RUGGED AND RELIABLE COTS STORAGE SOLUTIONS FOR DATA ACQUISITION SYSTEMS

Ofer Tsur, BSc, MBA
M-Systems
Kfar Saba, Israel

ABSTRACT

Due to the rotating mechanism in mechanical disks, they cannot provide the top-level reliability required for operation in harsh military environments. This paper describes three COTS alternatives to mechanical disks: ruggedized mechanical disks, solid-state flash disks and stacked PC Cards. It discusses their cost-effectiveness and aspects such as environmental specifications, endurance and data reliability. It highlights several methods used by flash disks to enhance endurance and reliability, as well as flash pricing and density trends. It presents data security requirements in actual emergency situations, and how flash disks can meet these requirements. It concludes with a feature-by-feature comparison of ruggedized disks, flash disks and stacked PC Cards.

KEYWORDS

Solid-state flash disk, NAND flash, SCSI, IDE, Sanitize

INTRODUCTION

Reliable data storage is a major concern for engineers designing military data acquisition systems in equipment such as data recorders, telemetrics and reconnaissance. All of these systems must operate under harsh environmental conditions. Although rotating mechanical disks provide very high storage capacity of more than 200GB and are priced at less than $500, inherent design limitations make them unsuitable to operate reliably under environmental extremes.

Their mechanical design, based on spinning platters and head-arms that read/write the information from/to the disk platters, cannot insure data integrity. In addition, mechanical disks do not operate reliably in temperatures outside the range of +5°C to +55°C. Military applications must guarantee reliable performance within the industrial temperature range of -40°C to +85°C. Another drawback of mechanical disks is their low tolerance for shock and vibration. They operate under maximum shock levels of 125G for 2.5” IDE/ATA disk (common for laptops) and up to 65G for SCSI disks (common for servers/desktops), and maximum vibration levels of up to 1G. These shock and vibration levels do not comply with MIL-STD requirements for tracked and wheeled vehicles, nor do they meet MIL standards for airborne and shipboard use. To overcome the limited reliability of mechanical disks, COTS solutions have been introduced to the marketplace.

RUGGEDIZED MECHANICAL DISKS

A rotating mechanical disk can be ruggedized by sealing it in a rigid cartridge. This cartridge encompasses the entire disk mechanism, including electronics, protecting it from high humidity and altitude fluctuations. Advanced sealed cartridges are used for embedded closed-loop servo systems, automatically compensating for temperature variations to ensure reliable head positioning over the entire operating temperature range. Ruggedized mechanical disks are available in capacities of up to 160GB, and deliver sustained read/write performance rates as high as 40MB/s. The cost of ruggedizing mechanical disks ranges from hundreds to thousands of dollars, depending on the ruggedization level required.

Although a sealed cartridge improves the ability of mechanical disks to withstand high altitude, humidity and higher level of shocks and vibrations, additional factors need to be considered. Sealing a mechanical disk doubles and sometimes even triples the size of the unit, in addition to adding excessive weight. A larger and heavier unit for airborne applications within helicopters and fighters translates into a high dollar premium, which is measured per square mm.

In order to improve the shock and vibration specification of a ruggedized mechanical disk, some designers mount it on a shock absorber. If the casing is insufficient to meet the required operating temperature range, a heating and cooling device is used. But the addition of these two components add weight and size, while consuming more power and making the total solution more costly. Based on current pricing for high capacity disks of more than 30GB, ruggedized mechanical disks are cost-effective. But for smaller capacities, the cost of ruggedization is too high per MB/GB to make them a viable option.

SOLID-STATE FLASH DISKS

Flash disks are solid-state with no moving parts, thereby eliminating seek time, latency and other electro-mechanical delays inherent in conventional disk drives. Since flash is a non-volatile memory technology (vs. volatile memory technology such as DRAM and SRAM), it retains data when power is off, as do mechanical disks.

Flash memory has become an ideal solution for replacing mechanical disks where reliability is a key requirement. For flash disks to provide true “drop-in replacements” for mechanical disks, they have identical dimensions, the same mounting holes and the same interfaces. The most common flash disk form factors are: 2.5” (laptop size disk) with IDE/ATA and Narrow SCSI interfaces and a 3.5” (desktop size disk) with Narrow SCSI and Wide SCSI interfaces.

