A Solid State of Affairs
Memory revolution or technological Flash in the pan?

Class: Digital Preservation
Lecturer: Kara Van Malssen
Student: David Neary

Introduction: To protect and save

Media preservation has traditionally placed a high priority on durability. Records in stone tablets survive from Sumerian times, more than 5,000 years ago; if not shattered with extreme force they can be still read today (provided the language contained on them has survived). Bookbinding is a thousand years older than the printing press, and serves as much to encapsulate a volume of pages as it does to protect them from exposure and physical stress. Paintings and photographs have largely always been framed whenever finances allowed. However the birth of audiovisual media since the late 1800s changed this emphasis on durability, largely due to the complexity of their recording. Wax cylinders and shellac discs could shatter with little effort. Outside of their protective metal cans, nitrate film reels could be set alight or ripped apart by the very apparatus used to view them. Plastic cassettes would come to the rescue of magnetic media, but even they could find their ribbons mangled in machines needed for playback.

Computer memory, the rise of which coincided with the birth of new media formats in the postwar era, similarly saw durability as an issue. Early concepts such as paper tape and punch cards were quickly done away in favor of magnetic media, which proved to be capable of storing vastly larger amounts of information and was sufficiently durable, from reels of tape to hard disk drives (HDDs) and 3½-inch floppies. Rugged and internally housed circuit-based memory such as RAM (random-access memory) and its dynamic cousin DRAM are volatile in their own way, in that a lack of power will cause a rapid loss of all saved data. RAM is thus useless for storage without a hard drive to back it up.

The need for speed

If durability and data retention are key aspects of all media, speed of access is another. (LTO-6 is only expected to last 30 years, while the Kish Tablet is 5,500 years old – and you only need to glance at the latter to find your place). Computer users often need to access vast amounts of data, much of which is active programs that must function alongside one another, processing documents, videos, or other functions. Under heavy workloads, even the fastest HDDs can suffer bottlenecks as their needles struggle to read all the information required of them. What is needed therefore is a non-volatile computer memory that is as fast as DRAM, as reliable as magnetic tape and as durable as stone. On paper, Flash memory looks like it could be just that memory type. Let's see just how durable that paper is under scrutiny.
Wait, so what is Flash?

And so, with a firm grasp of the need for durable, lasting and fast computer memory, we finally reach the issue at hand. Flash, in short, is a non-volatile storage medium that stores persistent data.\(^1\) Quite literally it stores individual electronic charge (actual electrons) in memory cells; if there is no charge a binary reading of 1 is recorded, if there is a charge, it reads as a 0. It does not need a steady power source to hold this charge. Flash memory is completely electronic, so unlike HDD or tape memory, it has no moving parts (you could say it is "un-revolutionary"). This means that it is largely silent and also less prone to physical breakdown in the way that HDD often is – although this comes with its own drawbacks as shall be detailed later on. While there have been many attempts to develop lasting solid-state storage over the past decades, from magnetic-core memory to experiments in ROM, and there are many new technologies in trial stages that would also quality as SSS, today when we speak about solid-state storage we are speaking about Flash, and when we speak about Flash, we are speaking of NAND Flash.\(^2\)

NOR and NAND. NAND or NOR?

Flash memory was developed in the 1980s by Dr. Fujio Masuoka, for Toshiba, and NOR Flash was first unveiled at the IEEE International Electron Devices Meeting in San Francisco in 1984.\(^3\) Named for the NOR logic gate, which registers a 0 bit under charge from any electron input (reading a 1 only when the charge comes from neither one input \textit{NOR} the other), NOR memory is extremely fast at reading back information, because it has true random-access memory. However, because it has far less storage space and considerably slower write times than NAND (and is also considerably more expensive), it is not commonly used for storage but rather as an internal replacement for RAM.\(^4\)

