QLC NAND Technology Is Ready for Mainstream Use in the Data Center

Intel® QLC 3D NAND SSDs can help reduce storage costs through consolidation, while providing the reliability and low-latency performance needed for common, read-heavy workloads.

The growing need to access more warm data

We've all seen the stats on the exponential growth of data. According to IDC, the volume of worldwide data is expected to swell to 175 zettabytes (ZB) by 2025.[1] These numbers are impressive, but the full story isn't just about the growing volumes of data. It's also about the need to have more of that data readily available, to keep it "warm," for use in today's business-critical applications such as analytics, artificial intelligence (AI), machine learning (ML), and content-delivery networks (CDN). For example, the accuracy of AI and analytics results typically increases with the volume and variety of data fed to those applications. Feeding those workloads with large and growing quantities of warm data requires high capacity storage that can keep up.

Many organizations deploy inexpensive hard disk drives (HDDs) in their data centers as a way to grapple with growing capacity needs. But HDDs can't keep pace with the faster read access required by modern workloads. HDDs also require a significant footprint in the data center, which adds to space, power, cooling, and replacement costs.

Another tactic some organizations take is to replace slower, bulkier HDDs with faster, denser triple-level cell (TLC) NAND solid-state drives (SSDs). TLC NAND SSDs are a good fit for mixed and write-heavy workloads, such as caching applications, but these drives typically are not cost- and capacity-optimized for large-scale, read-centric data needs.

The more practical option is to replace HDDs with Intel® QLC 3D NAND SSDs, which combine large, cost-effective capacity with read performance that is far greater than HDD read performance. In fact, the 30.72TB Intel® SSD D5-P5316 can provide up to 20x reduction of warm data storage footprint compared to HDDs.[2] Intel QLC 3D NAND SSDs also offer sequential read performance up to 25x above the sustained transfer rates of HDDs.[3] And because Intel QLC 3D NAND SSDs deliver density optimizations along with TLC-like read performance, these QLC NAND SSDs provide an opportunity to significantly lower capital expenditures (CapEx) for read-intensive storage workloads.

QLC NAND architecture

Quad-level cell (QLC) NAND technology expands data capacity beyond single-level cell (SLC), multi-level cell (MLC), and TLC NAND technologies by increasing density to four bits per cell, delivering 33 percent more bits per cell than TLC NAND. As density increases, cost per gigabyte (GB) drops (see Figure 1).

Intel QLC 3D NAND SSDs bring proven endurance

Intel QLC 3D NAND SSDs rely on several generations of improvements on a proven technology to deliver the right level of endurance. For example, the first generation Intel QLC NAND drive for the data center, the Intel SSD D5-P4320, offers up to 4x higher endurance than a competitor's QLC NAND SSD.[5] The newer Intel SSD D5-P5316 takes endurance levels even further by providing up to a 5x increase in random write endurance over previous-generation Intel QLC NAND SSDs.[6]

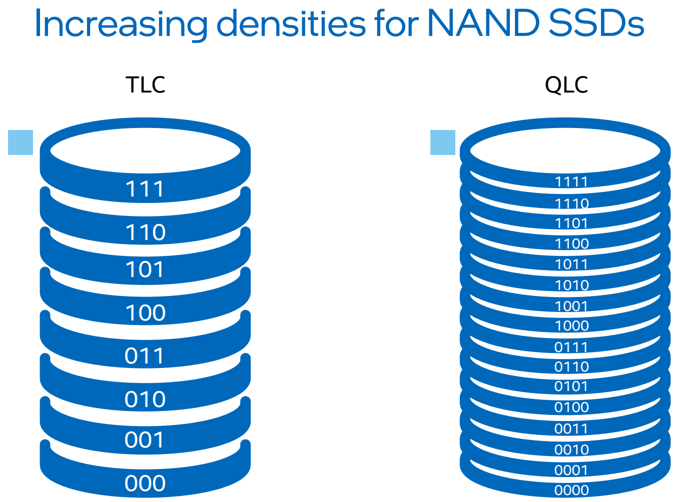

Figure 1. PCIe Intel QLC 3D NAND SSDs include four bits per cell, offering 33 percent more bits per cell than TLC SSDs, while delivering equivalent levels of sequential read performance.[4]

SSD and HDD endurance compared

HDD endurance is primarily a factor of physical limitations of the design, which relies on many moving parts. As these parts wear over time, errors can occur that slow performance or render the drive unusable. For example, vibration or shock can cause head misalignment. Because endurance is primarily a physical limitation with HDDs, these drives are equally affected by both reads and writes.

