

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

The Prelude

I spent about three weeks working on my review of Intel's X25-M SSD. Most of that was research and writing and re-writing to not only review the product but also to talk about the rest of the SSDs in the marketplace and their deficiencies. Truth be told, I spent more time working on SSDs that weren't the X25-M than the Intel drive itself. The Intel drive just worked as it should, the rest of them didn't.

If you read the article, you know I was pretty harsh on some of the SSDs out at the time and if you've ever used any of those SSDs, you know why. Needless to say, there was some definite fallout from that review. I'm used to negative manufacturer response after a GPU review, but I'm always a bit surprised when it happens in any other segment.

I took a day or two off after that review went live, I think it was a day. Afterwards, I immediately started working on a follow-up. There was a strange phenomenon a few people noticed, something I unfortunately picked up on after the review went live; if you filled the X25-M up and re-benchmarked it, it got slower. And I had no idea why.

A few weeks later, I had it figured out. But then Nehalem was right around the corner. I'd tackle it after that. But then a new batch of SSDs from OCZ and other vendors were almost ready. I told myself I'd do them all at the same time. Then CES happened.

The More Things Change, The More They Stay The Same

If you look at the SSD market today, you'd assume that it's very different from what it was just six months ago when the X25-M launched. People are worried that the Intel drive has issues with degrading performance over time. Some vendors are now shipping "revised" JMicron drives with multiple controllers, supposedly fixing all of the problems I talked about last year.

I hate to break it to you guys. As different as the world may seem today, it's all very much the same.

The Intel drive is still the best of the best. Yes, it, and other SSDs do get slower over time and later in this article I'll explain why it happens and why it's not as big of a deal as you'd think. The issues I complained about with the JMicron drives from last year are still alive and well today; they're just somewhat occluded.

Delay after delay kept me from writing this article, but I believe it's for the best. What went in to what you're about to read is nearly six months of research, testing and plain old work with SSDs.

The timing couldn't be more perfect. We're about to see a new wave of SSDs hit the market and it's time to separate the fact from the fiction, the benchmarks from reality and the men from the boys. The last time I wrote an article about SSDs I ruffled quite a few feathers. That's never my aim, but we'll see what comes of this one.

Bringing You Up to Speed: The History Lesson

Everyone remembers their first bike right? Mine was red. It had training wheels. I never really learned how to ride it, not that I didn't go outdoors, I was just too afraid to take those training wheels off I guess. That was a long time ago, but I remember my first bike.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I also remember my first SSD. It was a 1.8" PATA drive made by Samsung for the MacBook Air. It was lent to me by a vendor so I could compare its performance to the stock 1.8" mechanical HDD in the Air.

The benchmarks for that drive didn't really impress. Most application tests got a little slower and transfer speeds weren't really any better. Application launch times and battery life both improved, the former by a significant amount. But the drive was expensive; \$1000 from Apple and that's if you bought it with the MacBook Air. Buying it from a vendor would set you back even more. It benchmarked faster than hard drive, but the numbers didn't justify the cost. I pulled the drive out and sent it back after I was done with the review.

The next time I turned on my MacBook Air I thought it was broken. It took an eternity to boot and everything took forever to launch. Even though the benchmarks showed the SSD shaving off a few seconds of application launch time here and there, in the real world, it was noticeable. The rule of thumb is that it takes about a 10% difference in performance for a user to notice. The application tests didn't show a 10% difference in performance, but the application launch tests, those were showing 50% gains. It still wasn't worth \$1000, but it was worth a lot more than I originally thought.

It was the MacBook Air experience that made me understand one important point about SSDs: you don't think they're fast, until I take one away from you.

My second SSD was a 60GB SuperTalent drive. I built a HTPC using it. It was my boot drive and I chose it because it drew less power and was silent; it helped keep my HTPC cool and I wouldn't have to worry about drive crunching while watching a movie. My movies were stored elsewhere so the space didn't really matter. The experience was good, not great because I wasn't really hitting the drive for data, but it was problem-free.

SuperTalent was the first manufacturer to sell a SSD in a 3.5" enclosure, so when they announced their 120GB drive I told them I'd like to do a review of their SSD in a desktop. They shipped it to me and I wrongly assumed that it was the same as the 60GB drive in my HTPC just with twice the flash.

This drive did have twice the flash, but it was MLC (Multi-Level Cell) flash. While the 60GB drive I had was a SLC drive that used Samsung's controller, the MLC drive used a little known controller from a company called JMicron. Samsung had a MLC controller at the time but it was too expensive than what SuperTalent was shooting for. This drive was supposed to be affordable, and JMicron delivered an affordable controller.

After running a few tests, the drive went in my Mac Pro as my boot/application drive. I remembered the lesson I learned from my first SSD. I wasn't going to be able to fairly evaluate this drive until I really used it, then took it away. Little did I know what I was getting myself into.

The first thing I noticed about the drive was how fast everything launched. This experience was actually the source of my SSD proof-of-value test; take a freshly booted machine and without waiting for drive accesses to stop, launch every single application you want to have up and running at the same time. Do this on any system with a HDD and you'll be impatiently waiting. I did it on the SuperTalent SSD and, wow, everything just popped up. It was like my system wasn't even doing anything. Not even breaking a sweat.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I got so excited that I remember hopping on AIM to tell someone about how fast the SSD was. I had other apps running in the background and when I went to send that first IM and my machine paused. It was just for a fraction of a second, before the message I'd typed appeared in my conversation window. My system just paused.

Maybe it was a fluke.

I kept using the drive, and it kept happening. The pause wasn't just in my IM client, it would happen in other applications or even when switching between apps. Maybe there was a strange OS X incompatibility with this SSD? That'd be unfortunate, but also rather unbelievable. So I did some digging.

Others had complained about this problem. SuperTalent wasn't the only one to ship an affordable drive based on this controller; other manufacturers did as well. G.Skill, OCZ, Patriot and SiliconPower all had drives shipping with the same controller, and every other drive I tested exhibited the same problem.

I was in the midst of figuring out what was happening with these drives when Intel contacted me about reviewing the X25-M, its first SSD. Up to this point Intel had casually mentioned that their SSD was going to be different than the competition and prior to my JMicron experience I didn't really believe them. After all, how hard could it be? Drive controller logic is nowhere near as complicated as building a Nehalem, surely someone other than Intel could do a good-enough job.

After my SuperTalent/JMicron experience, I realized that there was room for improvement.

Drive vendors were mum on the issue of pausing or stuttering with their drives. Lots of finger pointing resulted. It was surely Microsoft's fault, or maybe Intel's. But none of the Samsung based drives had these problems.

Then the issue was cache. The JMicron controller used in these drives didn't support any external DRAM. Intel and Samsung's controllers did. It was cache that caused the problems, they said. But Intel's drive doesn't use the external DRAM for user data.

Fingers were pointed everywhere, but no one took responsibility for the fault. To their credit, OCZ really stepped up and took care of their customers that were unhappy with their drives. Despite how completely irate they were at my article, they seemed to do the right thing after it was published. I can't say the same for some of the other vendors.

The issue ended up being random write performance. These "affordable" MLC drives based on the JMicron controller were all tuned for maximum throughput. The sequential write speed of these drives could easily match and surpass that of the fastest hard drives:

If a company that had never made a hard drive before could come out with a product that on its first revision could outperform WD's VelociRaptor and be more reliable thanks to zero moving parts...well, you get the picture. Optimize for sequential reads and writes!

The problem is that modern day OSes tend to read and write data very randomly, albeit in specific areas of the disk. And the data being accessed is rarely large, it's usually very small on the order of a few KB in size. It's these sorts of accesses that no one seemed to think about; after all these vendors and controller manufacturers were used to making USB sticks and CF cards, not hard drives.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

	Sequential Read Performance
JMicron JMF602B MLC	134.7 MB/s
Western Digital VelociRaptor 300GB	118 MB/s

The chart above shows how much faster these affordable MLC SSDs were than the fastest 3.5" hard drive in sequential reads, but now look at random write performance:

	Random Write Latency	Random Write Bandwidth
JMicron JMF602B MLC	532.2 ms	0.02 MB/s
Western Digital VelociRaptor 300GB	7.2 ms	1.63 MB/s

While WD's VelociRaptor averaged less than 8ms to write 4KB, these JMicron drives took around 70x that! Let me ask you this, what do you notice more - things moving very fast or things moving very slow?

The traditional hard drive benchmarks showed that these SSDs were incredible. The real world usage and real world tests disagreed. Storage Review was one of the first sites to popularize real world testing of hard drives nearly a decade ago. It seems that we'd all forgotten the lessons they taught us.

Random write performance is quite possibly the most important performance metric for SSDs these days. It's what separates the drives that are worth buying from those that aren't. All SSDs at this point are luxury items, their cost per GB is much higher than that of conventional hard drives. And when you're buying a luxury anything, you don't want to buy a lame one.

	Cost Per GB from Newegg.com
Intel X25-E 32GB	\$12.88
Intel X25-M 80GB	\$4.29
OCZ Solid 60GB	\$2.33
OCZ Apex 60GB	\$2.98
OCZ Vertex 120GB	\$3.49
Samsung SLC 32GB	\$8.71
Western Digital Caviar SE16 640GB	\$0.12
Western Digital VelociRaptor 300GB	\$0.77

Why You Should Want an SSD

For the past several months I've been calling SSDs the single most noticeable upgrade you can do to your computer. Whether desktop or laptop, stick a good SSD in there and you'll notice the difference.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I'm always angered by the demos in any Steve Jobs keynote. Not because the demos themselves are somehow bad, but because Jobs always has a perfectly clean machine to run the demos on - and multiple machines at that. Anyone who has built a computer before knows the glory of a freshly installed system; everything just pops up on your screen. Applications, windows, everything - the system is just snappy.

Of course once you start installing more applications and have more things running in the background, your system stops being so snappy and you tend to just be thankful when it doesn't crash.

A big part of the problem is that once you have more installed on your system, there are more applications sending read/write requests to your IO subsystem. While our CPUs and GPUs thrive on being fed massive amounts of data in parallel, our hard drives aren't so appreciative of our multitasking demands. And this is where SSDs truly shine.

Before we go too far down the rabbit hole I want to share a few numbers with you.

This is Western Digital's VelociRaptor. It's a 300GB drive that spins its platters at 10,000RPM and is widely considered the world's fastest consumer desktop hard drive.



The 300GB VelociRaptor costs about \$0.77 per GB.

This is the Intel X25-M. The Conroe of the SSD world, the drive I reviewed last year. It costs about \$4.29 per GB; that's over 5x the VelociRaptor's cost per GB.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



The VelociRaptor is the dominant force in the consumer HDD industry and the X25-M is the svelte bullfighter of the SSD world.

Whenever anyone mentions a more affordable SSD you always get several detractors saying that you could easily buy 2 VelociRaptors for the same price. Allow me to show you one table that should change your opinion.

This is the Average Read Access test from Lavalys' Everest Disk benchmark. The test simply writes a bunch of files at random places on the disk and measures how long it takes to access the files.

Measuring random access is very important because that's what generally happens when you go to run an application while doing other things on your computer. It's random access that feels the slowest on your machine.

	Random Read Latency in ms
Intel X25-M	0.11 ms
Western Digital VelociRaptor	6.83 ms

The world's fastest consumer desktop hard drive, Western Digital's 300GB VelociRaptor can access a random file somewhere on its platters in about 6.83ms; that's pretty quick. Most hard drives will take closer to 8 or 9ms in this test. The Intel X25-M however? 0.11ms. The fastest SSDs can find the data you're looking for in around 0.1ms. That's an order of magnitude faster than the fastest hard drive on the market today.

The table is even more impressive when you realize that wherever the data is on your SSD, the read (and write) latency is the same. While HDDs are fastest when the data you

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

want is in the vicinity of the read/write heads, all parts of a SSD are accessed the same way. If you want 4KB of data, regardless of where it is, you'll get to it at the same speed from a SSD.

The table below looks at sequential read, sequential write and random write performance of these two kings of their respective castles. The speeds are in MB/s.

	Sequential Read (2MB Block)	Sequential Write (2MB Block)	Random Write (4KB Block)
Intel X25-M	230 MB/s	71 MB/s	23 MB/s
Western Digital VelociRaptor	118 MB/s	119 MB/s	1.6 MB/s

If you're curious, these numbers are best case scenario for the VelociRaptor and worst case scenario for the X25-M (I'll explain what that means later in the article). While the VelociRaptor is faster in the large block sequential writes look at the sequential read and random write performance. The X25-M destroys the VelociRaptor in sequential reads and is an order of magnitude greater in random write performance. If you're curious, it's the random write performance that you're most likely to notice and that's where a good SSD can really shine; you write 4KB files far more often than you do 2MB files while using your machine.

If the table above doesn't convince you, let me share one more datapoint with you. Ever play World of Warcraft? What we're looking at here is the amount of time it takes to get from the character selection screen into a realm with everything loaded. This is on a fully configured system with around 70GB of applications and data as well as real time anti-virus scanning going on in the background on every accessed file.

	WoW Enter Realm Time in Seconds
Intel X25-M	4.85 s
Western Digital VelociRaptor	12.5 s

The world's fastest hard drive gets us into the game in 12.5 seconds. The Intel X25-M does it in under 5.

SSDs make Vista usable. It doesn't matter how much background crunching the OS is doing, every application and game launches as if it were the only thing running on the machine. Everything launches quickly. Much faster than on a conventional hard drive. If you have the ability, try using your system with a SSD for a day then go back to your old hard drive; if that test doesn't convince you, nothing will. That's just a small taste of why you'd want an SSD, now let's get back to finding a good one.

Hey, There's an Elephant in the Room

When the first X25-M reviews went live a few people discovered something very important, something many of us (myself included) missed and should've addressed: the drive got

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

slower the more you filled it up. It's no great mystery why this happened, but it seemed odd at the time because it went against conventional thinking.



LegitReviews was one of the first to spot the SSD slowdown phenomenon, good work Nate.

It's worth mentioning that hard drives suffer from the same problem; just for a different reason.

Hard drives store data on platters; the platters rotate while an arm with read/write heads on it hovers over the surface of the platter and reads data while the platter spins. The diameter of the platter is greater the further out on the platter you go, that's just how circles work. The side effect is that for the same amount of rotation, the heads can cover more area on the outside of the platter than on the inside.

The result is that transfer speeds are greater on the outer sectors of the platter than on the inner ones. OSes thus try to write as much data to the outer sectors as possible, but like beachfront property - there's only a limited amount of space. Eventually you have to write to the slower parts of the drive and thus the more full your drive is, the slower your transfer rates will be for data stored in the innermost sectors.

Fragmentation also hurts hard drive performance. While modern day hard drives have gotten pretty quick at transferring large amounts of data stored sequentially, spread the data out all around the platter and things get real slow, real fast.

Randomness is the enemy of rotational storage.

Solid state drives aren't supposed to have these issues. Data is stored in flash, so it doesn't matter where it's located, you get to it at the same speed. SSDs have +5 armor

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

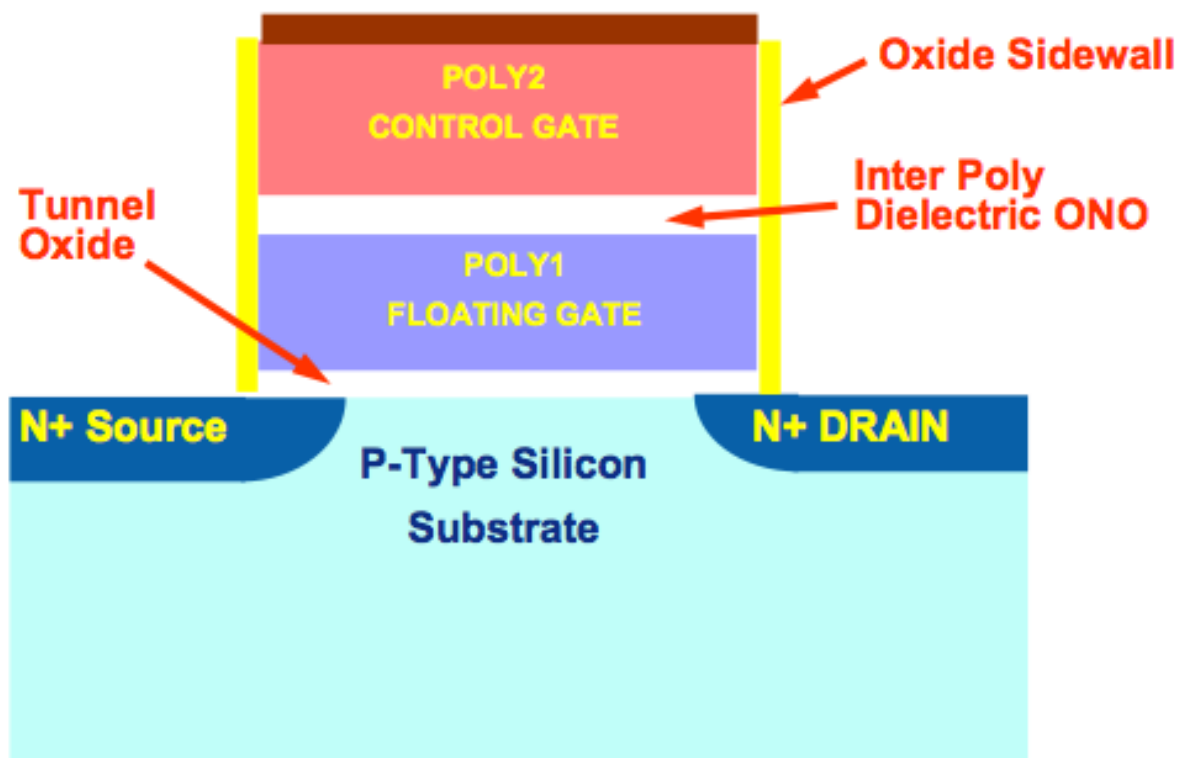
immunity to random access latency (that's got to be the single most geeky-sounding thing I've ever written, and I use words like latency a lot).

So why is it that when you fill up a SSD like Intel's X25-M that its performance goes down? Even more worrisome, why is it that when you delete data from the drive that its performance doesn't go back up?

While SSDs are truly immune to the same problems that plague HDDs, they do also get slower over time. How can both be true? It's time for another lesson in flash.