NAND Flash on the other hand is everywhere, in SSDs and Flash (thumb) drives, memory cards, cellphones and now laptops. The structure of its memory cells are considerably smaller than NOR’s, so its cost per bit is considerably lower. The name comes from the NAND logic gate, which records a 1 to memory unless electron input comes from two sources (so \textit{Not} from both one \textit{AND} the other). Toshiba debuted NAND Flash in 1991,\(^5\) with SanDisk developing the first marketed SSD later that year, at a price of more than $1,000 and a capacity of only 20MB. It has taken off at an astonishing rate since; contemporary SD cards are smaller than a postage stamp and can carry half a terabyte of data, while enterprise-level SSDs can hold 4TBs. Industry experts predict that by 2017 more than 227 million SSDs will have been sold, accounting for one third of all stored

\(^1\) Margaret Rouse, ‘Flash-based solid-state drive (SSD)’, January 2002, \url{http://searchsolidstatetorage.techtarget.com/definition/flash-based-solid-state-drive-SSD}

\(^2\) Ibid.

\(^3\) ‘NOR Flash Memory Invented’, \url{http://www.flash25.toshiba.com/#learn}

\(^4\) Rouse, op. cit.

\(^5\) ‘NAND Flash Memory Invented’, \url{http://www.flash25.toshiba.com/#learn}
The rapid rise of Flash can be seen in the following graph, which also demonstrates its various uses and their markets.

Figure 1: NAND Flash usage has more than doubled every two years since 2007

Flash is fast, with writing speeds of up to 5GBps. And Flash devices can certainly take a beating – the internal cells are on too minute a scale to be damaged by external forces, as opposed to HDDs which can crack internally if dropped. Anecdotal evidence goes further, with archivist Michael Grant telling me that a Flash drive of his, left in the pocket of a pair of trousers, once survived a washing machine cycle with no apparent loss of data. The challenge seems to be to try and break a Flash drive, rather than to keep one from breaking.

So if we’ve accounted for the high performance of Flash’s speed and its physical durability, what then for its data retention? In order to question this properly and draw fair comparison to other media types, we first need to gain a better understanding of how NAND Flash works.

**NAND structure: cells and blocks**

A NAND chip is structured on a progressively miniaturized scale. A chip is made up of planes, which are rows blocks; the blocks contain several thousand cells. Each cell contains a floating-gate transistor, which stores the charge, capped by a control gate. If the floating gate is charged it registers the charge as a 0. To recall this information, each control gate is pulsed by a signal from the NAND

---


9 Interview conducted by David Neary in person, New York, December 11, 2014
controller which measures the conductivity of the device channel, relaying the collective assembled data in a veritable 'flash'.

Figure 2: Structure of a Flash cell

Flash cells only held one bit of data per cell until the late '90s when Intel developed the capability for a Flash cell to hold more than one charge. Multi-level cell (MLC) Flash (as opposed to single-level cell (SLC)) revolutionized the market by making Flash considerably cheaper, more compact and also able to store more data. While SLC has been shown to be faster, last longer, and be less susceptible to data retention issues than MLC, it also costs 10 times as much. The number of read/write cycles permitted by Flash are also 10 times fewer for MLC. Tom Coughlin of the Solid-State Storage Initiative (SSSI) says: “Two bits per cell is very common in client and enterprise computing and three bits per cell is mainstream in consumer products such as USB sticks and SD cards and some client SSDs. There are controllers introduced to make three level cells common in more computer applications.” The potential for growth in Flash becomes increasingly great as the technology "addresses nano technology issues", although the increasing compaction of cells can lead to further problems.

NAND SSDs do not use random-access memory like NOR Flash, rather it uses a virtual mapping scheme to map the data to Physical Block Addresses (PBAs). “SSD virtual mapping is used for several reasons. For example, wear leveling algorithms distribute newly recorded data to new cell locations to promote even wear on the memory cells and thus improve the memory cell life

12 Interview conducted by David Neary via email, December 6, 2014
or endurance.”¹⁴ This is pivotal to the speed at which Flash works. Jon Jacobi writes that whereas in hard drives “data is basically written serially... Inside an SSD, data is written in a scattershot, parallel fashion down multiple channels to the multiple NAND chips at once. The more NAND chips an SSD has, the more channels it has to write/read across, and the faster the drive will be.”¹⁵