SSDs, on the other hand, do not have moving parts and are therefore not subject to the same patterns of wear. For SSDs, endurance is a function of individual cell degradation, which is caused primarily by writes, not reads, and is reduced through various mitigation techniques in controller firmware.

The degree of write wear to a NAND SSD depends partly on the state of data already on the drive, because data is written in pages but erased in blocks. When writing sequential data to a relatively new SSD, data can be efficiently written to successive, free pages on the drive. However, when small blocks of data need to be updated (as in revising documents or numerical values), the old data is read into memory, revised, and then re-written to a new page on the disk. The old page, containing deprecated data, is marked invalid.

When free pages are no longer available, those "invalid" pages need to be freed up for use in a background process called "defragmentation" or "wear leveling." All existing valid pages in a given block must first be copied to other free locations on the drive so that the original block only contains invalid, deprecated pages. The original block can then be erased to free up space for new data to be written. Internal NAND housekeeping processes like wear leveling lead to write amplification, where the total internal writes on an SSD are greater than the writes required to simply place new data on the drive. Since every write slightly degrades individual NAND cells, write amplification is a primary cause of wear.

Built-in processes help NAND SSDs distribute wear evenly across the drive. But the bottom line is that write-heavy workloads (random writes, in particular) cause NAND SSDs to wear out faster than other input/output (I/O) patterns because they result in greater write amplification.

QLC NAND SSDs offer exceptional endurance for modern workloads

For most modern workload deployments, QLC NAND SSDs provide more than sufficient levels of endurance because:

- Intel QLC NAND real-world endurance exceeds perceptions and theoretical limitations.
- The larger capacities of QLC NAND drives distribute wear across a larger area.
- Most SSDs use a small fraction of their rated life.

Actual endurance versus expected endurance

The Intel SSD D5-P5316 provides industry-leading levels of endurance for QLC NAND SSDs.[7] In addition, the real-world endurance of QLC NAND SSDs exceeds expectations in several ways. SSD drive controllers distribute writes evenly across the drive to avoid repeated impact on some cells over others. Large capacity QLC NAND drives offer much greater "surface area" to distribute writes, reducing overall degradation.

In addition, there is a notable gap between perceptions of endurance needs versus real-world usage patterns, as emphasized in a large-scale study presented at a USENIX conference in February 2020.[8] The research paper demonstrated that the endurance requirements of real-world applications are often significantly below what enterprises expect. As the authors state, "Based on our data, we predict that for the vast majority of enterprise users, a move towards QLC's PE cycle limits poses no risks, as 99% of systems use at most 15% of the rated life of their drives."

Lifetime endurance of large capacity QLC NAND equals smaller capacity TLC NAND

SSD endurance is commonly rated in drive writes per day (DWPD), which measures the amount of data that can be written to a drive per day over its warrantied lifetime. Another way to express endurance is in terabytes written (TBW) or petabytes written (PBW), which describes the total amount of data that can be written to a drive over its stated lifespan. The total endurance for a drive depends on its capacity. For example, Figure 2 shows a theoretical 8TB TLC SSD rated at 1

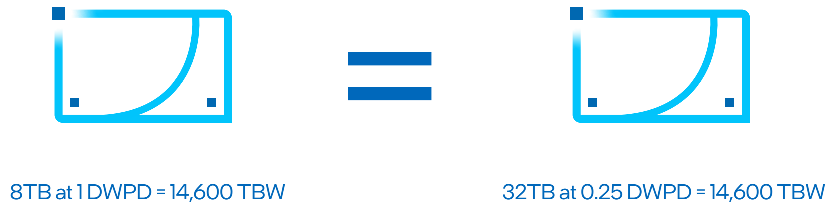

Figure 2. Example of how DWPD and capacity determine total bytes written.

