# StorageReview

SK hynix Develops PCIe Gen4 Enterprise SSD Portfolio

SK hynix is not only the world's third-largest semiconductor company and the second-largest memory chipmaker, they also use their technological acumen to make a well-regarded line of SATA and NVMe flash storage devices. In this article, we will provide an overview of the PE8010, SK hynix's latest PCIe Gen4 Non-Volatile Memory express (NVMe) drive, and highlight its impressive performance. We will first provide some background and explain why drives that use the PCIe Gen4 interface are such a game-changer for the datacenter.
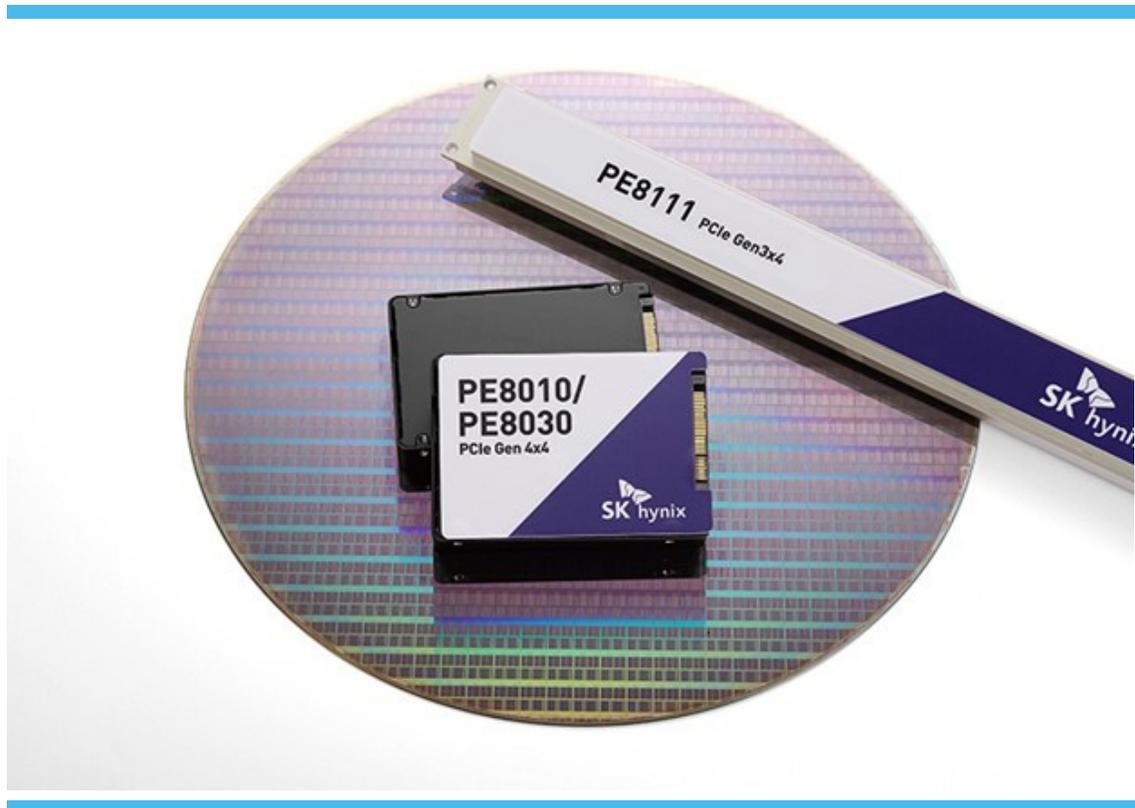
In June of 2019, SK hynix announced the launch of the PE6011, their latest NVMe enterprise SSD which we reviewed. The PE6011 is a low-powered NVMe device with 72-layer TLC 3D NAND flash and an SK hynix NVMe controller that was designed for the enterprise market. While not the top performer of comparable drives, the PE6011 gave good numbers all around and we found that it struck a nice balance between power consumption, performance, and cost.

SK hynix PE8000 NVMe Drives

In April of 2020, SK hynix announced the availability of its PE8000 series of NVMe SSD drives. Each has a maximum capacity of 8TB and uses the company's 96-layer triple-level cell (TLC) 4D NAND flash, as well as a controller that was developed in-house. There are currently two drives in this series: the PE8010, which was designed for read-intensive workloads, and the PE8030, which was designed for mixed-use workloads. The most interesting aspect of these two drives is that they both have a PCIe Gen4 interface and are compatible with the U.2 and newer U.3 form factors.

