

SSD endurance

HP Workstations



Solid-state devices (SSDs) are quickly being adopted by many market segments due to performance benefits, more rugged design, and lower power consumption when compared to traditional hard disk drives (HDDs). However, one major difference between SSDs and HDDs is that the NAND used to store information in SSDs wears out. Unlike reliability, endurance is a measure of how long an SSD will last before it is worn out. Many factors contribute to how long or how short this time can be. In the sections below, we will explore some of these factors and recommend techniques for optimizing your SSD's endurance.

Table of contents

Why do SSDs wear out?	2
How is endurance specified?	2
How is life expectancy calculated?	2
Why are there different endurance ratings?	2
What do SSDs currently do to improve endurance?	3
How can I increase my SSD's lifetime?	3

Why do SSDs wear out?

The wear-out in an SSD occurs due to the way information is stored in the NAND. Before the NAND can be written, the page must first be empty. For a page to be empty, an entire block consisting of multiple pages must be erased. When data is modified, the data in the page is read. The data to be written is merged with the existing page and then written into another location in the NAND array. At some future time the remaining pages within the NAND block are cleaned up, the block is erased, and then returned to the free pool. This process from initially writing a page within a block to clearing the block so that it can be reused is called the program-erase (P/E) cycle. It is important to note that the P/E cycle rating of NAND is a lower limit. The number of P/E cycles a NAND block can sustain before failure is represented by the write endurance. Each NAND storage element has a finite number of P/E cycles. Over time, as these elements degrade, their ability to hold a charge diminishes. This characteristic of NAND flash memory is due to the physical properties of the semiconductor material and the slow erosion of the insulator as a result of repeated writes to the storage element.

How is endurance specified?

The endurance of a NAND block is specified by the number of P/E cycles which the architecture can tolerate. There are several ways that manufacturers specify the endurance of an SSD. The two most common methods are to specify drive writes per day (DWPD) or total bytes written (TBW). Usually, these values take into consideration the drive's rated lifetime. So, you may see a listing on a datasheet that specifies "40 GB/day for 3 years" or "1 DWPD for 5 years." Often times, it is not possible to directly convert from one method to another because we are missing some important details. Fortunately, there is a methodology to help calculate the life expectancy of an SSD.

How is life expectancy calculated?

The P/E cycle rating for NAND is only one piece of the total write endurance puzzle for an SSD.

When the host executes a write to an SSD, the controller of the drive must actually perform multiple writes to different parts of the NAND. The number of internal writes for every host write is known as write amplification. Write amplification can vary from 1x to 10x or higher depending on what is being written to the drive and what has previously been written to the drive.

To calculate the life expectancy of an SSD, the following formula can be used:

$$\text{years} = \left(\frac{\text{capacity} * \text{P/E cycles}}{\text{host writes per day} * \text{write amplification}} \right) / 365$$

For example, a 240 GB¹ SSD with a 5,000 P/E cycle NAND, a user workload of 120 GB a day, and a write amplification of 5x would have a life expectancy of 5.4 years before the device is worn out. However, a similar SSD with only a 500 P/E cycle NAND would have an expected life of just over 6 months.

Why are there different endurance ratings?

Different endurance ratings exist due to different use models for SSD storage. Since write workloads are not severe in a typical client desktop environment, lower-endurance devices are acceptable in this cost-sensitive market. Extremes of temperature and high sequential write workloads exist in the data center, justifying the higher cost of a higher endurance part. Knowing your workloads can help you choose the best endurance/performance mix for a given storage budget.

Another reason for different endurance ratings, even within a device class (enterprise, client, etc.), relates to the type of NAND technology used in the device. There are three primary classes of NAND: single-level cell (SLC), multi-level cell (MLC), and triple-level cell (TLC). Each class of NAND has a different P/E cycle rating. Generally, cost tends to be directly proportional to the number of P/E cycles. HP Workstation SSDs are strictly MLC devices at this time.

