



Solid State Drive (SSD) FAQ

Santosh Kumar
Rajesh Vijayaraghavan

List of Questions

[Why SSD?](#)

[Why Dell SSD?](#)

[What are the types of SSDs?](#)

[What are the best Use cases & applications for SSDs?](#)

[Why I might notice a decrease in write performance when I compare a used drive to a new drive?](#)

[Data Retention: I have unplugged my SSD drive and put it into storage. How long can I expect the drive to retain my data without needing to plug the drive back in?](#)

[What is Overprovisioning?](#)

[What is Wear Leveling?](#)

[What is Garbage Collection?](#)

[What is Error Correction Code \(ECC\)?](#)

[What is Write Amplification Factor \(WAF\)?](#)

[What steps do SSD drives take to limit the likelihood of damaging cells due to excessive writes?](#)

[How is the SSD Useful Life Span calculated?](#)

[What is TRIM/UNMAP and do Dell enterprise SSD drives support it?](#)

[How do SSDs maintain Data Integrity?](#)

[How is SSD Sanitized?](#)

[What are the recommended application tuning & OS settings?](#)

[What is Endurance Management?](#)

[What Warranty do Dell SSDs carry?](#)

Glossary

Data Retention:

Data retention is the timespan over which a ROM remains accurately readable. It is how long the cell would maintain its programmed state when the chip is not under power bias. Data retention is very sensitive to number of P/E cycle put on the flash cell and also dependent on external environment. High temperature tends to reduce retention duration. Number of read cycles performed can also degrade this retention.

P/E (Program/Erase) Cycle:

In NAND flash, storage is achieved using floating-gate transistors that form NAND gates. As such, the non-programmed state of a bit is 1, while the programming operation injects charge into the floating gate and its resultant bit becomes 0. The opposite operation, erase, extracts the stored charge and reverts the state to 1. The erase and program operations inherently cause degradation of the oxide layer isolating the floating gate; this is the reason for NAND flash's finite lifespan (30K-1M program/erase cycles for SLC typically, 2.5K-10K program/erase cycles for MLC, 10K-30K program/erase cycles for eMLC).

Flash Translation Layer (FTL):

Flash Translation Layer, is a software layer used in computing to support normal file systems with flash memory. FTL is a translation layer between the sector-based file system and NAND flash chips. It enables the operating system and file system access NAND flash memory devices as access disk drives. A FTL hides the complexity of flash by providing a logical block interface to the flash device. Since flash does not support overwriting flash pages in place, an FTL maps logical blocks to physical flash pages and erase blocks.

Metadata:

The metadata is used for the management of the stored information or data in the NAND flash memory. The metadata generally includes, a logical-to-physical address mapping table of the stored information, information of attributes of the stored information, and any other data that can assist in the management of the stored information.

Virtual Pool:

Virtual pool is a set of NAND erased blocks ready to be programmed.

[Click here to go to list of questions](#)

1. Why SSD?

Unlike hard disk drives (HDDs) which use a spinning platter to store data, solid state drives (SSDs) use solid state memory NAND chips. HDDs have several different mechanical moving parts which make them susceptible to handling damage. Solid state drives, on the other hand have no moving parts and are therefore much less susceptible to handling damage even when impacted during use.

SSDs deliver ultra-high performance input/output operations per second (IOPS), and very low latency for transaction-intensive server and storage applications. Properly used in systems with HDDs, they reduce total cost of ownership (TCO) through low power consumption and low operating temperature.

[Click here to go to list of questions](#)

2. Why Dell SSD?

Dell closely manages all the steps necessary to supply its customers with the high-quality Solid State Drives required for demanding Enterprise applications. This includes: -

- Initial supplier qualification and continuous quality testing ;
- Specific firmware creation;
- Bill of Material control and extensive reliability testing;
- Ongoing product quality certifications

All Dell Enterprise Solid State Drives are developed to precisely match the Dell Enterprise systems and to provide customers with an optimal production environment.

The hard drive industry has recently seen consolidation of suppliers and standardization of drives. This has not been the case for Solid State drives. There are many SSD manufactures and Dell cannot guarantee any level of functionality or compatibility on Dell servers using SSD's that were not purchased from Dell.

[Click here to go to list of questions](#)

3. What are the types of SSDs?

Solid state drives (SSDs) based on flash memory generally demonstrate lower latencies than the hard disk drives (HDDs), often enabling faster response times. For random read workloads, SSDs deliver higher throughput relative to HDDs.

a. Based on Nand Flash

- I. SLC, or Single Level Cell, allows for the storage of one bit of information per NAND memory cell. SLC NAND offers relatively fast read and write capabilities, high endurance, and relatively simple error correction algorithms. SLC is typically the most expensive NAND technology. With SLC drives each cell is spec'd to last for around 100K writes.

