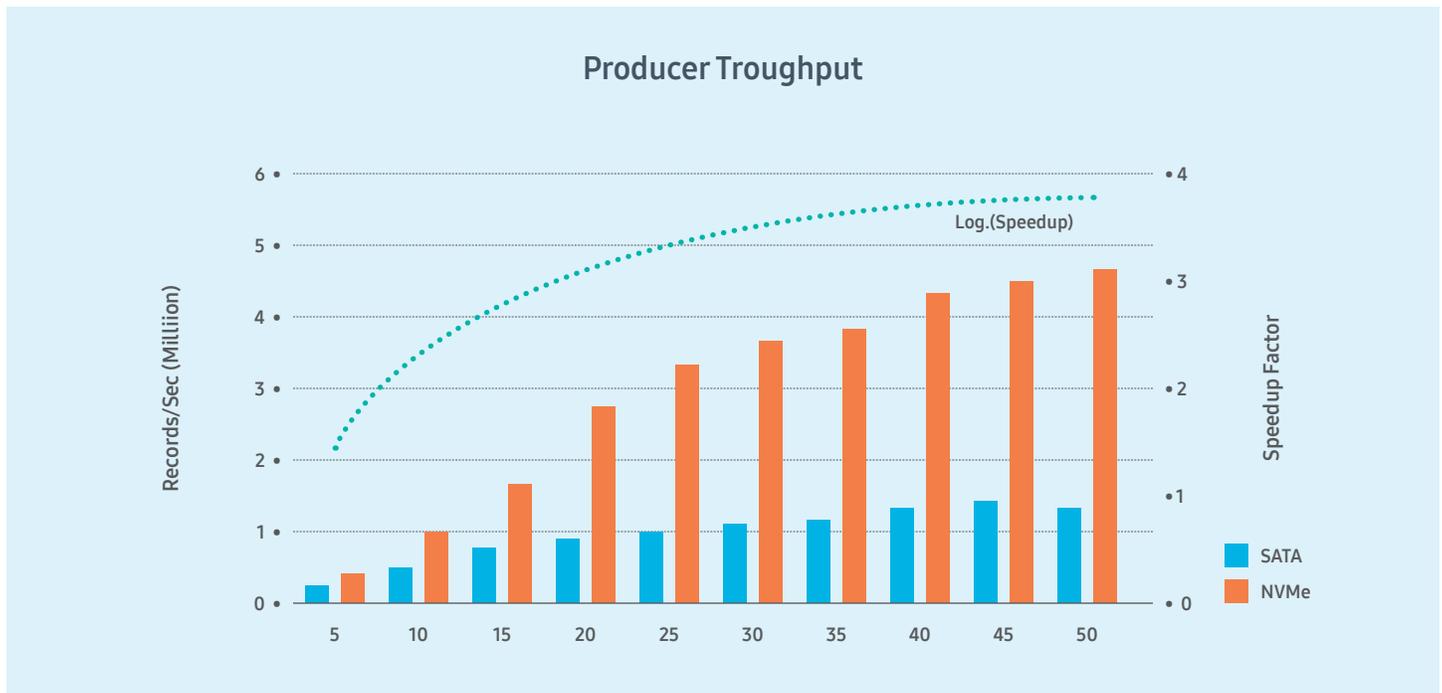


Figure 5. Kafka Producer Throughput with PM963 NVMe SSD

Applications (continued)

Spark

The Transaction Processing Performance Council (TPC) has recently released its first big-data benchmark called TPC-xHS V2. We have recently completed an internal study of this workload that indicates that the Samsung PM963 data center-class device delivers a 45% improvement over current state-of-the-art SATA SSDs (e.g. the Samsung SM863) for a scale factor of 1TB using four data nodes with Broadwell-class Intel CPUs. Specifically, the maximum read and write throughput of the PM963 NVMe SSD is higher than that of the SM863 SATA SSD by 84% and 78% respectively. Samsung's data center-class NVMe SSDs (PM963) are well-suited for Hadoop clusters, delivering superior performance when compared to any SATA device.