The complexity of Flash, and the reason even casual users find their drives fluctuating in memory allowances, is rooted in the relationship between cell and block. If we comprehend the minute scale of a Flash cell, in which electron charge can be written to and read from a single cell, it becomes clear that releasing that charge would not be possible on a cell by cell basis – it would require an electric signal to tell the cell to release its electric charge. This is why the cell/block relationship is so central to the technology: only by wiping an entire block of cells can you dissipate the charge in a single cell, returning every bit in that block to 1. Taking a 2GB Flash drive as an example, we see how defining an action this is. Each block in a 2GB drive contains 128KB of information¹⁶ – if we assume we are talking SLC (for simplicity’s sake, but it’s less likely), that is slightly more than a million cells (and bits) per block, and indeed there are several thousand blocks per drive. The scale is dizzying.

So in order to drop one bit of information, the Flash controller (the firmware that allows it to function¹⁷) must wipe the entire 128KB in that block. If there is information in that block that is to be kept, the controller copies it to another block or blocks, before wiping the first block completely. The constant writing, rewriting and deleting that block erase requires takes its toll on the cells, slowly wearing them out, and ultimately is what puts limitations on the numbers of input/output (I/O) operations allowed by Flash memory. These read/write limitations are countered by “wear leveling”, an algorithm used by the Flash controller¹⁸ to distribute data evenly amongst all of the device’s blocks. “Wear leveling ensures that no cell endures heavy use while another sits virgin next to it.”¹⁹

Memory limitations

The constant wear on the cells also affects the device’s speed, as tests by the SSSI have shown, with evidence suggesting that NAND can slow down to HDD speeds if the wear leveling controller is not efficient.²⁰ In order to test the performance

---

¹⁵ Jacobi, op. cit.
¹⁶ Micron, op. cit., p. 2
¹⁷ All NAND chips carry around 104% of their stated memory. This extra space is where the controller data resides. Jacobi, op. cit.
¹⁹ Jacobi, op. cit.
of various drives, the SSSI wrote to the drives repeatedly on a cycle for several hours and saw a rapid drop in their IOPS (input/output per second) before leveling off at a performance speed often less than 20% of that of its ‘fresh out of the box’ (FOP) status. The graph below\(^{21}\) demonstrates the swift fall in IOPS for the series of (anonymous) SSDs the SSSI tested – it is particularly worth noting that the one SLC they tested did not escape this performance decline.

![SSD Performance States - Normalized IOPS](image)

**Figure 3: Steady writing/rewriting reveals a rapid decline in SSDs’ IOPS performance**

The issue of read/write limitations is perhaps not as great a concern as might be first believed. It must be emphasized that the above graph demonstrating IOPS decline was under circumstances that no commercial user would ever subject their SSD to. SLC is rated to last 100,000 cycles of 100% block writing, while MLC will last 10,000 – this appears to drop by approximately a factor of 10 for every additional bit cells are designed to hold. This may seem to make MLC sound a risky investment, but you would still need to write the entire capacity of the SSD once every single day for 25 years to wear out all its cells.\(^{22}\)

There are major concerns for users of solid-state drives, but read/write limitations are not amongst them.

**Read Disturb and Data Retention issues**

Read disturb is a concern in Flash that has resulted in data corruption. As we will recall from earlier, Flash data is read via an electronic pulse sent to the control gate of the memory cell. That is to say that the act of reading Flash memory is a similar signal to the act of writing to Flash. Read disturb is what happens when this electronic pulse accidentally charges neighboring Flash cells, causing 1s to

---

21 Ibid., p. 2
22 Jacobi, op. cit.
read as 0s. Read disturb is thus equivalent to bit flipping. As Flash geometries continue to shrink, cells are compacted closer together and so the number of ranges of electrical charges required increase as cell densities increase.\(^{23}\) However, Paul Wassenberg of SSSI tells me that a well-designed controller can be used to address read disturb and that it is actually becoming less of a problem as new firmware developments are made.\(^{24}\) The fact that it is the act of regularly reading the same data that causes read disturb, Wassenberg posits, is one of the primary arguments for using SSDs as archival storage, since such materials are "seldom accessed".\(^{25}\)