Innovation is at the core of SK hynix, they've long been an engineering-focused organization. They have plans to continually meet changing demands from the industry by embracing new NAND technologies. On the roadmap are several products, like the PE8110 which is an SSD family that will leverage 128-layer, 512Gb NAND. Additionally, SK hynix is working on M.2 and E1.S (EDSFF Short) drives with this NAND technology as well.

In addition to new NAND, there are new form factors to explore. SK hynix will soon be offering the PE8111, which uses the emerging E1.L (EDSFF Long) form factor, commonly referred to as "ruler." This form factor is becoming increasingly popular with hyperscalers and is starting to creep into mainstream server designs as well.

The PE8111 uses 128-layer 1Tb TLC NAND, a first in the industry. SK hynix has just recently begun sampling the drive to customers, but this highlights the increasingly flexible breadth of their flash offerings. A 16TB capacity version of the drive is the current focus, but we're told that a 32TB capacity drive is on the roadmap as well.


## SK hynix PE8010 Specifications (PCIe Gen4x4)

Form Factor	U.2/U.3 2.5" (15mm)
Interface	PCIe Gen4x4 / NVMe 1.3c
NAND	4D V5 TLC

Capacity	960 GB	1.92 TB	3.84 TB	7.68 TB
Performance				
Sequential Read (128KB)	Up to 6,000MB/s	Up to 6,500MB/s	Up to 6,500MB/s	Up to 6,500MB/s
Sequential Write (128KB)	Up to 1,450MB/S	Up to 2,800MB/s	Up to 3,700MB/s	Up to 3,700MB/s
Random Read (4KB, QD64)	Up to 500K IOPS	Up to 850K IOPS	Up to 1,100K IOPS	Up to 1,100K IOPS
Random Write (4KB, QD64)	Up to 65K IOPS	Up to 120K IOPS	Up to 140K IOPS	Up to 145K IOPS
QOS 4KB Random Read	70µs	70µs	70µs	70µs
QOS 4KB Random Write	20µs	20µs	20µs	20µs
Security	AES 256-bit encryption			
Endurance				
MTBF	2 million hours			
UBER	1 sector per 10^{17} bits read			
Power consumption				
Active Read/Write Typical	Up to 17.0W			
Idle	Up to 5.0W			
Allowable voltage	12.0V ± 8%			
Operating Temperature	0-70°C			
Shock	1500G, duration 0.5ms			
Physical				
Dimensions (WxDxH)	69.85 x 100 x 15 mm			
Weight	135.6g (± 5%)			

