A Comparison of Client and Enterprise SSD Data Path Protection

Doug Rollins, Senior Strategic Applications Engineer
Micron Technology, Inc.
Technical Marketing Brief

Data Path Protection Overview

This document describes the general concept of data path protection in SSDs and details the key features of data path protection that are designed into Micron's RealSSD™ products. It will describe how data moves from the SATA PHY, through the data path, into the NAND, and back, and how each transition is protected. It also explains the differences between Micron's client and enterprise data path protection methods.

The Data Path

Data that is written is referred to as data; and the logical block address [LBA] for that data—the location associated with it—is called metadata, which literally means data about data. The term data path refers to the logical course or path that user data and metadata follow as they move throughout the SSD. The path encompasses the complete, end-to-end course taken by data as it is either written to or read from the underlying NAND media. Some of the protection is hardware-based and some is software-based; each type is designed and deployed where it is most effective.

Host data WRITE commands are shown in blue. Host READ commands take the reverse path and are shown in green.

1. Data moves from SATA transport into the host FIFO.
2. After exiting the FIFO, it migrates to the buffer manager.
3. Next, the buffer manager sends the data to a second FIFO.
4. Upon exiting the second FIFO, the data enters the NAND FIFO.
5. After exiting the NAND FIFO, the data moves to the NAND Flash controller.
6. Finally, the data exits the Flash controller and is written to the NAND media.

All Micron SSDs share some common data path protection features:

- DRAM logical-to-physical (L2P) table parity check
- Host LBA embedding and read page LBA check

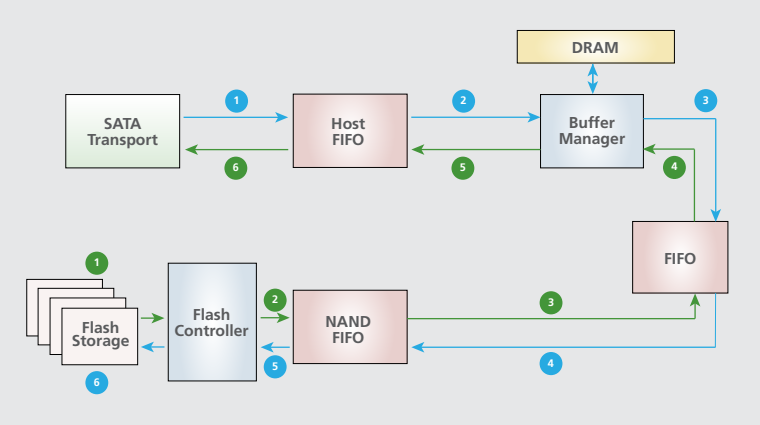

Figure 1: Typical data path for a SATA SSD

DRAM L2P Table Parity Check

The NAND media in an SSD is written in pages, whereas computer systems typically write data to mass storage devices like SSDs in sectors. A page is the smallest unit of NAND storage that can be addressed. Although it varyies by NAND design, a page is typically 4KB to 8KB in size. A simplified diagram of an SSD is shown below.

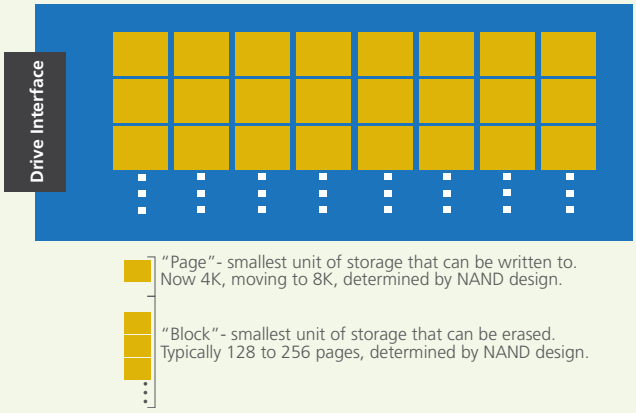

Figure 2: Simplified SSD diagram

Figure 2 shows the SSD PCB in blue and the NAND pages in gold. The NAND blocks, which are logical groupings of pages, are the vertical columns of the NAND pages. The actual number of pages and blocks vary, based on the SSD and NAND design.

The SSD interface, shown in black in the figure, could be SAS, SATA, PCI Express, or some other interface.

Operating systems address mass storage devices in sectors. With hard drives, rotating disks store data on magnetic platters, each logically laid out in a series of concentric rings called tracks. Tracks are further divided into smaller logical storage elements called sectors, typically 512 bytes each. SSDs emulate this logical addressing method.