DWPD with a five-year warrantied lifespan can handle 8TB of data writes per day for five years. But note that a theoretical 32TB QLC NAND SSD rated at 0.25 DWPD with a five-year warranty actually has similar overall endurance. That's because the total amount of data that can be written to this drive is the same as the smaller TLC drive (32TB capacity times 0.25 DWPD).

Additionally, storage administrators can boost effective endurance even further by reserving spare drive space—in the theoretical example above, 20 percent. By increasing the reserved area on the drive, wear from write amplification is reduced because there is more available space for efficient defragmentation and wear leveling.

For NAND SSD endurance, workloads matter

As described above, real-world endurance for NAND SSDs is much higher than HDDs, and it can be substantially greater than stated DWPD, despite common misconceptions. But how endurance plays out for an individual organization also depends on usage patterns and the nature of its workloads.

As shown in Figure 3, HDDs tend to have consistent but low endurance, while NAND SSDs exhibit a range of endurance, depending on the pattern and block size of the data.

For example, if an application is dependent on frequent random writes of small block sizes, it will tax NAND SSDs more heavily than sequential reads of large block data. This is why QLC NAND SSDs are an ideal fit for read-heavy workloads that also require large capacities for fast access to more data.

Although it is optimized for read performance, QLC NAND can also be a good fit for other data-usage patterns. Ample write performance on large block sizes enables QLC NAND to span into mixed workloads in certain segments.

See the "Supporting modern workloads with Intel QLC 3D NAND SSDs in the data center" section of this paper for specific examples of workload placement.

Quality and reliability of QLC NAND

Because QLC NAND technology is complex, many in the industry perceive that compromises must be made in areas such as data reliability, data retention, and overall drive reliability. Intel's three decades of floating gate architecture experience have delivered a no-compromise solution. Intel QLC 3D NAND SSDs meet all Joint Electron Device Engineering Council (JEDEC) requirements and are equivalent to widely adopted TLC NAND SSDs.