Flash disks operate in the harshest environmental conditions, as defined in MIL-STD 810F. They are used as rugged data storage within military and aerospace systems, delivering significantly higher reliability and a maintenance-free solution as opposed to traditional mechanical disks that are susceptible to mechanical failures and performance degradation in harsh conditions. Solid-state flash disks operate within -40°C to +85°C temperature ranges, absorbing shock conditions at 1500G and random vibrations of 16G at 80,000 feet altitude.

Table 1: Mechanical Disk vs. Solid-State Flash Disk

Category		Mechanical Disk	Solid-State Flash Disk
Environmental Specifications	Operating Temperature Range	+5°C to +55°C	-40°C to +85°C
	Non-Operating Temperature Range	-40°C to +70°C	-55°C to +95°C
	Operating Shock	20G - 25G	1500G
	Operating Vibration	1G (22-500Hz)	16G (20-2000Hz)
	Humidity	5%-90%	5%-95%
	Operating Altitude	15,000 ft	80,000 ft
	Acoustics, Idle/Ready	2.9 (Bels)	0
Environmental Standards	Shock and Vibration MIL-STD 810F	Does not comply	Complies, MIL-STD 810F
Reliability	Actual/Fielded MTBF	<70,000	>700,000
Performance	Average Seek	7.0-3.0msec	0.5-0.02msec
	Average Latency	5.0-2.0msec	None
	Sustained Read Rate	15.0-50.0MB/sec	2.0-30.0MB/sec
	Sustained Write Rate	10.0-40.0MB/sec	2.0-20.0MB/sec
Power	Power Idle	5.0-0.8 Watts	1.0-0.035 Watts
	Power Read/Write	10.0-5.0 Watts	3.0-0.325 Watts
Security	Delete disk data in emergency without retrieving the data	Does not comply, requires degaussers or total destruction	Complies with: - Fast Secure Erase - Sanitize (Purge)
Capacity	3.5” Form Factor	40GB-200GB	128MB-100GB
	2.5” Form Factor	20GB-60GB	128MB-40GB
Cost	Procurement	$70-$600	<$600/GByte

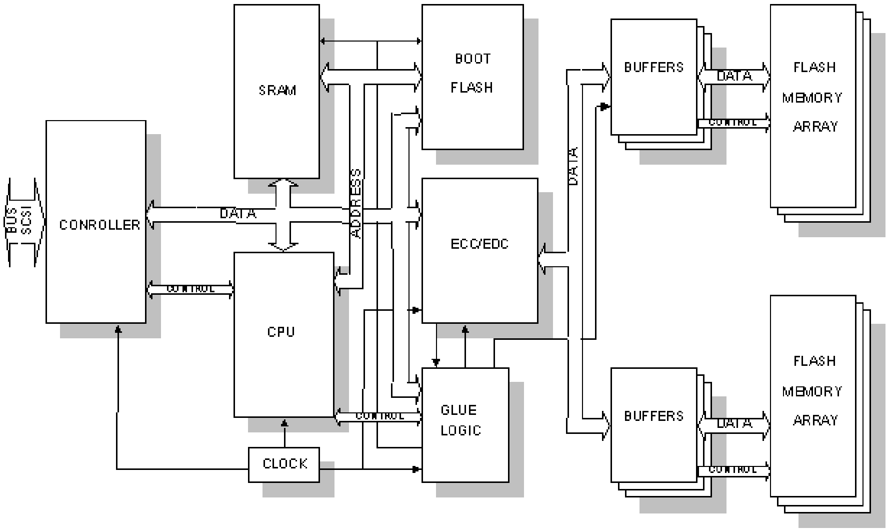

Figure 1: Solid-state flash disk based on SCSI or ATA controller, Error Detection and Correction Code and a CPU, enabling the disk read/write operations to the flash array.

Figure 2: Solid-state flash disks operate within -40°C to +85°C temperature ranges, absorb shock conditions at 1500G and random vibrations of 16G at 80,000 feet altitude.

