When Good Memories Go Bad – Data Recovery of Flash Media

Sean Barry, Kroll Ontrack, www.krollontrack.com

Computer hard disk drive storage has an impressive history of technological advances, from the RAMAC storage systems in the 1950s to perpendicular recording introduced a few years ago. The advancement of these technologies has increased the density of today's hard disk drives to the terabyte range. Along the way frustrations with new storage technology have been proprietary equipment formats, initial high investments, and failed expectations of reliability.

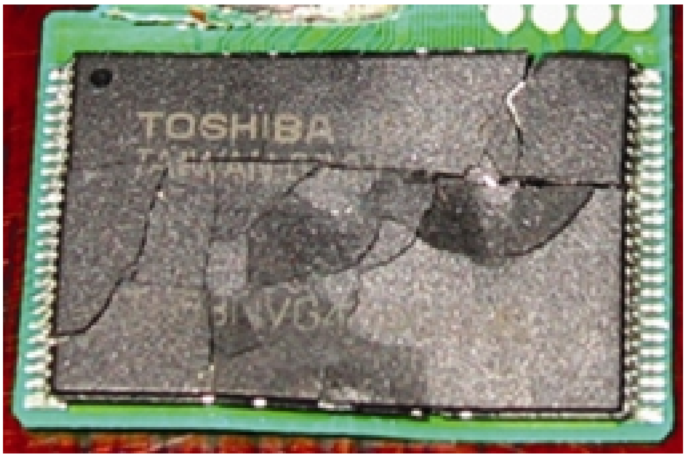
Even though there are many options available for the preservation of data, many users fail to make consistent use of data protection technologies. Digital data is delicate regardless of how it is stored and data loss at any level is frustrating for all computer users, and for a large number of users it is devastating. As computer users, we have come to accept that we will have some sort of data loss sometime during our lives. Yet, the user's perception of the reliability of that storage device changes after a data loss. Users come to believe that the storage media is not reliable and that it will never reliably hold data again.

Over the past decade, users have become increasingly reliant on flash memory technology as a means to store data. Due to its small design and wide application, users have access to more flash storage than perhaps they appreciate. Personal music devices with embedded flash technology have replaced the portable 'boom-box' from the 1980s. Flash memory has revolutionized the photography field to the extent that photographic film is no longer the mainstream method of taking or storing photos. In addition, the storage capacity contrast from removable magnetic floppy disks to flash 'jump drives' as a means of affordable yet portable storage has changed user data transportation habits.

It is clear that flash memory storage has become a viable alternative to legacy data storage devices and media because of its low power usage, reduced heat displacement, and the complete lack of mechanical parts, giving it a kind of 'solid-state' structure. Despite its many benefits, data loss is still a threat and the same is true of flash memory storage. In fact, data recovery of flash memory storage has been on the rise as more consumers use the technology to store their electronic files.

Data Recovery of Flash Media Storage

Although it's 'solid state structure' does limit breakage on some level, flash storage can suffer the same traditional causes of data loss as traditional storage. Data recovery service companies regularly see many cases of data loss from flash drives. Physical damage often affects flash devices because of their small size. Figure 1.1 shows examples of physical damage to portable flash memory. The first picture shows the result of being at the bottom of a travel bag or purse, , The 'jump drive' simply got crushed. The middle picture shows a camera flash memory card described by the user as being, "attacked by a large animal." The third picture is an x-ray of another camera SD card that had sustained damage to the circuit board traces.



An example of physical damage to flash media.

The good news is that not all levels of data loss, even those resulting from physical damage, are terminal to flash media. If a flash device has suffered any level of physical damage then a reputable data recovery service company should be sought for an evaluation. The goal of the recovery effort is get the device operational so that the data can be retrieved. In some cases this requires small, temporary repairs to the device. In other cases, the recovery effort may need to go further.