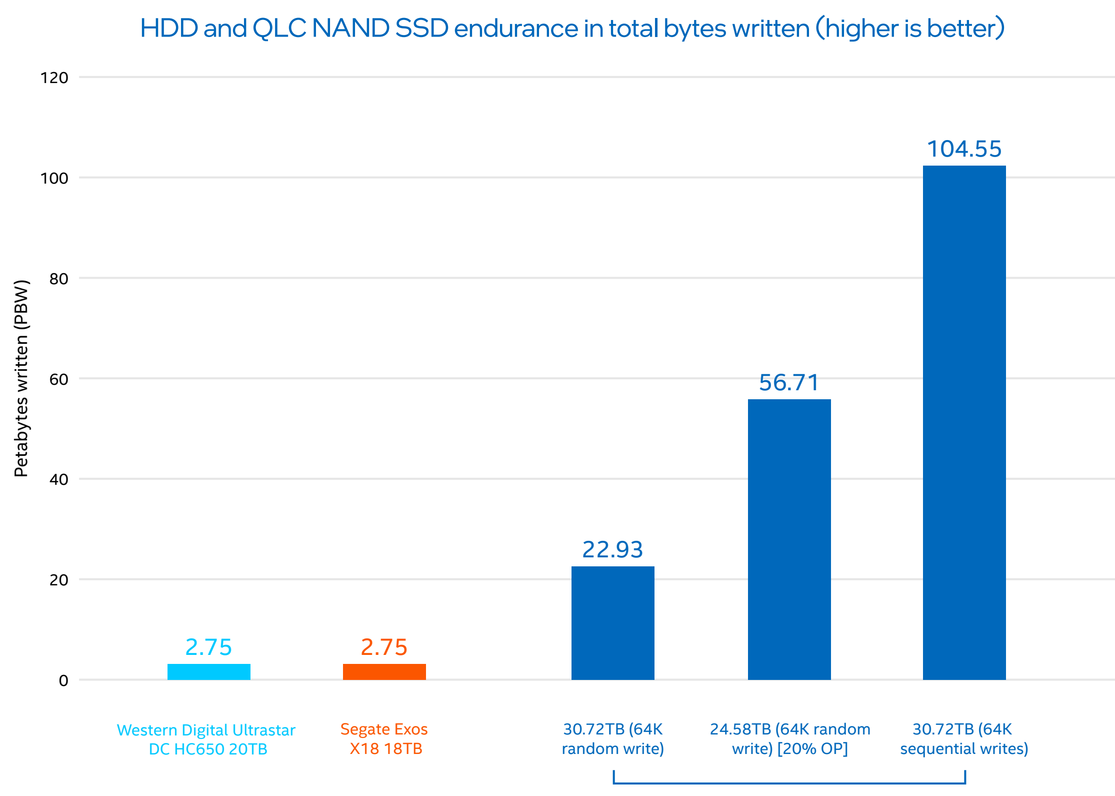

Figure 3. Actual QLC NAND endurance of HDDs compared to the Intel SSD D5-P5316, in petabytes written (PBW).[9]

Table 1 shows how an Intel QLC NAND SSD stacks up against a TLC NAND SSD. Both have five-year warranties, in addition to identical ratings for failure rate, uncorrectable-bit error rate (UBER), operating vibration, and temperature range.

Table 1. Comparison of quality and reliability specifications for an Intel QLC 3D NAND SSD and an Intel TLC NAND SSD.

Specification	TLC NAND SSD: Intel SSD D7-P5510[10]	QLC NAND SSD: Intel SSD D5-P5316[7]
Quality		
Mean Time Between Failures (MTBF)	2M hours	2M hours
Annualized Failure Rate (AFR)	<0.44%	<0.44%
Reliability		
Warranty	5 years	5 years
UBER	1 per 10^{17} bits read	1 per 10^{17} bits read
Data Retention (Temperature at 40°C)	3 months	3 months
Environment		
Operating Vibration	2.17 Grms	2.17 Grms
Temperature Range	0°C to 70°C	0°C to 70°C

Differences are even more pronounced when comparing Intel QLC 3D NAND SSDs to the quality and reliability ratings of many enterprise HDDs. Table 2 compares the Intel SSD D5-P5316 to two common enterprise HDDs. The Intel QLC 3D NAND SSD delivers an UBER two orders of magnitude higher than the HDDs shown. The QLC NAND drive also excels in a wide range of operating conditions, with high specifications for vibration, temperature, and magnetism.

Table 2. Comparison of quality and reliability specifications for an Intel QLC 3D NAND SSD compared to common enterprise HDDs.

	Intel SSD D5-P5316[7]	Western Digital Ultrastar DC HC650[11]	Seagate Exos X18[12]
Drive Type	**QLC NAND SSD**	**HDD**	**HDD**
Warranty	5 years	5 years	5 years
UBER	1 per 10^{17} bits read	1 per 10^{15} bits read	1 per 10^{15} bits read
Drive Reads/Writes per Day (DxPD)	0.41	0.08 (20TB capacity)	0.094
Program/Erase (P/E) Cycle	3,000	Not applicable (N/A)	N/A
Operating Vibration	2.17 Grms	0.67 Grms	0.70 Grms
Affected by Magnetism	No	Yes	Yes

In addition to the advantages highlighted in Table 2, NAND SSDs also exhibit lower real-world failure rates compared to HDDs. The same USENIX study referenced earlier found that average annual replacement rates (ARRs) vary from as little as 0.07 percent to nearly 1.2 percent for NAND SSDs and 2 to 9 percent for HDDs. Comparing the two ranges shows NAND SSDs delivering between 7.5x and 28x lower ARRs than their HDD counterparts.[8]

Supporting modern workloads with Intel QLC 3D NAND SSDs in the data center

With high capacity, read-optimized Intel QLC 3D NAND SSDs, you can derive more value from more of your warm data, while helping to reduce total cost of ownership (TCO) on a proven technology.

Accelerate access to more of your data

QLC NAND excels at both sequential and random read performance for massive quantities of warm data. Those characteristics make the drive a good fit for modern enterprise applications, such as AI or analytics.

These use cases rely on fast read performance with predictable low latency to support faster access to data and insights, and scalability to handle future needs.