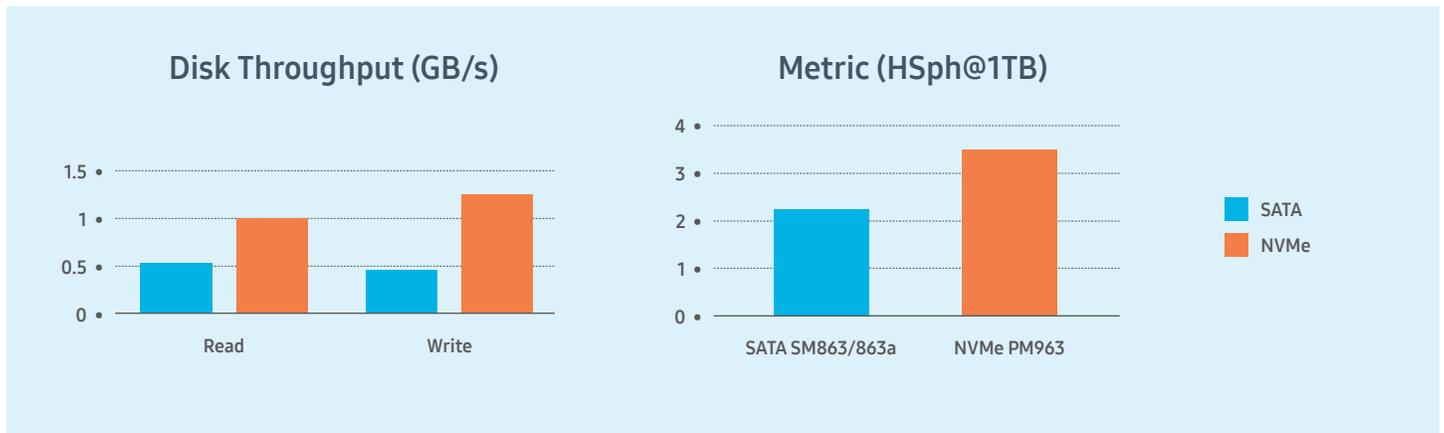


Figure 6. TPCx-HS Disk Bandwidth and HSph@1TB

Applications (continued)

Alluxio

Alluxio, formerly Tachyon, is a memory-speed, virtual distributed storage system. It allows application compute frameworks to access data stored in disparate storage systems at memory-speed via a single interface. That interface is able to seamlessly access data residing in Alluxio and any underlying file system (underFS) that it supports, such as HDFS or Amazon S3. The referenced publication demonstrates the effectiveness of high-performance NVMe SSDs; often, a few NVMe SSDs can be cost-effectively used for performance very close to that of DRAM. This enables Alluxio-specific storage to scale to 10s of terabytes - multiples of NVMe SSD capacities. In the referenced publication, we describe a change made to Alluxio that enables applications to directly use the NVMe SSD at the storage target.

For more information about Alluxio:

<http://www.samsung.com/us/labs/pdfs/collateral/Alluxio-plus-NVMe-WP-v6.pdf>

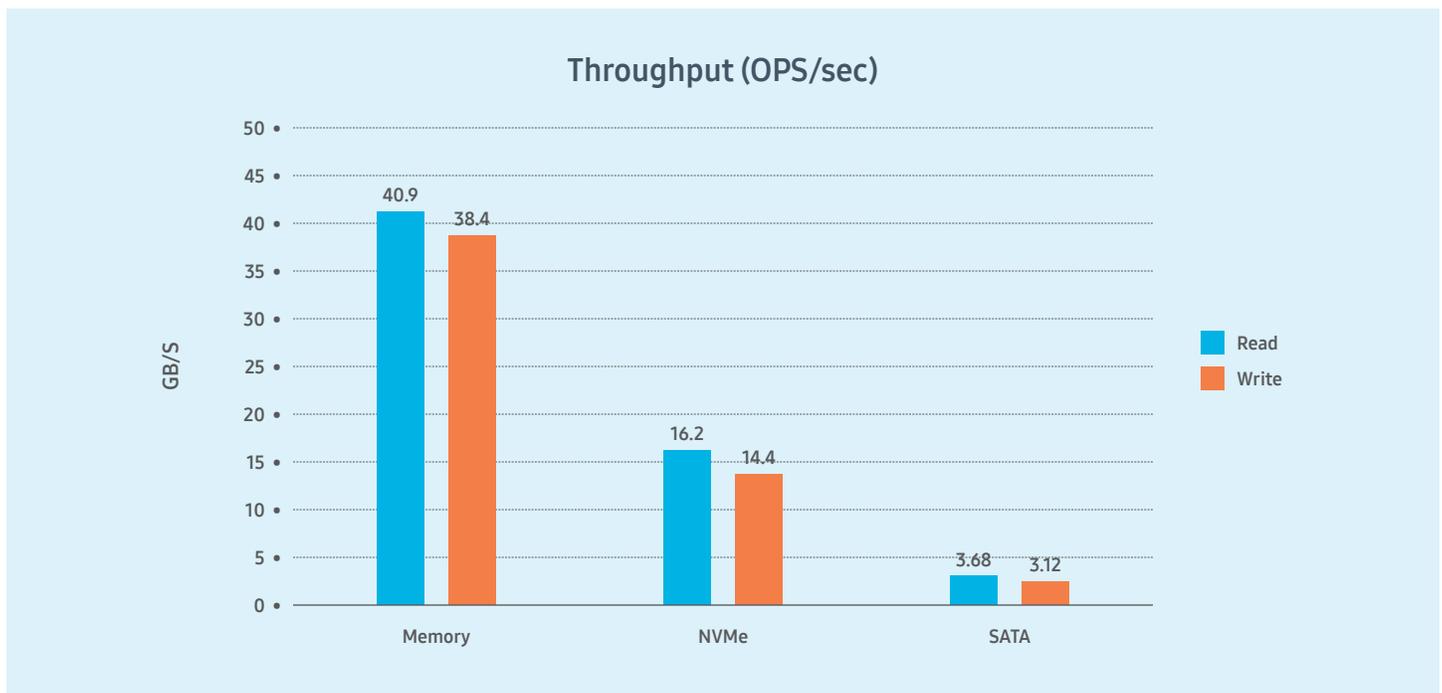


Figure 7. Alluxio Throughput with NVMe SSDs

Applications (continued)

VoltDB

The VoltDB has advanced the role of an OLTP database into a real-time DSS (Decision Support System). Using In-Memory architecture, it has built a database that provides extremely high throughput for today's real-time applications such as fraud and anomaly detection. However, many application require high durability that volatile DRAM cannot deliver. VoltDB uses NVMe SSDs to automatically backup and restore databases during failures while continuing to deliver much higher throughput compared to traditional databases. The referenced paper shows that NVMe SSDs are ideal candidates if applications require a high degree of durability in an in-memory database. The reference publication shows that the Samsung PM1725 delivers 80% more throughput, with a 40% reduction in latency, for TPC-C like workloads.

For more information about VoltDB:

<http://www.samsung.com/us/labs/pdfs/collateral/A-White-Paper-by-Samsung-Memory-Solutions-Lab-and-VoltDB.pdf>

