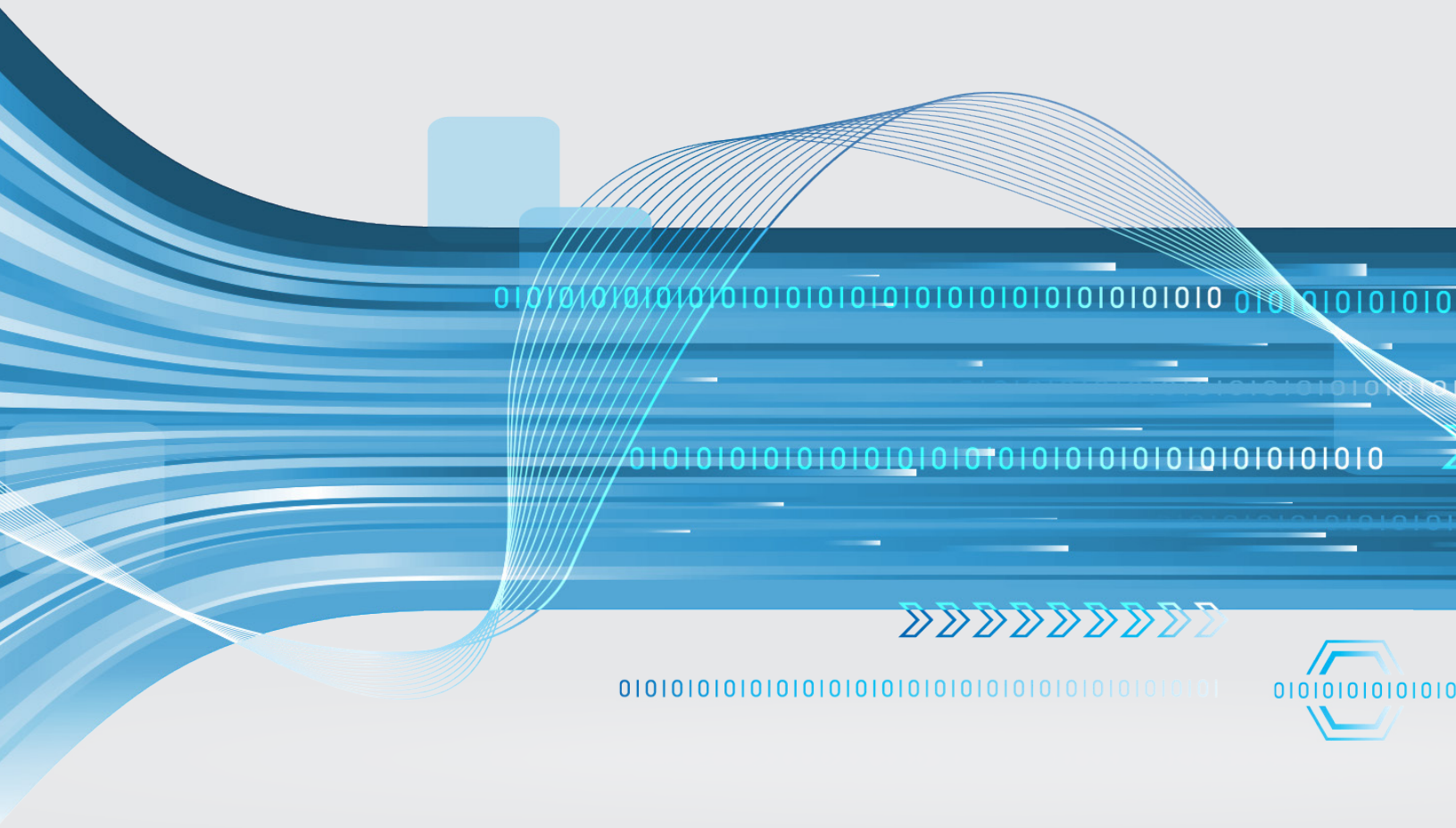
PNY®



The Benefits of Over-Provisioning on Solid-State Drives

Abstract

Some solid-state drive companies are releasing higher capacity SSDs giving the illusion that the buyer is receiving more product for their money, however this is not the whole story. All SSDs require what is known in the industry as Over-Provisioning (OP). Through our research at PNY Technologies, we have found that reducing the OP in order to achieve "Binary" capacities (128GB, 256GB, 512GB, etc.) sacrifices performance, endurance, and reliability for a small increase in capacity. In this white paper, we detail the reasons why a small sacrifice in capacity is worth it in the long run.

Introduction

A Solid-State Drive (SSD) is a storage device that typically uses NAND flash memory in order to store data. While this has significant advantages in terms of speed and reliability, NAND is not a perfect storage mechanism. It introduces some challenges that are solved by technological advancements in flash management work to ensure that your data is safe and secure. A critical aspect of the efficacy of standard SSD functionality is Over-Provisioning (OP).

The primary challenge of an SSD is that the NAND has a finite amount of Program/Erase cycles (P/E cycles) which adds difficulty in maintaining the endurance of the drive. At some point, the drive fills up with host writes and we need to make space for future writes by erasing the invalid data locations. However, since an erase on NAND can only be done at a block level, any valid pages on that block need to be moved to a new location before the block can be erased. This process of moving valid data to a new location to free up previously written blocks is called Garbage Collection. Additionally, some of the physical blocks may be written more frequently than the others thereby making it tough to predict the P/E cycles for any given location. SSDs address this concern by moving data around the NAND pool evenly such that no particular physical blocks are written to significantly more than any others. This process is known as Wear Leveling and is an integral part of the SSD's endurance.