Reads are unlimited. SLC drives are more suited for enterprise environments because of their durability. They can be cost prohibitive in consumer applications.

- II. MLC, or Multi Level Cell, technology in general is less robust than SLC as there are two bits stored in each cell. If one cell is lost two bits will be lost. With MLC drives each cell is spec'd to last between 3,000 to 5,000 writes. The drives are usually available in larger capacities and are usually less expensive. MLC based SSDs are being used in enterprise applications deploying smart management techniques such as overprovisioning and endurance management (defined later in document).
- III. eMLC, or enterprise MLC is a variant of MLC technology that is harvested from the highest quality portion of the NAND wafer and programmed uniquely to increase erase cycles. eMLC achieves endurance levels of 30,000 write cycles, whereas some of the newest MLC only has 3,000 write cycles. eMLC makes a tradeoff to enable this endurance by giving up [data retention](#). eMLC addresses that problem by lengthening the flash memory chips' internal page programming (tProg) cycle which creates a better, more lasting data write, but slows write performance. Since eMLC SSDs are somewhere between MLC and SLC on write endurance, their price is usually between the two types. By adding advanced endurance management techniques, this technology can be successfully used in general purpose enterprise applications.

b. Based on Host Interface

- I. SATA SSD: SATA SSDs are based on the industry standard SATA interface. SATA SSDs provide reasonable performance for enterprise servers.
- II. SAS SSD: SAS SSDs are based on the industry standard SAS interface. SAS SSDs combine superior reliability, data integrity, and data fail recovery making them suitable for enterprise applications.

[Click here to go to list of questions](#)

4. What are the best Use cases & applications for SSDs?

SSDs are best suited to applications that require the highest performance. I/O-intensive applications such as databases, data mining, data warehousing, analytics, trading, high-performance computing, server virtualization, Web serving and email system are most suitable for SSD use.

- SLC SSD is the preferred technology for write caching, and read caching applications where reads are random and write intensive.
- eMLC SSD will increasingly become the preferred option when handling a mix of both reads and writes, and especially advantageous when budgets are tight.
- MLC SSD is the most cost effective solution for read intensive applications such as accessing a database table.

SSD Types/Applications-Use cases

Flash Technology	Application Kind	Applications
MLC/eMLC	Web based & client computing	Front End Web
		Streaming Media
		Web Applications
		Email/Messaging
		Collaboration
eMLC/SLC	DSS/HPC/OLTP/Storage	HPC/Supercomputing
		Data Warehousing/Mining
		Infrastructure
		Virtual Desktop
		OLTP/Database/Business Processing
		Data Caching

[Click here to go to list of questions](#)

5. Why I might notice a decrease in write performance when I compare a used drive to a new drive?

SSD drives are intended for use in environments that perform a majority of reads vs. writes. In order for drives to live up to a specific warranty period, MLC drives will often have an endurance management mechanism built into the drives. If the drive projects that the useful life is going to fall short of its warranty, the drive will use a throttling mechanism to slow down the speed of the writes.

[Click here to go to list of questions](#)

6. I have unplugged my SSD drive and put it into storage. How long can I expect the drive to retain my data without needing to plug the drive back in?

It depends on the how much the flash has been used ([P/E cycle](#) used), type of flash, and storage temperature. In MLC and SLC, this can be as low as 3 months and best case can be more than 10 years. The retention is highly dependent on temperature and workload.

NAND Technology	Data Retention @ rated P/E cycle
SLC	6 Months
eMLC	3 months
MLC	3 Months

[Click here to go to list of questions](#)

7. What is Overprovisioning?

Over provisioning is a technique used in the design of flash SSDs and flash media cards. By providing extra memory capacity (which the user can't access) the SSD controller can more easily create pre-erased blocks ready to be used in the [virtual pool](#). Overprovisioning improves:

- Write performance & IOPS
- Reliability & endurance

[Click here to go to list of questions](#)

8. What is Wear Leveling?

NAND flash memory is susceptible to wear due to repeated program and erase cycles that are commonly done in data storage applications and systems using [Flash Translation Layer](#) (FTL). Constantly programming and erasing to the same memory location eventually wears that portion of memory out and makes it invalid. As a result, the NAND flash would have limited lifetime. To prevent scenarios such as these from occurring, special algorithms are deployed within the SSD called wear leveling. As the term suggests, wear leveling provides a method for distributing program and erase cycles uniformly throughout all of the memory blocks within the SSD. This prevents continuous program and erase cycles to the same memory block, resulting in greater extended life to the overall NAND flash memory.