The Anatomy of an SSD

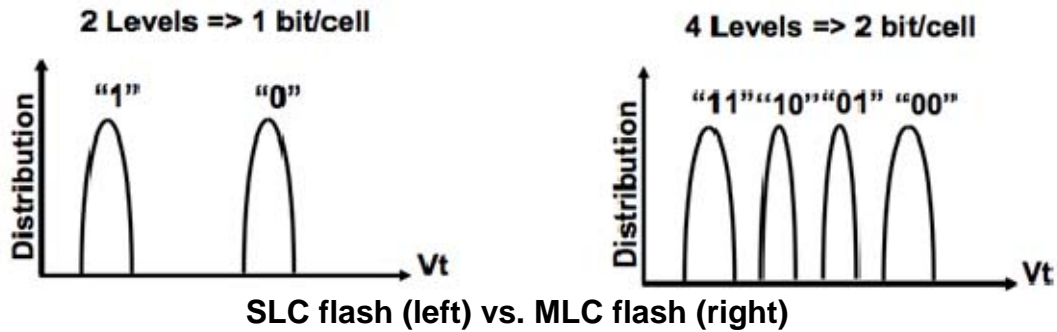
Let's meet Mr. N-channel MOSFET again:



This is the building block of NAND-flash; one transistor is required per cell. A single NAND-flash cell can either store one or two bits of data. If it stores one, then it's called a Single Level Cell (SLC) flash and if it stores two then it's a Multi Level Cell (MLC) flash. Both are physically made the same way; in fact there's nothing that separates MLC from SLC flash, it's just a matter of how the data is stored in and read from the cell.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Flash is read from and written to in a guess-and-test fashion. You apply a voltage to the cell and check to see how it responds. You keep increasing the voltage until you get a result.

	SLC NAND flash	MLC NAND flash
Random Read	25 μ s	50 μ s
Erase	2ms per block	2ms per block
Programming	250 μ s	900 μ s

With four voltage levels to check, MLC flash takes around 3x longer to write to as SLC. On the flip side you get twice the capacity at the same cost. Because of this distinction, and the fact that even MLC flash is more than fast enough for a SSD, you'll only see MLC used for desktop SSDs while SLC is used for enterprise level server SSDs.

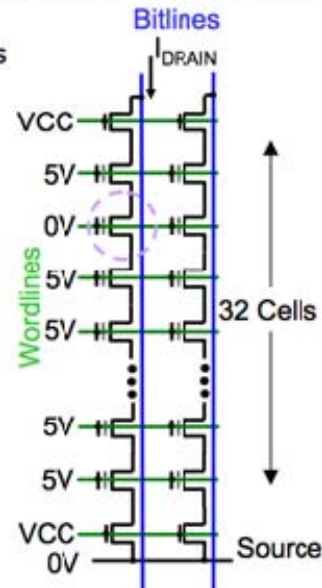
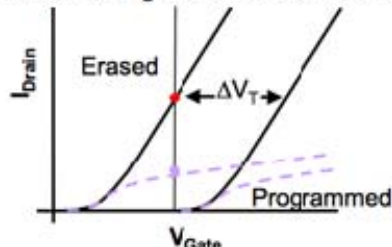
Array Architecture & Read Operation

Array Architecture:

- Cells are in rectangular arrays of wordlines (gates), bitlines (drains), and common source
- Product = many "blocks" of ~1M cells each
- In NAND block, cells on a bitline are in *strings* of typ. 32 cells in series, as with NMOS transistors in a NAND gate
- NAND has NMOS select devices at source and drain, to isolate one block from another

Read:

- Cells are read by applying a gate voltage and testing the drain (bitline) current
- NAND: 31 cells are on, so the drain current is zero if the selected cell is OFF and is determined by the series resistance of the string if the selected cell is ON



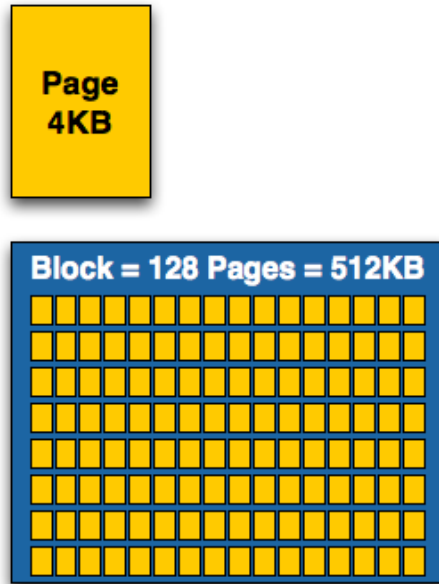
Cells are strung together in arrays as depicted in the image to the right

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

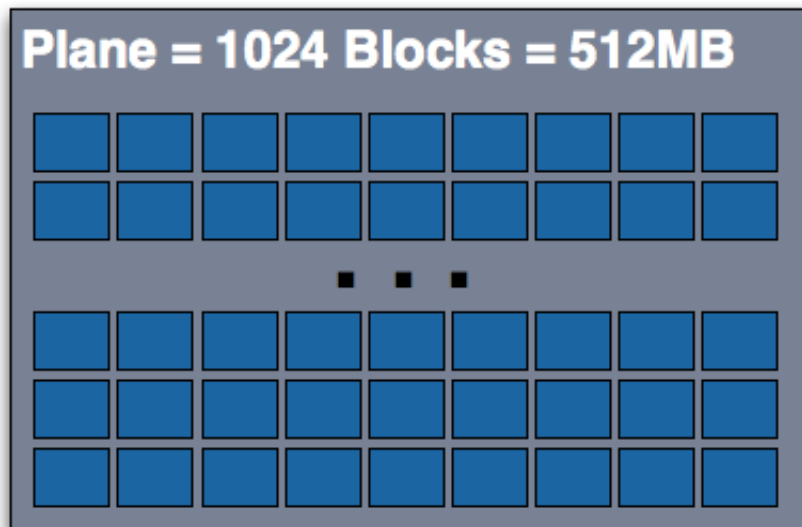
So a single cell stores either one or two bits of data, but where do we go from there? Groups of cells are organized into pages, the smallest structure that's readable/writable in a SSD. Today 4KB pages are standard on SSDs.

Pages are grouped together into blocks; today it's common to have 128 pages in a block (512KB in a block). A block is the smallest structure that can be erased in a NAND-flash device. So while you can read from and write to a page, you can only erase a block (128 pages at a time). This is where many of the SSD's problems stem from, I'll repeat this again later because it's one of the most important parts of understanding SSDs.



Arrays of cells are grouped into a page, arrays of pages are grouped into blocks

Blocks are then grouped into planes, and you'll find multiple planes on a single NAND-flash die.



The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

The combining doesn't stop there; you can usually find either one, two or four die per package. While you'll see a single NAND-flash IC, there may actually be two or four die in that package. You can also stack multiple ICs on top of each other to minimize board real estate usage.

Strength in Numbers, What makes SSDs Fast

Given the way a single NAND-flash IC is organized one thing should come to mind: parallelism.

Fundamentally the flash that's used in SSDs cut from the same cloth as the flash that's used in USB drives. And if you've ever used a USB flash drive you know that those things aren't all that fast. Peak performance to a single NAND-flash IC is going to be somewhere in the 5 - 40MB/s range. You get the faster transfer rates by reading/writing in parallel to multiple die in the same package.

The real performance comes from accessing multiple NAND ICs concurrently. If each device can give you 20MB/s of bandwidth and you've got 10 devices you can access at the same time, that's 200MB/s of bandwidth. While hard drives like reads/writes to be at the same place on the drive, SSDs don't mind; some are even architected to prefer that data be spread out all over the drive so it can hit as many flash devices as possible in tandem. Most drives these days have 4 - 10 channel controllers.

The Recap

I told you I'd mention this again because it's hugely important, so here it is:

A single NAND flash die is subdivided into blocks. The typical case these days is that each block is 512KB in size. Each block is further subdivided into pages, with the typical page size these days being 4KB.

Now you can read and write to individual pages, so long as they are empty. However once a page has been written, it can't be overwritten, it must be erased first before you can write to it again. And therein lies the problem, the smallest structure you can erase in a NAND flash device today is a block. **Once more, you can read/write 4KB at a time, but you can only erase 512KB at a time.**

It gets worse. Every time you erase a block, you reduce the lifespan of the flash. Standard MLC NAND flash can only be erased 10,000 times before it goes bad and stops storing data.

Based on what I've just told you there are two things you don't want to do when writing to flash: 1) you don't want to overwrite data, and 2) you don't want to erase data. If flash were used as a replacement for DVD-Rs then we wouldn't have a problem, but it's being used as a replacement for conventional HDDs. Who thought that would be a good idea?

It turns out that the benefits are more than worth the inconvenience of dealing with these pesky rules; so we work around them.

Most people don't fill up their drives, so SSD controller makers get around the problem by writing to every page on the drive before ever erasing a single block.

If you go about using all available pages to write to and never erasing anything from the drive, you'll eventually run out of available pages. I'm sure there's a fossil fuel analogy

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

somewhere in there. While your drive won't technically be full (you may have been diligently deleting files along the way and only using a fraction of your drive's capacity), eventually every single block on your drive will be full of both valid and invalid pages.

In other words, even if you're using only 60% of your drive, chances are that 100% of your drive will get written to simply by day to day creation/deletion of files.

The Blind SSD

Modern OSes talk to hard drives using logical block addressing. While hard drives are rotational media, logical block addressing organizes sectors on a hard drive linearly. When you go to save a file, Windows simply issues a write command for your file at a specific logical block address, say LBA 15 for example.

Your OS knows what LBAs are available and which ones are occupied. When you delete a file, the LBAs that point to that file on your hard disk are listed as available. The data you've deleted hasn't actually been removed and it doesn't get wiped until those sectors on the drive are actually overwritten.

Believe it or not, SSDs actually work the same way.

The flash translation layer in a SSD controller maps LBAs to pages on the drive. The table below explains what happens to the data on the SSD depending on the action in the OS:

Action in the OS	Reaction on a HDD	Reaction on an SSD
File Create	Write to a Sector	Write to a Page
File Overwrite	Write new data to the same Sector	Write to a Different Page if possible, else Erase Block and Write to the Same Page
File Delete	Nothing	Nothing

When you delete a file in your OS, there is no reaction from either a hard drive or SSD. It isn't until you overwrite the sector (on a hard drive) or page (on a SSD) that you actually lose the data. File recovery programs use this property to their advantage and that's how they help you recover deleted files.

The key distinction between HDDs and SSDs however is what happens when you overwrite a file. While a HDD can simply write the new data to the same sector, a SSD will allocate a new (or previously used) page for the overwritten data. The page that contains the now invalid data will simply be marked as invalid and at some point it'll get erased.

Putting Theory to Practice: Understanding the SSD Performance Degradation Problem

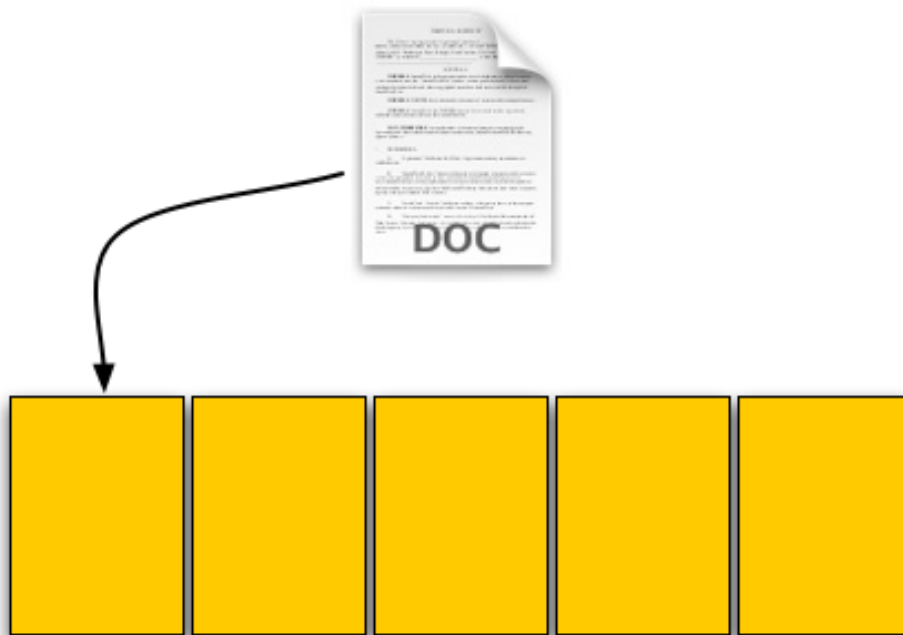
Let's look at the problem in the real world. You, me and our best friend have decided to start making SSDs. We buy up some NAND-flash and build a controller. The table below summarizes our drive's characteristics:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

	Our Hypothetical SSD
Page Size	4KB
Block Size	5 Pages (20KB)
Drive Size	1 Block (20KB)
Read Speed	2 KB/s
Write Speed	1 KB/s

Through impressive marketing and your incredibly good looks we sell a drive. Our customer first goes to save a 4KB text file to his brand new SSD. The request comes down to our controller, which finds that all pages are empty, and allocates the first page to this text file.



Our SSD. The yellow boxes are empty pages

The user then goes and saves an 8KB JPEG. The request, once again, comes down to our controller, and fills the next two pages with the image.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



The picture is 8KB and thus occupies two pages, which are thankfully empty

The OS reports that 60% of our drive is now full, which it is. Three of the five open pages are occupied with data and the remaining two pages are empty.

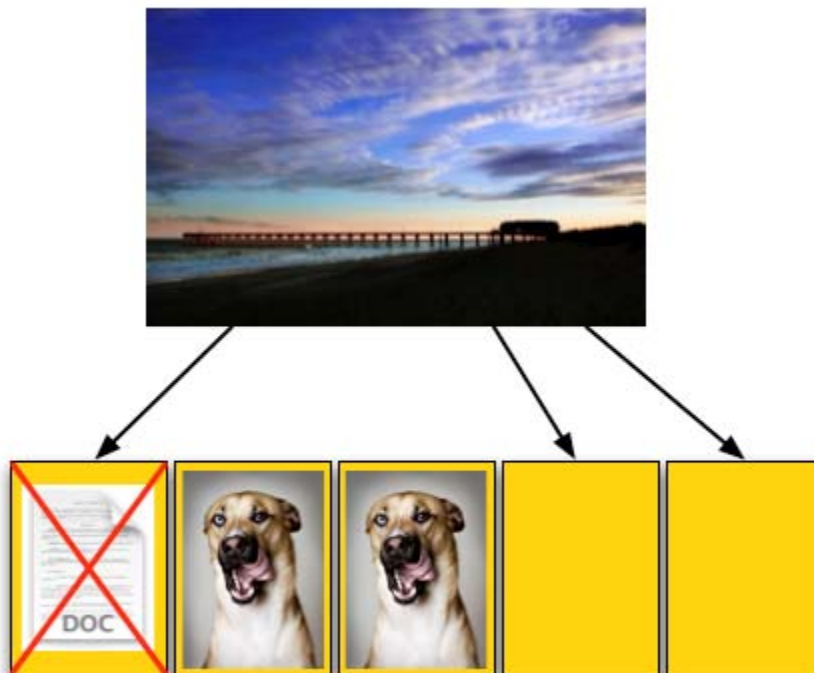
Now let's say that the user goes back and deletes that original text file. This request doesn't ever reach our controller, as far as our controller is concerned we've got three valid and two empty pages.

For our final write, the user wants to save a 12KB JPEG, that requires three 4KB pages to store. The OS knows that the first LBA, the one allocated to the 4KB text file, can be overwritten; so it tells our controller to overwrite that LBA as well as store the last 8KB of the image in our last available LBAs.

Now we have a problem once these requests get to our SSD controller. We've got three pages worth of write requests incoming, but only two pages free. Remember that the OS knows we have 12KB free, but on the drive only 8KB is actually free, 4KB is in use by an invalid page. We need to erase that page in order to complete the write request.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Uhoh, problem. We don't have enough empty pages.

Remember back to Flash 101, even though we have to erase just one page we can't; you can't erase pages, only blocks. We have to erase all of our data just to get rid of the invalid page, then write it all back again.

To do so we first read the entire block back into memory somewhere; if we've got a good controller we'll just read it into an on-die cache (steps 1 and 2 below), if not hopefully there's some off-die memory we can use as a scratch pad. With the block read, we can modify it, remove the invalid page and replace it with good data (steps 3 and 4). But we've only done that in memory somewhere, now we need to write it to flash. Since we've got all of our data in memory, we can erase the entire block in flash and write the new block (step 5).

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Now let's think about what's just happened. As far as the OS is concerned we needed to write 12KB of data and it got written. Our SSD controller knows what really transpired however. In order to write that 12KB of data we had to first read 12KB then write an entire block, or 20KB.

Our SSD is quite slow, it can only write at 1KB/s and read at 2KB/s. Writing 12KB should have taken 12 seconds but since we had to read 12KB and then write 20KB the whole operation now took 26 seconds.

To the end user it would look like our write speed dropped from 1KB/s to 0.46KB/s, since it took us 26 seconds to write 12KB.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

Are things starting to make sense now? This is why the Intel X25-M and other SSDs get slower the more you use them, and it's also why the write speeds drop the most while the read speeds stay about the same. When writing to an empty page the SSD can write very quickly, but when writing to a page that already has data in it there's additional overhead that must be dealt with thus reducing the write speeds.

Free Space to the Rescue

There's not much we can do about the scenario I just described; you can't erase individual pages, that's the reality of NAND-flash. There are some things we can do to make it better though.

The most frequently used approach is to under provision the drive. Let's say we only shipped our drive with 20KB of space to the end user, but we actually had 24KB of flash on the drive. The remaining 4KB could be used by our controller; how, you say?