Table 1. The three types of SSD NAND

NAND type	Single-level cell (SLC)	Multi-level cell (MLC)	Triple-level cell (TLC)
Data storage	1 bit	2 bits	3 bits
Write endurance	~100,000 P/E cycles	3,000-10,000 P/E cycles	300-3,000 P/E cycles
Cost	High	Moderate	Low

What do SSDs currently do to improve endurance?

SSD and SSD controller designers have several methods at their disposal to help improve the endurance of a drive. Most of these methods involve minimizing the number of P/E cycles that the NAND sees. Some of the most common methods for increasing endurance are explained below. There are many others used such as Digital Signal Processing (DSP), improved Error Correction Code (ECC) and write amplification mitigation techniques which are beyond the scope of this paper.

Better wear leveling

Flash blocks start to suffer bit failures after a certain number of erase/write cycles. Without wear leveling, host applications would constantly write and rewrite a small number of blocks which would quickly wear them out. Wear leveling spreads the NAND flash erase/write cycles out evenly across all NAND locations.

There are two primary categories for wear leveling: static and dynamic.

In dynamic wear leveling, two types of blocks are included in the wear-leveling algorithm: those that have undergone a change and those that are unused. Blocks of data that don't change are left alone by dynamic wear-leveling routines. Dynamic wear leveling is most often employed in USB flash drives and flash cards.

In static wear leveling, all blocks across all available flash in the device participate in the wear-leveling operations. This ensures all blocks receive the same amount of wear. Static wear leveling is most often used in desktop and notebook SSDs.

External data buffering

Every erase/write cycle will weaken NAND flash. In order to reduce the erase/write cycles, drive designers employ a volatile data buffer (similar in architecture to main system memory) as a temporary collection and staging area for data. This RAM can reside inside of the drive's controller or in external DRAM chips mounted on the drive's printed circuit board. This RAM is used to gather several short writes into one larger write that is then sent to the NAND. This minimizes the amount of writes to the NAND and also increases the drive's performance. This process is called "write coalescing."

Another function of external RAM is to buffer several successive writes to the same sector without actually writing to the flash until the final "version" of the data is ready for placement into the NAND. From the host's perspective, the data is repeatedly writing to the same locations on the drive. In actuality, the host is writing over and over to the RAM. In this mode, several hundred writes to a single sector may turn into only a single write to the NAND.

Overprovisioning

Overprovisioning is one of the most common methods that SSD engineers employ to help an SSD have a longer life than the endurance rating the NAND supports. Overprovisioning is the employment of more flash storage than presented to the device's interface. This extra flash is included in the wear-leveling operation but does not add to the user data capacity. The more overprovisioned space an SSD has (up to a practical limit), the longer it will last.

It is easy to spot an overprovisioned drive by looking at its advertised user capacity. Typically, drives that are aligned on a binary boundary (e.g. 128 GB, 256 GB and 512 GB)¹ do not have any overprovisioned space. They can have a rudimentary pool of spare blocks, which may be included in wear-leveling operations. Drives that report non-binary aligned capacities (e.g. 180 GB, 240 GB and 480 GB)¹ are overprovisioned by the difference between the raw binary capacity and the advertised user capacity. For example, a drive with a user capacity of 240 GB will have 256 GB of total NAND and is thus overprovisioned by about 7%. Some high-endurance enterprise drives will take this a step further and report a 256 GB¹ raw capacity as 200 GB¹—essentially a 20% overprovisioning.

The main reason that not all drives are overprovisioned is simply cost. The end user is paying for storage space they cannot access in exchange for a longer drive life.

TRIM

A different way that blocks are freed up is through the use of the TRIM command. The TRIM command allows the operating system (OS) to tell the SSD which sectors are no longer being used. The SSD can then erase the blocks that the OS has no more use for and free them up for inclusion into wear-leveling operations. The SSD will free these blocks as part of a background operation during idle times. This also helps to improve drive performance similar to a defrag operation on a rotational hard drive. If you are using an SSD, do not perform a defragmentation operation on it. This will drastically shorten the drive's life and it is not necessary for SSDs.