There are two types of wear leveling, dynamic and static. The dynamic wear algorithm guarantees that data program and erase cycles will be evenly distributed throughout all the blocks within the NAND flash. The algorithm is dynamic because it is executed every time the data in the write buffer of the drive is flushed and written to flash memory. Dynamic wear leveling alone cannot insure that all blocks are being wear-leveled at the same rate. There is also the special case when data is written and stored in flash for long periods of time or indefinitely. While other blocks are actively being swapped, erased and pooled, these blocks remain inactive in the wear-leveling process. To insure that all blocks are being wear-leveled at the same rate, a secondary wear-leveling algorithm called static wear leveling is deployed. Static wear leveling addresses the blocks that are inactive and have data stored in them.

Dell SSD drives incorporate both static and dynamic wear leveling algorithms to make sure the NAND blocks are wearing evenly for greater extended life of the SSD.

[Click here to go to list of questions](#)

9. What is Garbage Collection?

Flash memory is made up of cells which store one or more bits of data each. These cells are grouped into pages, which are the smallest discrete locations to which data can be written. The pages are collected into blocks, which are the smallest discrete locations that can be erased. Flash memory cannot be directly overwritten like a hard disk drive; it must first be erased. Thus, while an empty

page in a block can be written directly, it cannot be overwritten without first erasing an entire block of pages.

As the drive is used, data changes, and the changed data is written to other pages in the block or to new blocks. At this point, the old (stale) pages are marked as invalid and can be reclaimed by erasing the entire block. To do this, however, any still-valid information on all of the other occupied pages in the block must be moved to another block.

The requirement to relocate valid data and then erase blocks before writing new data into the same block causes **write amplification**; the total number of writes required at the flash memory is higher than the host computer originally requested. It also causes the SSD to perform write operations at a slower rate when it is busy moving data from blocks that need to be erased while concurrently writing new data from the host computer.

SSD controllers use a technique called **garbage collection** to free up previously written blocks. This process also consolidates pages by moving and rewriting pages from multiple blocks to fill up fewer new ones. The old blocks are then erased to provide storage space for new incoming data. However, since flash blocks can only be written so many times before failing, it is important to also wear-level the entire SSD to avoid wearing out any one block prematurely.

[Click here to go to list of questions](#)

10. What is Error Correction Code (ECC)?

The deterioration of the flash memory cell over time and the disruptions from neighboring flash memory pages can lead to random bit errors in the stored data. While the chances of any given data bit being corrupted is quite small, the vast number of data bits in a storage system makes the likelihood of data corruption a very real possibility.

Error detection and correction codes are used in flash memory storage systems to protect the data from corruption. Dell SSD drives are equipped with the industry's most advanced ECC algorithm to achieve enterprise level of uncorrectable bit error rate of 10^{-17} .

[Click here to go to list of questions](#)

11. What is Write Amplification Factor (WAF)?

The write amplification factor is the amount of data the SSD controller has to write in relation to the amount of data that the host controller wants to write. A write amplification factor of 1 is perfect, it means you wanted to write 1MB and the SSD's controller wrote 1MB. A write amplification factor greater than 1 isn't desirable, but is an unfortunate fact of life. The higher your write amplification, the quicker your drive will wear out and the lower its performance will be.

$$\frac{\text{data written to the Flash memory}}{\text{data written by the host}} = \text{write amplification}$$

12. What steps do SSD drives take to limit the likelihood of damaging cells due to excessive writes?

Dell uses following methods to avoid damaging flash cells and extend the life of the SSD drive:

- **Overprovisioning:** The process of increasing the spare area on a solid-state drive. It increases the available “ready to be written” resource pool which decreases write amplification. Since there is less background data movement required, performance and endurance increases. As an example, a drive with 100 GB usable capacity would have an extra 28 GB hidden capacity. The remaining capacity would be used for wear leveling.
- **Wear leveling:** Dell SSD drives uses both static & dynamic wear leveling techniques. Wear leveling allows for data to be mapped to different locations on the drive to avoid writing too often to the same cell.
- **Garbage Collection:** Dell SSD drives are equipped with sophisticated advanced level garbage collection technique. The "Garbage Collection Process" eliminates the need to perform erasure of the whole block prior to every write. It accumulates data marked for erase as "Garbage" and performs whole block erase as space reclamation in order to reuse the block, often doing this as a background process when the drive is not busy with I/O.
- **Data Buffering & caching:** Dell SSD drives use DRAM for data buffer, caching to minimize write amplification ensuring the likelihood of damaging cells due to excessive writes.

[Click here to go to list of questions](#)

13. How is the SSD Useful Life Span calculated?

The useful life of an SSD is governed by three key parameters: SSD NAND flash technology, capacity of the drive, and the application usage model. In general the following life cycle calculator can be used to figure how long the drive will last.