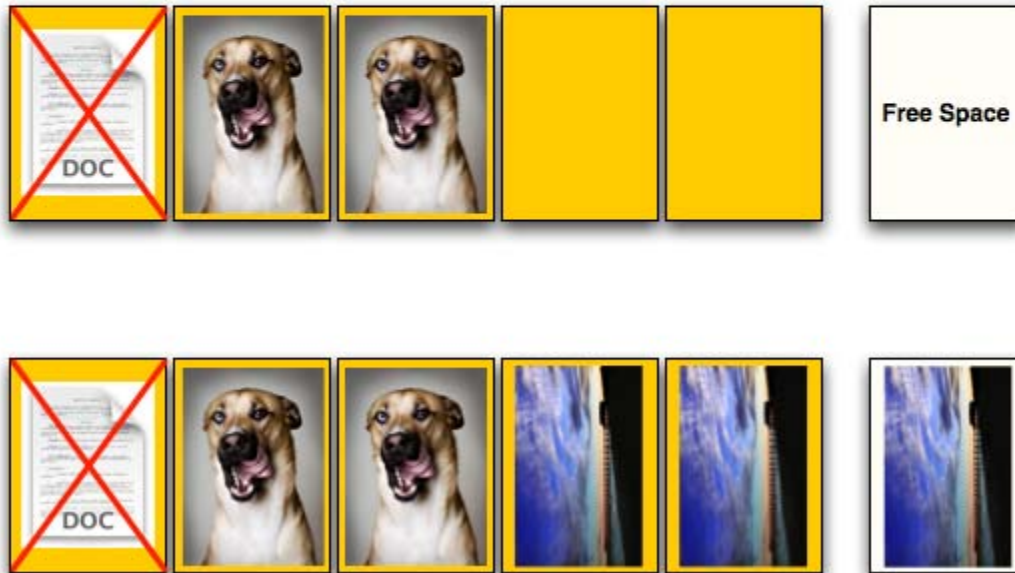
In the scenario from the last page we had to write 12KB of data to our drive, but we only had 8KB in free pages and a 4KB invalid page. In order to write the 12KB we had to perform a read-modify-write which took over twice as long as a 12KB write should take.

If we had an extra 4KB of space our 12KB write from earlier could've proceeded without a problem. Take a look at how it would've worked:

We'd write 8KB to the user-facing flash, and then the remaining 4KB would get written to the overflow flash. Our write speed would still be 12KB/s and everything would be right in the world.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Now if we deleted and tried to write 4KB of data however, we'd run into the same problem again. We're simply delaying the inevitable by shipping our drive with an extra 4KB of space.

The more spare-area we ship with, the longer our performance will remain at its peak level. But again, you have to pay the piper at some point.

Intel ships its X25-M with 7.5 - 8% more area than is actually reported to the OS. The more expensive enterprise version ships with the same amount of flash, but even more spare area. Random writes all over the drive are more likely in a server environment so Intel keeps more of the flash on the X25-E as spare area. You're able to do this yourself if you own an X25-M; simply perform a secure erase and immediately partition the drive smaller than its actual capacity. The controller will use the unpartitioned space as spare area.

The Trim Command: Coming Soon to a Drive Near You

We run into these problems primarily because the drive doesn't know when a file is deleted, only when one is overwritten. Thus we lose performance when we go to write a new file at the expense of maintaining lightning quick deletion speeds. The latter doesn't really matter though, now does it?

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

There's a command you may have heard of called TRIM. The command would require proper OS and drive support, but with it you could effectively let the OS tell the SSD to wipe invalid pages before they are overwritten.

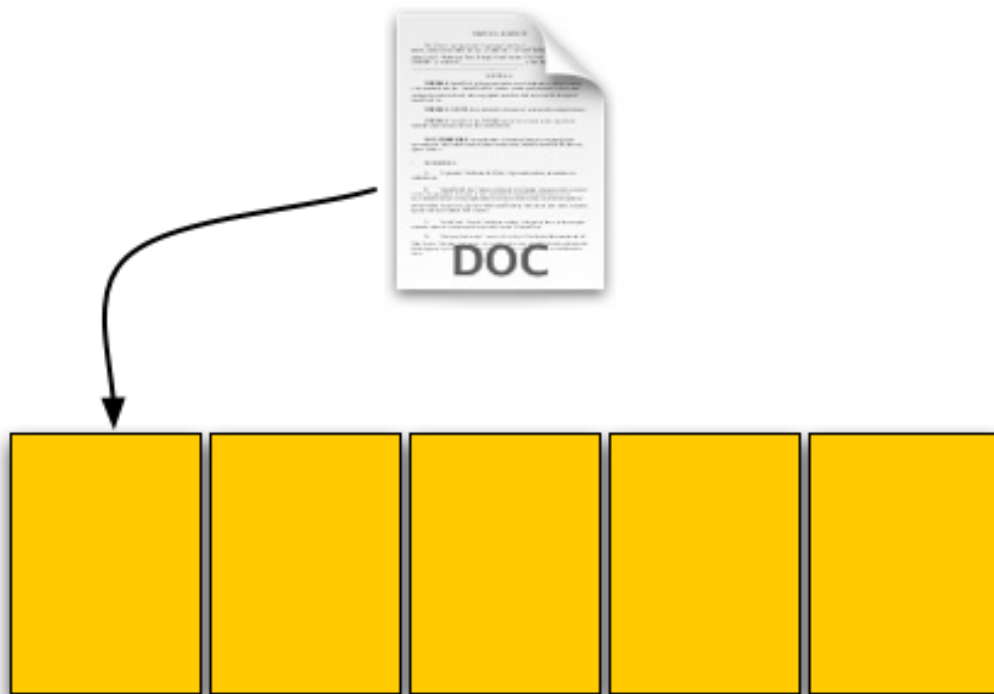
The process works like this:

First, a TRIM-supporting OS (e.g. Windows 7 will support TRIM at some point) queries the hard drive for its rotational speed. If the drive responds by saying 0, the OS knows it's a SSD and turns off features like defrag. It also enables the use of the TRIM command.

When you delete a file, the OS sends a trim command for the LBAs covered by the file to the SSD controller. The controller will then copy the block to cache, wipe the deleted pages, and write the new block with freshly cleaned pages to the drive.

Now when you go to write a file to that block you've got empty pages to write to and your write performance will be closer to what it should be.

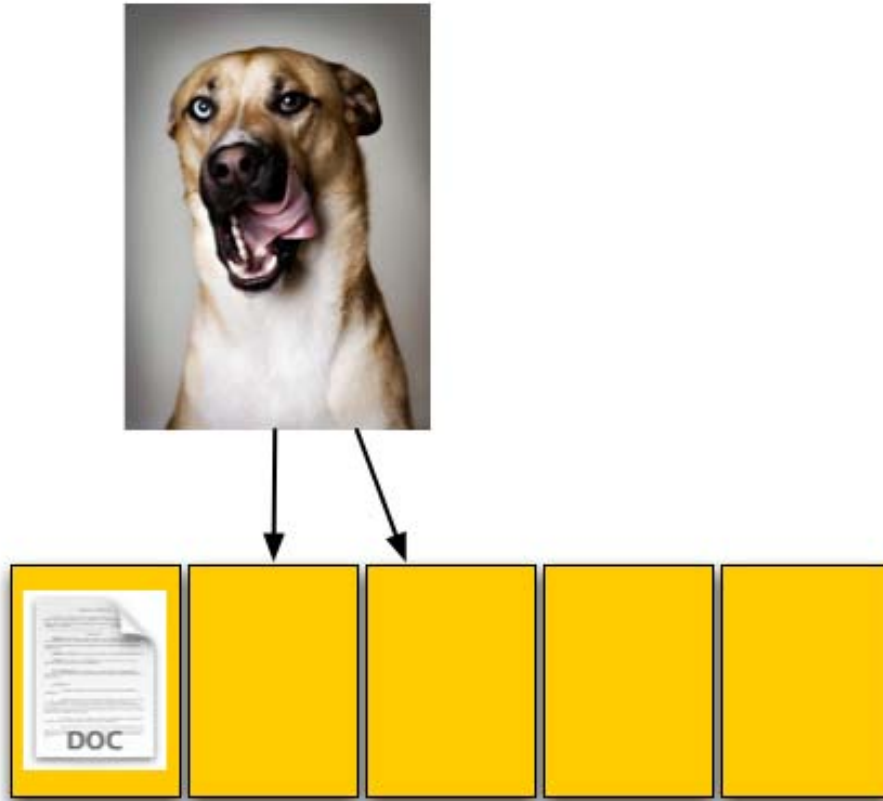
In our example from earlier, here's what would happen if our OS and drive supported TRIM:



Our user saves his 4KB text file, which gets put in a new page on a fresh drive. No differences here.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

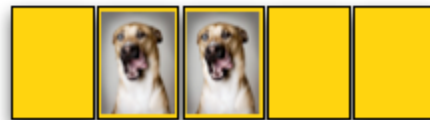
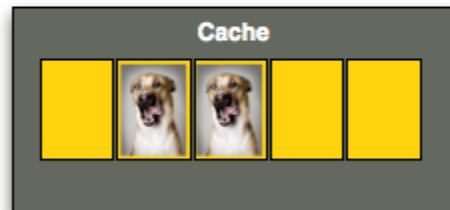
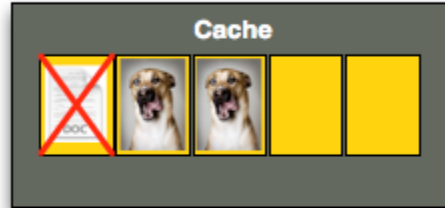


Next was a 8KB JPEG. Two pages allocated; again, no differences.

The third step was deleting the original 4KB text file. Since our drive now supports TRIM, when this deletion request comes down the drive will actually read the entire block, remove the first LBA and write the new block back to the flash:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



The TRIM command forces the block to be cleaned before our final write. There's additional overhead but it happens after a delete and not during a critical write.

Our drive is now at 40% capacity, just like the OS thinks it is. When our user goes to save his 12KB JPEG, the write goes at full speed. Problem solved. Well, sorta.



While the TRIM command will alleviate the problem, it won't eliminate it. The TRIM command can't be invoked when you're simply overwriting a file, for example when you save changes to a document. In those situations you'll still have to pay the performance penalty.

Every controller manufacturer I've talked to intends on supporting TRIM whenever there's an OS that takes advantage of it. The big unknown is whether or not current drives will be firmware-upgradeable to supporting TRIM as no manufacturer has a clear firmware upgrade strategy at this point.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I expect that whenever Windows 7 supports TRIM we'll see a new generation of drives with support for the command. Whether or not existing drives will be upgraded remains to be seen, but I'd highly encourage it.

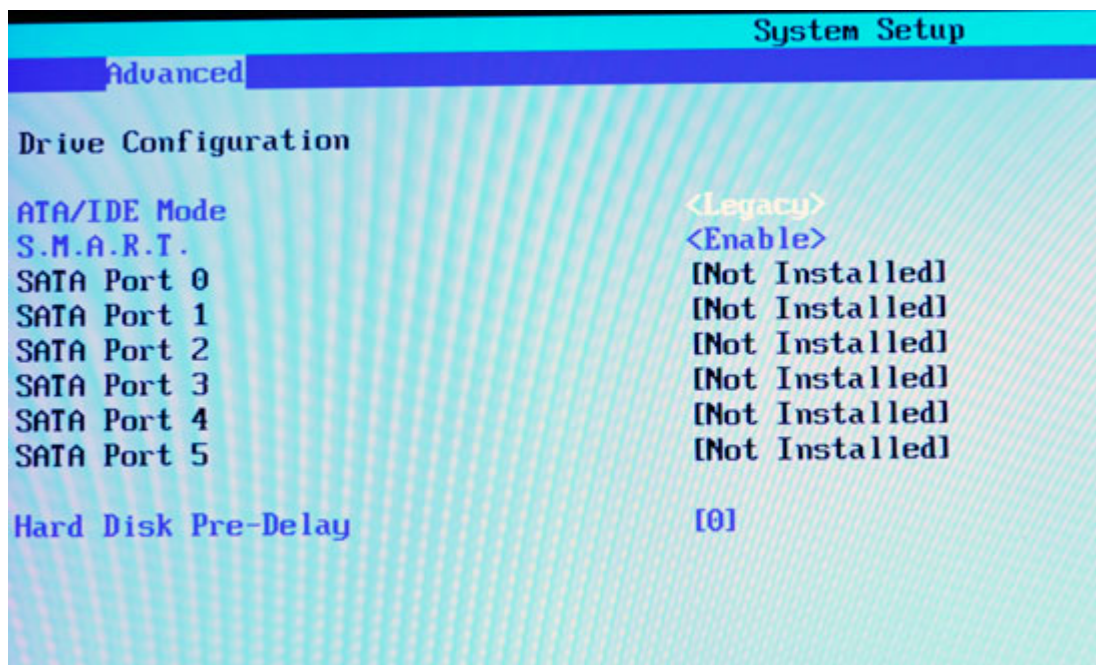
To the manufacturers making these drives: your customers buying them today at exorbitant prices deserve your utmost support. If it's possible to enable TRIM on existing hardware, you owe it to them to offer the upgrade. Their gratitude would most likely be expressed by continuing to purchase SSDs and encouraging others to do so as well. Upset them, and you'll simply be delaying the migration to solid state storage.

Restoring Your Drive to Peak Performance

Based on my explanation there's one sure-fire way to make your drive fast again. Formatting/deleting everything on the drive won't work because those pages on the drive will remain full of data.

If you are doing a clean setup of your machine and want to restore your drive to its native state you'll have to perform a secure erase. Intel distributed a tool with the first X25-M review kits called HDD ERASE. This tool will take any SSD and free every last page on the drive. Obviously you'll lose all of your data but your drive will be super fast again!

In order for HDDERASE to work you need to have your SATA controller running in Legacy IDE mode, you can select this in your BIOS. Your drive will have to be connected to one of the first four SATA ports off of the controller.



Boot to a command prompt (I just use a bootable DOS image on my USB stick) and run the executable. Answer its questions carefully and with honor.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

```
*****
| SECURE ERASE FREEWARE: COMPLETELY ERASE YOUR HARD DISK DRIVE |
| Version 3.3 |
*****

MAXIMUM AND FASTEST SECURITY ERASE IN NEWER DRIVES (GREATER THAN ABOUT 15GB)
IS BY ISSUING THE IN-DRIVE SECURE ERASE COMMAND.

DO YOU WANT TO PROCEED? (Y/N):
```

```
*****
| WELCOME TO SECURE ERASE FREEWARE: COMPLETELY ERASE YOUR HARD DISK DRIVE |
| Version 3.3 |
*****

P0 is NONE
P1 is NONE
S0 is ATA device DCZ-VERTEX 1199
S1 is ATA device SAMSUNG MCCOE64G5MPP-0UA

PLEASE SELECT A DRIVE
  type S0 for secondary master
  type S1 for secondary slave
  type EX to exit the program

PLEASE ENTER YOUR SELECTION:
```

```
*****
| Continuing will permanently erase your drive |
*****

Press C to continue or E to exit: c

Enhanced security erase...
This may take anywhere from 20 minutes to a few hours
depending on the size of your drive
```

Tell it you would like to perform a secure erase (an extended one isn't necessary) and when it asks you if you wish to view LBA 0 you can say no. Reboot your machine (don't forget to set your SATA controller back to whatever mode you had it in before, e.g. RAID or AHCI), partition your drive and you're back in business.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

On a good SSD I don't believe there's a need to secure erase periodically, but whenever you format or re-image your drive, I'd recommend a secure erase since you're killing all of your data anyway.

Simulating a Used Drive

Since SSD performance degrades over time, it's important to not only look at how well these drives perform new - but also the worst they'd perform over their lifetime. In order to do so we'd need a repeatable way of "seasoning" a drive to reduce its performance to the worst it could possibly get. The most realistic worst-case scenario is one where every single block on the drive is full of data. If a secure erase wipes all LBAs, that's the best place to start. To simulate a well seasoned drive I first secure erased the drive.

After the secure erase, I used iometer to write one contiguous file across the disk - filling up the entire drive with 128KB blocks. In the case of the 80GB Intel X25-M, that's 74.5GB of data on the drive before I run a single benchmark. The spare area is left untouched.

Next, I take my test image and I restore it onto the partition with a sector by sector copy. The sequential file write made sure that data is stored in every page of the SSD, the test image restore adds a twist of randomness (and realism) to the data.

There are other ways to produce a drive in its well-used state, but this ends up being the most consistent and repeatable. To confirm that my little simulation does indeed produce a realistically worn drive I ran PCMark on three different drives: 1) a freshly secure-erased Intel X25-M, 2) an Intel X25-M setup using the method I just described and 3) the Intel X25-M used in my CPU testbed that has been through hundreds of SYSMark runs.

The benchmark of choice is PCMark Vantage; it simulates the real world better than most drive benchmarks. The results are below:

Intel X25-M State	PCMark Vantage Overall Test	PCMark Vantage HDD Test
Fresh Secure Erase	11902	29879
Simulated Used Drive	11536	23252
Actual Testbed Used Drive	11140	23438

The secure erased system loses about 3% of its overall performance and 22% of its hard drive specific performance compared to my testbed drive. The seasoning method I described above produces a drive with nearly identical drops in performance.

The method appears to be sound.

Now that we have a way of simulating a used drive, let's see how the contestants fared.

