Micron® Quad-Level Cell Technology Brings Affordable Solid State Storage to More Applications

QLC Empowers Immense, Read-Focused Workloads

Overview

For years, read-focused workloads were relegated to legacy storage like HDDs, with affordability being the remaining obstacle to flash adoption. Quad-level cell (QLC) NAND flash technology stores four bits in each NAND cell, which is 33 percent more bits per cell than triple-level cell (TLC) NAND. As a result, QLC helps drive a more approachable price point, narrowing the affordability gap between HDDs and solid state storage. Micron is the first to make QLC technology available in an enterprise-class SSD.[1]

QLC brings these amazing benefits:

- Speeds read-focused workloads
- Enables emerging applications
- Brings critical data closer to CPUs
- Reduces rack space requirements

This technical brief highlights how Micron® QLC technology is a great fit for applications that read more than they write, brings data closer to the CPU, and packs more density per cell for storage-dense servers and racks to reduce the space demands for high-priority, read-intensive data sets.

Fast Facts

- Micron is a leader in QLC NAND and is the first SSD manufacturer to bring QLC benefits to enterprise-class SSDs.[1]
- QLC packs 33 percent more bits in each NAND cell than TLC, enabling immense gains at the system, rack and data center levels.
- The affordability of four-bits-per-cell NAND enables more applications and more workloads to transition to SSDs more easily.

1. Micron 5210 ION enterprise SSD

Micron

Many Applications Read Far More Than They Write

When applications and workloads are accessing storage, they are either reading data, writing data or some combination of both. With four bits per cell, QLC storage enables us to move more read-intensive workloads to flash, freeing them from the constraints of legacy storage.

When selecting which workloads to move to QLC, we must be aware of the amount of data applications write. This is especially important when moving applications to QLC because flash wears when written and different flash types have different write endurance. QLC has lower write endurance than other types of flash. Learn more in the NAND Storage Basics section at the bottom of this brief.

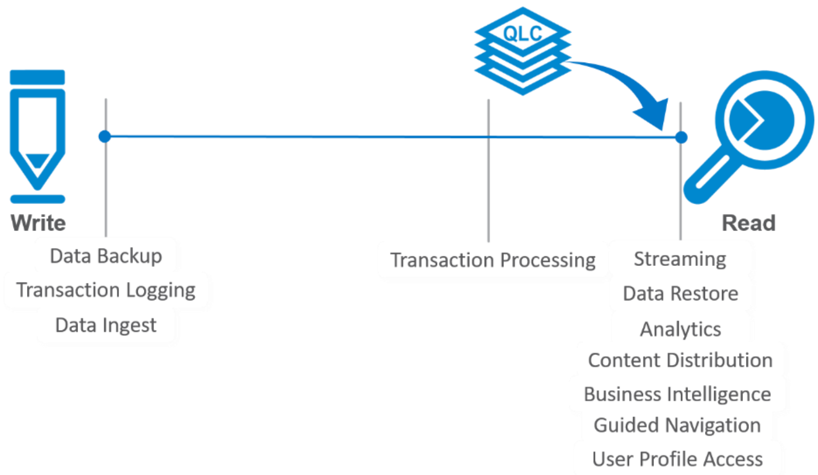

Figure 1: Application Storage Access Type Distribution

Knowing how much our applications write data helps identify those that read far more than they write. An application with mostly read I/O is a good candidate for QLC migration. We can take advantage of QLC by moving these read-focused workloads away from more expensive types of flash or from slower legacy HDDs (and by using other types of flash for mixed-use workloads).

Figure 1 shows how some applications typically access storage. Applications that are write-intensive are on the left, read-intensive are on the right and mixed-access (some read, some write) are in between. Most of the example applications read far more data than they write, making QLC a good fit. Table 1 below shows how these applications could benefit when migrated to QLC.