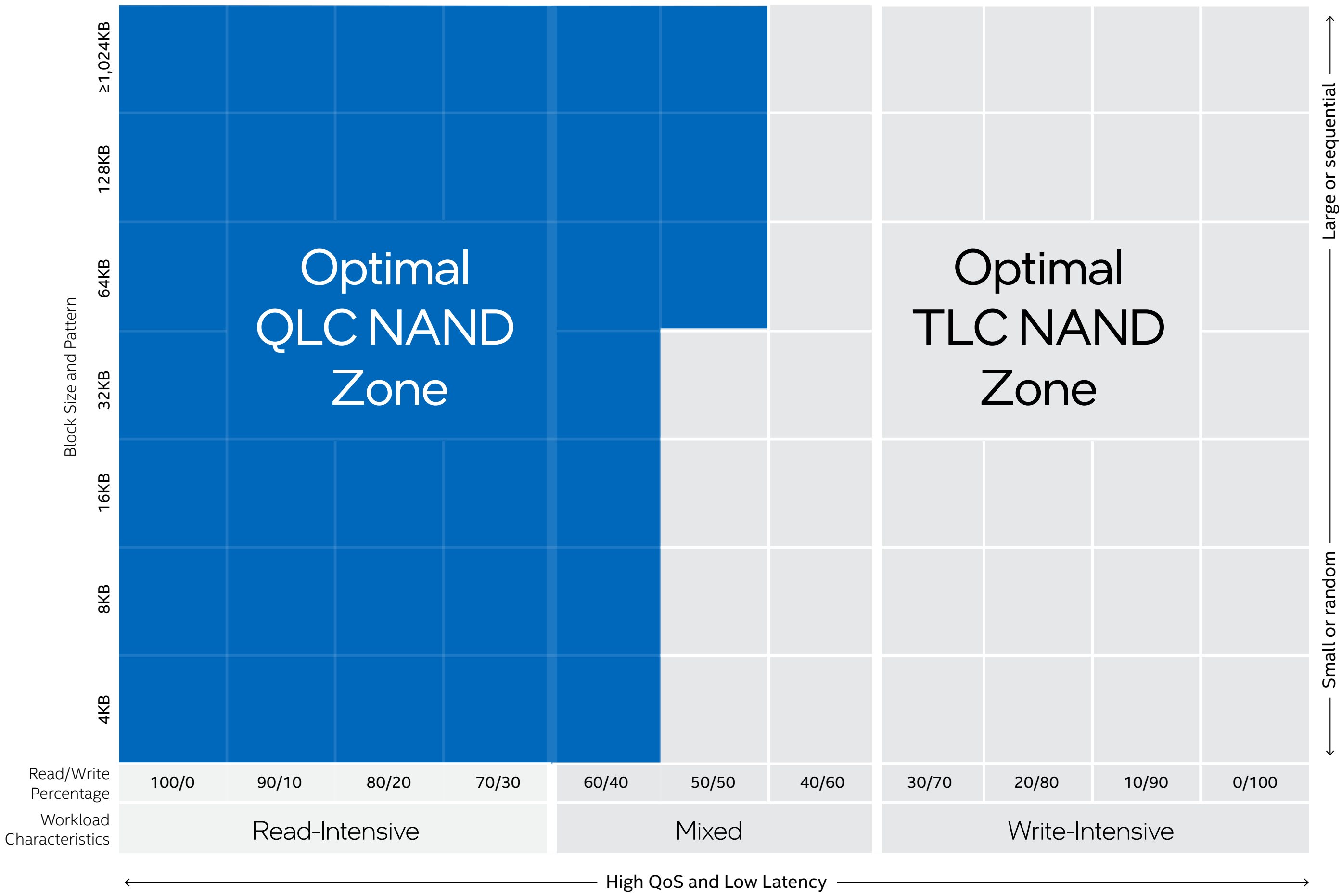

Figure 4. Characteristics for matching workloads to QLC NAND SSDs.

To determine if QLC NAND SSDs are a good match for your workloads, look for the following characteristics:

- Bandwidth: High read bandwidth to move massive data volumes to compute
- Quality of service (QoS): Low latency and high QoS to help optimize compute utilization and reduce time-to-results
- Usage pattern: Read-intensive (sequential or random) workloads, even in the presence of large-block write pressure
- Block size: Small to large datasets

Figure 4 shows which workload segments align well with QLC NAND SSDs, based on usage patterns.