When the drive fills up, the controller needs extra capacity to write new data to the SSD before garbage collecting the old data. The drive uses this additional space, called OP, in order to improve functionality of various processes. The larger the OP, the more available capacity there is to move the valid data without having to re-write multiple blocks. This means less garbage collection, which in turn means better performance, higher endurance, and a more reliable SSD.

Note: *Due to the binary/decimal conversion of storage capacity in the OS, all SSDs have what is known as "hidden" OP. For the sake of simplicity, this is removed from the calculation of OP and does not change the outcome of the white paper.*

Performance

One of the most readily visible impacts of lower OP is the performance of the drive. While the SSD is in a new, fresh-out-of-the-box state, the controller can write anywhere on the drive without having to incur Garbage Collection. There is generally no impact to the performance of an SSD while it is in this state, but it will not last for long. In a world where a single video game can take up to 80GB of data, the drive fills up with data quickly, and starts to run out of valid space. In our tests, a competitor's 128GB drive with 0% OP starts out with significantly lower performance to our 120GB drive with 7% OP, which is not a result of the amount of OP but the competitor's drive settings. As the drives move to a steady state after around two to three hours of writes, the performance starts to suffer. In our tests, reducing the capacity by 7% gives us over 80% performance improvement once the drive starts to fill up.

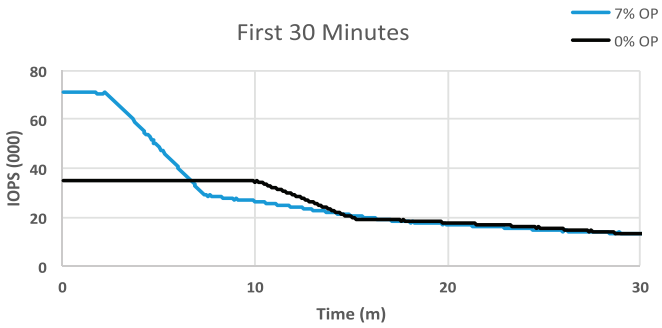


Figure 1: Performance in the first 30 minutes

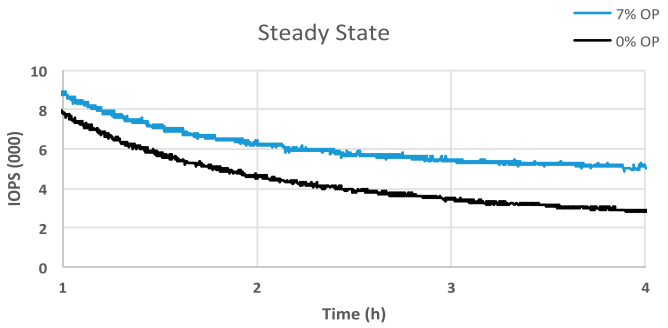


Figure 2: Steady state performance

Procedure (Implemented on both SSDs in the same order): *Secure erase the SSD, initialize the SSD in Windows drive management, using IOMeter ver. 1.1.0 with one worker, 32 outstanding I/Os, 4k transfer size, random workload, and Full Random IO data pattern, write for 4 hours.*

Endurance and Reliability

The endurance of the drive is impacted by a phenomenon known a Write Amplification. After Garbage Collection begins, saving a single 1GB file causes a chain reaction of data being moved that can ultimately write over 3GB to the NAND on a drive with no OP. This is quantified by a Write Amplification Factor (WAF), which in this example would be 3.0 as a single write causes 3x the data to be written to the NAND. Naturally, the more data that is written to the NAND, the shorter the drive life gets. When comparing a drive with 0% OP vs a drive with 7% OP, the 0% OP drive writes up to 50% more data to the NAND, which over the rated life of the drive can amount to over 10TB of unnecessary writes.

We define reliability as how many drives will reach their endurance and performance specifications. Simulated workloads can get us to a certain point in quantifying reliability, but in the real world, each drive is subjected to different workloads. Some are lighter workloads that are easy for the controller to manage, while heavier workloads are more difficult and incur a greater performance and endurance hit. Ultimately, over-provisioning will lighten the workload for the controller and make a more reliable drive.

$$\text{WAF} = \frac{\textit{Total Bytes Written to NAND}}{\textit{Total Bytes Written from Host}}$$

	120GB (7% OP)	128GB (0% OP)
NAND Writes	**42,112 GB**	**14,976 GB**
Host Writes	**22,558 GB**	**5,257 GB**
WAF	**1.87**	**2.85**

Procedure (Implemented on both SSDs in the same order): *Secure erase the SSD, run the JEDEC JESD219A workload for a large number of GB, calculate WAF on both drives.*

Conclusions

When analyzing solid-state drives, the outcome is rarely black-and-white. Some companies will argue that sacrificing performance, endurance, and reliability is worth the small increase in storage. However, at PNY we strive to make a drive that will outlive the system it's encased in. In reserving a small percentage of the total capacity of the drive, we can create a better, faster, and longer lasting product. Over the long term, PNY believes that over-provisioning will continue to be a leading technique to safeguard SSD performance and endurance. For the highest quality SSDs that boast longevity and quality at their core, OP is essential as NAND industry continues to shrink memory components to reduce cost, extend efficiencies, and scale over time