New vs Used SSD Performance

We begin our look at how the overhead of managing pages impacts SSD performance with iometer. The table below shows iometer random write performance; there are two rows for each drive, one for "new" performance after a secure erase and one for "used" performance after the drive has been well used.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

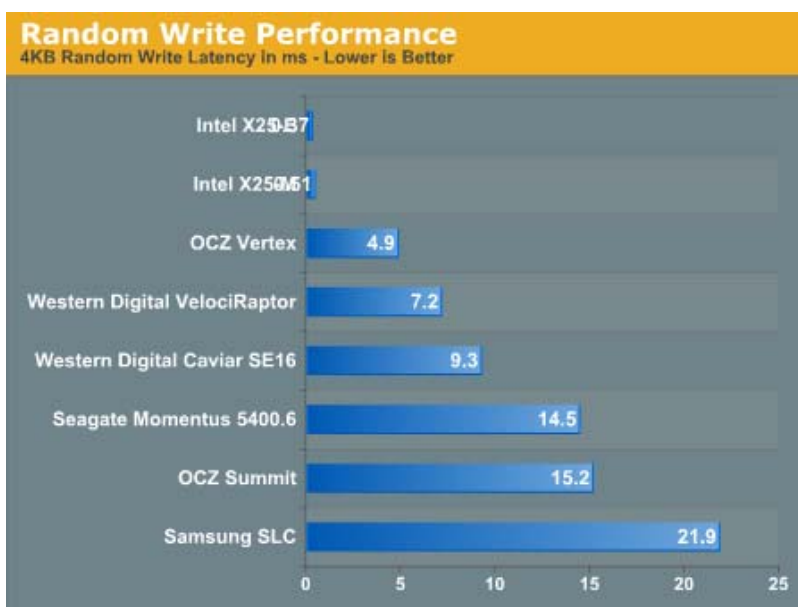
4KB Random Write Speed	New	"Used"
Intel X25-E		31.7 MB/s
Intel X25-M	39.3 MB/s	23.1 MB/s
JMicron JMF602B MLC	0.02 MB/s	0.02 MB/s
JMicron JMF602Bx2 MLC	0.03 MB/s	0.03 MB/s
OCZ Summit	12.8 MB/s	0.77 MB/s
OCZ Vertex	8.2 MB/s	2.41 MB/s
Samsung SLC	2.61 MB/s	0.53 MB/s
Seagate Momentus 5400.6	0.81 MB/s	-
Western Digital Caviar SE16	1.26 MB/s	-
Western Digital VelociRaptor	1.63 MB/s	-

Note that the "used" performance should be the slowest you'll ever see the drive get. In theory, all of the pages are filled with some sort of data at this point.

All of the drives, with the exception of the JMicron based SSDs went down in performance in the "used" state. And the only reason the JMicron drive didn't get any slower was because it is already bottlenecked elsewhere; you can't get much slower than 0.03MB/s in this test.

These are pretty serious performance drops; the OCZ Vertex runs at nearly 1/4 the speed after it's been used and Intel's X25-M can only crunch through about 60% the IOs per second that it did when brand new.

So are SSDs doomed? Is performance going to tank over time and make these things worthless?



"Used" SSD performance vs. conventional hard drives.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

Pay close attention to the average write latency in the graph above. While Intel's X25-M pulls an extremely fast sub-0.3ms write latency normally, it levels off at 0.51ms in its used mode. The OCZ Vertex manages a 1.43ms new and 4.86ms used. There's additional overhead for every write but a well designed SSD will still manage extremely low write latencies. To put things in perspective, look at these drives at their worst compared to Western Digital's VelociRaptor. The degraded performance X25-M still completes write requests in around 1/8 the time of the VelociRaptor. Transfer speeds are still 8x higher as well.

Note that not all SSDs see their performance drop gracefully. The two Samsung based drives perform more like hard drives here, but I'll explain that tradeoff much later in this article.

How does this all translate into real world performance? I ran PCMark Vantage on the new and used Intel drive to see how performance changed.

PCMark Overall Score	New	"Used"	% Drop
Intel X25-M	11902	11536	3%
OCZ Summit	10972	9916	9.6%
OCZ Vertex	11253	9836	14.4%
Samsung SLC	10143	9118	10.1%
Seagate Momentus 5400.6	6817	-	-
Western Digital VelociRaptor	7500	-	-

The real world performance hit varies from 0 - 14% depending on the drive. While the drives are still faster than a regular hard drive, performance does drop in the real world by a noticeable amount. The trim command would keep the drive's performance closer to its peak for longer, but it would not have prevented this from happening.

PCMark Vantage HDD Test	New	"Used"	% Drop
Intel X25-M	29879	23252	22%
JMicron JMF602Bx2 MLC	11613	11283	3%
OCZ Summit	25754	16624	36%
OCZ Vertex	20753	17854	14%
Samsung SLC	17406	12392	29%
Seagate Momentus 5400.6	3525	-	-
Western Digital VelociRaptor	6313	-	-

HDD specific tests show much more severe drops, ranging from 20 - 40% depending on the drive. Despite the performance drop, these drives are still much faster than even the fastest hard drives.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

SSD Aging: Read Speed is Largely Unaffected

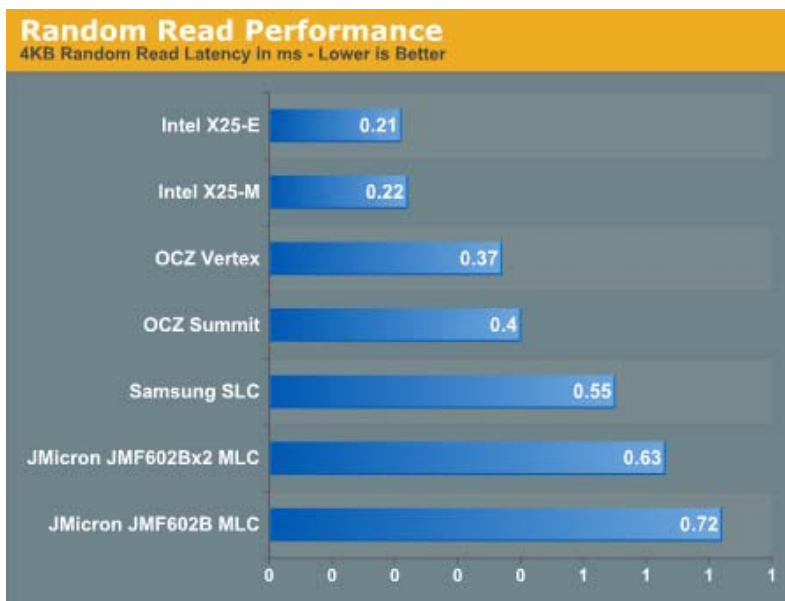
Given the nature of the SSD performance-over-time “problem” you’d expect to only pay the performance penalty when writing files, not reading. And for once, I don’t have any weird exceptions to talk about - this is generally the case.

The table below shows sequential read performance for 2MB blocks on new vs. “used” SSDs. I even included data for a couple of the hard drives in the "Used" column; for those numbers I'm simply measuring transfer rates from the slowest parts of the platter:

2MB Sequential Read Speed	New	"Used"
Intel X25-E		240.1 MB/s
Intel X25-M	264.1 MB/s	230.2 MB/s
JMicron JMF602B MLC	134.7 MB/s	134.7 MB/s
JMicron JMF602Bx2 MLC	164.1 MB/s	164.1 MB/s
OCZ Summit	248.6 MB/s	208.6 MB/s
OCZ Vertex	257.8 MB/s	250.1 MB/s
Samsung SLC		101.4 MB/s
Seagate Momentus 5400.6	77.9 MB/s	-
Western Digital Caviar SE16	104.6 MB/s	54.3 MB/s
Western Digital VelociRaptor	118.0 MB/s	79.2 MB/s

The best SSDs still transfer data at over 2x the rate of the VelociRaptor.

Read latency is also extremely good on these worn SSDs:



The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I left the conventional hard drives out of the chart simply because they completely screw up the scale. The VelociRaptor has a latency of 7.2ms in this iometer test with a queue depth of 3 IOs; that's an order of magnitude slower than the slowest SSD here.

Since you only pay the overhead penalty when you go to write to a previously-written block, the performance degradation only really occurs when you're writing - not when you're reading.

Now your OS is always writing to your drive, and that's why we see a performance impact even if you're just launching applications and opening files and such, but the penalty is much less tangible when it comes to read performance.

The Verdict

There's no skirting the issue: even the best SSDs lose performance the more you use them. Eventually their performance should level off but what matters the most is how their performance degrades.

In using the X25-M I'd say that the performance drop was noticeable but not a deal breaker - and the data tends to agree with me. With average write latencies still well under 1ms, the drive maintained its most important performance characteristic - the ability to perform random accesses much faster than a conventional hard drive.

Keep in mind that with the cost per GB being as high as it is, these SSDs aren't going to be used for large file storage in a desktop or notebook. You're far more likely to use one as your boot/applications drive. As such, what matter the most aren't peak transfer rates but rather fast access times. On a well designed drive with a good controller, peak transfer rates may fall over time, but latency remains good.

You end up with a drive that still manages to be much faster than the fastest 3.5" hard drives, but slower than when you first got it.

If, however, you aren't ok with the performance drop over time then it's worth considering what your options will be. When drives ship with ATA-TRIM support, hopefully late this year, they will do a better job of staying closer to their maximum performance. But the problem won't be solved completely. Instead, what we'll need to see is a more fundamental architectural change to eliminate the problem.

I still believe that a SSD is the single most effective performance upgrade you can do to your PC; even while taking this behavior into account. While personally I wouldn't give up a SSD in any of my machines, I can understand the hesitation in investing a great deal of money in one today.

Intel's X25-M: Not So Adaptive Performance?

The Intel drive is in a constant quest to return to peak performance, that's what its controller is designed to do. The drive is constantly cleaning as it goes along to ensure its performance is as high as possible, for as long as possible. A recent PC Perspective investigation unearthed a scenario where the X25-M is unable to recover and is stuck at a significantly lower level of performance until the drive is secure erased once more.

There's not much I can say about the issue other than I've been working with Intel on it very closely and it's not something I'm overly concerned about at this point. I can replicate

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

the PC Perspective data but not by using my machine normally. Through the right combination of benchmarks I can effectively put the drive in a poor performance state that it won't recover from without a secure erase. I should also mention that I can do the same to other drives as well.

I'll be able to say more soon but for now just hang tight. I'd venture a guess that Intel would not leave its most loyal customers out in the cold after spending \$400 - \$800 on a SSD. I can't speak for Intel but like I said, stay tuned.

Latency vs. Bandwidth: What to Look for in a SSD

It took me months to get my head wrapped around it, but I think I finally get it. We often talk about the concepts of bandwidth and latency but rarely are they as tangible as they are here today.

When I speak of latency I'm talking about how long it takes to complete a request, or fetch a block of data. When I mention bandwidth, I'm talking about how much you can read/write at once. Think of latency as the speed limit and bandwidth as the number of lanes on a highway.

If you're the only car on the highway, you're going to notice the impact of latency more than bandwidth. A speed limit of 70 mph instead of 35 is going to impact you much more than if you added more lanes to the road.

If you're a city planner however and your only concern is getting as many people to work and back, you're going to notice the impact of bandwidth more than latency. It doesn't matter how fast a single car can move, what matters is how many cars you can move during rush hour traffic.

I'd argue that if you're a desktop user and you're using an SSD as a boot/application drive, what will matter most is latency. After you've got your machine setup the way you want it, the majority of accesses are going to be sequential reads and random reads/writes of very small file sizes. Things like updating file tables, scanning individual files for viruses, writing your web browser cache. What influences these tasks is latency, not bandwidth.

If you were constantly moving large multi-gigabyte files to and from your disk then total bandwidth would be more important. SSDs are still fairly limited in size and I don't think you'll be backing up many Blu-ray discs to them given their high cost per GB. It's latency that matters here.

Obviously I'll be testing both latency and bandwidth, but I wanted to spend a moment talking about the synthetic latency tests.

Iometer is a tool that can simulate any combination of disk accesses you can think of. If you know how an application or OS hits the disk, iometer can simulate it. While random disk accesses are the reason that desktop/notebook hard drives feel so slow, the accesses are generally confined to particular areas of the disk. For example, when you're writing a file the OS needs to update a table mapping the file you're writing to the LBAs it allocated for the file. The table that contains all of the LBA mapping is most likely located far away from the file you're writing, thus the process of writing files to the same area can look like random writes to two different groups of LBAs. But the accesses aren't spread out across the entire drive.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

In my original X25-M article I ran a 4KB random write test over the entire span of the drive. That's a bit more ridiculous than even the toughest user will be on his/her desktop. For this article I'm limiting the random write test to an 8GB space of the drive; it makes the benchmark a little more realistic for a desktop/notebook workload.

The other thing I've done is increased the number of outstanding IOs from 1 to 3. I've found that in a multitasking user environment Vista will generally have a maximum of 3 or 4 outstanding IOs (read/write requests).

The combination of the two results in a 100% random file write of 4KB files with 3 outstanding IOs to an 8GB portion of the drive for 3 minutes. That should be enough time to get a general idea of how well these drives will perform when it comes to random file write latency in a worst case, but realistic usage scenario.

The Return of the JMicron based SSD

With the SSD slowdown addressed it's time to start talking about new products. And I'll start by addressing the infamous JMicron JMF602 based SSDs.

For starters, a second revision of the JMF602 controller came out last year - the JMF602B. This controller had twice the cache of the original JMF602A and thus didn't pause/stutter as often.

The JMicron JMF602B is the controller found in G.Skill's line of SSDs as well as OCZ's Core V2, the OCZ Solid and the entire table below of SSDs:

JMicron JMF602B Based SSDs
G.Skill FM-25S2
G.Skill Titan
OCZ Apex
OCZ Core V2
OCZ Solid
Patriot Warp
SuperTalent MasterDrive

All I need to do is point to our trusty iometer test to tell you that the issues that plagued the original JMicron drives I complained about apply here as well:

Iometer 4KB Random Writes, IOqueue=3, 8GB sector space	IOs per second	MB/s	Average Latency	Maximum Latency
JMF602B MLC Drive	5.61	0.02 MB/s	532.2 ms	2042 ms

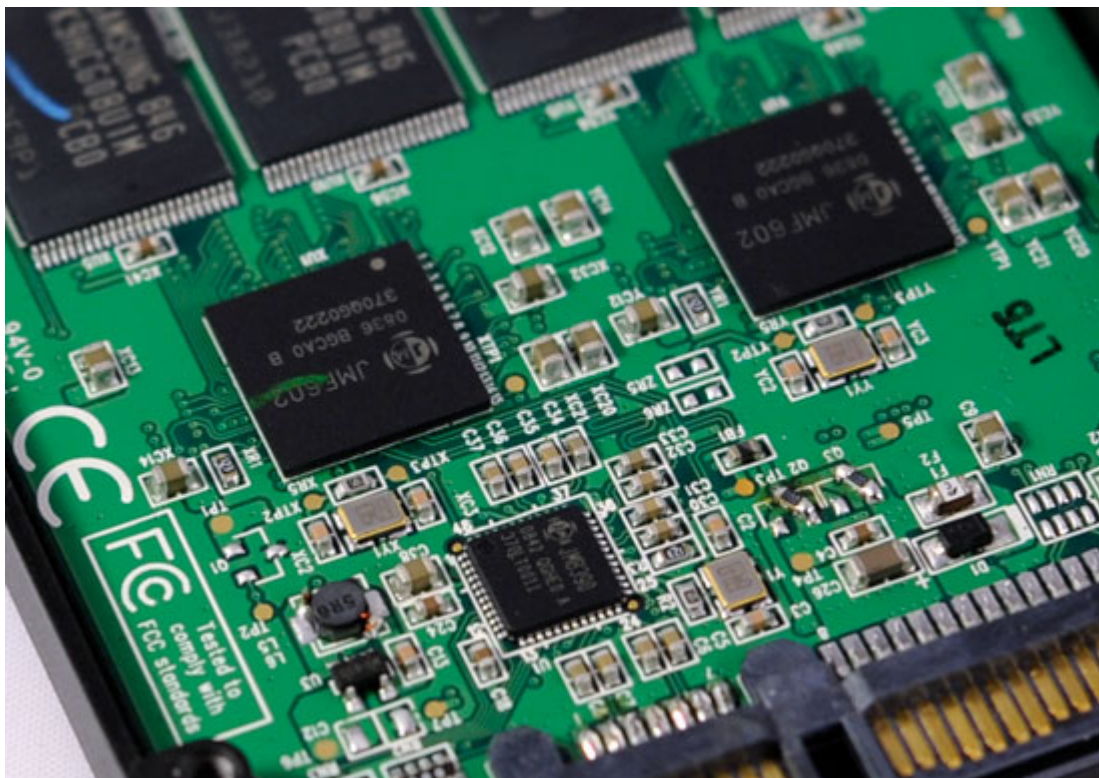
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

On average it takes nearly half a second to complete a random 4KB write request to one of these drives. No thanks.

The single-chip JMF602B based drives are now being sold as value solutions. While you can make the argument that the pausing and stuttering is acceptable for a very light workload in a single-tasking environment, simply try doing anything while installing an application or have anti-virus software running in the background and you won't be pleased by these drives. Save your money, get a better drive.

The next step up from the JMF602B based drives are drives based on two JMF602B controllers. Confused? Allow me to explain. The problem is that JMicron's next SSD controller design won't be ready anytime in the near future, and shipping mediocre product is a better option than shipping no product, so some vendors chose to take two JMF602B controllers and put them in RAID, using another JMicron controller.



Two JMF602B controllers and a JMicron RAID controller

The problem, my dear friends, is that the worst case scenario latency penalty - at best, gets cut in half using this approach. You'll remember that the JMF602 based drives could, under normal load, have a write-latency of nearly 0.5 - 2 seconds. Put two controllers together and you'll get a worst-case scenario write latency of about one second under load or half a second with only a single app running. To test the theory I ran the now infamous 4K random write iometer script on these "new" drives:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

Iometer 4KB Random Writes, IOqueue=3, 8GB sector space	IOs per second	MB/s	Average Latency	Maximum Latency
JMF602B MLC Drive	5.61	0.02 MB/s	532.2 ms	2042 ms
Dual JMF602B MLC Controller Drive	8.18	0.03 MB/s	366.1 ms	1168.2 ms

To some irate SSD vendors, these may just be numbers, but let's put a little bit of thought into what we're seeing here shall we? These iometer results are saying that occasionally when you go to write a 4KB file (for example, loading a website, sending an IM and having the conversation logged or even just saving changes to a word doc) the drive will take over a second to respond.

I don't care what sort of drive you're using, 2.5", 3.5", 5400RPM or 7200RPM, if you hit a 1 second pause you notice it and such performance degradation is not acceptable. Now these tests are more multitasking oriented, but if you run a single IO on the drive you'll find that the maximum latency is still half a second.

The average column tells an even more bothersome story. Not only is the worst case scenario a 1168 ms write, on average you're looking at over a quarter of a second just to write 4KB.

The G.Skill Titan has recently garnered positive reviews for being a fast, affordable, SSD. Many argued that it was even on the level of the Intel X25-M. I'm sorry to say it folks, that's just plain wrong.