How can I increase my SSD's lifetime?

In addition to the hard work that SSD designers put into prolonging the lifetime of their products, there are several things you can do to further increase your SSD's life.

Begin by giving your SSD some breathing room. If your application requires 120 GB to be consumed, opt for a 256 GB device over a 128 GB device.¹ Typically, smaller SSDs will wear out sooner than larger ones. This is due to two primary factors. First, smaller SSDs are usually filled very close to their maximum user data capacity. As a result, there is less space available for wear leveling and more erase/write cycles are required to perform wear leveling in the limited space. Second, smaller-capacity SSDs usually have fewer NAND channels to spread wear across multiple NAND die. Although there is additional cost for a larger-capacity device, the cost will be greatly offset by a lengthened replacement cycle.

Next, overprovision your drive. You can increase the lifetime by only partitioning a portion of the device's total capacity. For example, if you have a 256 GB drive— only partition it to 240 GB. This will greatly extend the life of the drive. A 20% overprovisioning level (partitioning only 200 GB) would extend the life further. A good rule of thumb is every time you double the drive's overprovisioning you add 1x to the drive's endurance.

Another way to extend the lifetime of your drive is to better understand your workloads. By performing an analysis of the intended workload environment, you can make better decisions about which type of drive to use and how much overprovisioning is required. For instance, If you are editing 4K video that requires many large sequential reads and writes, you would want an SSD with higher endurance. In these environments, an enterprise-class SSD would be a much more cost-effective choice.

If your workloads are primarily desktop productivity (email, spreadsheet, light engineering), the majority of disk accesses will be random and 4K or smaller in size. In these environments, a client-class device would perform very well. HP Workstations offer only MLC devices and carefully chooses only those drives that will offer you the best endurance and highest quality.

Finally, optimize your environment to enhance drive endurance. There are many operating system modifications that can be made to greatly increase the life of your SSD. Here are a few to consider:

- On a system with 8 GB or more of main memory, consider shutting off (or setting to minimum) the page file.
- Turn off system hibernation. Every time the system hibernates, it has to write the contents of system memory to the SSD.
- If you have an SSD as the boot drive, consider installing a HDD as a data drive and move your "My Documents" folder to the HDD. This will reduce the amount of writes to the SSD tremendously.
- Consider disabling both Superfetch and Prefetch. These acceleration tools put an increased write load on your SSD, especially if you have not disabled hibernation.
- In Windows 8/8.1², ensure that "optimize drives" is enabled. This will ensure that the drive is being "TRIMed."
- In Windows 7², ensure that "disk defrag" is disabled. Defragging an SSD prematurely wears it out due to repeated read-modify-write operations.
- Regularly run a monitor application to track the drive's wear. This will help you anticipate when to replace the drive before it wears out.

While endurance is an important consideration in selecting an SSD, the SSD selection process for HP Workstations considers a long list of demanding needs of our customers. Heavy workloads, extreme performance, enterprise quality and reliability, and end-to-end data protection all play a role in the SSD selection process. HP tests every storage device against industry, as well as HP, internal quality and reliability standards. As new technologies flood the industry, customers can be assured that their performance, quality, and reliability demands continue to take center stage with HP Workstations.

To read more about SSD technology for HP Workstations, go to h20195.www2.hp.com/V2/GetPDF.aspx/4AA3-8500ENW.pdf

¹ For hard drives and solid state drives, 1 GB = 1 billion bytes. TB = 1 trillion bytes. Actual formatted capacity is less. Up to 10 GB of system disk (for Windows 7) is reserved for system recovery software.

² Not all features are available in all editions or versions of Windows. Systems may require upgraded and/or separately purchased hardware, drivers and/or software to take full advantage of Windows functionality. See microsoft.com.

Sign up for updates
hp.com/go/getupdated