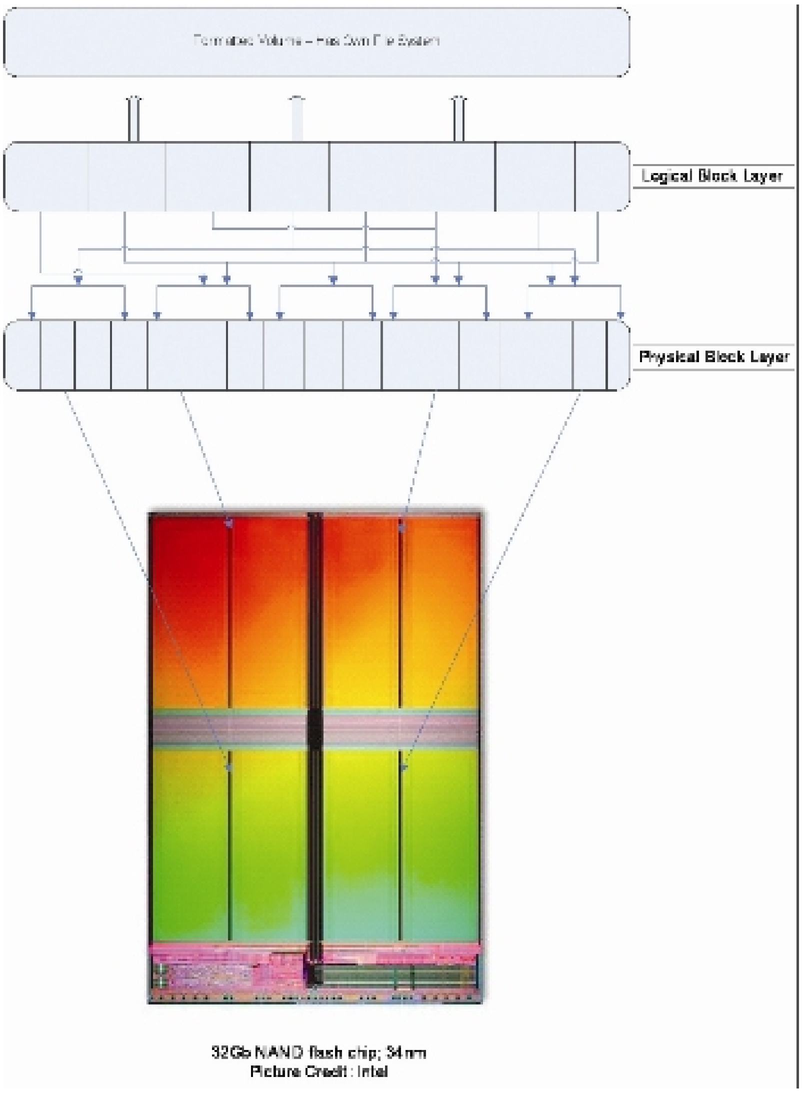
Most flash storage devices are comprised of a few semiconductor chips. One or two of the chips will be the actual flash memory chips. The remaining chip will be the flash controller chip which is the gate-keeper for all of the data written to the flash memory. The flash controller chip is an extremely complex microprocessor that not only manages how data flows in and out of the device, but it also controls how the data is written to the flash chips. The controller even corrects any errors that may develop during data transfer and storage.

If the flash controller chip of the device becomes inoperable due to damage, then the only way to recover the data from the flash memory chips is to remove the flash memory chips and attempt a raw extraction of the data. This is a highly technical task requiring special tools and expertise. Do not try this at home! Because they are so small, flash memory chips have been soldered with extremely thin interfaces and connections. Too much force exerted to remove the chips or too much heat applied during the extraction process may damage the flash memory chips, or compromise the inner substrate containing the microscopic floating gate memory cells.

Due to the proprietary nature of controller chip design, the extracted chips cannot be reattached to a similar flash memory device. Even devices made by the same manufacturer have a wide variety of data communication protocols and data management. Data recovery engineers place the flash memory chips in special equipment and use custom software to gain access to the data.

The next challenge of the data recovery effort is to organize the extracted data.

Traditional hard disk drives store data in a linear fashion with available space starting at the outer edge of the platters and concentrically moving toward the center of the platters. Flash memory storage is much more dynamic. A way to understand this is illustrated in figure 1.2. Flash data storage is similar to how earth’s continents are floating on top of a layer of molten magma. As in earth’s geophysics, data in a flash memory array is moved and redistributed while the upper layer, the file system, remains intact. The reason for this low-level data management is so that the floating gate cells will have an equal amount of data written to them, also known as programming cycles.



An Illustration of layers from the flash die to the formatted file system

The method employed to inject electrons into the floating gate causes the gate to fail to hold an electron charge over time. If a block of flash memory cells is constantly written to, then those cells will quickly wear out. In an effort to manage the endurance of the flash memory array, a process called ‘wear-leveling’ is used by the controller chip. This process equalizes the number of times specific areas are written to as well as manages the frequency of when a group of flash cells was last written. For example, if a file has been copied to the flash device and has not been changed or moved for a specified period of time, the flash controller will reprogram another area of the flash memory array with the same data, then will perform an erase function on the original area, freeing those electrons and preventing those cells from becoming stagnant. Therefore, the data on any flash storage device is in a constant state of movement—regardless of whether the user recopies files to another folder on the device.