One of the most popular dual JMF602B drives

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

If you focus exclusively on peak transfer rates then these drives work just fine. You'll find that, unless you're running a Blu-ray rip server, you don't spend most of your time copying multi-GB files to and from the drive. Instead, on a normal desktop, the majority of your disk accesses will be small file reads and writes and these drives can't cut it there.

Some vendors have put out optimization guides designed to minimize stuttering with these JMFB02B based drives. The guides generally do whatever they can to limit the number and frequency of small file writes to your drive (e.g. disabling search indexing, storing your temporary internet files on a RAM drive). While it's true that doing such things will reduce stuttering on these drives, the optimizations don't solve the problem - they merely shift the cause of it. The moment an application other than Vista or your web browser goes to write to your SSD you'll have to pay the small file write penalty once more. Don't settle.

But what option is there? Is Intel's X25-M the only drive on the market worth recommending? What if you can't afford spending \$390 for 80GB. Is there no cheaper option?

OCZ Tries Again with the Vertex

This is Ryan Petersen:



He's the CEO of OCZ Technology. He wasn't too happy after my Intel X25-M SSD review.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Although that review was about the X25-M, it addressed a major shortcoming with a number of other SSDs in the market - the most popular at the time being OCZ's Core line.

The problem with the Core line was that although they were competitively priced, they had a major performance issue: small file write latency. The problem manifested itself in system-wide stuttering during disk intensive multitasking. It wouldn't take a lot to make those drives stutter, and until that article went live, everyone assumed that the problem was Windows, the disk controller, or something else.

I placed the blame on the drives and Ryan wasn't all too happy.

We argued. A lot. He claimed it was unfair that I was comparing the Core series to Intel's far more expensive X25-M; I argued that the user experience offered by his drives was inferior to that of a regular hard drive. We never really saw eye to eye on SSDs after that article.

Ryan told me the problem was a lack of cache, and that they'd have a future version of their SSD with a large off-chip DRAM cache. Ryan did not want to send me any more SSDs that didn't have a cache on them. It wasn't just OCZ, we had difficulty getting JMicron based SSDs from most manufacturers after that article. Manufacturers were pissed.

I offered to look at any new SSDs they had coming, regardless of who made the controller. But I made one thing very clear - I did not believe that the JMicron based drives were fit for sale. I felt that they were hurting the image of SSDs in the market and doing more harm than good. Ask any of the major players in the flash and OS industries and you'll hear the

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

same complaint. We need good SSDs in the market, not a mixture of great ones and bad ones.

In early December I got word from OCZ that their first drive with a DRAM cache was nearing final production. The controller was from a Korean company called Indilinx, with a number of ex-Samsung engineers. JMicron was out for this product, they didn't have a suitable controller - the JMF602B was it. I won't draw too much attention to the fact that if the JMF602B drives were indeed fine then there would be no need to consider another controller manufacturer.

The Indilinx Barefoot controller would support up to 64MB of off-chip DRAM, but have no on-die cache. All user data would be buffered in this memory. Now 64MB is more than enough to handle anything being written to the drive, but the memory is also used for the ARM7 based processor for its work in dealing with wear leveling and flash block cleaning (removing invalid pages).

There's a key distinction here between the Barefoot and Intel's controller - the X25-M doesn't store any user data in its DRAM. Technically the Barefoot approach is no less prone to dataloss than a standard hard drive, but the Intel approach is a bit better. In the event of sudden power loss there's a chance that some of the data in the Barefoot's 64MB buffer could be lost before it's committed to flash.

The OCZ Vertex drive would have a more heavy duty controller, with an ARM processor and a large off-chip DRAM to improve performance. To top it all off, OCZ would sell it at a price much lower than Intel's X25-M. Could this finally be the affordable SSD we've been looking for?

OCZ Sends Me SSDs, Once More

As I mentioned before, after the X25-M article I was somewhat blacklisted from getting more SSDs to test. Since the Core V2, I hadn't tested a single SSD from OCZ or anyone else for that matter. Everything on the market was either based on Samsung's SLC drive, JMicron MLC or the Intel X25 series.

Needless to say, I was excited when I got a box from OCZ.

I got the drives early in the morning. Excited, I opened them up. Inside the box were three drives. The OCZ Apex (I'd never reviewed it, and OCZ reluctantly sent another JMicron drive to me), the Vertex and the Summit.

The Summit was based on Samsung's latest MLC SSD controller, which I'd heard great things about. It's supposed to compete with Intel's drive.

The Vertex is the drive I was most interested in. A value SSD that didn't suck, or at least that's what it seemed to be on paper. Ryan even left me a little note on the box:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



O RLY?

It's worth noting that although other SSD makers will be making drives based on the Indilinx and Samsung controllers, OCZ was the first to get me drives based on both of these controllers. In fact, I was done testing the OCZ Summit based on Samsung's latest controller before Samsung ever offered to ship me the drive.

I pushed the Apex and Summit drives aside for now. What I wanted to know was how the Vertex performed.



The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I cloned my system drive and used the Vertex in my personal machine. As soon as I hit the desktop I knew there was a problem; all of my icons took longer than they should've to load. It took about 30 minutes of actual usage for the drive to stutter and within a couple of hours performance got so unbearable that I had to pull it out.

Sigh, the Vertex was broken.

Disappointed, I went back to OCZ

OCZ was worried. The last time I reviewed one of their SSDs I was truthful about it, and it hurt their sales considerably. Customers were returning drives, and to OCZ's credit, they stepped up and even accepted some returns themselves - something that most manufacturers wouldn't have done. Regardless what they had told me, there was some admission of fault there. Those JMicron drives were nothing short of crap.

As soon as OCZ started getting word that I wasn't pleased with Vertex, they went into a state of panic. These drives all do very well in synthetic HDD tests like HDTach and ATTO, that's generally all they're reviewed in, so that's all they're tested in. But now OCZ was hearing that the Vertex wasn't passing some of my tests and they had no idea what it was failing or why.

I tend to do a good job of keeping what tests I run secret until the review is published, so there isn't any unfair optimization. I purposefully introduce new tests to our performance suites to help keep manufacturers honest and optimizing for real world usage scenarios rather than specific benchmarks. OCZ had no idea what I was running, but they knew that the Vertex wasn't doing well.

Summit on the other hand was performing just fine, but that's an expensive drive. Vertex was supposed to be good, it should've been good, there's no reason for it to be performing this poorly. I ran the infamous iometer test to see what was going on:

Iometer 4KB Random Writes, IOqueue=1, 8GB sector space	IOs per second	MB/s	Average Latency	Maximum Latency
Original Pre-release OCZ Vertex	20.7	0.08 MB/s	48.2 ms	484.5 ms

How on earth is this acceptable at all? Average latency of 48.2ms and a maximum latency as bad as the Apex and G.Skill Titan drives? I've heard some SSD vendors dismiss the iometer results but let me caution you against that. What these numbers are telling us is that on average, when your OS goes to write a 4KB file somewhere on your drive, it'll take nearly 50ms. That's 4.5x longer than a 5400 RPM 2.5" notebook drive and that's the average case. What part of that sounds acceptable? Anyone who tells you otherwise is delusional.

I thought for sure that the drive was broken and that we'd made no progress since last fall. But the drive hadn't launched yet, while there were glowing reviews of it, no one had wasted any money. I wrote an email to Ryan Petersen, OCZ's CEO. I described my findings and told him that while the Vertex's performance was better than any of the

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

JMicron solutions, it was unacceptable for anything other than perhaps extremely light, single-tasking usage.

I told him it sucked. He said that wasn't fair. We argued over email but he came back and asked me what I needed to see to make the drive better.

I told him I'd need an average response time in the sub-1ms range and a max latency no worse than Intel's 94ms. I didn't think it would be possible. I was prepared for OCZ to hate me once more. He told me to give him a couple of days.

Once More, With Feeling

Ryan said we'd lose some sequential write performance. The drive would no longer be capable of 230MB/s writes, perhaps only down to 80 or 90MB/s now. I told him it didn't matter, that write latency needed to come down and if it were at the sacrifice of sequential throughput then it'd be fine. He asked me if I was sure, I said yes. I still didn't think he could do it.

A couple days later and I got word that OCZ just got a new firmware revision back from Korea with the changes I'd asked for. They were going to do a quick test and if the results made me happy, they'd overnight a drive to me for Saturday delivery.

He sent me these iometer results:

Not sure were syncing up exactly on how we tested but these response times are in general much faster than our own tests

```
IO METER TEST Method  IOps  MBps  Average Response Time  Maximum Response Time
Physical Drive Test   4784  18.69  0.208  83.590
5Mil Sector Logical Drive Test  2488  9.72  0.401  7.272
5 to 25Mil Sector Logical Drive Test  2394  9.35  0.417  134.169
5 to 25Mil Sector Logical Drive Test (2nd)  2430  9.49  0.411  132.999
```

-----Original Message-----

From: anand.shimpi@anandtech.com [<mailto:anand.shimpi@anandtech.com>]

Sent: Friday, February 20, 2009 10:44 AM

To: Ryan M. Petersen

Subject: Re: RE: The Skinny on Vertex

The New Vertex was looking good, but was it too good to be true?

I couldn't believe it. There was no way. "Sure", I said, "send the drive over". He asked if I'd be around on Saturday to receive it. I would be, I'm always around.

This was the drive I got:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



No markings, no label, no packaging - just a black box that supposedly fixed all of my problems. I ran the iometer test first...it passed. I ran PCMark. Performance improved. There's no way this thing was fixed. I skipped all of the other tests and threw it in my machine, once again cloning my system drive. Not a single pause. Not a single stutter.

The drive felt slower than the Intel or Summit drives, but that was fine, it didn't pause. My machine was usable. Slower is fine, invasive with regards to my user experience is not.

I took the Vertex back out and ran it through the full suite of tests. It worked. Look at the PCMark Vantage results to see just what re-focusing on small file random write latency did to the drive's performance:

The Vertex went from performing like the OCZ Apex (dual JMicron JMF602B controllers) to performing more like an Intel X25-M or OCZ Summit. I'll get to the full performance data later on in this article, but let's just say that we finally have a more affordable SSD option. It's not the fastest drive in the world, but it passes the test for giving you the benefits of a SSD without being worse in some way than a conventional hard drive.

As the Smoke Cleared, OCZ Won Me Over

Now let's be very clear what happened here. OCZ took the feedback I gave them, and despite it resulting in a product with fewer marketable features implemented the changes. It's a lot easier to say that your drive is capable of speeds of up to 230MB/s than it is to say it won't stutter, the assumption is that it won't stutter.

As far as I know, this is the one of the only reviews (if not the only) at the time of publication that's using the new Vertex firmware. Everything else is based on the old firmware which did not make it to production. Keep that in mind if you're looking to compare numbers or wondering why the drives behave differently across reviews. The old

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

firmware never shipped thanks to OCZ's quick acting, so if you own one of these drives - you have a fixed version.

While I didn't really see eye to eye with any of the SSD makers that got trashed in the X25-M review, OCZ was at least willing to listen. On top of that, OCZ was willing to take my feedback, go back to Indilinx and push for a different version of the firmware despite it resulting in a drive that may be harder to sell to the uninformed. The entire production of Vertex drives was held up until they ended up with a firmware revision that behaved as it should. It's the sort of agility you can only have in a smaller company, but it's a trait that OCZ chose to exercise.

They were the first to bring an Indilinx drive to the market, the first to produce a working drive based on Samsung's latest controller, and the company that fixed the Indilinx firmware. I've upset companies in the past and while tempers flared after the X25-M review, OCZ at least made it clear this round that their desire is to produce the best drive they could. After the firmware was finalized, OCZ even admitted to me that they felt they had a much better drive; they weren't just trying to please me, but they felt that their customers would be happier.

I should also point out that the firmware that OCZ pushed for will now be available to all other manufacturers building Indilinx based drives. It was a move that not only helped OCZ but could help every other manufacturer who ships a drive based on this controller.

None of this really matters when it comes to the drive itself, but I felt that the backstory was at least just as important as the benchmarks. Perhaps showing you all a different side of what goes on behind the scenes of one of these reviews.

The OCZ Summit: First with Samsung's New Controller

Well before Intel ever introduced the X25-M, there was one chip company who had already been making well behaved SSD controllers: Samsung.

The fit just made sense; Samsung makes NAND flash, Samsung makes microprocessors, thus Samsung should make SSDs.

Samsung is a unique company in that it is very well known in both the OEM and consumer spaces. You'll find Samsung ICs in nearly everything; DRAM, smartphones, even the flash on the JMicron drives I'm always complaining about - Samsung makes it. With strong OEM ties as well as a good-sized HDD business, it didn't take long for Samsung to get into the SSD market.

Because it was primarily selling SSDs to OEMs like Apple and Dell, the controller had to be perfect. Stuttering, pausing and strange reliability problems wouldn't cut it. Apple wouldn't dare ship a MacBook Air with a SSD that would deliver anything less than a flawless usage experience.

With that sort of pressure, Samsung's SSDs and its controllers always just worked. Even before the JMF602A ever shipped, Samsung's controllers were doing just fine. They had to. Their customers were Apple and Lenovo, there's no room for silliness.

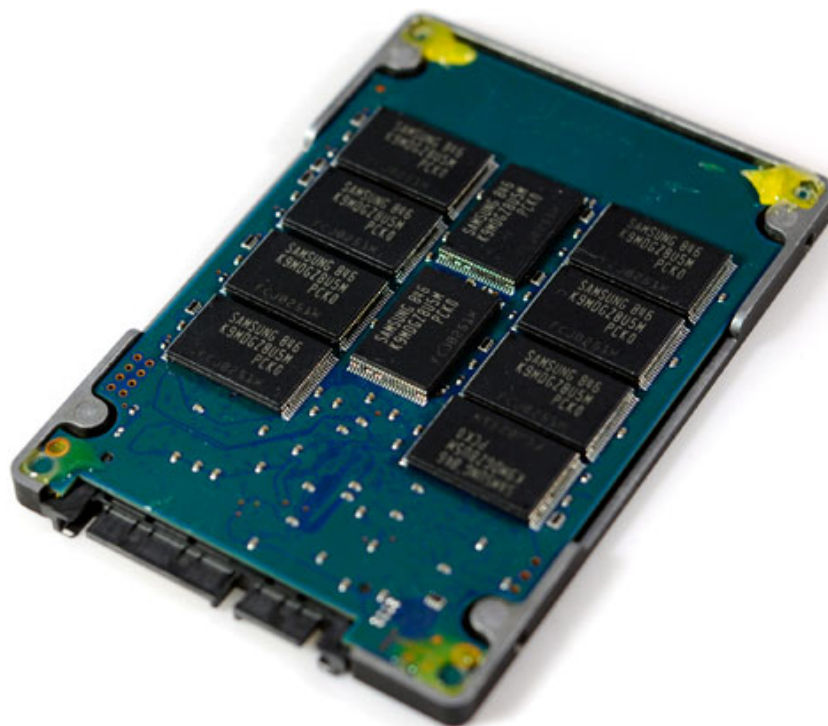
There were two problems with Samsung's controllers: 1) they were expensive, and 2) they weren't very fast.

The SSD Anthology Understanding SSDs and New Drives

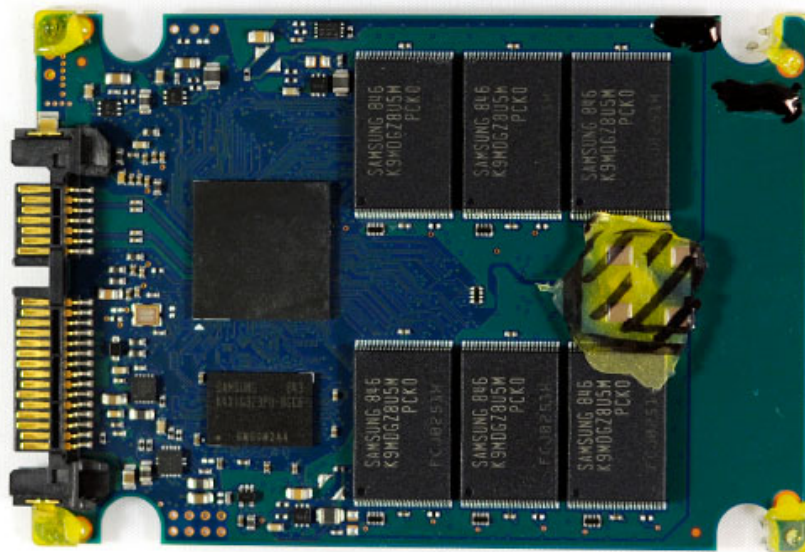
Anand Lal Shimpi

The cost drove SSD makers to companies like JMicron, to drive SSD prices down faster than Samsung would allow. The performance made it so that the very first SSDs weren't much faster than the fastest hard drives, and in some cases they were slower.

Late last year Samsung announced a new version of its MLC controller that would slowly replace all of the existing Samsung MLC drives in the market. OCZ was the first to get us a sample based on this new controller, even faster than Samsung and the drive is called the Summit.



OCZ's Summit, it's an early sample



Lots of glue and no screws on the pre-release Summit. The final drive should be ready soon.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I'll save a detailed investigation into the Summit and Samsung's controller for another article, this one is already long enough. The Summit should be priced similarly to the Intel drive, although OCZ is trying to make it cheaper. The performance level is designed to be greater than the Vertex and competitive with the Intel X25-M. We'll see about that shortly.

The Test

If you've managed to make it this far in the article then I owe you a great deal of gratitude. While I'd offer to make you dinner it'd probably be a little impractical, so instead I'll help you find out how these SSDs perform.

Note that every single benchmark here was run with the drive in a "used" state. Again, I did so by performing a secure erase on the drive, filling it to capacity, then restoring my test bed image over the partition. I can definitely make the drives benchmark faster, but I'm trying to provide performance data that shows you how your drive will behave after you've owned it for a while.

To make the benchmarks all that more stressful and realistic, I kept Avast Anti-Virus running and scanning every file accessed in the background of all of my non-suite and non-synthetic tests. Basically everything that was hand-timed. Firmware 0122 is utilized for this review. OCZ recently released firmware revision 1199 that we are testing currently and another release (1275) is expected shortly. We will provide an update if there is a significant difference in performance or the user experience routines.