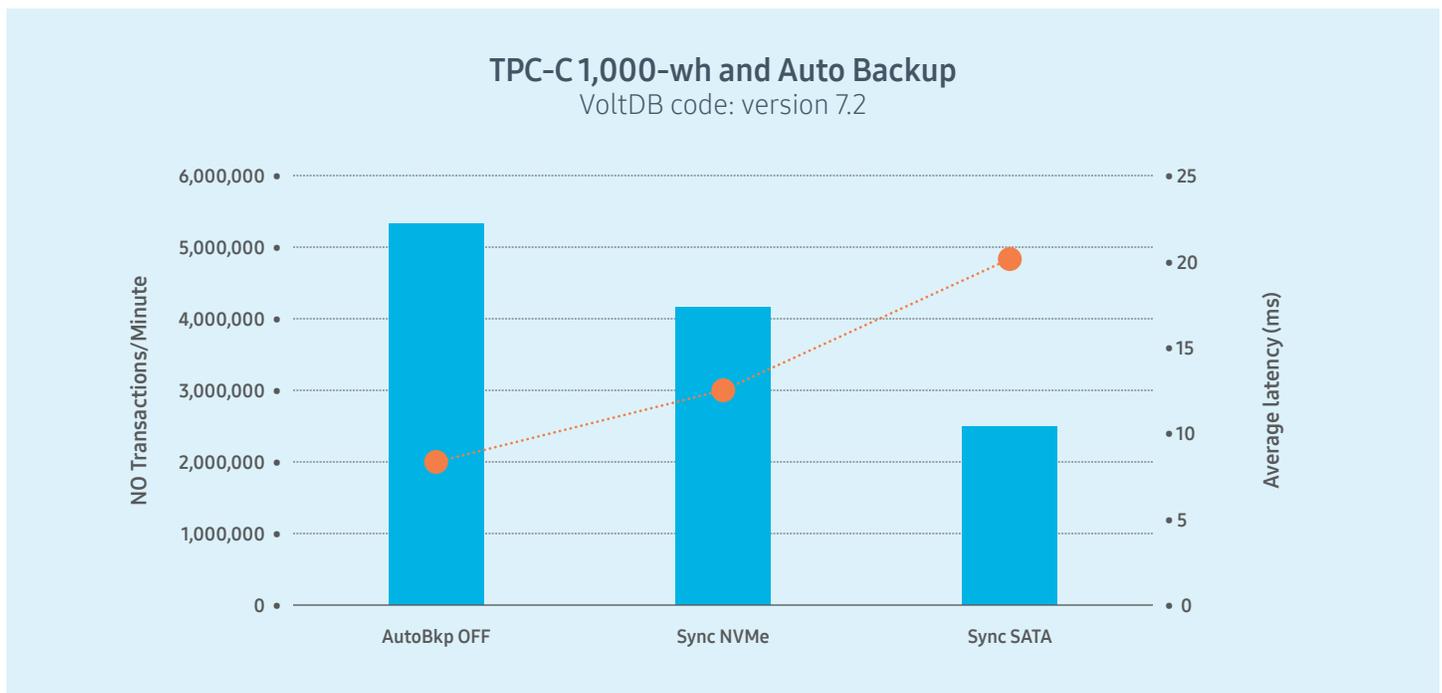


Figure 8. Throughput and Latency Gains with NVMe

Applications (continued)

NVMe-Over-Fabrics Use Cases

Data center applications with high-performance storage requirements continue to increase. Consequently, NVMe SSDs are increasingly deployed to cater to the ever-growing demand for throughput and low-latency access. As NVMe SSDs become more prevalent in the data center, Flash Disaggregation has been proposed to improve resource efficiency and reduce the overall cost of NVMe deployment. NVMe-over-Fabrics (NVMe-OF) is a recent NVMe standard extension that enables remote NVMe device access over different network fabrics. It eliminates unnecessary protocol translations (such as SCSI) along the I/O path to the remote device, exposing the NVMe multiple-paired queue design.

Using FIO and RocksDB, the referenced article below demonstrates that NVMe-OF adds very little overhead compared to direct attached SSDs.

For more information about NVME-Over-Fabrics:

<http://www.samsung.com/us/labs/pdfs/nvmf-disaggregation-preprint.pdf>

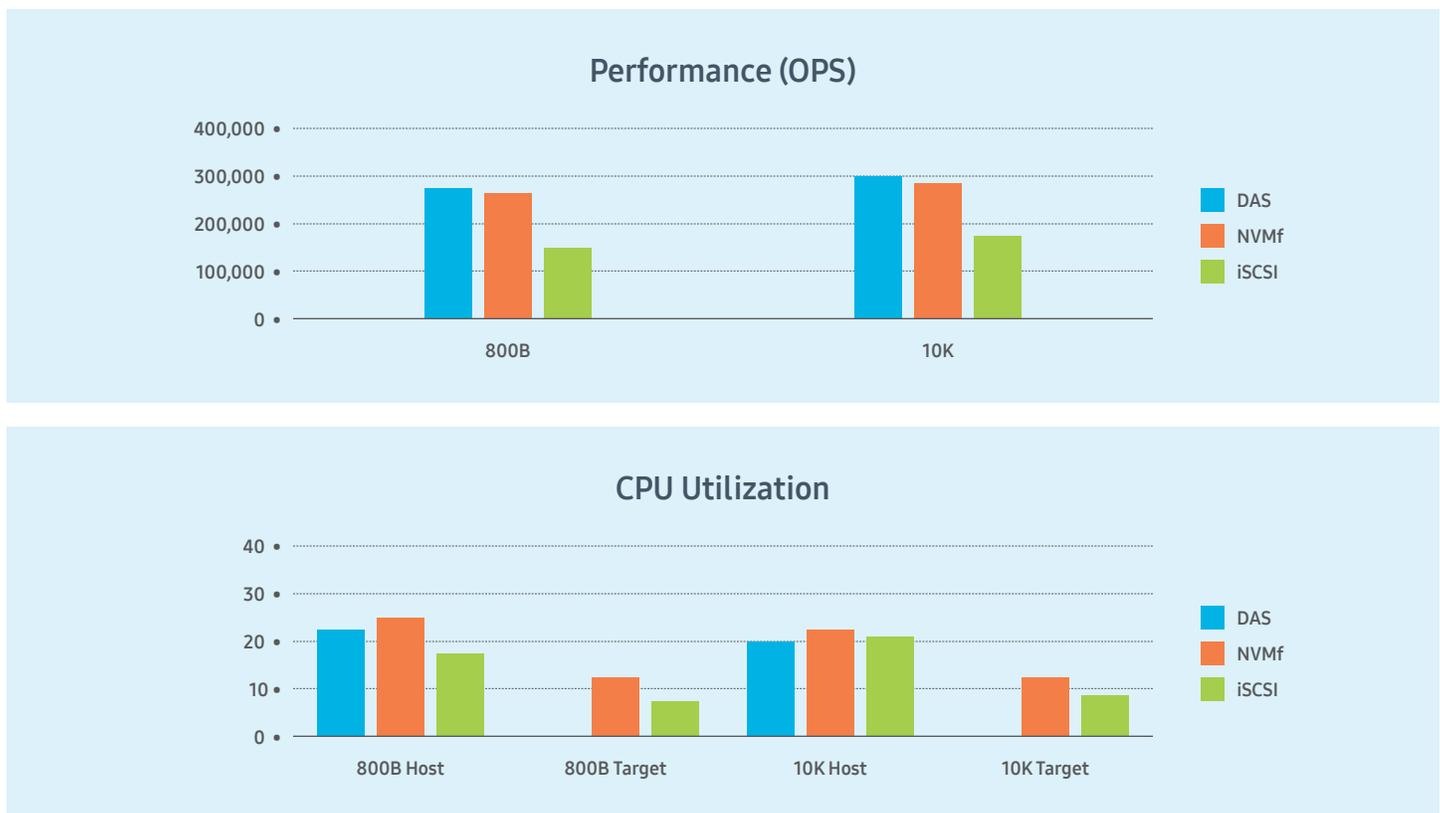


Figure 9. Throughput and CPU Utilization of RocksDB Benchmark

Applications (continued)

We recently completed testing MySQL database workloads using NVMe-OF storage devices. The graph below shows that NVMe-OF can enable high performance NVMe SSDs to be shared across a network. This allows better NVMe SSD performance utilization, while simultaneously enabling disaggregation that delivers additional data center management benefits.