Workload	How QLC Builds Value
Real-time Analytics and Big Data	Provides a performance uplift for the big data “back end.” For example, a Hadoop Distributed File System can deliver more value with high-capacity storage that’s affordable and lightning-quick.
Business Intelligence and Decision Support Systems	Quickly mine massive data sets using faster, deeper queries to build more responsive, more detailed analytics for better insights.
Active Archives and Large Block Storage	Move active archives from a storage resource into a strategic data asset and deliver massive large-block data streams with ease.
Read-Intensive Artificial Intelligence (AI)	Provides the kind of speed that AI Algorithms depend on to quickly identify patterns in large data sets with fast, high-capacity storage.
Machine and Deep Learning	QLC delivers speed with an approachable price point for immense data sets — because machines can only learn as fast as they can read and analyze data.
Content Delivery, Video on Demand, Content Streaming	Supports massive, parallel requests and streams to deliver more assets to more users more consistently.
NoSQL Databases	Breathe fresh life into data-driven workloads like content classification and tagging as well as user profile acceleration.
User Authentication	Quick storage means quick authentication

Table 1: Example Workloads and QLC Benefits

The applications in Table 1 all share a common thread that makes QLC a good option: they read far more data than they write and quick, efficient reads are critical to their value.

Bring Data Closer to the CPU

Moving data closer to the CPU enables us to process it, learn from it and derive value from it much faster. While traditional HDDs are limited by their physical design (spinning platters, physical heads moving back and forth across those platters, magnets, servos and so on), QLC (and other flash technologies) do not require these moving parts. As a result, this allows data to move closer to the CPU (which lowers latency so the CPU does not have to wait so long for data).

To illustrate the latency differences flash can make, Figure 2 shows a comparison between a QLC SSD[2] and two legacy enterprise HDDs: a 10K RPM (middle) and a 7200 RPM (bottom).

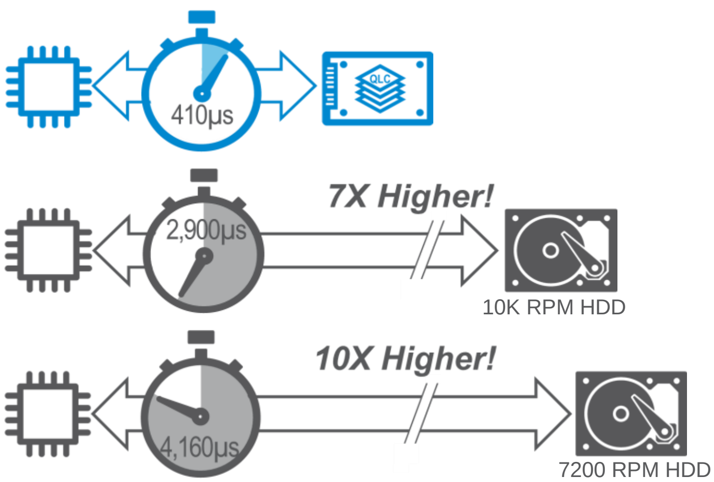

Figure 2: QLC vs. HDD Latency Comparision

The legacy HDDs have much higher latency. In the QLC example, latency is just 410µs while the 10K RPM HDD latency is 2900µs and the 7200 RPM HDD latency balloons to 4160µs.[2] Note that the HDD latencies are so much higher that they are shown on a broken scale in Figure 2.

Much of an HDD's excessive latency is due to its mechanical nature (for example, random access on an HDD may require its heads to swing across its spinning platters). The exact latency difference depends on the HDD design.

2. Micron 5210 ION enterprise SSD. HDD information from public datasheets for HDDs from a reputable, well-known supplier. 10K RPM stated average latency = 2.90ms while 7200 RPM stated latency = 4.16ms. No additional HDD latency measurement details provided. QLC latency based on 4K transfer with queue depth = 32.