CPU:	Intel Core i7-965 (3.2GHz)
Motherboard:	Intel DX58SO (Intel X58)
Chipset:	Intel X58
Chipset Drivers:	Intel 9.1.1.1010 (Intel)
Memory:	Qimonda DDR3-1066 4 x 1GB (7-7-7-20)
Video Card:	eVGA GeForce GTX 280
Video Drivers:	NVIDIA ForceWare 180.43
Desktop Resolution:	1920 x 1200
OS:	Windows Vista Ultimate 64-bit

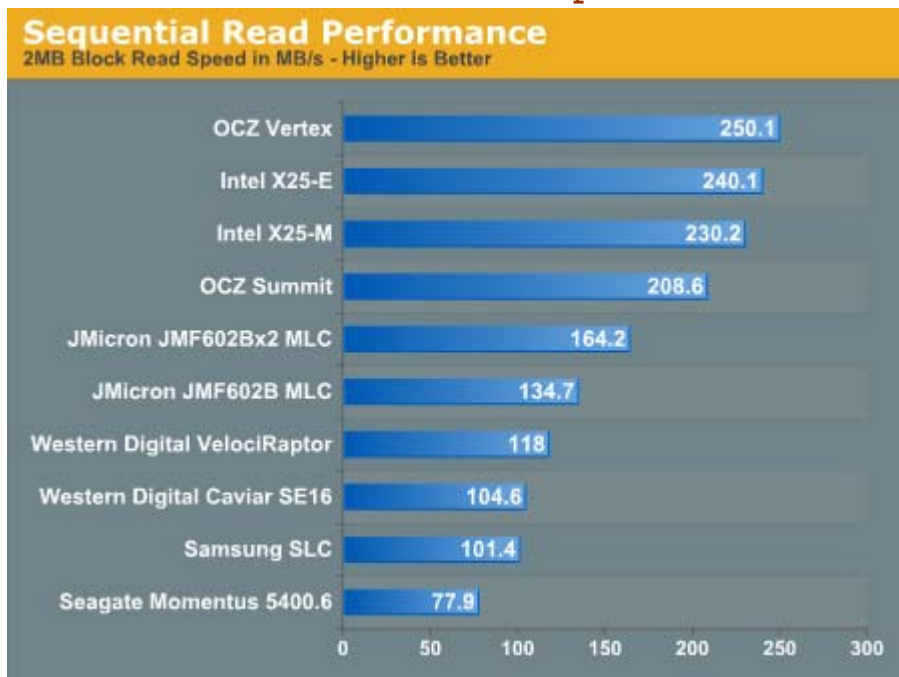
Sequential Read/Write Performance

To measure sequential read and write performance I created an iometer test script to read and write 2MB files to the drives. The writes are 100% sequential and are performed on an unpartitioned drive for 3 minutes.

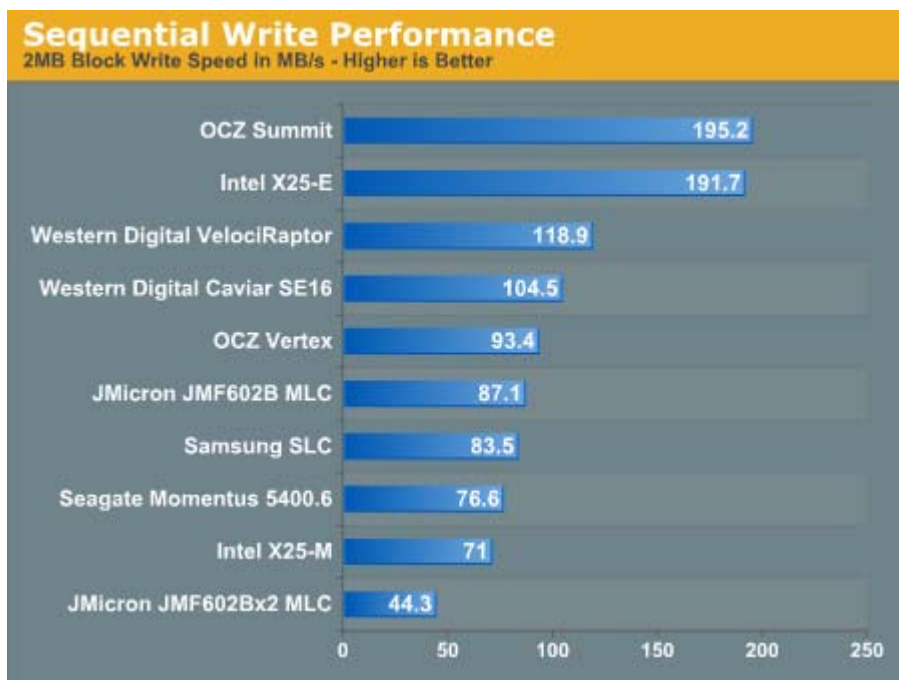
Performance is reported in MB/s, higher numbers are better.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



The sequential read performance of the latest generation of SSDs is nothing short of amazing. The OCZ Vertex pulls through with 250MB/s, over twice the speed of the VelociRaptor and 2.5x the speed of the Samsung SLC drive that OCZ used to sell for \$1000.



Looking at write performance we see things change a bit. The OCZ Summit and Intel X25-E are the only two drives that can outperform the VelociRaptor, and they do so handsomely. The rest of the drives fall below even the 7200 RPM Caviar SE16. They are by no means slow, they just don't destroy the fastest hard drives as they do in other tests.

The SSD Anthology Understanding SSDs and New Drives

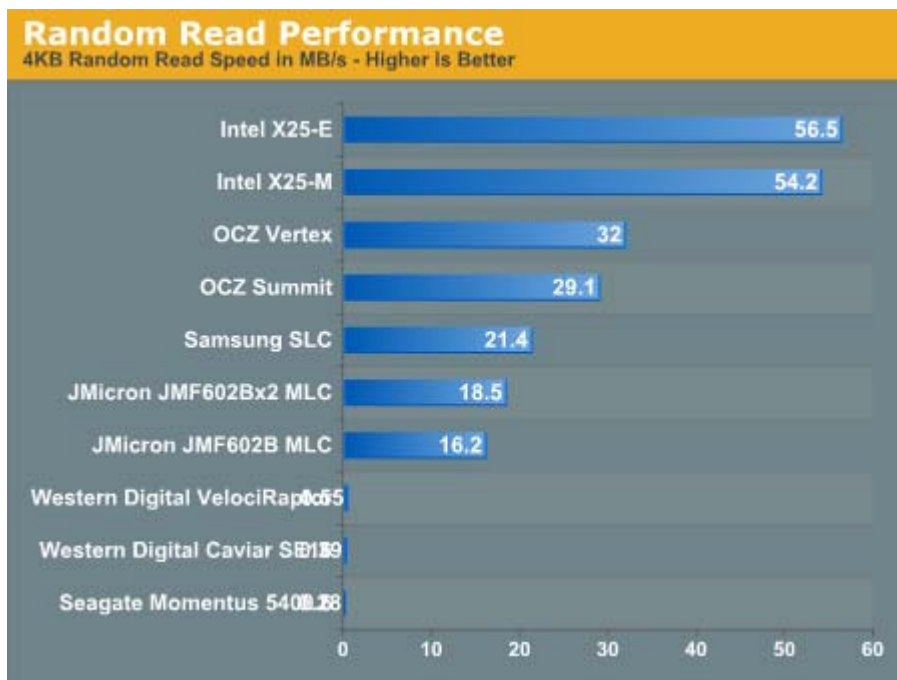
Anand Lal Shimpi

While the X25-E should perform the way it does, the OCZ Summit uses MLC flash yet it performs like an SLC drive. This is a very important benchmark as it shows the sort of performance Samsung has optimized for with its controller. This drive is designed to shatter bandwidth barriers, but what about latency?

Random Read/Write Performance

Arguably much more important to any PC user than sequential read/write performance is random access performance. It's not often that you're writing large files sequentially to your disk, but you do encounter tons of small file reads/writes as you use your PC.

To measure random read/write performance I created an iometer script that peppered the drive with random requests, with an IO queue depth of 3 (to add some multitasking spice to the test). The write test was performed over an 8GB range on the drive, while the read test was performed across the whole drive. I ran the test for 3 minutes.



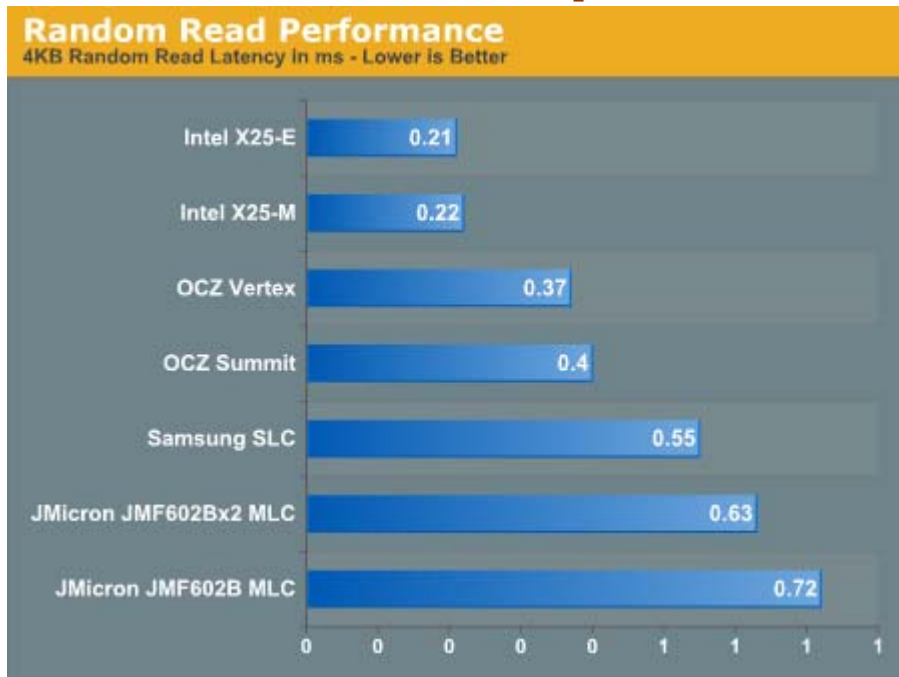
The three hard drives all posted scores below 1MB/s and thus aren't visible on our graph above. This is where SSDs shine and no hard drive, regardless of how many you RAID together, can come close.

The two Intel drives top the charts and maintain a huge lead. The OCZ Vertex actually beats out the more expensive (and unreleased) Summit drive with a respectable 32MB/s transfer rate here. Note that the Vertex is also faster than last year's Samsung SLC drive that everyone was selling for \$1000. Even the JMicron drives do just fine here.

If we look at latency instead of transfer rate it helps put things in perspective:

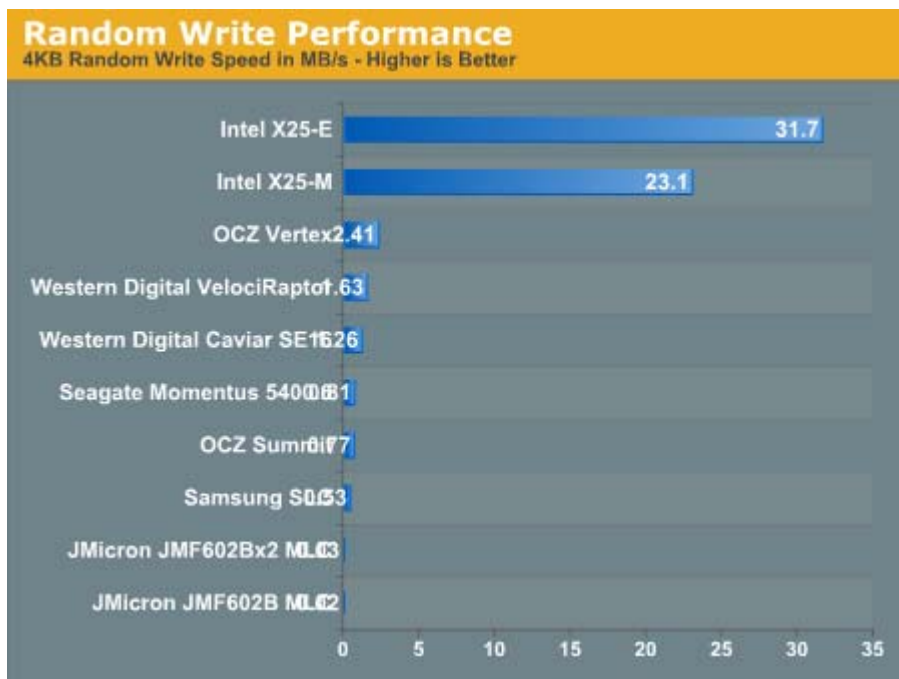
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Read latencies for hard drives have always been measured in several ms, but every single SSD here manages to complete random reads in less than 1ms under load.

Random write speed is where we can thin the SSD flock:



Only the Intel drives and to an extent, the OCZ Vertex, post numbers visible on this scale. Let's go to a table to see everything in greater detail:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

4KB Random Write Speed	
Intel X25-E	31.7 MB/s
Intel X25-M	23.1 MB/s
JMicron JMF602B MLC	0.02 MB/s
JMicron JMF602Bx2 MLC	0.03 MB/s
OCZ Summit	0.77 MB/s
OCZ Vertex	2.41 MB/s
Samsung SLC	0.53 MB/s
Seagate Momentus 5400.6	0.81 MB/s
Western Digital Caviar SE16	1.26 MB/s
Western Digital VelociRaptor	1.63 MB/s

Every single drive other than the Intel X25-E, X25-M and OCZ's Vertex is slower than the 2.5" Seagate Momentus 5400.6 hard drive in this test. The Vertex, thanks to OCZ's tweaks, is now 48% faster than the VelociRaptor.

The Intel drives are of course architected for the type of performance needed on a desktop/notebook and thus they deliver very high random write performance.

Random write performance is merely one corner of the performance world. A drive needs good sequential read, sequential write, random read *and* random write performance. The fatal mistake is that most vendors ignore random write performance and simply try to post the best sequential read/write speeds; doing so simply produces a drive that's undesirable.

While the Vertex is slower than Intel's X25-M, it's also about half the price per GB. And note that the Vertex is still 48% faster than the VelociRaptor here, and multiple times faster in the other tests.

SYSMark 2007

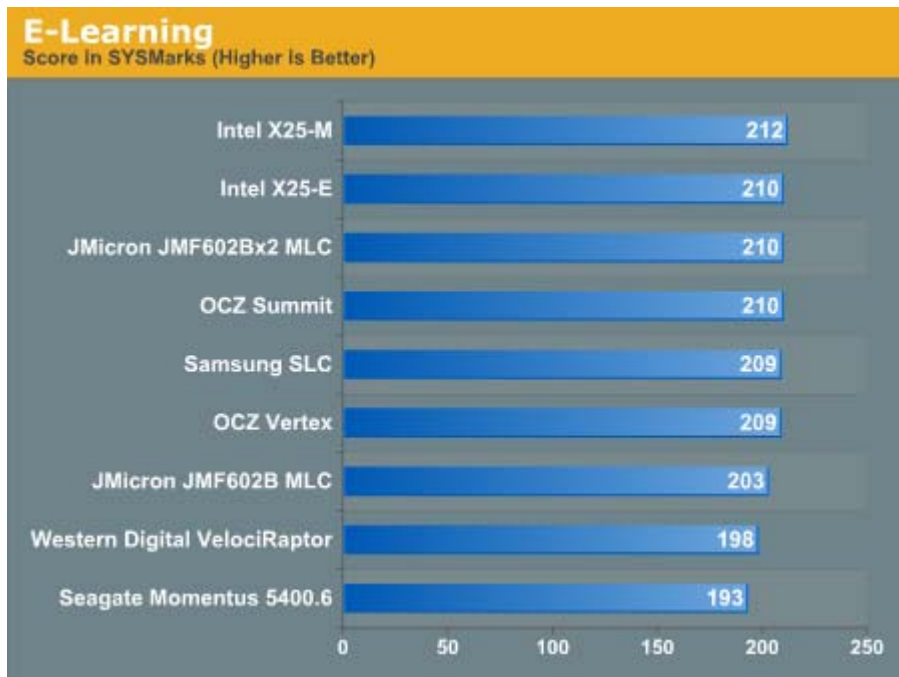
Our first test is the full SYSMark 2007 benchmark suite. It's a system-level performance test that measures performance based on average response time, making it a great candidate for our SSDs, which should provide pretty low response times, at least compared to standard mechanical disks.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

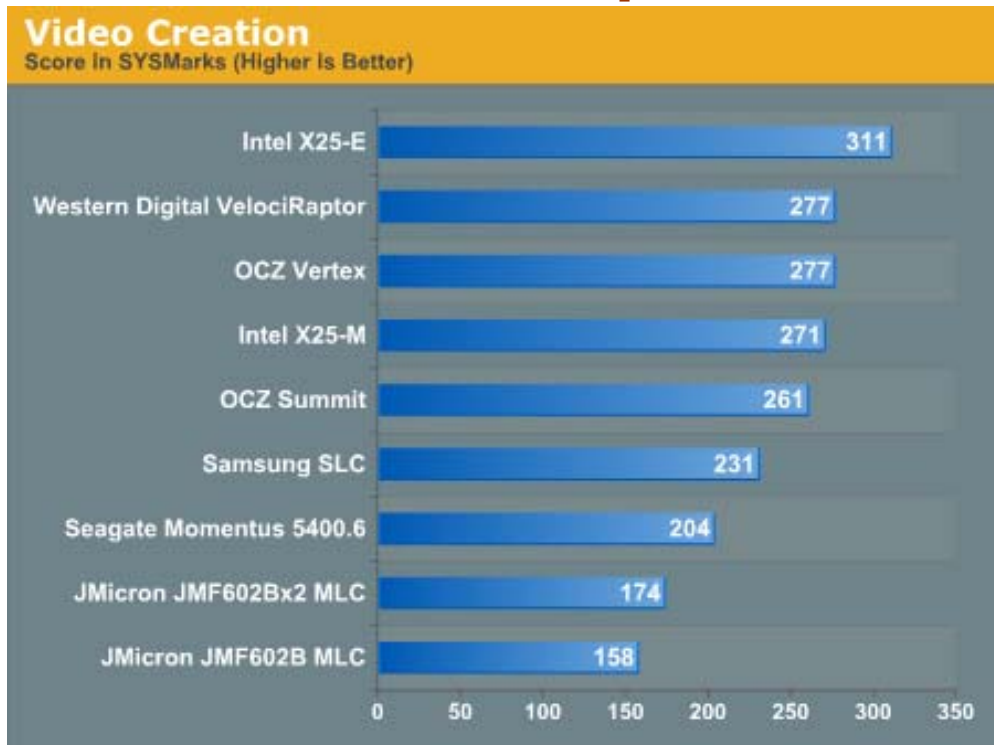


SYSMark is a better CPU test than a storage test, but it still shows that all of these SSDs are easily capable of delivering performance greater than or equal to the VelociRaptor. Note that the JM micron drives are both slower than the 5400RPM notebook hard drive in this test.