Intel QLC NAND performance

The Intel SSD D5-P5316 includes modern firmware to enable high capacity storage with exceptional read performance and scalability. This drive is also the industry's first QLC NAND SSD to include a PCIe 4.0 controller. Serial ATA (SATA)-based QLC NAND SSDs can limit throughput, in comparison to the multi-lane I/O capabilities of the PCIe interface and efficient NVM Express (NVMe) protocol support offered by Intel QLC 3D NAND SSDs. As a result, Intel QLC 3D NAND SSDs are better suited to the massive throughput requirements of modern workloads.

Architectural and feature improvements designed to accelerate read performance also benefit the latest-generation Intel QLC 3D NAND SSDs. Compared to the previous generation, the Intel SSD D5-P5316 exhibits up to 38 percent better random read performance and up to 2x better sequential read performance.[13,14] The drive also includes intelligent firmware that delivers 48 percent better latency performance at 99.999 percent QoS, compared to the previous generation.[15]

QLC NAND performance benefits are even more pronounced in comparison to HDDs. Despite their apparent cost/capacity advantage, HDDs limit the value of your stored data because these slower drives can't support workloads that require fast access to warm data. As Table 3 demonstrates, the latest-generation Intel QLC 3D NAND SSD provides up to 25x better performance for sequential reads, compared to common enterprise HDDs.[3]

Table 3. Comparison of read performance for an Intel QLC 3D NAND SSD compared to two commonly used HDDs.

	Intel SSD D5-P5316[7]	Western Digital Ultrastar DC HC650[11]	Seagate Exos X18[12]
Drive Type	**QLC NAND SSD**	**HDD**	**HDD**
Sequential Read (SSD)/ Sustained transfer rate (HDD) (Up to)	7,000 MB/s	250 MB/s (sustained transfer rate)	270 MB/s (sustained transfer rate)
Random Read (Up to)	800K input/output operations per second (IOPS) (4K random reads)	Not specified	170 (queue depth 16)
Average Latency	120 µs/180 µs (random read/write)	4.16 ms	4.16 ms

Lower your storage TCO

TCO is a way of understanding how to make purchasing decisions based on acquisition costs and the ongoing expenses of running systems. Because business objectives and configurations vary significantly, there is no such thing as a one-size-fits-all TCO calculation. But there is, among leading storage hyper-scalers and innovators, alignment on the factors and modeling of those factors needed to derive a robust TCO measurement based on CapEx and operating expenses (OpEx). Some of these factors are less broadly used, or even understood, than others. The inherent advantages of SSDs compared to HDDs—and, more broadly, QLC SSDs compared to both HDDs and TLC SSDs—make the understanding of these factors critical as enterprises look to modernize storage in the data economy.

CapEx begins with effective capacity. Measured as terabytes effective (TBe), effective capacity is the actual usable storage space after replication, capacity utilization, and data-reduction methods are accounted for. Effective capacity has a large impact on TCO due to a multiplying effect caused by the high cost of the total raw storage purchased to provide redundancy and meet performance requirements. Storage solution providers typically advertise effective capacity when describing their cost per TB of storage.

OpEx looks at the cost of power, cooling, and drive failures. Industry leaders combine the TBe approach for CapEx with OpEx to derive a total bottom-line metric of TCO dollars per TBe per rack.

QLC NAND SSD cost savings over HDDs

It is well understood that SSDs perform significantly better than HDDs. Almost as well understood is the reliability advantage of SSDs. Given these intrinsic advantages, SSDs do not need replication for performance, and they generally require much less replication for reliability.

Higher SSD performance also lends itself to much more efficient data-reduction methods than HDDs. Data reduction is the ratio of host data stored to physical storage required; a 50 percent ratio would be equivalent to a 2:1 data-reduction ratio. Because data reduction allows the user to store more data than is on the physical hardware, the resulting effective capacity is increased. Compression and deduplication technologies can greatly decrease the required raw storage capacity needed to meet a "usable capacity" requirement.

Modern algorithms are optimized for SSDs, taking advantage of their performance to enable a high data-reduction ratio (DRR) while delivering high application performance. By example, the Zstandard compression algorithm from Facebook achieves compress and decompress speeds much faster than HDDs can read/write, thus allowing the use of the algorithms on SSDs in real time. Another example is VMware vSAN, where compression and deduplication are only offered in all-flash configurations.

