SEAGATE

NVMe—Performance for the SSD Age

Technology Paper

Non-Volatile Memory Express optimizes system's devices for significantly better performance

Non-Volatile Memory Express (NVMe) is a scalable host controller interface protocol that operates over a PCI Express (PCIe) bus. Interface protocols govern the way host systems interact with storage devices. And systems clearly perform better when the interfaces and their protocols are optimized for the characteristics of the system's devices.

SATA is the most widely used interface for solid state drives (SSD) today. Throughout the consumer world and in much of the enterprise market, SATA SSDs that use the legacy Advanced Host Controller Interface (AHCI) protocol still reign. But despite their ubiquity in the SSD space, SATA and the AHCI protocol were originally designed with the latency and throughput characteristics of HDDs. As a result, they aren't ideal for taking full advantage of an SSD's strengths. NVMe, in contrast, is designed as an optimal interface protocol for NVM technologies like flash, so NVMe-based SSDs are better able to live up to the technology's potential.

NVMe—Performance for the SSD Age

High Sequential Throughput and Random IOPS

PCIe Bus

With NVMe, SSDs connect to host systems via the PCIe bus. This connection still requires an SSD controller chip but it no longer needs the SATA/SAS controller (Figure 1), a move that can double—and sometimes triple—sequential throughput rates, owing to the elimination of several processing latencies and the constriction of the SATA/SAS pipeline.

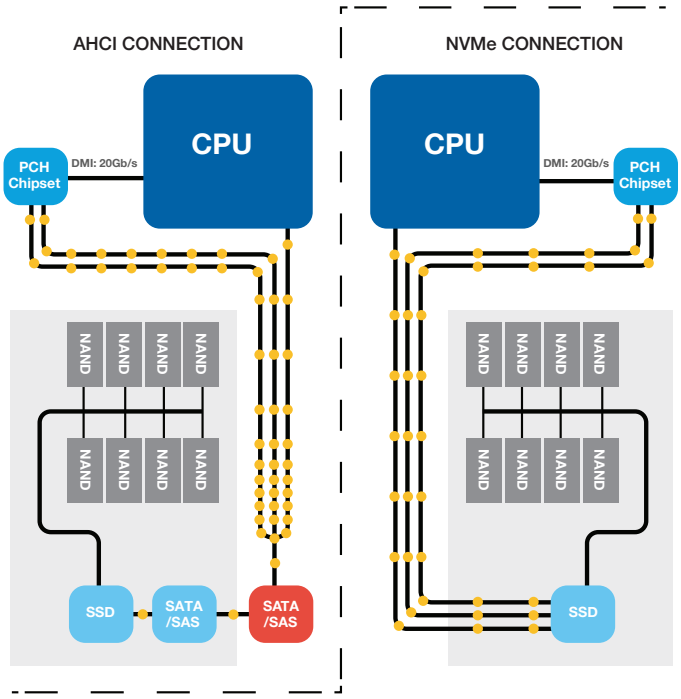

Figure 1. AHCI vs. NVMe paths

Throughput also flows in parallel along each available PCIe lane. In a four-lane PCIe 3.0 connection, for example, with each lane offering 8Gb/s of bandwidth, raw bandwidth would be nearly 4GB/s. More of this potential speed can be preserved on system platforms that allow a direct PCIe pipeline into the CPU, rather than having to add overhead by passing through a chipset.

Command Queue

NVMe also has a much more robust command-queue structure with a significantly larger queue depth than AHCI/SATA. Whereas AHCI/SATA has one command queue with a depth of 32 commands, NVMe is designed to have up to 65,535 queues with as many as 65,536 commands per queue. The much higher queue depth for NVMe allows for a greater number of commands that can be executed simultaneously.

This capability, along with flash memory's die structure (described further below), helps to drastically improve the random IOPS performance of NVMe SSDs, compared with AHCI/SATA-based SSDs. Accessing internal components is as easy as sliding out the tray and clicking a small latch to release the top panel. To replace a drive, simply undock it from the SATA port and connect the new one. No cables means no tangling or confusion. You can remove and replace a fan, as needed, in the same way. Similarly, the power supply units (PSU) slide out with ease. These types of chassis innovations make it dramatically easier to replace user-serviceable components.