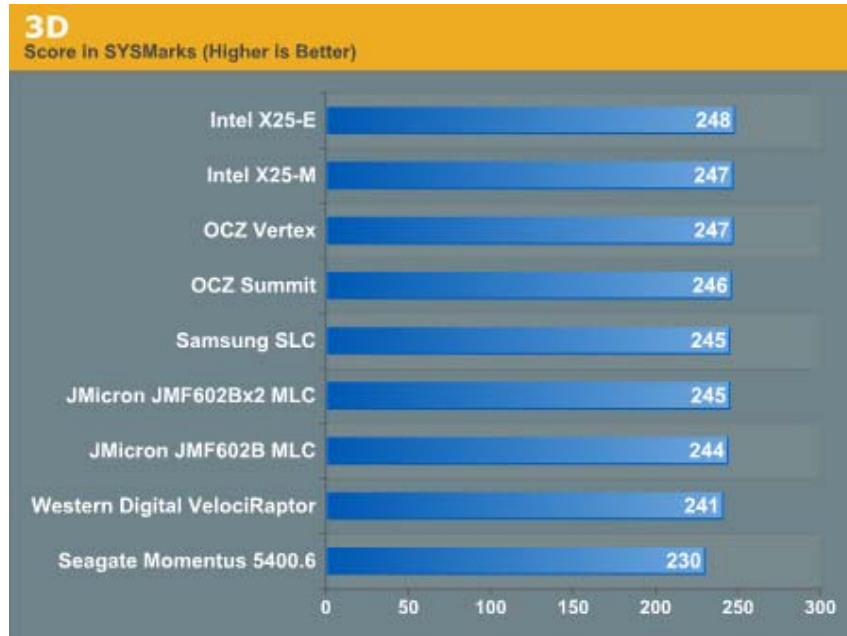
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



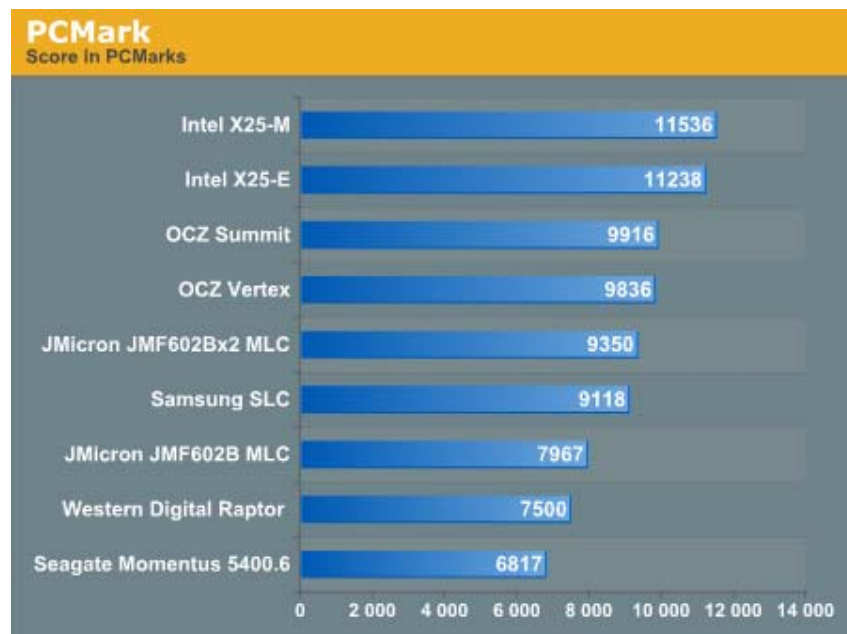
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



PCMark Vantage

Next up is PCMark Vantage, another system-wide performance suite. For those of you who aren't familiar with PCMark Vantage, it ends up being the most real-world-like hard drive test I can come up with. It runs things like application launches, file searches, web browsing, contacts searching, video playback, photo editing and other completely mundane but real-world tasks. I've described the benchmark in great detail before but if you'd like to read up on what it does in particular, take a look at Futuremark's whitepaper on the benchmark; it's not perfect, but it's good enough to be a member of a comprehensive storage benchmark suite. Any performance impacts here would most likely be reflected in the real world.



The SSD Anthology Understanding SSDs and New Drives

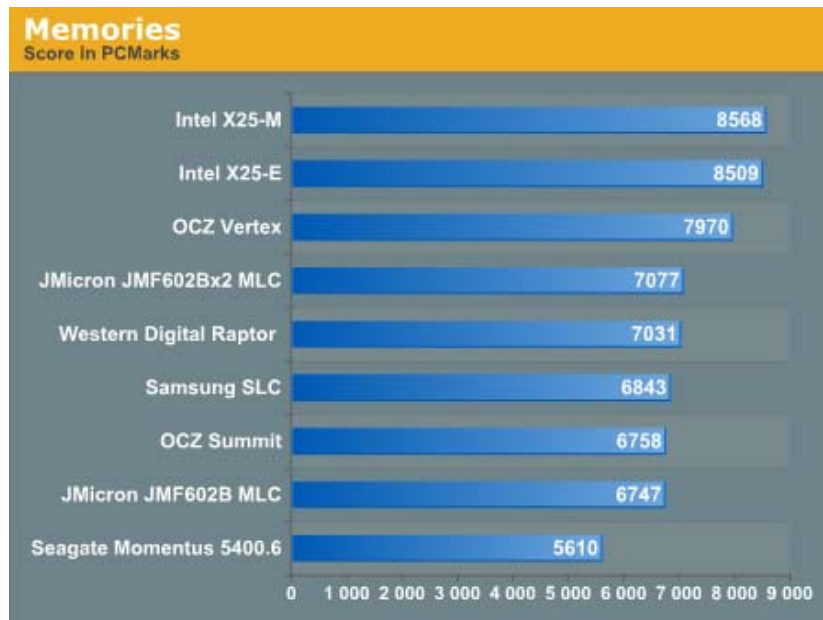
Anand Lal Shimpi

If you've paid attention to the synthetic tests from the previous pages, the results here should make sense. The Intel drives take the top two spots followed by the two OCZ drives, then the JMicron and conventional HDDs take up the rear.

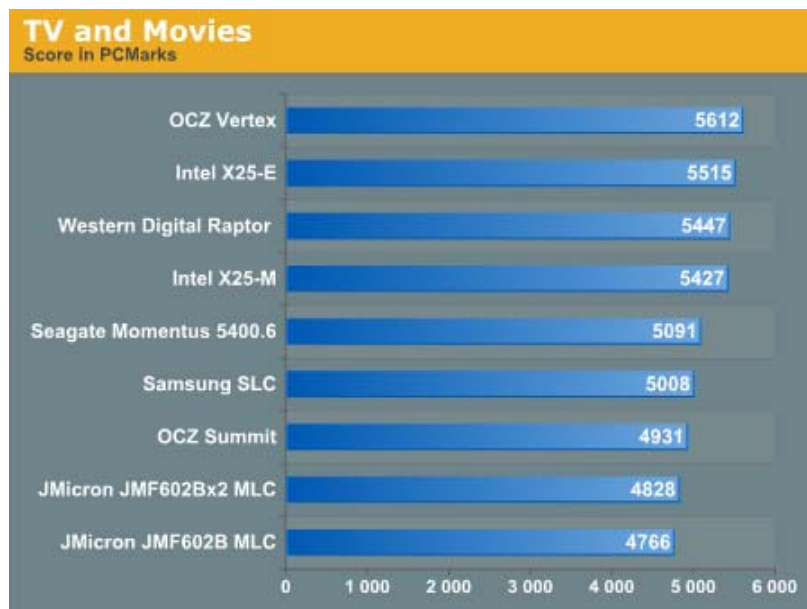
While PCMark does do a great job of measuring disk performance, it doesn't seem to stress random write performance as much, allowing the JMicron drives to relax a bit.

Now let's look at the individual test suites:

The memories suite includes a test involving importing pictures into Windows Photo Gallery and editing them, a fairly benign task that easily falls into the category of being very influenced by disk performance.



The TV and Movies tests focus on on video transcoding which is mostly CPU bound, but one of the tests involves Windows Media Center which tends to be disk bound.

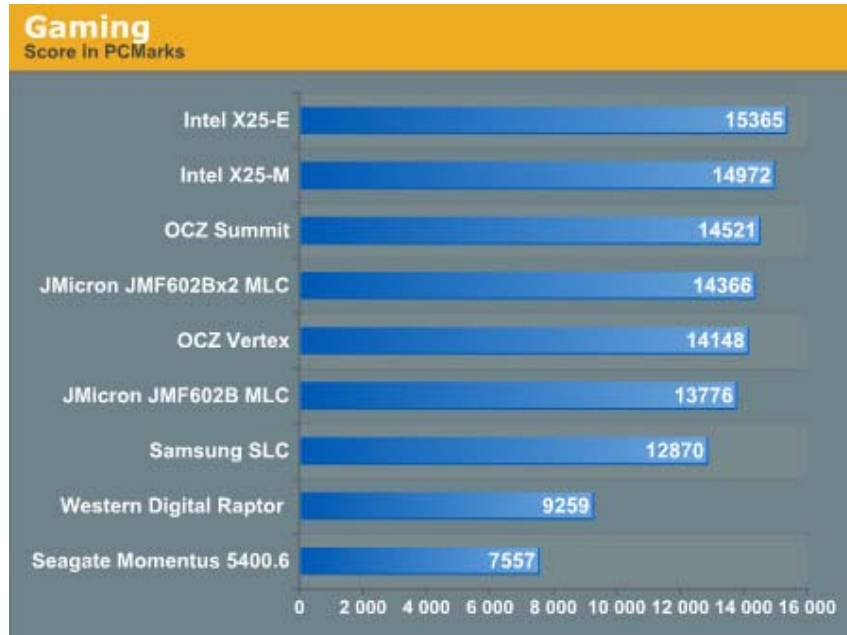


The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

SSDs won't always dominate and in many cases they won't offer tangible improvements over a fast hard drive.

The gaming tests are very well suited to SSDs since they spend a good portion of their time focusing on reading textures and loading level data. All of the SSDs dominate here, but as you'll see later on in my gaming tests the benefits of an SSD really vary depending on the game. Take these results as a best case scenario of what can happen, not the norm.



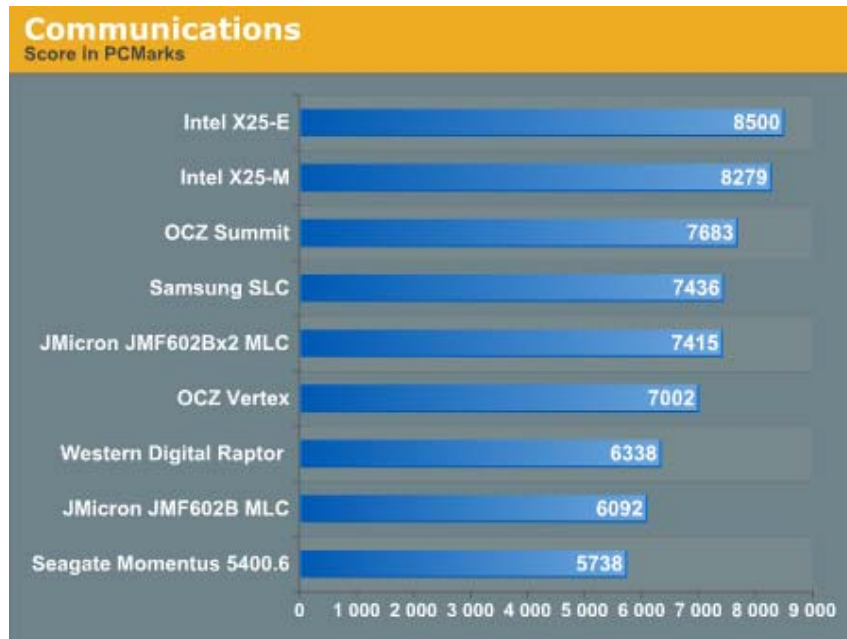
In the Music suite the main test is a multitasking scenario: the test simulates surfing the web in IE7, transcoding an audio file and adding music to Windows Media Player (the most disk intensive portion of the test).



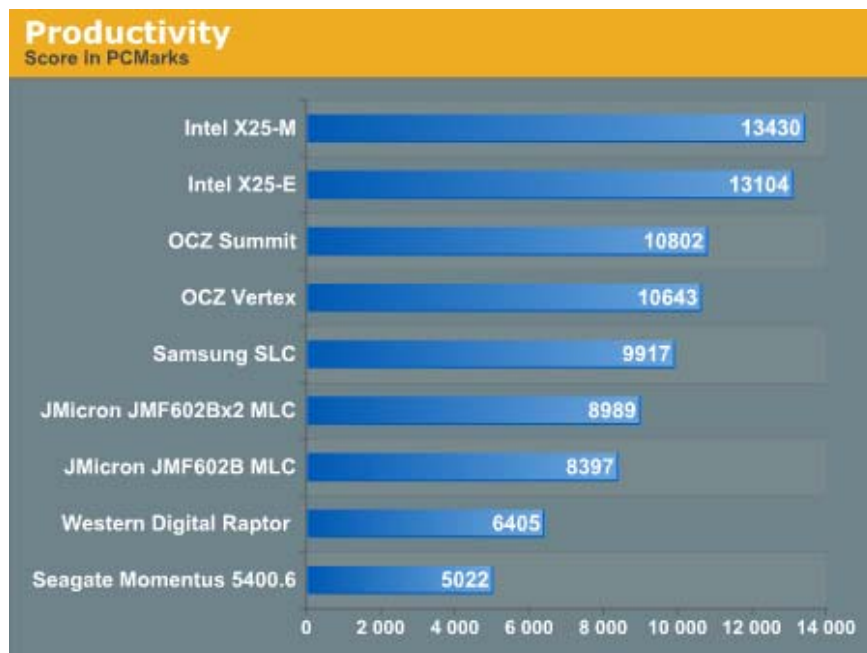
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

The Communications suite is made up of two tests, both involving light multitasking. The first test simulates data encryption/decryption while running message rules in Windows Mail. The second test simulates web surfing (including opening/closing tabs) in IE7, data decryption and running Windows Defender.



I love PCMark's Productivity test; in this test there are four tasks going on at once, searching through Windows contacts, searching through Windows Mail, browsing multiple webpages in IE7 and loading applications. This is as real world of a scenario as you get and it happens to be representative of one of the most frustrating HDD usage models - trying to do multiple things at once. There's nothing more annoying than trying to launch a simple application while you're doing other things in the background and have the load take seemingly forever.

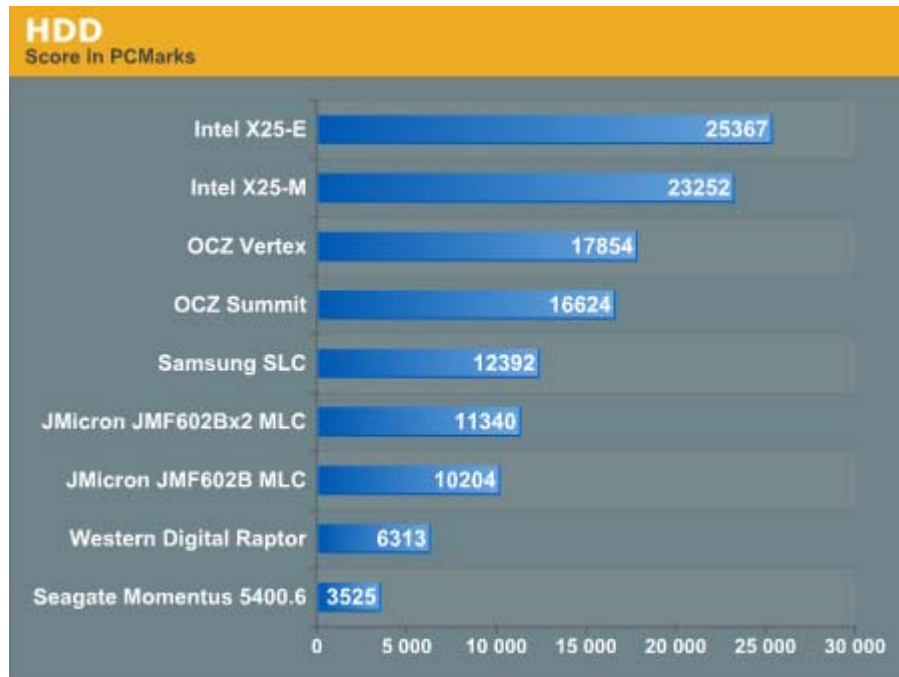


The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

The results here are the best characterization of my personal experience with the drives. The Intel drives are the fastest, a good 25% faster than the Summit or Vertex. Next up are the OCZ drives, with the Vertex dangerously close to the Summit. The older Samsung SLC is next in the standings, followed by the JMicron drives. There's a healthy combination of reads and writes going on here which benefits all of the SSDs, including the less desirable ones.

The final PCMark Vantage suite is HDD specific and this is where you'll see the biggest differences between the drives:



Again we're seeing the sort of breakdown we'd expect. The Intel drives come out ahead, while the Vertex is the best bang for your buck.