Because the NAND page and the host sectors are different sizes, an SSD has to build and maintain a data structure that enables it to translate between the host writing data to or reading data from a sector, and the physical NAND page on which that data is actually placed. It creates virtual sectors on the SSD, organizing the NAND media exactly like the physical sectors on a rotating disk drive, albeit much faster. This table structure is typically maintained in the SSD's DRAM; for speed, the structure is updated frequently. It stores all the logical-to-physical pointer data, enabling the host to interact with the NAND-based SSD the same way it would with a conventional rotating disk drive.

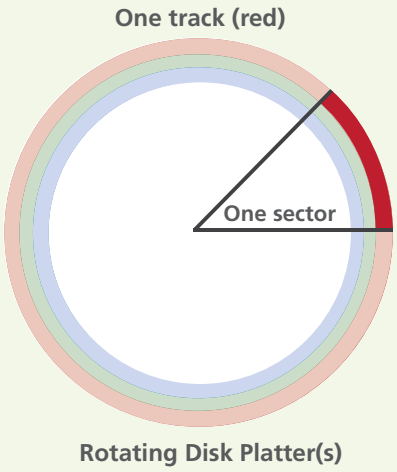

Figure 3: Rotating disk sectors vs. SSD sectors

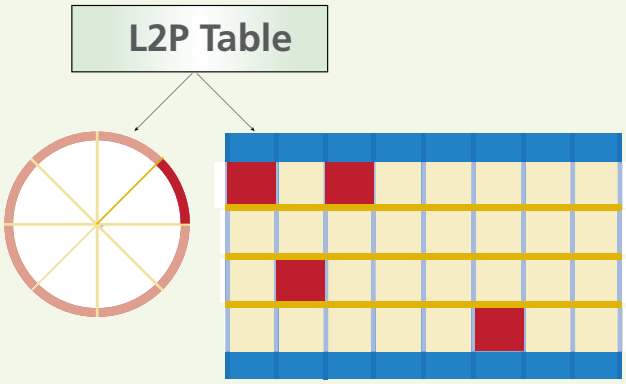

Figure 4: L2P table

This data structure is referred to as the L2P table.

Because the L2P table is stored in DRAM, the potential exists for data corruption within the table itself. The consequence of such corruption can be severe, with data loss being paramount.

Because of this potential, the L2P tables in all of Micron's RealSSD products are protected by parity. This parity check protection method provides an extremely robust structure.

When the host writes data to the SSD, two key elements combine: the actual data to be written and the LBA from which it came. For example, if the host data was a building, the LBA would be that building's street address. All data written to the SSD has an LBA associated with it. Micron embeds the host LBA, along with the user data it references, logically, joining them before writing anything to the NAND, as shown on the right.

The LBA data is stored in the L2P table for retrieval when the host requests that the data be read.

When the READ command is issued, the SSD looks to the L2P table to determine which NAND page(s) contain the data.

When the data is read from the NAND, the embedded LBA is also read and compared against the value in the L2P table to ensure they match.

This process for storing the LBA when data is written, and then reading it back and comparing it to the data stored in the L2P table, ensures that the SSD returns the correct data when requested.

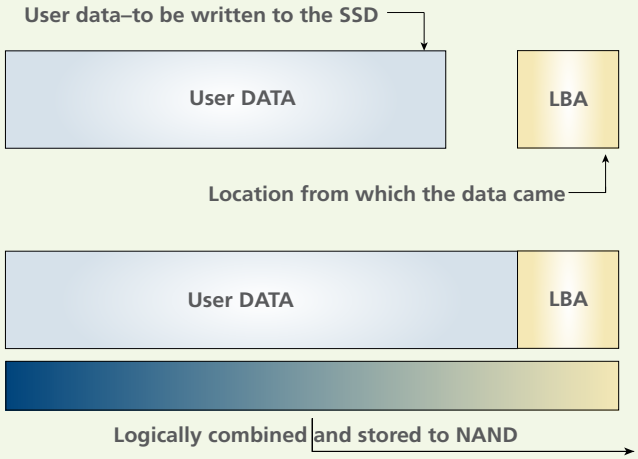

Figure 5: Writing data to the SSD

Client-Specific Data Path Protection

In addition to the L2P parity protection and the host LBA embedding and read page check, Micron client SSDs employ parity generation and check functions as well as cyclic redundancy checksum (CRC) and BCH encoding and checking to protect user data as shown below.

As the data passes from the SATA transport to the host FIFO, parity is generated. As the data exits the host FIFO, parity is checked.

Next, CRC and BCH error correction codes are generated and stored with the data.