Improved Latency and Consistent Response Times

NVMe takes advantage of an SSD's die structure and beefed-up interrupt handling capabilities to attain significant improvements in latency and response times. As noted earlier, AHCI was developed for HDDs, which access data sequentially. It takes time for the actuator to move from one location on a spinning platter to another when it is reading or writing data. This delay is called *seek time* and contributes to the overall latency in a drive's performance. NVMe, however, recognizes that the die structure of flash memory inherently suits parallel execution and allows data on any part of the flash array to be accessed instantaneously and simultaneously. This capability means any latency associated with seek time is eliminated.

NVMe—Performance for the SSD Age

Reduced CPU Utilization

NVMe SSDs require fewer CPU cycles to complete the same number of IOPS and use approximately 33% less CPU resources than their AHCI/SATA counterparts (Figure 2).

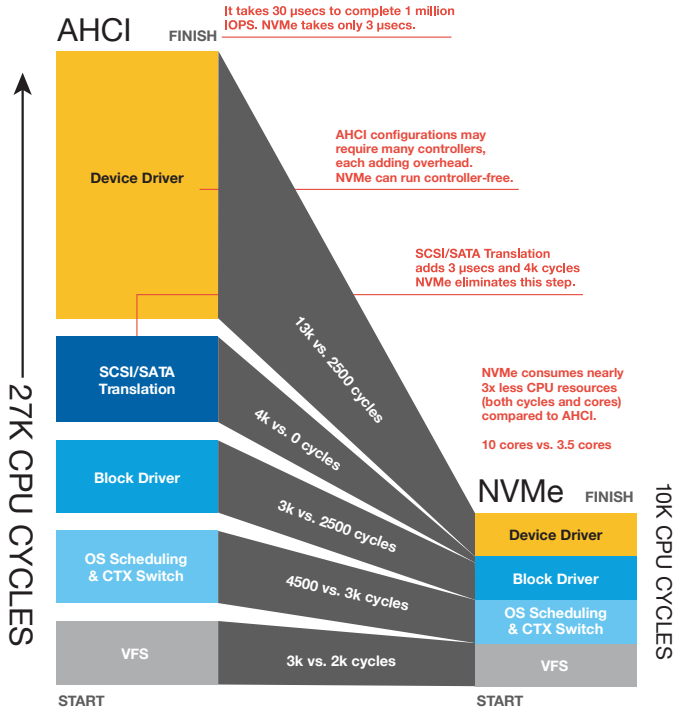

Figure 2. AHCI vs. NVMe: What it takes to reach 1 million IOPS

Fewer Clock Cycles

NVMe requires far fewer CPU cycles to execute a given number of IOPS than does AHCI. This reduction is achieved by, among other things, the complete removal of the SATA/SCSI translation requirement, as well as the elimination of a number of controllers that AHCI configurations might require.

Fewer Cores

Fewer cores are necessary to support NVMe and allow for a smaller capital investment in enterprise compute resources. This means significantly less power consumption—up to 35% less—translating into potential savings. The energy advantage expands further once multiple AHCI bus controllers for handling SATA/SAS loads are eliminated.

Conclusion

As developments in the SSD space continue to push and surpass the limits of AHCI technologies, NVMe over PCIe is looking more and more like a winner. Unlike AHCI, which was developed with HDD technology in mind, NVMe was designed specifically as an interface protocol for flash. SSDs that use NVMe over IP protocol are able to make full use of the advantages that flash technology has to offer over standard storage technologies. With higher sequential throughput, parallel processing, faster IOPS and lower power expenditures, NVMe-based SSDs like those in Seagate's broad portfolio of flash drives are the future of storage technology.

seagate.com

AMERICAS Seagate Technology LLC 10200 South De Anza Boulevard, Cupertino, California 95014, United States, 408-658-1000
ASIA/PACIFIC Seagate Singapore International Headquarters Pte. Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, 65-6485-3888
EUROPE, MIDDLE EAST AND AFRICA Seagate Technology SAS 16–18, rue du Dôme, 92100 Boulogne-Billancourt, France, 33 1-4186 10 00

TP692.1-1610US, October 2016