

SSD Makers Wrestle with Performance Degradation

September 30, 2009

By [Herman Mehling](#)

[EnterpriseStorageForum.com](#)

Solid state drives ([SSD](#)) are known for performance that is many times that of hard disk drives (HDDs), but what's not so well known is that SSD performance tends to degrade over time, and benchmarks show that SSDs can perform much better new than when heavily used.

The issue is an important one when making enterprise storage buying decisions, and it's created an opportunity for vendors that can develop SSDs that perform more consistently over time. Among the vendors that claim to have solved the problem are Fusion-io, Pliant Technology and STEC (NASDAQ: STEC).

SSDs suffer from a difficulty that doesn't exist in HDDs — the flash must be erased before new data can be written into it, said Jim Handy, an analyst at Objective Analysis, a market research firm specializing in SSDs and semiconductors.

"This erase, which can take up to a half second, would bring the SSD to its knees were it not for some clever work-arounds that SSD makers build into their controllers," said Handy. "One of these is to over-provision, to build more flash into the SSD than appears to the outside world."

SSD technologies typically suffer significant performance degradation over time — by as much as 50 percent or more — as more data is written to the [NAND](#) flash memory and as applications accessing the device vary the read-to-write ratio, said Greg Goelz, vice president of marketing at Pliant Technology.

"This 'performance droop' causes big issues for mission-critical, [I/O](#)-intensive data center and high-performance computing environments, which require consistent, predictable performance over time and across a wide range of workloads," said Goelz.

Problems with NAND Flash

A NAND flash cell is a small electrical storage device with a finite number of uses due to the effects of programming (removing the charge) and erasing the cell. During a program/erase event, the NAND flash cell can degrade to a point where too much energy is trapped in the cell.

"This means the cell cannot be drained and is stuck in a full state," said Lance Smith, senior vice president of product marketing for Fusion-io. "In other words, bits will remain a '0' for NAND flash."

Unlike traditional hard disk drives, SSDs must avoid writing repeatedly in the same location. Otherwise, a cell will wear out. SSD designers can avoid this problem by writing across the entire capacity of the drive before writing to the same location twice. This is called wear leveling.

A good design will attempt to perform the erasure well ahead of time to ensure the write event is not held up due to the lengthy amount of time it takes to perform the erasure, said Smith. Otherwise, write performance will be limited by the rate of erasures, which is much slower.

To handle these issues, SSD makers, including Fusion-io, have implemented a wear-leveling algorithm that creates an abstraction layer.

"Here, you have a logical block and a physical block," said Smith.

The logical block points data to a different physical cell with every write, ensuring that information is not erased and that cells experience consistent wear over time. A background maintenance application — called a groomer — reclaims erased blocks of data and moves data around the NAND flash chip as needed, maximizing the use of the space and ensuring that no data is erased.

The grooming process itself, however, can lead to reductions in data speed as data is coalesced to accommodate newer data and ensure data integrity on the NAND flash chips.

Over-Provisioning Offers a Solution

Smith said Fusion-io has addressed the issue by allowing customers to over-provision, giving them extra grooming space for data depending on their write needs.

Fusion-io's 80GB ioDrives are factory configured with 20 percent over-provisioning to accommodate typical usage in the enterprise environment.

However, for write-heavy applications, users can increase the amount of over-provisioning to 40, 50 or 60 percent to suit their needs. In this way, users can use exactly the amount of space they need for the write cycle to perform with maximum proficiency.

Consistent SSD Performance

Both Pliant and STEC have created SSDs with proprietary controllers and firmware designed to deliver consistent performance over time.

Pliant's new [Enterprise Flash Drives](#) (EFDs) have a number of unique features and techniques to eliminate performance droop, said Goelz.

"Our EFD delivers two to four times greater sustained I/O performance than today's fastest SSDs, providing consistent, predictable system performance across a wide range of workloads over an extended period of time," said Goelz.

The EFDs maintain this performance level whether reading or writing data, and even as enterprise applications vary the read-to-write ratio.

Goelz said Pliant's EFD is the only solution able to perform common tasks such as on-going memory reclaim and other data integrity management functions transparently in the background, without affecting I/O performance.

It also transparently manages a host of more advanced tasks, including background Patrol Read, triple redundant ECC (Error Correction Code) protected [metadata](#), and extended ECC to ensure data integrity without affecting performance.

Pliant's [SAS](#) interface enables EFDs to perform concurrent reading and writing operations at four times the link bandwidth of the single-port, half-duplex [SATA](#) interface, which is commonly used by competitive SSD products, according to Goelz.

STEC's Approach

"STEC's drives appear in almost all major storage OEM sockets for many reasons, with performance being one of the most important ones," said Scott Shadley, STEC's senior manager for SSD technical marketing.

"Our drives are designed to eliminate the performance problems that exist within many SSD products," said Shadley.

"How do we do this? Our drives are developed in such a way that they do not expect any type of idle time from the host system. This is vital, as a drive that expects 'free time' will have very different performance parameters."

Shadley said STEC drives operate under specific IP and technology that allow for all activities within the drive to work simultaneously without affecting the host.

STEC's drives also have built-in controllers and in-house firmware focused on allowing the drives to accept host commands, move data within the drive, perform ECC and other background activities under all workloads with no affect on throughput or performance to the host system. This is accomplished by significant design effort focused around the transactions within the host-drive interfaces.

The raw media functions within the drives are intentionally separated and buffered from the host signaling in case any slowing or issues presented on the media interfaces are not pushed through the controllers to the host. This prevents any slowing in drive performance.