The Fresh Boot Test

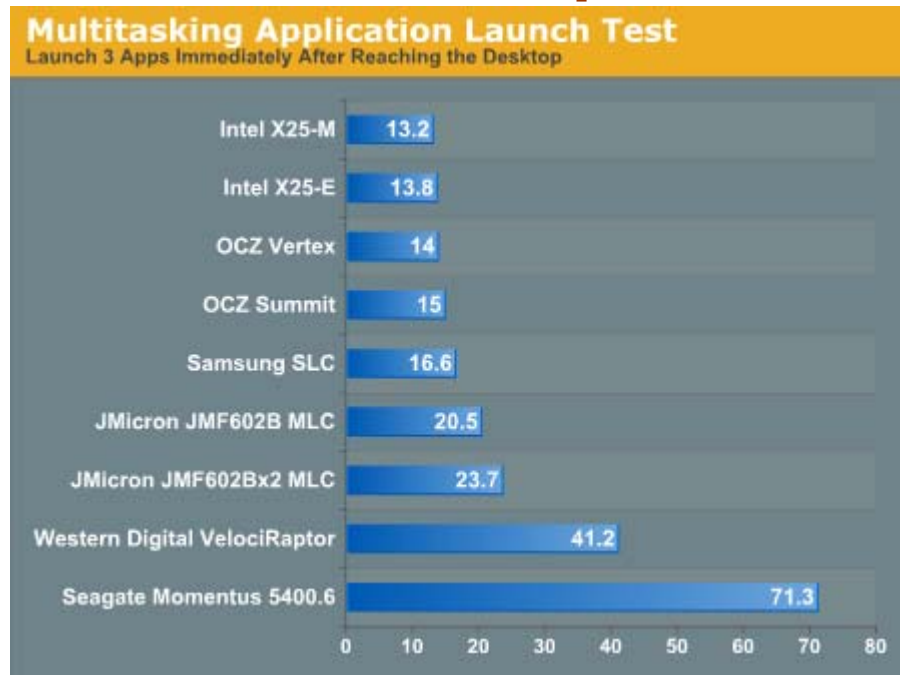
Allow me to set the stage. You just turn on your PC. I'm talking about a well used PC with tons of applications and data on the drive, not a clean test image. The moment you hit the Windows desktop you go and fire up the three applications you need to start working with right away.

If you ever wanted to know why SSDs are so much better, this is your reason. I ran through that exact scenario on our SSD testbeds. As soon as I hit the Vista desktop I ran Internet Explorer, Adobe Photoshop CS4 and Pinnacle Studio 12; I waited for all three to load, in the case of Pinnacle Studio I waited for my HD video project to load before stopping the timer.

The results speak for themselves:

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Everyone's beloved posterchild, the Western Digital VelociRaptor took 41.2 seconds to fully launch all three applications. Normal hard drives will fare much worse. The Seagate Momentus 5400.6, a high performance 5400RPM notebook drive took another 30 seconds on top of the WD time.

Now look at the SSDs; the worst SSDs we've got launch these applications in half the time of the VelociRaptor. The Intel X25-M will load the apps in about 13 seconds, barely a second longer than how long it takes to run Pinnacle Studio alone on an idle machine.

A good SSD makes Vista usable. All of the background tasks are nothing for these drives. If you ever sit there at an idle desktop and hear Vista go to town on your hard drive, those are IO operations that will bring any normal drive to its knees - or at least keep it busy enough to make all other IO requests take much longer than they should.

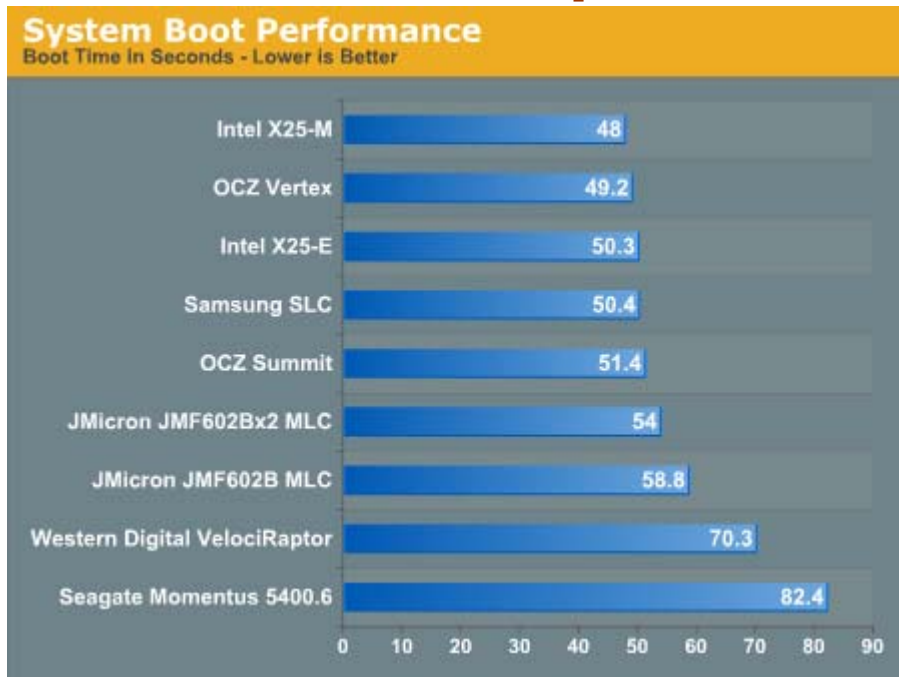
The SSDs that are worth recommending all deliver anywhere from 2x to 40x the number of IOs per second for small, random file writes compared to the Raptor. It doesn't matter how many Raptors you RAID together, you'll never achieve this sort of performance.

Application Launch Times

You know the drill; grab a stop watch and let's see how fast some of these apps run. The first on the list was just how long it takes to boot my application testbed. This is a fully configured machine with over 50GB of applications and data. I started timing as soon as I hit the power button and stopped it as soon as I got to the Windows desktop. Note that around 15 seconds of the boot takes place before the machine ever finishes its POST:

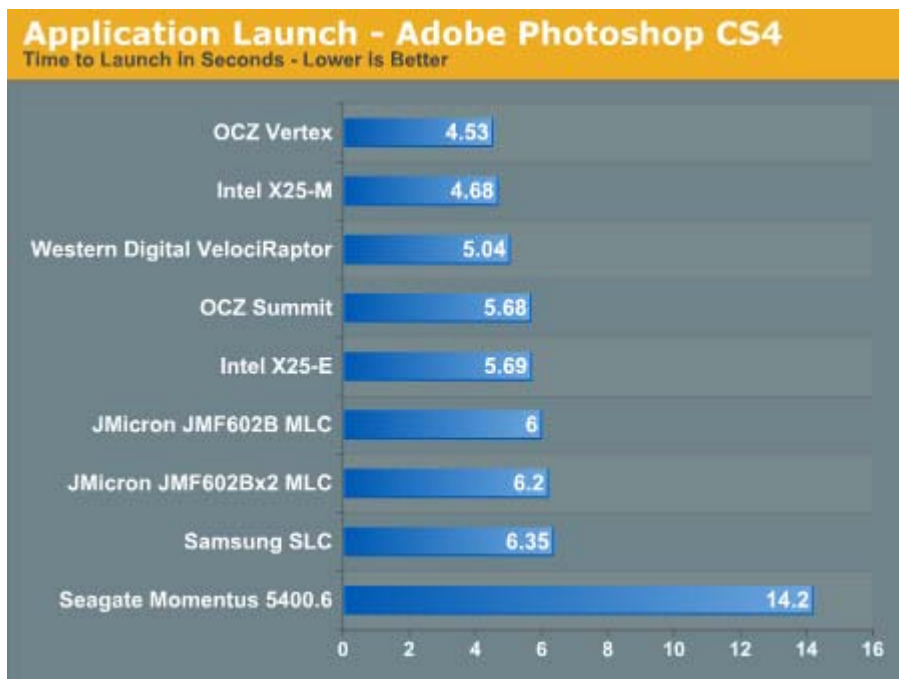
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Most of the SSDs perform similarly here and all of them are a good 20 - 30% faster than the WD VelociRaptor.

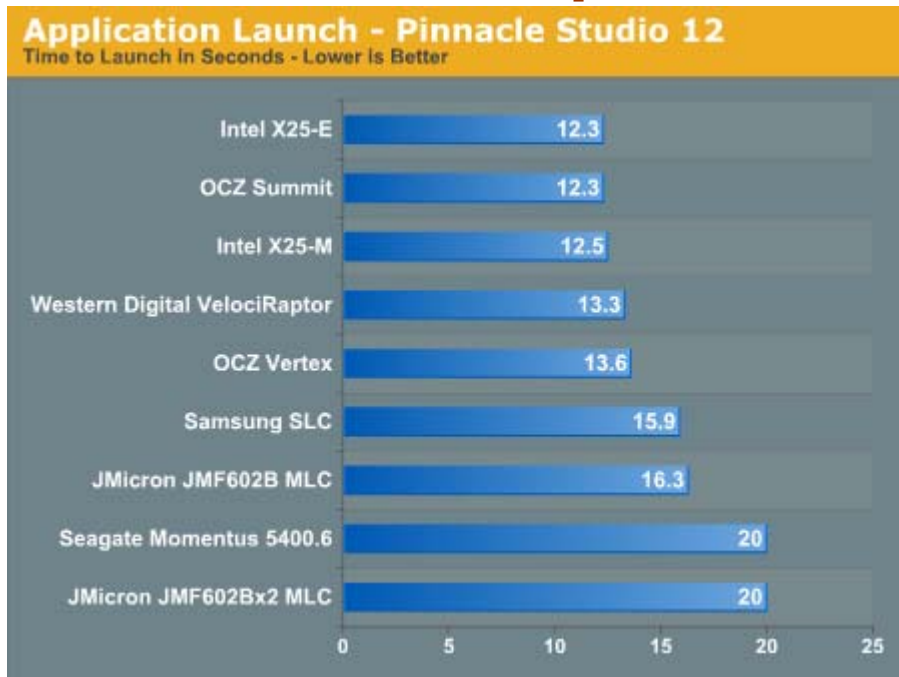
I tried launching Adobe Photoshop CS4:



Not all applications will launch faster than a VelociRaptor on a SSD, but let's not forget that the VelociRaptor is the world's fastest consumer desktop hard drive. Put one of these SSDs in a notebook and you'll completely redefine the word performance.

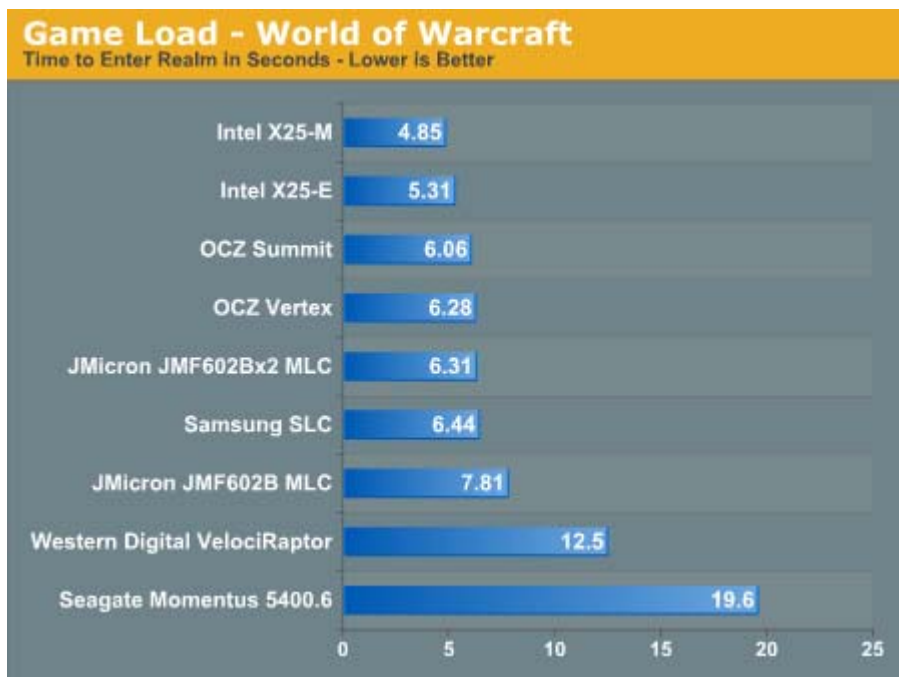
The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi



Game Loading Performance

I chose two games to show both ends of the load time spectrum. The first test is World of Warcraft, I'm simply timing how long it takes from the character selection screen to a fully loaded scene in the realm I've chosen. Fully loaded means no more popping textures or waiting for anything else to load:

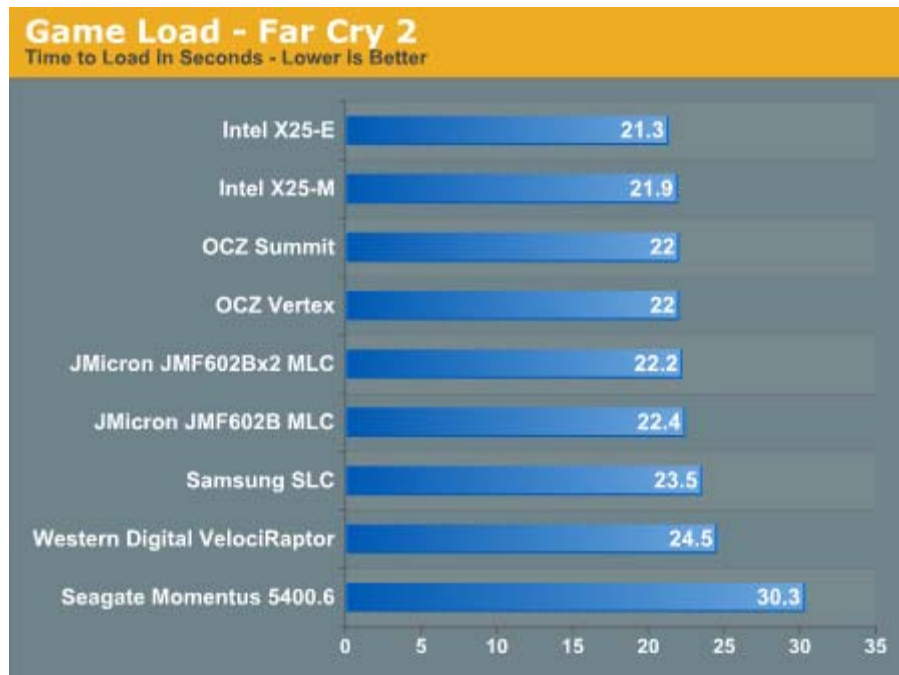


The Intel X25-M does the best here, even slightly outperforming the X25-E thanks to a more consumer-optimized firmware. The two OCZ drives occupy the third and fourth places, followed by the rest of the pack. No hard drive can keep up here thanks to the superior random read speed of a SSD.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

The Far Cry 2 test is simply running the default benchmark and timing how long it takes to get from clicking launch to the loaded level:



While the SSDs take the cake here, the VelociRaptor isn't far behind. Once again the two Intel drives take the lead, followed by the two OCZ drives. The Vertex and the Summit tie in performance.

Final Words

I began this article with a recap of my history with SSDs, stating that the more things change, the more they stay the same. Honestly, today, the SSD world isn't much different.

Drives are most definitely cheaper today; the Intel X25-M originally sold at close to \$600 for 80GB and is now down in the \$340 - \$360 range. The Samsung SLC drives have lost their hefty price tags and are now just as affordable as the more mainstream MLC solutions.

But the segmentation of the SSD market still exists. There are good drives and there are bad ones.

Ultimately it all boils down to what you optimize for. On its desktop drives, Intel chose to optimize for the sort of random writes you'd find on a desktop. The X25-E is much more resilient to the workload a multi-user environment would throw at it, such as in a server and thus carries a handsome price tag.

At first glance it would appear that Samsung's latest controller used in the preview OCZ Summit drive I tested optimizes for the opposite end of the spectrum: sacrificing latency for bandwidth. As the Summit was used more and more, its random write latency went up while its sequential write speed remained incredibly high. Based on these characteristics I'd venture that the Summit would be a great drive for a personal file server, while the Intel X25-M is better suited as a boot/app drive in your system.

The SSD Anthology Understanding SSDs and New Drives

Anand Lal Shimpi

I'd argue that Intel got it "right". Given the limited sizes of SSDs today and the high cost per GB, no one in their right mind is using these drives for mass storage of large files - they're using them as boot and application drives, that's where they excel after all.

Over the past year Intel continually claimed that its expertise in making chipsets, microprocessors and generally with the system as a whole led to a superior SSD design. Based on my tests and my own personal use of the drive and literally every other one in this article, I'd tend to agree.

OCZ and Indilinx initially made the mistake of designing the Vertex and its Barefoot controller similarly to the Samsung based Summit. It boasted very high read/write speeds but at the expense of small file write latency. In the revised firmware, the one that led to the shipping version, OCZ went back to Indilinx and changed approaches. The drive now performs like a slower Intel drive; rightfully so, as it's cheaper.

While I wouldn't recommend any of the JMicron based drives, with the Vertex I do believe we have a true value alternative to the X25-M. The Intel drive is still the best, but it comes at a high cost. The Vertex can give you a similar experience, definitely one superior to even the fastest hard drives, but at a lower price. And I'll spare you the obligatory reference to the current state of the global economy. The Samsung SLC drives have come down in price but they don't seem to age as gracefully as the Intel or OCZ Vertex drives. If you want price/performance, the Vertex appears to be the best option and if you want all-out performance, snag the Intel drive.

The only potential gotcha is that both OCZ and Indilinx are smaller companies than Intel. There's a lot of validation that goes into these drives and making sure they work in every configuration. While the Vertex worked totally fine in the configurations I tested, that's not to say that every last bug has been worked out. There are a couple of threads in OCZ's own forums that suggest compatibility problems with particular configurations; while this hasn't been my own experience, it's worth looking into before you purchase the drive.

While personally I'm not put off by the gradual slowdown of SSDs, I can understand the hesitation. In the benchmarks we've looked at today, for the most part these drives perform better than the fastest hard drives even when the SSDs are well worn. But with support for TRIM hopefully arriving close to the release of Windows 7, it may be very tempting to wait. Given that the technology is still very new, the next few revisions to drives and controllers should hold tremendous improvements.

Drives will get better and although we're still looking at SSDs in their infancy, as a boot/application drive I still believe it's the single best upgrade you can do to your machine today. I've moved all of my testbeds to SSDs as well as my personal desktop. At least now we have two options to choose from: the X25-M and the Vertex.