If read disturb is a reason to consider SSDs for archival use, issues with data retention should strike fear into the digital archivists heart. The permanent storage of electrical charge is simply impossible. Turn your cellphone off for a week without plugging it in and it will not have the same amount of power remaining when you turn it back on. Charge dissipates, or leaks, over time. SSDs and other Flash memory devices are not immune to this, and the leakage from individual cells (turning 0s into 1s) is unpredictable and largely inevitable. Wassenberg has suggested that high temperatures can exacerbate this loss of charge, but there is as yet insufficient evidence to confirm this.\(^{26}\)

### High costs and crashes

Since speed of retrieval is certainly Flash’s highest selling point, it is the primary reason it sells at such a high price. David Rosenthal says that since Flash has many market benefits over HDD, “flash will command a premium over hard disk prices so that the market directs the limited supply of flash to those applications, such as tablets, smartphones, and high-performance servers, where its added value is highest”.\(^{27}\) While the cost of SSDs has dropped from $3/GB in 2005 to $0.50/GB in 2014,\(^{28}\)\(^{29}\) there is still no way for them to compete as storage-only against HDDs. Even if current trends continue, Henry Newman points out that SSDs will still cost around $0.15/GB in 2020;\(^{30}\) meanwhile just this week Seagate announced the shingled magnetic recording drive (SMR), which it claims can store 8TB at a cost of $0.03/GB.\(^{31}\) EMC’s graph, below, demonstrates that despite


\(^{24}\) Ibid. Interviewed conducted by David Neary via email, December 5, 2014

\(^{25}\) Ibid.

\(^{26}\) Ibid.


\(^{28}\) Newman, op. cit.


\(^{30}\) Newman, op. cit.

falling SSD prices, the steady decline in the cost of HDD makes it unfathomable the two could ever match prices.\footnote{32}

\textbf{Flash vs HDD : Industry Cost Trends}

MLC 8X More Expensive Than Performance HDD Through 2016

A final concern to raise regarding SSD is spontaneous failure. Technology fails, sometimes for no apparent reason. As was noted earlier, SSDs are silent devices, with no spinning, whirring parts. While this keeps them a lot cooler than their hard disk cousins, it also means they provide no audible cues that the end may be nigh. And when an SSD crashes, it crashes hard. While information can be salvaged from HDDs using various rebooting techniques, for the most part “a dead SSD is a dead SSD”.\footnote{33} Jacobi points out that companies such as DriveSavers have a very high success rate at salvaging data from bricked SSDs, but admits their prices are very high also.\footnote{34}

\textbf{Archiving and industry}

Can Flash be used to store data long-term? The major concern here is not read disturb or data retention, but simply the lack of testing. The technology is evolving at such a head-spinningly rapid rate that the research simply doesn’t exist. “SSDs are not a place you want to keep data for 10 years,” says Howard Marks of DeepStorage LLC, “as the charge will dissipate and you will lose data.”\footnote{35} The question is not ‘how long do SSDs hold data?’ It’s ‘can we trust them to store the data for as long as they claim?’ All preservation is an active process, transferring will always be required. Indeed LTO tape may be guaranteed for 30


\footnote{33} Rick Broida, ‘With an SSD, Backups Are More Important than Ever’, May 2013, \url{http://www.pcworld.com/article/2039540/with-an-ssd-backups-are-more-important-than-ever.html}

\footnote{34} Jacobi, op. cit.

\footnote{35} Schwartz, op. cit.
years archival, but with a generation going obsolete every four to five year due to backwards-incompatibility, the higher number feels largely redundant. If SSDs could be guaranteed for five years, accompanied by an active preservation plan they would serve as well as LTO and provide much faster retrieval of data – it’s the cost, then, that gets in the way.