Life [years] = (Endurance [[P/E cycles](#)] * Capacity [physical, bytes] * Overprovisioning Factor) / (Write Speed [Bps] * Duty Cycle [cycles] * Write % * WAF) / (36 *24* 3,600)

Parameters:

Endurance, NAND [P/E Cycle](#): 100K SLC, 30K eMLC, 3K MLC

Capacity: Usable capacity of the SSD

Overprovisioning Factor: Over provision NAND percentage

Write Speed: Speed of write in Bytes per second

Duty Cycle: Usage duty cycle

Write %: percentage of writes during SSD usage

[WAF](#): Controller Write Amplification factor computed based on application use case

[Click here to go to list of questions](#)

14. What is TRIM/UNMAP and do Dell enterprise SSD drives support it?

Certain Operating Systems support the TRIM function, which translates deleted files to the associated LBA (logical block address) on the storage device (SSD). For SATA, the command is also called TRIM, for SAS, the command is called UNMAP. The TRIM/UNMAP command notifies the drive it no longer needs data in certain LBAs (Logical Block Address) which then free up a number of NAND pages.

The TRIM/UNMAP command needs to be supported by the OS, the drive, and the controller in order to work. The TRIM/UNMAP command could result in higher SSD performance from both the reduced data needed to be rewritten during garbage collection and the higher free space resulting on the drive. Current shipping Dell enterprise drives have high enough performance and endurance so they do not yet support these commands even if the OS supports them. These features are being investigated for subsequent Dell SSD offerings.

[Click here to go to list of questions](#)

15. How do SSDs maintain Data Integrity?

a. Dell SSD drives data integrity is maintained using the following methods:

- Robust ECC
- Data path CRC protection
- Multiple [metadata](#) & FW copy
- [metadata](#) checksum protection
- Robust voltage rail design to ensure stable power to NAND flash memory

b. Sudden Power Loss protection

Compared with hard-disk drives (HDDs), solid-state drives (SSDs) are more robust against shock, consume less power, faster access times, and better read performance. However, certain SSD designs have data and file system corruption challenges in the event of sudden power loss. An effective power failure data protection mechanism needs to function before and after a disruptive power failure in order to provide comprehensive data protection.

Dell Enterprise SSDs contain hardware and firmware based power failure data protection features. They include a power failure detection circuit that monitors the voltage supply and sends a signal to the SSD controller if the voltage drops below a predefined threshold. This would trigger the SSD to disconnect from the input power and start the necessary steps to move the temporary buffer data and [metadata](#) to NAND flash. An on-board power hold-up circuitry & capacitor are implemented to provide enough energy for this operation. The hold-up capacitor is overprovisioned multifold to guarantee enough energy for the life of the drive.

[Click here to go to list of questions](#)

16. How is SSD Sanitized?

SSDs can be sanitized by writing over the entire drive capacity several times. Dell is currently investigating the secure erase and self-encrypting features on SED (Self Encrypting Drive) SSDs for future releases. These techniques enable a faster and efficient way to sanitize a SSD.

[Click here to go to list of questions](#)

17. What are the recommended application tuning & OS settings?

- Aligned IO: Aligned IO can have a tremendous impact on SSD performance and endurance. Aligned IO for an SSD gives efficiency to the device for managing the NAND writes and can also boost SSD endurance by reducing the number of Read-Modify-Write operations that cause extra writes to occur in the background on the SSD
- Varying Queue Depths: Queue depth is an important factor for systems and storage devices. Efficiencies can be gained from increasing queue depth to the SSD devices which allow for more efficient handling of write operations and may also help reduce write amplification that can affect the endurance life of the SSD.
- Use TRIM: Refer section 15
- Disable disk defragmentation: On a magnetic drive, defragmentation organizes the drive in such a way that data sectors are close to one another to improve performance. However, on Solid State Drives, having the data close together makes no difference, since SSDs can access data at the same speed no matter where it is. Thus, defragmentation of SSDs isn't necessary and can actually cause additional unnecessary NAND wearing.
- Disable indexing: Indexing usually speeds up searching on HDDs. However, it is not advantageous on SSDs. Because indexing constantly tries to maintain a database of the files on the system and its properties, it causes a lot of small writes, at which SSDs do not excel. But, SSDs do excel at reading, and thus the drive will be able to access the data quickly, even without an index.

[Click here to go to list of questions](#)

18. What is Endurance Management?

The use of an endurance management algorithm ensures that sufficient Program/Erase (P/E) cycles are available for the warranty time period of the drive. The firmware will limit writes if a drive is written heavily. However, customers will rarely see performance throttling when an SSD is used under the intended application.

[Click here to go to list of questions](#)

19. What Warranty do Dell SSDs carry?

Dell SSDs come with a standard three years warranty (extendable up to five years on certain high performance enterprise SSDs). Enterprise Value, Read Intensive and Slim SSD Class Drives are not eligible for extended warranty coverage beyond the standard three year warranty.

[Click here to go to list of questions](#)