Finally, a CRC is generated just before the data enters the NAND FIFO, and then it's checked when exiting.

When data is read from the NAND, the process occurs in reverse order: CRC is generated as the data moves from the NAND Flash controller to the NAND FIFO and is checked on exit; the CRC and BCH codes generated on write are read and verified; and finally parity is generated before the data enters the host FIFO and is checked upon exit.

This sequence of parity and CRC generation and checks ensures that any potential data corruption inside the SSD is detected (bits reversed or flipped), so the host can take corrective action.

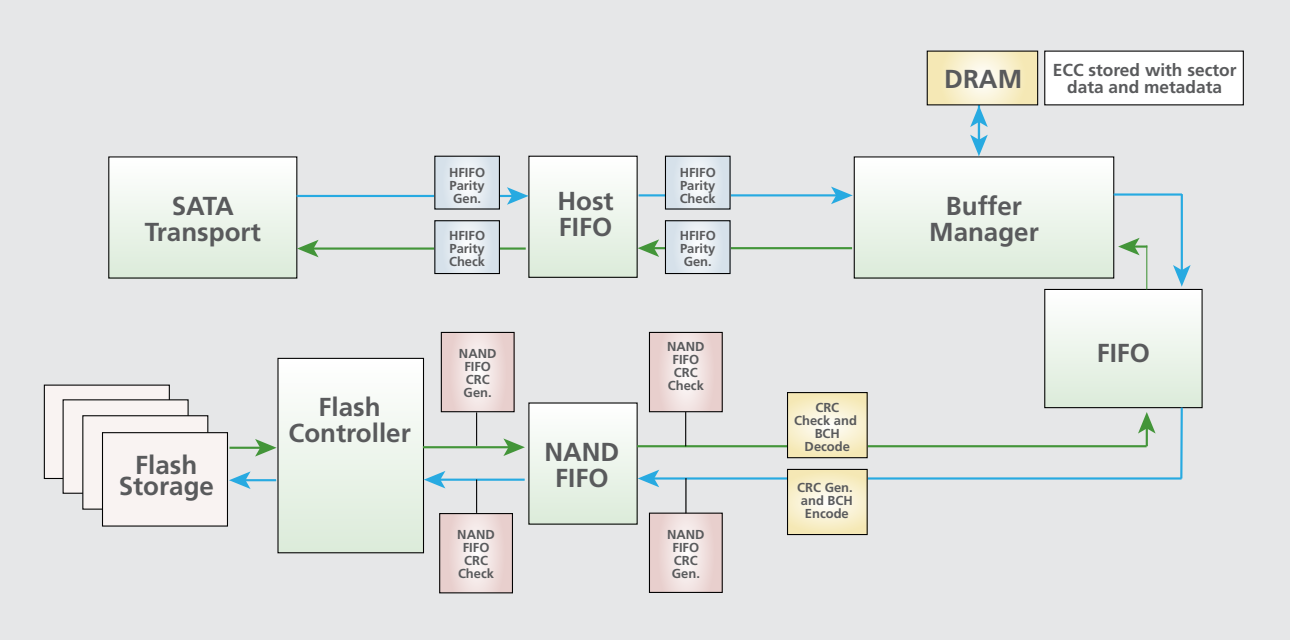

Figure 6: Client-specific data path protection

Enterprise-Specific Data Path Protection

Our enterprise SATA drives build on the foundation of proven data path protection for client drives, but go one step further, adding protection in the form of memory path error correction as shown below.

All the data path protection mechanisms described for Micron's client-focused SSDs are also used in our enterprise drives. However, an additional memory protection ECC (MPECC) is added. MPECC is designed to protect the host data by adding ECC coverage to the data as it enters the SSD.

A 12-byte MPECC is generated on the host data in the SATA PHY and is independent of any ECC provided by the NAND devices themselves.

This additional MPECC follows the host data through the SSD.

As the MPECC and user data enter the host FIFO, parity is generated. As the data exits the host FIFO, that parity is checked.

In the buffer manager, further MPECC protection is generated on the associated metadata. By adding MPECC protection to the metadata, both host data and meta data are protected. As the host data, its metadata, and the MPECC generated for both types of data exit the FIFO adjacent to the buffer manager, both are checked.

Next, CRC and BCH codes are generated (as with client drives).

Finally, parity is generated before the data enters the NAND FIFO, and that parity is checked upon exit.