TLC is NAND flash memory that stores three bits of information per cell, rather than one bit of memory or two bits in single-level cell (SLC) and multi-level cell (MLC) flash memory. SK hynix 4D technology is basically Periphery Under Cell (PUC) technology with conventional 3D NAND technology stacked on top of it. By using 96-layer 512Gb (Gigabit) TLC 4D NAND flash, the PE8010 is priced competitively with the last generation of smaller capacity drives.
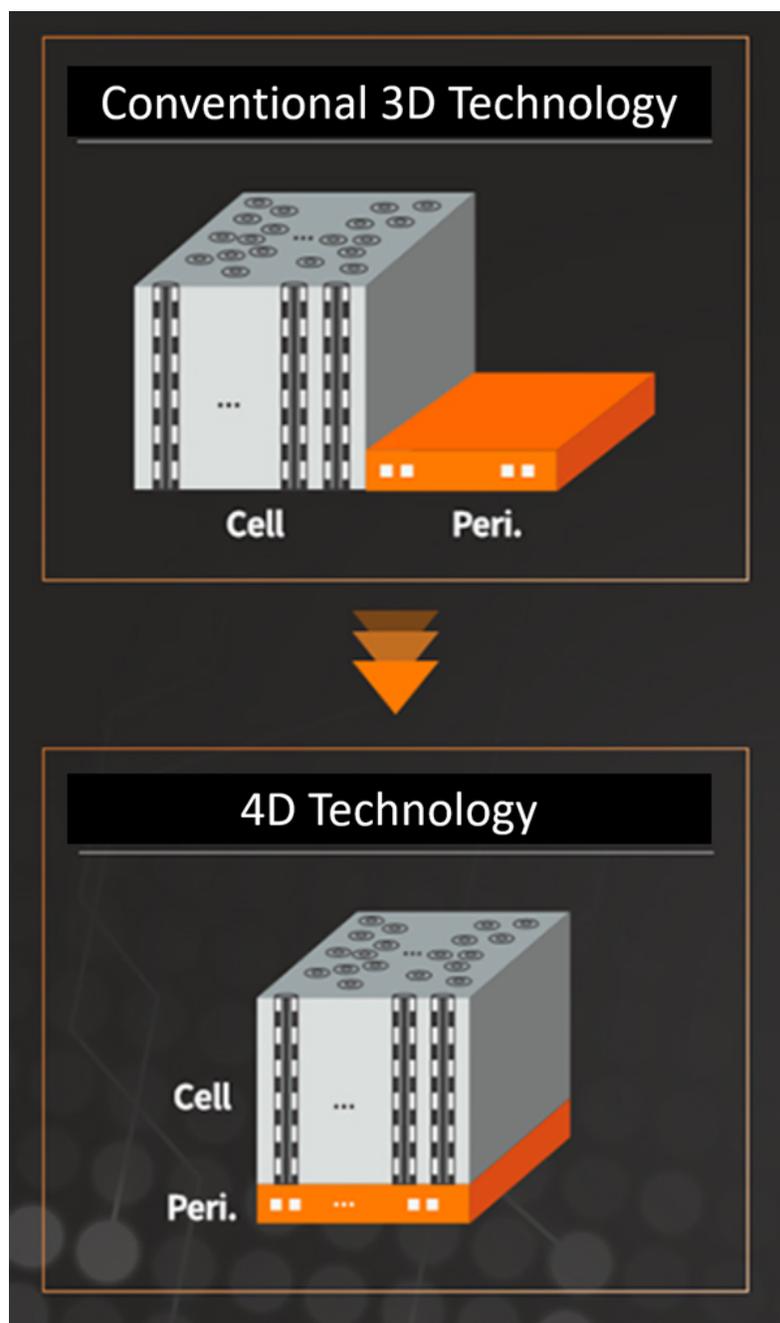
The other interesting feature of this drive is that it supports both the older U.2 and newer U.3 form factors. The U.2 (FKA SFF-8639) is a computer interface for connecting SAS and SATA drives to a computer and can use up to four PCIe lanes. The U.3 is built on top of the U.2 spec and uses a U.2 connector; however, it supports NVMe, SAS, and SATA drives, as well as hot-swapping between these three different drive types if the firmware supports it. U.3 drives are backward compatible with U.2, but U.2 drives are not compatible with pure U.3 hosts.

Image © OpenCompute
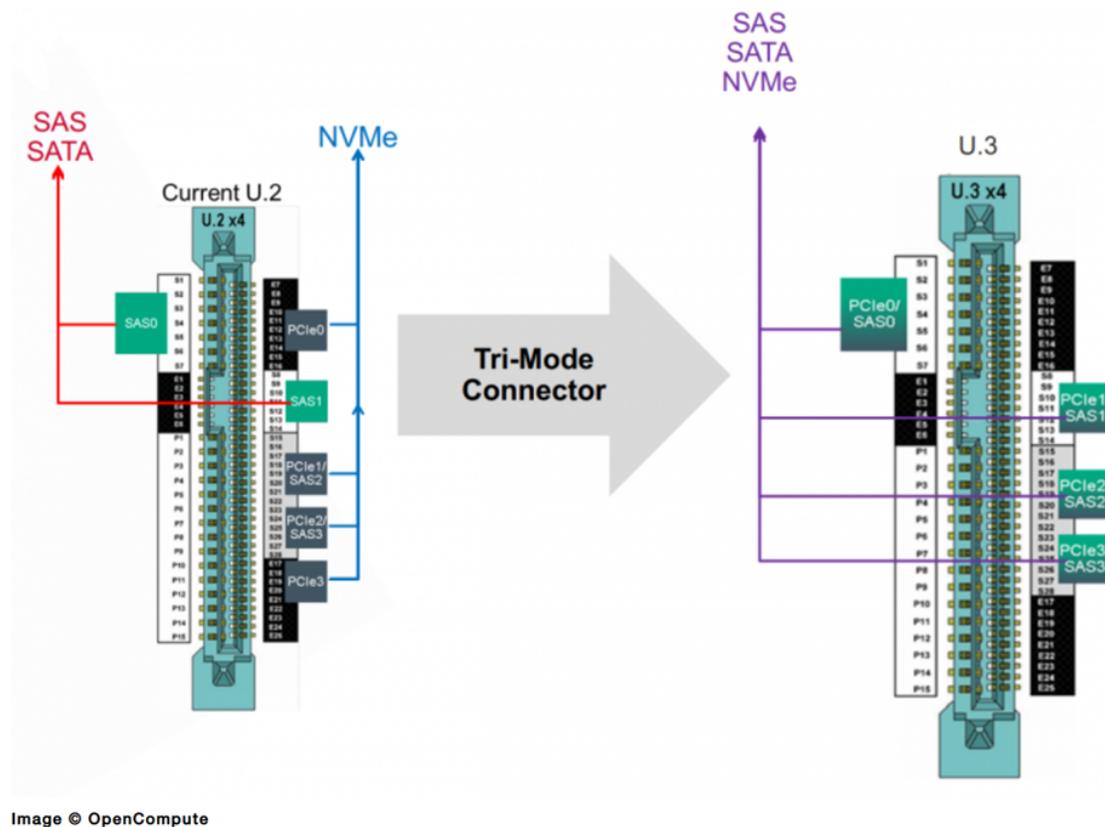
However, the biggest feature and advantage of the PE8010 drive is its support of the PCIe Gen4 bus standard. The Gen4 bus doubles the bandwidth of the PCIe Gen3 bus while maintaining backward and forward compatibility in both software support and mechanical interface with the Gen3 bus. This allows the Gen4 device to have up to 16 GT/s (gigatransfers/second) vs. the 8 GT/s for Gen3 devices. This equates to data transfer rates of up to 2GB/s (gigabytes/second) vs. 1GB/s per lane.