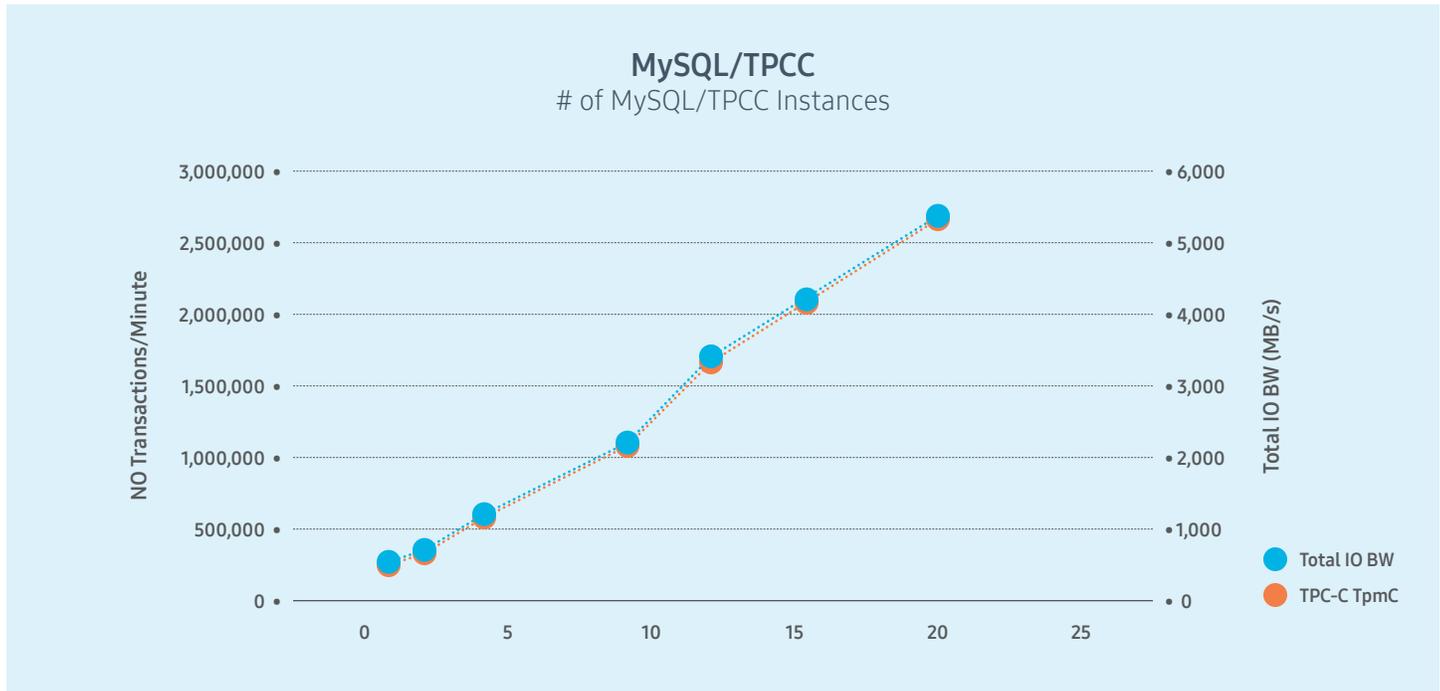


Figure 10. Linear Scaling of MySQL w/ TPC-C using a NVMe-OF Target using NVMe SSD's

Conclusion

The challenge of selecting the most appropriate storage device and storage architecture can be quite vexing. The wide variety of applications and a continually evolving software ecosystem requires detailed analysis and tuning. Simply swapping NVMe SSDs for legacy storage devices is not enough to truly realize the full benefits of the new generation of SSDs with their dramatic latency and throughput improvements. With end-to-end performance engineering matched to the most advantageous NAND-based storage devices, we have shown that there is a Samsung SSD for almost any data center application.

About Samsung Electronics Co., Ltd.

Samsung Electronics Co., Ltd. inspires the world and shapes the future with transformative ideas and technologies. The company is redefining the worlds of TVs, smartphones, wearable devices, tablets, cameras, digital appliances, printers, medical equipment, network systems, and semiconductor and LED solutions.

For the latest news, please visit the Samsung Newsroom at news.samsung.com.

Copyright 2017 Samsung Electronics, Co. Ltd. and VoltDB, Inc. All Rights Reserved. All brand, product, service names and logos are trademarks and/or registered trademarks of their respective owners and are hereby recognized and acknowledged. Specifications and designs are subject to change without notice. All data were deemed correct at time of creation. Samsung and VoltDB are not liable for errors or omissions. PID 2017-07