QLC NAND SSD cost savings over TLC NAND SSDs

While not as massive as the HDD replacement opportunity, savings over TLC SSD deployments can still be significant. This opportunity is concentrated on CapEx, because the primary factor is the intrinsic storage-density advantage of QLC NAND and the resulting cost-per-TB savings opportunity compared to TLC NAND. With read performance equivalent to TLC, QLC NAND is a TCO opportunity for read-intensive workloads.

The changing nature of storage

Enterprises are rethinking their approach to data. Historically, HDDs delivered a cost-effective way to store large quantities of cold data. But as organizations shift to modern applications, such as AI, analytics, and big data, they are facing a new reality: the need for fast access to massive and growing quantities of data.

At the same time, HDD innovation is decelerating as SSD innovation accelerates. According to analysts at Wikibon, HDD production and use is dropping, and "… volumes will continue to decline rapidly by a factor of 10 by the end of this decade."[16]

As businesses strive to deliver more products and services faster, the need for reliable, fast access to massive quantities of data keeps growing. Organizations can't rely on aging technologies to support modern applications. Intel QLC 3D NAND SSDs are built on innovative generation-over-generation improvements to fill the capacity, performance, and TCO gaps between slower, less-efficient HDDs and more costly TLC SSDs.

In fact, many industry leaders—spanning hyper-scaler CSPs, storage solution providers, innovative start-ups, and others—have already made the shift to QLC NAND technology.

Modernize on a proven technology

Intel offers a strong history of technical expertise and industry leadership in storage. By building on proven technology, Intel has been able to provide innovative improvements that address growing demands for fast access to more data.

Floating gate NAND technology provides a strong voltage threshold window and cell isolation that enable it to scale confidently to higher bits per cell. With this third generation of QLC NAND technology, Intel has extended the vertical floating gate design to a 144-layer QLC 3D NAND that delivers industry-leading density with levels of quality and reliability that match TLC SSDs and vastly exceed slower legacy HDDs.[17]

That performance increases with each successive generation. Compared to the previous generation Intel SSD D5-P4326, the Intel SSD D5-P5316 provides:

- Up to 38 percent better random read performance[13]
- Up to 2x better sequential read performance[14]
- Up to 48 percent lower latency at 99.999 percent QoS[15]

Extract more value from your data with Intel QLC 3D NAND SSDs

If your business relies on reliable, fast read access to warm data, Intel QLC 3D NAND SSDs can fill the gap left between slow, inefficient HDDs and costly TLC NAND SSDs. Built on proven technology, Intel QLC 3D NAND SSDs combine high-density storage with exceptional low-latency read performance to support modern business-critical workloads, including ML, AI, CDNs, analytics, and big data. In addition, Intel QLC 3D NAND SSDs help lower TCO by consolidating the storage footprint in the data center.

Learn more

About Intel 3D NAND SSDs

Intel® SSD D5-P5316 product brief

QLC NAND SSDs Are Optimal for Modern Workloads

Replace Legacy Storage in Content-Delivery Networks with Efficient, Cost-Effective Intel® QLC 3D NAND SSDs

[1] Data taken from the IDC report: "Data Age 2025: The Evolution of Data to Life-Critical."

[2] "Up to 20x reduction of warm storage footprint" claim is based on comparing 4TB HDDs, which require 10 (2U) of rack space to fill up 1PB or storage, against 30.72TB Intel SSD D5-5316 E1.L or U.2 drives, which take 1U of rack space to fill up 1PB of storage. That's up to 20x greater rack consolidation.

[3] Sequential read performance based on Intel SSD D5-P5316 (intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds/d5-series/d5-p5316-series.html) compared to Seagate Exos X18 (seagate.com/files/www-content/datasheets/pdfs/exos-x18-channel-DS2045-1-2007GB-en_SG.pdf).

[4] Equivalent read performance between Intel QLC NAND and Intel TLC NAND is based on Intel measurements of up to 7,000 MB/s for sequential read workloads for the Intel SSD D5-P5316.