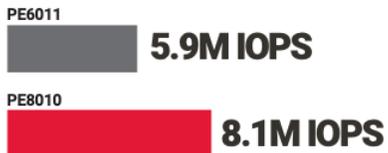| Bandwidth per PCIe Lane				
Number of Lanes	x1	x4	x8	x16
PCIe 3.0	1 GB/s	4 GB/s	8 GB/s	16 GB/s
PCIe 4.0	2 GB/s	8 GB/s	16 GB/s	32 GB/s

As PCIe Gen4 is the latest bus technology, the current servers that support it are still working their way through the pipeline. Advanced Micro Devices (AMD) has been more aggressive in their support, and we are expecting to see Intel systems with PCIe Gen4 support soon.

For our testing, we used a single proc Lenovo SR635 with an AMD 7742 CPU. We did a write-up and video on the Lenovo SR635 (located here and here) earlier this year, but here is a brief recap. The SR635 is a 1U, single-socket server that supports 2nd gen AMD EPYC CPUs and up to 2TB of 3200MHz of DRAM. Moreover, and most importantly for our testing, it supports PCIe 4.0 from end-to-end, depending on the backplane installed.

For our initial testing in CentOS 7 with ten 3.84TB PE6011 versus ten 3.84TB PE8010 NVMe SSDs, we look at traditional "four-corners" throughput and bandwidth tests using 4K random and 64K sequential workloads.
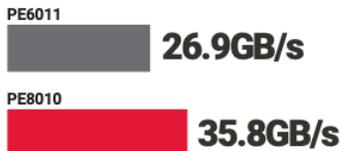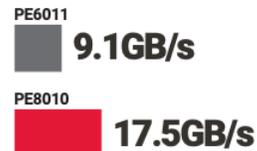
4K RANDOM 100% READ