“For workloads where storage capacity is critical and performance is not, such as backup and archiving, you want to make sure you are not using flash at all,” recommends Dell. It seems therefore impossible at present to recommend Flash for archival purposes. Tom Coughlin, however, adds: “There are organizations who have advocated Flash memory as a special type of archival media. While Flash will not be as low cost an archival media as magnetic tape or optical discs it might be used as a write once read many (WORM) media with higher performance than magnetic tape or optical discs. The additional cost might be acceptable for some niche applications.” At present the only conceivable argument for NAND storage over other archival data formats is the amount of space it takes up, as the graph below demonstrates. But as we have discussed already, the smaller Flash gets, the more prone it is to error.

![Volumetric Density (TB/in³) for best case scenarios](image)

Figure 5: NAND stores data in a considerably smaller volume of space than other memory formats

If archives do choose to use Flash, therefore, they are taking a gamble that no research has yet implied could pay off. Recommendations for Flash archival use are therefore largely inconsequential, but can be summed up thus:

- Store in a controlled cool environment: Evidence suggests colder temperatures may help prevent Flash cells from losing their charge and thus their data

---

37 Interview conducted by David Neary via email, December 6, 2014
38 Fontana, op. cit., p. 15
Plan for failure: Expecting SSDs to no longer function after 10 years means that yearly checking of data integrity will be required for as long as the device is in use. This could become very labor-intensive depending on the size of the archive.

Backups and backups: Have more than one backup, and frankly, one of those backups should be on a format other than SSD. This may seem to defeat the purpose of using SSD at all, but it’s better than having the SSD crash and defeat your data.

As Flash becomes more common in personal computing (I am typing up this paper on an all-Flash laptop), its key benefits clearly lie in enterprise, where its ability to store data and process it at rapid speeds with low latency cannot be beat. Mark Peters writes: “What was in 2008 just a “vitamin supplement” to the regular storage diet of traditional spinning disk is now available in a whole “flash-aisle” of options, spanning hybrid implementations, all-flash arrays, and server-side deployments.”

Hybrid drives are becoming increasingly common, supplementing reliable HDD storage with SSD processing, to “intelligently cache and put hot data on the much higher performing SSDs, leaving cold data on the hard drives.” It is because of Flash’s processing speed that enterprise revenues alone for SSDs are expected to hit $3.5 billion in 2016.

Conclusion: A future for SSD

Flash memory has changed the nature of digital storage, but also the nature of the transfer of information. A Flash drive can have files uploaded to it, be passed to a neighbor and have the files downloaded from it, faster than most broadband transfers. Until the rise of cloud storage it was for a decade the most reliable way to move large files from one computer to another. But while a Flash drive may survive its day in the washing machine, evidence still fails to present itself that Flash can safely hold data for periods in excess of 5-10 years while sitting comfortably on a shelf.

Advancements are coming however; Samsung are pioneering 3D V-NAND Flash, with claims that it will last 10 years longer than NAND and consume 50% less power. The system works by stacking Flash cells in a vertical direction as well as horizontal. This means more cells can be fit in each block, but with their actually being less compressed than they are currently in NAND – thus reducing read disturb risks.

---

40 Schwartz, op. cit.
41 Ibid.
42 Dodson, op. cit.
43 Ibid.
SSD technology is beginning to move away from NAND, with phase-change memory (PCM) an emerging technology that industry experts will want to watch closely. PCM works on the concept that chalcogenide glass has a different molecular structure depending on the speed at which it is cooled, reading its two different crystalline forms as a 1 or 0. Exceptional switching speed and high storage density are promised, but how long it will be able to hold its memory for has yet to be revealed. In addition, no suggestion of its cost has yet been given.

Moving forward, Flash will continue to chew up its third of the memory market, and provide users with faster access to their stored (and hopefully backed-up) materials. It seems unlikely Flash can ever be trusted for long-term archival use, as the dissipation of charge is random and unpredictable. In the meantime, for short-term storage and rapid backup, Flash is hard to beat. And as it’s shown over the past 30 years, it can only get better.

---

Webography