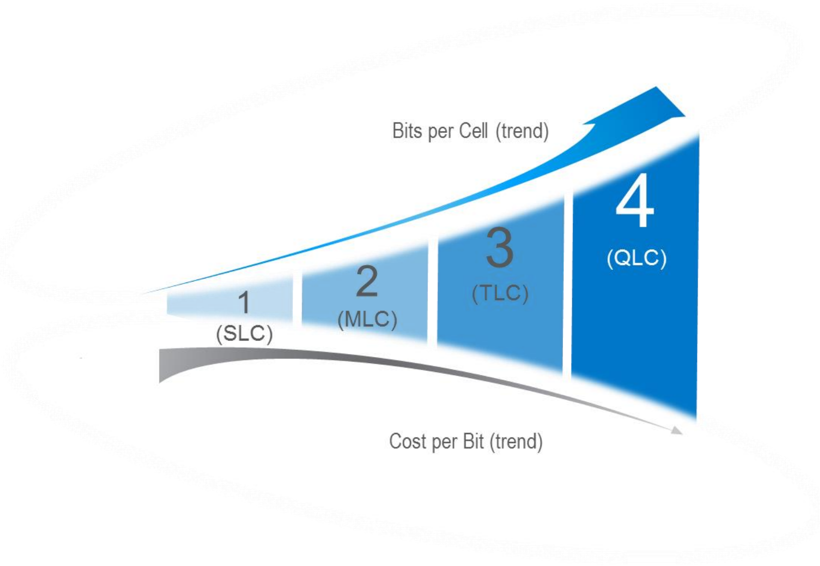

Figure 3 – NAND Bit Density by Generation

Pack 33% More Capacity per Cell

Single-level cell (SLC) (with one bit per cell) was the first broadly available flash technology. When introduced, SLC flash was expensive with limited capacity.

When flash transitioned from SLC to multi-level cell (MLC) technology, the number of bits in each cell doubled (from one to two). This milestone made flash more affordable, broadening its use.

As MLC moved to triple-level cell (TLC) technology, the number of bits per cell increased 50 percent (from two to three). Price per bit lowered and adoption broadened again. The migration from MLC to TLC saw flash become a mainstay in mainstream IT due to lower costs, solid reliability and high performance.

The introduction of QLC adds another bit per cell (from three to four), a 33 percent increase over prior-generation technology. This change enables improved SSD economics, moving flash prices ever closer to legacy HDDs.

QLC's introduction is different from previous generations. Prior transitions saw the new technology (MLC then TLC) largely replace the old. QLC, however, is a complementary technology to TLC. It fills a much-needed gap between TLC and legacy HDDs, enabling more read-focused workloads to move from legacy HDDs to flash.

QLC Builds Solid Storage Density by Server, Rack and Data Center

Ever-increasing data loads and more applications accessing that data mean IT departments are pressed to meet increasing storage and service demands. When we are building new servers to replace older, end-of-equipment-cycle servers, we typically have a choice: high per-drive capacity in the larger 3.5-inch form factor with relatively low read IOPS or smaller-capacity 2.5-inch drives with higher read performance.[3] While 2TB 3.5-inch HDDs are standard in many current HDD installations, major HDD manufacturers offer enterprise 3.5-inch HDDs up to 14TB and 2.5-inch 10K RPM HDDs up to 2.4TB.

We can get good per-unit storage density using newer HDDs and standard 2U servers, supporting up to twelve 3.5-inch or twenty-four 2.5-inch HDDs (some specialty designs may support more). Available raw capacity per server varies with configuration and intended use. Figure 4 shows three typical HDD configurations.

