KIOXIA

Technical Brief

Data Loss Mitigation with KIOXIA Enterprise and Data Center NVMe® SSDs

Solid state drives (SSDs) use flash memory cells to store data and have high reliability when compared to hard drives or other forms of memory. However, SSDs can fail over time, and when they fail in a data center, a large amount of data can be lost. Depending on IT policies and failure preparedness, the loss of data can be easily recoverable from a RAID configuration, off-site backup or other options. Without mitigation processes in place, data losses can be disastrous especially if intellectual property (IP), customer information and/or financial data is involved.

At the device level, the NVMe specification that enables a host system to communicate with an SSD across a PCIe® bus includes several features for IT administrators to monitor the health of SSDs in their data centers. These features include the ability for NVMe SSDs to issue warnings as each approaches their respective end of life. KIOXIA, one of the leaders in PCIe 4.0 enterprise and data center NVMe SSD, builds on the safety mechanisms within the NVMe specification with additional features that can extend NVMe SSD life, and enable data access after the drive reaches end of life.

This technical brief will cover these key capabilities within the current NVMe 1.4 specification and a high-level explanation of the KIOXIA built-in NVMe SSD failure mitigation features.

Monitoring SSD Health

In a data center, losing a drive or multiple drives without backup, redundancy or data protection can be devastating as customer records, critical business data, IP and other important files could be permanently gone within the blink of an eye, and its reach can affect other data centers and many users globally (Figure 1). SSDs in a fault tolerant configuration (such as RAID) are similar to tires on a vehicle – if one blows out, the vehicle can be navigated to safety, but to a lesser extent than if all tires operated properly. If multiple tires blow out simultaneously, then it becomes even harder to regain control of the vehicle.

Figure 1: a failed SSD can negatively affect a data center and its global reach

Before an actual drive failure occurs, the latest NVMe 1.4 specification keeps IT administrators informed regarding the health of these NVMe drives by issuing advanced warnings that enable personnel to take preventative measures. The specification uses Self-Monitoring Analysis and Reporting Technology (SMART) health logs for these purposes, as well as system notifications called ‘asynchronous events’ where the SSD notifies the host of status changes, errors and health information. These events are messaged to the host until it takes the necessary steps to clear the event. Notifications include events related to drive health when it reaches a critical state, as well as SSD health status and error codes provided to the host over the drive's life. Key codes and attributes include:

Attribute Name	Log Address	Description of SMART Health Log Event
Critical Error	Byte 00h (Bit 0)	The available spare capacity has fallen below threshold (the spare area may be used up so the drive can no longer be written to)
	Byte 00h (Bit 01)	The SSD has reached a temperature lesser/greater than the defined minimum/maximum threshold
	Byte 00h (Bit 02)	SSD reliability has been compromised due to significant media errors
	Byte 00h (Bit 03)	The SSD has been placed in 'Read-only' mode
	Byte 00h (Bit 04)	The volatile memory device is failing
Available Spare	Byte 03h	Shows the normalized percentage (0 to 100%) of the remaining spare capacity
Available Spare Threshold	Byte 04h	An asynchronous event may occur when the available spare capacity falls below the value specified in this field (can be user set)
Percentage Used	Byte 05h	Shows the vendor-specific estimate of the SSD's life based on actual usage and the manufacturer's prediction
Media and Data Integrity	Bytes 160-175	Shows the number of occurrences where the SSD detected an uncorrectable data integrity error

These health monitoring logs in the NVMe 1.4 specification provide exceptional insight into the health of an SSD, enabling IT administrators to monitor a drive's health directly and take preventative measures (such as replacing an SSD) prior to a drive failure.

Built-in SSD Failure Mitigation

KIOXIA SSDs feature a variety of built-in failure mitigation capabilities such as wear leveling, endurance throttling, end-of-life behavior, RAID 6 die failure recovery and dual-port functionality. These capabilities are featured in CM6 Series PCIe 4.0 enterprise NVMe SSDs and CD6 Series PCIe 4.0 data center NVMe SSDs, and include:

Failure Mitigation Feature	Description
Wear Leveling	The KIOXIA developed algorithm is 'Always On' and ensures that the SSD's media is wearing evenly so that premature drive failure can be prevented from overuse in a portion of the drive such as writing constantly to the same logical bus address (LBA) range.
Endurance Throttling	KIOXIA PCIe 4.0 enterprise and data center NVMe 1.4-compliant SSDs can throttle the amount of writes made to the SSD in order to meet the rated 5-year warranty period and life expectancy specification
End-of-Life Behavior	KIOXIA enables IT administrators the ability to set up behaviors as the drive reaches specified SSD Life Left thresholds. This includes behaviors such as reducing the performance of the drive, sending asynchronous events to the host, and turning on the 'Read-only' mode. This is a significant benefit to IT administrators who can customize how, when and what actions to take upon reaching thresholds.
RAID 6 Die Failure Recovery	KIOXIA PCIe 4.0 enterprise and data center NVMe 1.4-compliant SSDs implement die-failure recovery at the die-level with the use of internal RAID 6[1]. This enables an SSD to sustain a flash memory die failure, recover from it and still read all of the data. Once data has been recovered, a backup is made to a new address. This implementation enables data recovery without the loss of drive functionality or uptime, and crucial for server use.
Dual-Port Functionality	KIOXIA CM6 Series PCIe 4.0 enterprise SSDs include dual port functionality that enables a second path of access to the same drive in case a path failure occurs. This functionality is ideal for storage applications requiring redundancy, high-availability or protection against single-path failure.

Benefits of a Proactive Monitoring Strategy

With the tools provided by the NVMe specification, IT administrators can closely monitor the health of NVMe SSDs on a drive-by-drive basis, and can take preemptive measures against imminent failures or data losses before they happen through critical warning messages received from SMART health logs.

These features are extremely useful and further amplified by KIOXIA's failure mitigation methods (wear leveling, endurance throttling, end-of-life behavior, die failure recovery and dual-port functionality) that are built within its PCIe 4.0 NVMe 1.4 SSDs. Together, these capabilities can significantly reduce the possibility of data losses while preserving data integrity.

KIOXIA PCIe 4.0 Enterprise and Data Center NVMe 1.4 SSDs

The CM6 Series is KIOXIA's 3rd generation enterprise NVMe SSD product line. The series features significantly improved performance from PCIe Gen3 to PCIe Gen4, 30.72 terabyte[2] (TB) maximum capacity, dual-port for high availability, 1 Drive Write(s) Per Day[3] (DWPD) models for read-intensive applications (CM6-R Series) and 3 DWPD models for mixed use applications (CM6-V Series), up to a 25-watt power envelope and a host of security options.

The CD6 Series is KIOXIA's 3rd generation data center NVMe SSD product line. The series features significantly improved performance from PCIe Gen3 to PCIe Gen4, 15.36 TB maximum capacity, 1 DWPD models for read-intensive applications (CD6-R Series) and 3 DWPD models for mixed use applications (CD6-V Series), up to a 19-watt power envelope and a host of security options.

CM6 Series SSDs

PCIe 4.0 and NVMe 1.4 Specification Compliant

High-Performance

SeqRead = up to 6,900 MB/s
RanRead = up to 1.4M IOPS
SeqWrite = up to 4,200 MB/s
RanWrite = up to 350K IOPS

Endurance and Capacities

1 and 3 DWPD options
800 GB - 30,720 GB capacities

CD6 Series SSDs

PCIe 4.0 and NVMe 1.4 Specification Compliant

High-Performance

SeqRead = up to 6,200 MB/s
RanRead = up to 1.0M IOPS
SeqWrite = up to 4,000 MB/s
RanWrite = up to 250K IOPS

Endurance and Capacities

1 and 3 DWPD options
800 GB – 15,360 GB capacities

Summary

Through advanced technologies provided by the NVMe 1.4 specification, coupled with KIOXIA CM6 Series PCIe 4.0 enterprise NVMe SSDs or CD6 Series PCIe 4.0 data center NVMe SSDs, the potential of a drive failure can be lowered significantly. When combined, these technologies can strengthen a data center infrastructure and deliver failure mitigation tools to help prevent against data loss and unnecessary downtime.

For more information on KIOXIA PCIe 4.0 NVMe 1.4 SSDs:

Enterprise SSDs: https://business.kioxia.com/en-us/ssd/enterprise-ssd.html
Data Center SSDs: https://business.kioxia.com/en-us/ssd/data-center-ssd.html

NOTES:

[1] RAID 6 is similar to RAID 5 and is a secure and widely used RAID configuration and is different than RAID 5 as it enables the parity data to be written to two drives.

[2] Definition of capacity - KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes and a terabyte (TB) as 1,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1Gbit = 2^{30} bits = 1,073,741,824 bits, 1GB = 2^{30} bytes = 1,073,741,824 bytes and 1TB = 2^{40} bytes = 1,099,511,627,776 bytes and therefore shows less storage capacity. Available storage capacity (including examples of various media files) will vary based on file size, formatting, settings, software and operating system, and/or pre-installed software applications, or media content. Actual formatted capacity may vary.

[3] Drive Write(s) per Day: One full drive write per day means the drive can be written and re-written to full capacity once a day, every day, for the specified lifetime. Actual results may vary due to system configuration, usage, and other factors