In a recovery scenario where the flash controller chip is no longer managing the data, wear-leveling algorithms and logical block addressing schemes create highly complex challenges. In addition, if the flash device has two or more flash memory chips the data may be distributed, at a byte level, between chips to improve performance. This is referred to as a plane distribution. This distribution scheme also serves to evenly wear two chips at a time, instead of filling one up with data first and then going to the second chip. Finally, some manufacturers will divide areas of each chip into regions and organize them into groups of logical organization for multi-operational microprocessors. This equates to a third layer of distribution employed to balance the data storage. What keeps all of this data together?

The flash controller uses a translation scheme to manage the logical block addressing (LBA conversion) and this is called the Flash Translation Layer (FTL). This translation method is usually proprietary for each manufacturer. Due to the limitation of the floating gate cells, clever FTL data management techniques are employed so that the longevity of the flash memory is maintained. However, not all manufacturers make devices with the same data layout, FTL, or wear-leveling processes. In fact, manufacturers can change designs within a product-line very quickly. The average user would never notice these differences from one device to another.

The challenges of data recovery from flash media can be summed up as:
* Proprietary data organization
* Address line complexity (microprocessor communication routes)
* Flash memory chip density (number of flash chip configurations within an array)
* Implementation of customized wear-leveling algorithms

The complexity of these storage devices is by design. Performance, NAND cell density, and larger storage capacities are what are driving the research and development departments to explore different ways to meet consumer demand.

Consumer demand also requires data protection on these devices. Due to the highly transportable nature of flash devices, users want to ensure that whatever data files they have are protected. While the goal of encryption software is to protect the data, in a recovery situation, the goal is to retrieve the data with the user’s consent and appropriate decryption keys. However if the methods employed for encryption are just as device exclusive, complex, and in flux as the data management schemes, then the risk is that recovery of critical data may be impossible. Consumers of flash storage should insist that manufacturers provide a complete storage solution that includes a data recovery option not if, but when needed.

As computer users, we all are familiar with data loss and that it always happens at the wrong time. While the advantages of flash storage are strong (i.e. no moving parts, device is more rugged than a hard disk drive, etc.), flash technology has write limitations and it eventually wears out. Therefore, before adopting this technology at the enterprise level, those responsible for the decision need to have a discussion with product vendors regarding warranty coverage, and how data loss will be mitigated in the event of a data disaster.

Future of Flash Data Storage and Recoverability
Non-volatile flash-based memory has a long road ahead of it to completely replace hard disk drives as a primary storage medium. Manufacturers are working hard to produce new technology that will be in mainstream use within the next few years. The prices and manufacturing process for flash memory is improving and the cost to produce the chips is getting less expensive, making this technology more accessible to computer users. Furthermore, increasing the performance and reliability of flash memory will make it more viable for enterprise and mobile computing.

The flash memory industry will benefit from consolidation and standardization of technology and data management methods. Manufacturer sponsored organizations are trying to bring all manufacturers of flash devices together to agree on the storage mechanics, from the address line, internal bus architecture, AT command set, to reliability and testing. All of this is important because consumers want to know if their data storage devices are going to be reliable as well as meet the demand of performance and capacity.

The flash memory storage industry has grown out of the semiconductor manufacturing arena and has become a new technology that is already proving itself. For example, two large companies from the computing and storage segments have already begun offering flash based solid-state drives (SSD) as part of their enterprise storage offerings. In the mobile computing sector, two high profile laptop/notebook manufacturers have begun offering SSD drives as an option for their systems. More recently this year, ‘Netbooks’ have entered the retail market stream and are employing a complete flash memory storage system with an embedded operating system. These smaller versions of laptops may prove to make a big splash with early adopters of cloud computing.

The next 3-9 years in the flash industry will be interesting as engineers and manufacturers produce new ways for non-volatile memory to replace the aging, but still reliable, hard disk drive. However, it is unlikely that flash technology will completely overtake hard disk drive technology in the near term. Aforementioned issues of cost, reliability, and technology consistency need to be successfully addressed in order for mass replacement to occur.

For full adoption of non-volatile memory storage to take hold, the issue of data reliability is an area that requires a complete solution from product vendors. Consumers of this technology will require a simple, industry-wide metric of reliability for the flash products available. Finally, a consolidation within the data management process will provide a unified storage method that will provide equal comparisons among products. If the worst should happen and a data loss event affects a user or a large computer system, then a quick recovery will be possible and the original data recovered.