PE6011
5.9M IOPS

PE8010
8.1M IOPS

4K RANDOM 100% WRITE

PE6011
2.08M IOPS

PE8010
3.97M IOPS

64K SEQUENTIAL 100% READ

PE6011
26.9GB/s

PE8010
35.8GB/s

64K SEQUENTIAL 100% WRITE
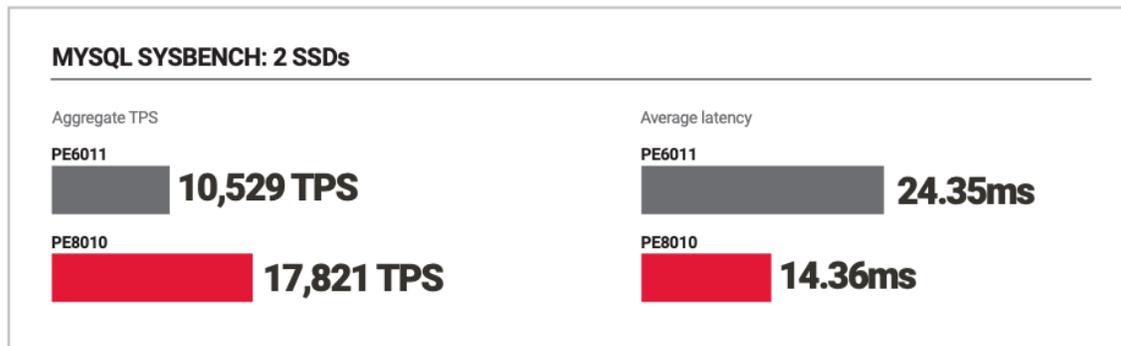
PE6011
9.1GB/s

PE8010
17.5GB/s

Next, we change our software environment to VMware ESXi 6.7u3 and compare the performance of a smaller SSD group size. For SQL Server we use 4 NVMe SSDs each, or one per VM, with our SQL Server 2014 instance running our 1,500 scale TPC-C workload profile.

SQL SERVER: 4 SSDs

Average latency

PE6011
5.25ms

PE8010
2.5ms

Lastly, we concentrate our MySQL Sysbench workload onto 2 SSDs from each group, with 8VMs placed evenly across the pair of SSDs. This gives us a workload footprint of 4VMs per SSD, focusing on the storage impact of the environment.

MYSQL SYSBENCH: 2 SSDs

Aggregate TPS		Average latency	
PE6011	10,529 TPS	PE6011	24.35ms
PE8010	17,821 TPS	PE8010	14.36ms

The PE8010 has the strongest performance of all the NVMe devices that we have tested to date. It can deliver sequential read speeds of up to 35.8GB/s, and sequential write speeds of up to 17.5GB/s, with random reads and writes of up to 8.1M and 3.97M input/output operations per second (IOPS), respectively. Compared to the PE6011, the company's previous PCIe-based Gen3 solution, the PE8010 offers a substantial performance improvement of 133% for sequential reads, and 192% for random writes.

The primary benefits of PCIe Gen4 that are shown throughout the tests are that businesses are able to better utilize system CPU and RAM resources, by keeping the bottleneck off the storage layer.

In application workloads such as SQL Server or MySQL, workload performance saw latency sliced in half. In the CPU-intensive MySQL Sysbench workload, which focused the I/O onto two SSDs, the PE8010 offered a 69% increase in workload performance.

As servers offer processors with more cores or support expanded DRAM capacities, having a storage device that can keep up with the next-generation components is paramount. In our testing we found the PE8010 to be capable of pushing our testbed system to the edge. Modern applications clearly stand to benefit greatly from these latest technologies.

Conclusion

Many times, new technologies fall far short of their performance expectations, but in this case, we were very impressed with the PE8010's actual performance. Given the fact that the PE8010 is expected to carry a small premium compared to PCIe Gen3 devices, we expect it will be a winner for SK hynix. We also expect Gen4 drives to push overall server performance to new levels, and SK hynix is out of the gate early by already shipping Gen4 devices and being a clear leader in this technology.

Undoubtedly, due to its presence with cloud providers where every watt used affects the bottom line, the PE8010 has been optimized for power usage and has best-in-class IOPS/Watt, which results in TCO and OPEX reduction.

We do have some caveats regarding the use of these drives that are worth mentioning. Due to their excellent performance, a server full of PCIe Gen4 devices will overwhelm a CPU; as such, you will need to be smart in the way you design systems. The best way to use these devices until servers are fully designed to make use of Gen4 drives is with an appropriate mix of high-end Gen4 with more traditional NVMe storage or even SATA for the cost-optimized flash tier.

It's important to keep in mind that the PE8010 is price competitive with NVMe Gen3 devices, so if you are planning a server refresh, you may want to consider purchasing them for your existing servers if needed, even if they only have Gen3 support. Because Gen4 drives are backward compatible, they can be used with older systems for now and then moved to a newer system later. Of course, the Gen4 devices will still have the Gen3 bandwidth limitations until newer Gen4 compatible systems are acquired.

With the continued growth of flash in everything from tier-0 workloads to archive-class storage, SSD adoption is showing no signs of slowing in the enterprise. With the new Gen4 products and all of the other SSD development in process, SK hynix is clearly a driving force for an all-flash future.