[5] "4x higher endurance than the competition" claim based on comparing a 7.68TB Intel SSD D5-P4320 (2,803 TBW) to a 7.68TB Micron 5210 ION SSD (700 TBW) (micron.com/solutions/technical-briefs/micron-5210-ion-ssd).

[6] "5x higher endurance generation-over-generation" claim based on comparing endurance (indirection-unit-aligned, random-write workload) between a 30.72TB Intel SSD D5-P5316 (22,930 TBW) and a 15.36TB Intel SSD D5-P4326 (4,400 TBW). For equivalent comparison, the Intel SSD D5-P5316 uses 64KB random write workload and the Intel SSD D5-P4326 uses 16KB random-write workload for endurance measurements.

[7] Intel. "Intel® SSD D5-5316 provides industry-leading endurance at 0.41 DWPD." intel.com/content/www/us/en/products/memory-storage/solid-state-drives/data-center-ssds/d5-series/d5-p5316-series.html.

[8] Stathis Maneas and Kaveh Mahdaviani, University of Toronto; Tim Emami, NetApp; Bianca Schroeder, University of Toronto. "A Study of SSD Reliability in Large Scale Enterprise Storage Deployments." February 2020. usenix.org/system/files/fast20-maneas.pdf.

[9] Intel SSD D5-P5316 drive endurance is based on 64KB random write and 64KB sequential write workload. 20% OP represents endurance for an over-provisioned drive. Over-provisioning can be done by creating a namespace with a size of 80% of the available area. Seagate Exos X18 endurance is derived from their datasheet: seagate.com/files/www-content/datasheets/pdfs/exos-x18-channel-DS2045-1-2007GB-en_SG.pdf. Western Digital Ultrastar DC HC650 endurance is derived from their datasheet: https://documents.westerndigital.com/content/dam/doc-library/en_us/assets/public/western-digital/product/data-center-drives/ultrastar-dc-hc600-series/data-sheet-ultrastar-dc-hc650.pdf. HDD DWPD calculation: (550TB/year)/365 days/capacity (TB). HDD PBW Calculation: ((550TB/year)*5 years)/1000.

[10] Intel. "Cloud Inspired. Performance Optimized." December 2020. intel.com/content/dam/www/public/us/en/documents/product-briefs/d7-p5510-product-brief-v2.pdf.

[11] Western Digital Data Sheet: Ultrastar DC HC650. https://documents.westerndigital.com/content/dam/doc-library/en_us/assets/public/western-digital/product/data-center-drives/ultrastar-dc-hc600-series/data-sheet-ultrastar-dc-hc650.pdf.

[12] Seagate Data Sheet: Exos X18. seagate.com/files/www-content/datasheets/pdfs/exos-x18-channel-DS2045-1-2007GB-en_SG.pdf.

[13] "Up to 38 percent better random read performance vs. previous generation" claim based on Intel product specifications comparing measured performance for 4KB random read, queue depth 256 (QD 256) performance between a 15.36TB Intel SSD D5-P5316 and a 15.36TB Intel SSD D5-P4326. Measured performance was 800K and 580K IOPS for the Intel SSD D5-P5316 and the Intel SSD D5-P4326, respectively.

[14] "Up to 2.1x sequential read performance vs. previous generation" claim based on Intel product specifications comparing measured performance for 128KB sequential read, QD 256 performance between a 15.36TB Intel SSD D5-P5316 and a 15.36TB Intel SSD D5-P4326. Measured performance was 7.0 and 3.2 GB/s for the Intel SSD D5-P5316 and the Intel SSD D5-P4326, respectively.

[15] "48 percent better latency performance at 99.999 percent" claim based on Intel product specifications comparing measured performance for 4KB random read, queue depth 1 (QD 1) latency performance at 99.999 percent between a 15.36TB Intel SSD D5-P5316 and a 15.36TB Intel SSD D5-P4326. Measured performance was 600 and 1,150 µs for the Intel SSD D5-P5316 and Intel SSD D5-P4326, respectively. Percentage change is 48 percent.

[16] Wikibon. "QLC Flash HAMRs HDD." January 2021. https://wikibon.com/qlc-flash-hamrs-hdd/.

[17] **Industry-leading capacity scaling.** Highest capacity in the industry with 30.72TB Intel SSD D5-P5316 drives.