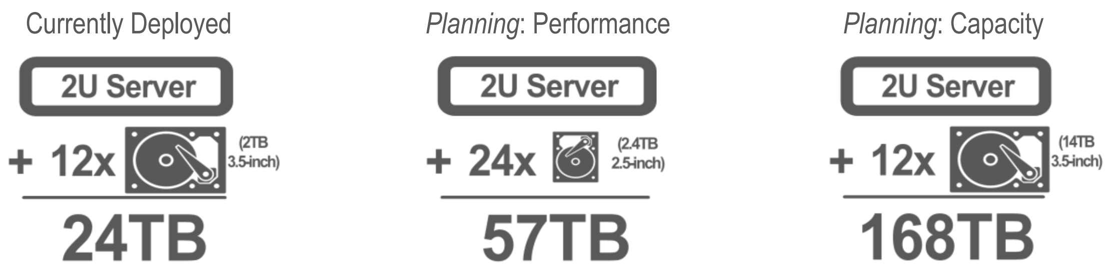

Figure 4: Typical Raw Capacity 2U HDD-Based Servers

The advent of 14TB HDDs immensely boosts per server capacity. Based only on the server density shown in Figure 4, if we want maximum HDD-based capacity, we would choose those 14TB HDDs. If planning for read performance and capacity, we would choose 10K RPM 2.4TB hybrid 2.5-inch form factor HDDs[2].

QLC Means Higher-Density Servers

Introducing QLC storage changes the picture for high-capacity and high-read-performance servers.

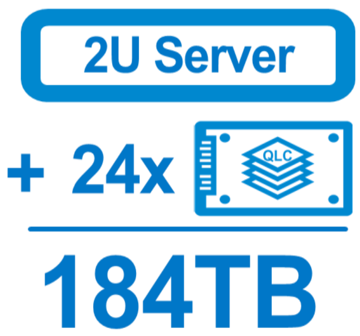

Figure 5: Raw Capacity per Server With QLC

For example, QLC easily enables 7.68TB per SSD in a 2.5-inch form factor.[4] Using the same 2U server with twenty-four drive slots, Figure 5 shows that QLC can enable as much as 184TB raw capacity.

QLC's larger raw capacity advantage is clear when compared to the example currently deployed server (with 2TB HDDs): the QLC-equipped server store is 7.7X as much.

QLC enables 3.2X the capacity of the example performance server (10K RPM 2.4TB Hybrid HDDs) and about 10 percent higher capacity than the 14TB HDD configuration.

3. IOPS and capacity data from public data sheets for 14TB and 10K RPM hybrid HDDs available from major manufacturers. At the time of this document's publication, 2.4TB was the largest commercially available 10K RPM Hybrid HDD available from a major manufacturer.
4. Calculations based on 7.68TB Micron 5210 ION.

Reduce Rack Space Demands of Large Data Sets

QLC enables per-server density improvements ranging from substantial (7.7X) to less dramatic (10 percent). Looking at these improvements at rack scale provides a different perspective.

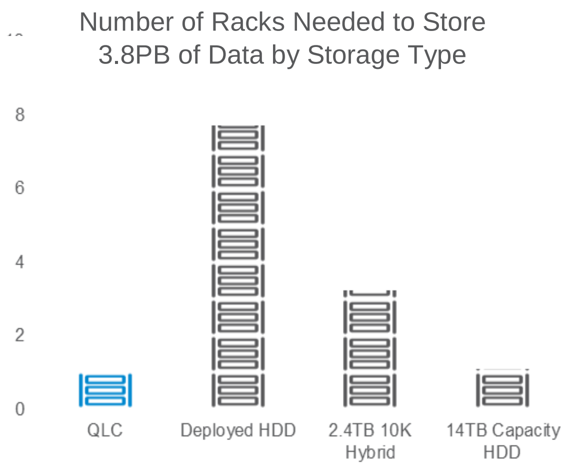

Figure 6: Rack Space Reductions

Figure 6 shows how many racks are needed to store an example 3.8PB data set when using a standard 42U rack and completely populating it with 2U servers as noted earlier (all 42U being occupied by servers).

QLC's density advantage is clear when compared to a typical deployed server and a new read-performance HDD-focused server.