On READ commands, the process is carried out in reverse order. As the data is read from NAND, parity is generated as the data enters the NAND FIFO and that parity is checked upon exit. The CRC is then checked and the BCH decoded. MPECC protection is generated (on both user data and metadata) before the data passes through the FIFO adjacent to the buffer manager. The MPECC on the metadata is checked in the buffer manager. Prior to entering the host FIFO, parity is generated. Upon exit of the host FIFO, that parity is checked. Next, the MPECC is checked, and finally the data is returned via the PHY.

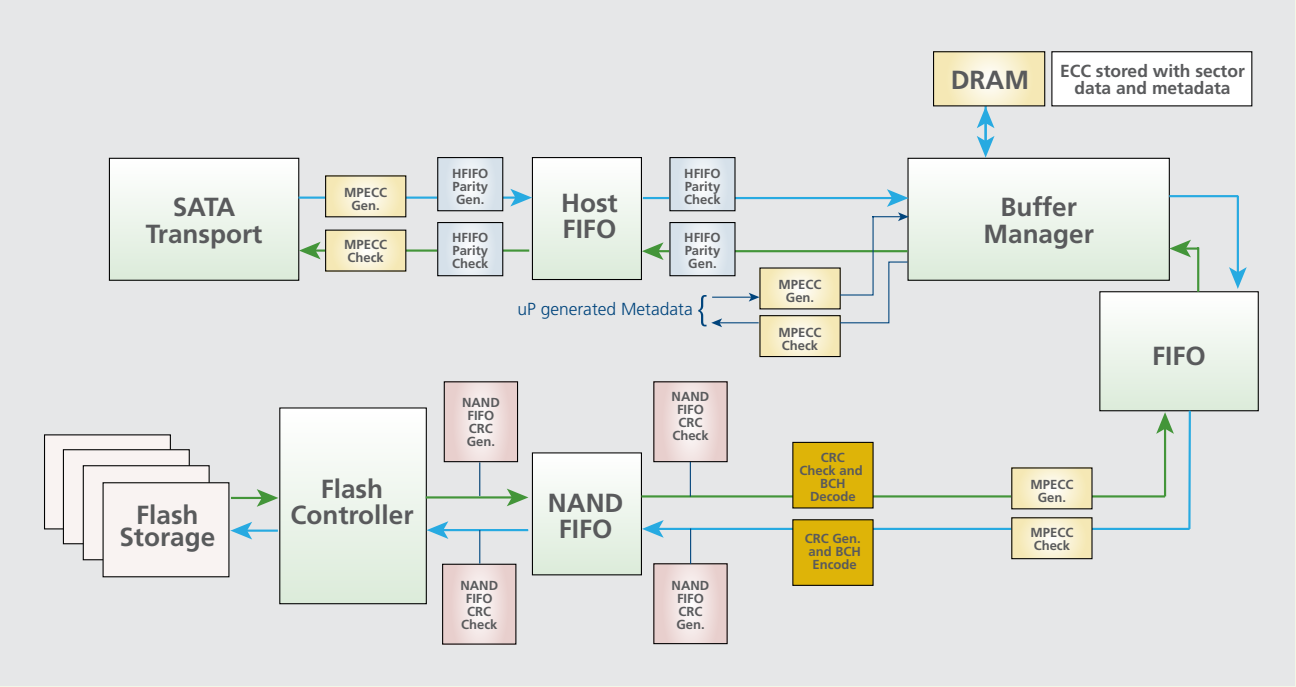

Figure 7: Enterprise-specific data path protection

Client and Enterprise SSDs: Key Data Path Protection Differences

Additional MPECC protection enables our enterprise SSDs to correct single-bit errors within the data path and report double bit errors such that the host can take appropriate action.

SSD Type	*Single-Bit Error Detection?*	*Single-Bit Error Correction?*	*Double-Bit Error Detection?*
Client	Yes	No	No
Enterprise	Yes	Yes	Yes

Table 1: Single-bit error detection/correction

Conclusion

Both client and enterprise SSDs must protect data inside the drive—from the connection to the host system, through the circuits of the SSD, to the NAND Flash memory. While the NAND Flash memory employs its own ECC protection, Micron uses additional state-of-the art data protection methods, such as parity protection on internal buffers and checksum generation and checking.

Our client SSDs employ these techniques to offer robust, optimized data integrity ideally matched to the rigorous demands of the client SSD market. Building on that solid client data protection design, our enterprise SSDs offer additional data correction capabilities by protecting both the host data and its metadata. This enhanced level of protection is required for true enterprise-class SSDs. By enabling single-bit correction in the data path, our enterprise SSDs can handle the most demanding server and storage workloads seen every day in world-class data centers.

micron.com