A closer inspection of Figure 6 also shows that the highest-density HDD configuration (twelve 14TB HDDs) does not fit into one rack; it requires a full rack plus about 4U in a second rack.

This comparison illustrates how QLC's ability to increase per-server density reduces rack space for an enormous effect at scale.

The Bottom Line

Micron's QLC introduction — the next generation of flash storage technology — packs 33 percent more bits in each cell enabling more applications and more workloads to transition to SSDs more easily and more affordably. Micron is the first SSD manufacturer to bring QLC benefits to enterprise-class SSDs.

Legacy HDD capacity has skyrocketed with enterprise drives reaching 14TB, but their 3.5-inch form factor limits per-server storage. Legacy 10K RPM hybrid HDDs (long the choice for read-focused applications demanding performance) have limited capacity. These limitations force a difficult choice: larger capacity but slow or lower capacity but fast?

QLC enables both capacity and speed.

Extra storage density enables higher per-server capacity, freeing IT groups to drive more data and reduce rack space (a real benefit for crowded data centers).

QLC brings the all-flash data center a step closer to broad adoption. While many workloads have already been migrated to flash, the holdouts have been read-focused applications like business intelligence and analytics, NoSQL databases and content delivery, video on demand and streaming, big data and active archives, and data center backup and restore. Micron QLC technology empowers these applications by making solid state storage more approachable for read-focused workloads.

Learn more about QLC technology and its transformative effect on your business at micron.com. Stay up to date on what's trending in storage by reading Micron's Storage Blog and following us on Twitter @MicronStorage.

NAND Storage Basics

The paper discusses several QLC advantages and suggests potential applications and workloads that should work well with QLC. To understand these advantages in a bit more detail, some background on NAND (in general) and QLC (specifically) may help.

NAND storage is different from many other types of storage. When we write data to a NAND cell (also known as programming), we must erase the data before writing new data to that same cell. This process is referred to as a program/erase (P/E) cycle.

NAND is programed and erased by applying a voltage to send electrons through an insulator. The location of those electrons (and their quantity) determine when current will flow between a source and a sink (called a voltage threshold), determining the data stored in that cell (the 1s and 0s).

When writing and erasing NAND, it sends the electrons through the insulator and back, and the insulator starts to wear (the exact number of these cycles in each individual cell varies by NAND design). Eventually, the insulator wears to the point where it may have difficulty keeping the electrons in their correct (programmed) location, which makes it increasingly more difficult to determine if the electrons are where we put them, or if they have migrated on their own.

SLC (Single-Level Cell) – One Bit per Cell

When we store one bit (SLC), we don’t need to keep close tabs on electron locations, so a few electrons migrating isn’t much of a concern. Because we are only storing a 1 or a 0, we only need to accurately determine if voltage flows or not.

MLC (Multi-Level Cell) – Two Bits per Cell

Perhaps misnamed, MLC stores two bits per cell, so we need more precision (determining voltage threshold is more complex). We must distinguish between 00, 01, 10 or 11. Migrating electrons have more of an impact, so we cannot wear the insulator as much as with SLC.

TLC (Triple-Level Cell) – Three Bits per Cell

This trend continues with TLC where we store three bits: 001, 010, 100, …110 and 111. Migrating electrons have more effect than in MLC, which further reduces tolerable insulator wear.

QLC (Quad-Level Cell) – Four Bits per Cell

QLC stores four bits (16 possible combinations of 1s and 0s). With QLC, migrating electrons have the most significant effect. Tolerable insulator wear is further reduced.

Therefore, QLC is a great fit for read-centric workloads because NAND cells are worn negligibly when reading data versus worn more when writing data (programming and erasing). When writing and rewriting a lot of data, the insulator wears more quickly. If a NAND cell can tolerate that wear, it is well suited to read/write mixed accesses. The less wear-tolerable NAND cells are, the better they are suited for read-centric workloads and applications.

micron.com