CHOOSING THE RIGHT SOLID-STATE FLASH DISK SOLUTION

Choosing the right flash disk is much more complicated than simply choosing the right storage interface (IDE/SCSI/Wide-SCSI/FC) and the required disk capacity (in MBytes/GBytes), or evaluating the product that will deliver the best disk performance (sustained read/write rate). Since NAND flash technology has some inherent limitations that affect reliability, the right flash disk must overcome these limitations to meet the extremely high reliability requirements of mission-critical applications.

Addressing the issue of reliability requires understanding how flash works. Flash is non re-writable and must be erased before it can be written to again. Flash is erased in blocks (typical block size is 4 to 64 KBytes), which are much larger than disk sectors (512 bytes). In addition, flash has a limited number of erase cycles of 100,000 to 1,000,000 (flash has no limitation for read operations). NAND flash can accumulate up to 2% of bad blocks during the manufacturing process, as well as additional bad blocks during flash operation. The number of bad blocks also accumulates due to the write/erase operation, during which electrons are captured in an oxide layer and create internal electrical stress. The stress is created when only one block is accessed repeatedly by an erase/write operation, while the remaining blocks are untouched. This stress increases the bad block accumulation rate during flash erase/write operation.

ENHANCING SOLID-STATE FLASH DISK ENDURANCE

Some flash disk manufacturers incorporate methods to enhance endurance. One of the most common techniques is called *erase before write* or *counter wear-leveling*. This method implements a counter for every flash block to count the number of write cycles. Every new write operation is preceded by an erase operation to the block where the data will be stored, and the block counter is updated. When the specific counter reaches the flash erase cycle limit, the block is marked as a bad block. Although the data in that block can be read an unlimited number of times, it cannot be written anymore. When the application tries to execute a new write operation to such a block, the new data is stored in a different block taken from a pool of spare blocks, and a pointer points to the new location. The "erase before write" algorithm intensively wears out the flash over time, as the whole erasable block is erased every time even though only part of the block is updated. If the application executes write operations to the same location over and over again, these blocks will reach the erase cycle limit, decreasing the total flash disk capacity available for write operations.

A more advanced method to enhance flash disk endurance, while keeping the entire media size available for write operations without a decrease in capacity over time, is called TrueFFS® (True Flash File System). TrueFFS technology uses *dynamic wear-leveling*, *virtual mapping*, *garbage collection* and *bad block mapping-out* (BBM) algorithms to optimize flash usage.

The dynamic wear-leveling algorithm guarantees that all blocks in the flash disk are erased the same number of times, preventing the application from repeatedly writing to the same location until flash blocks wear out. Dynamic wear-leveling enables the entire available capacity of the flash disk to be used for write operations. TrueFFS performs virtual mapping of logical sectors to physical blocks, transparent to the user's application. The TrueFFS garbage collection process eliminates the need to erase the whole block prior to every write. This process accumulates data marked for erase as garbage, and performs a whole block erase as space reclamation in order to re-use the block for the next write operation. Once a block becomes problematic, the BBM algorithm marks the block as a "bad block" so that TrueFFS no longer uses it. Large pools of spare blocks, up to 4% of the flash disk capacity, are used to increase disk endurance.

IMPROVING DATA RELIABLITY IN SOLID-STATE FLASH DISK

Some flash disk manufacturers use a volatile DRAM/SRAM data buffer to increase disk performance. Powering down when the disk is written, or when data resides in the cache, causes an incomplete write sequence. A DRAM/SRAM cache buffer also causes disk performance to decline when the cache buffer is full (during write operation) and when data does not reside in the cache (during read operation). Using a flash disk without a volatile data cache increases disk reliability under unstable power conditions, and also provides sustained read/write performance that is undisturbed by cache status.

Even if no volatile caching is used, power cycling may cause data corruption. It is important to verify that the flash disk does not tolerate "in-between" data states caused by a power failure, during which the data is only partially written to the disk during a write operation. To prevent this, the flash disk should use the following sequence: perform the write operation, verify that all data was written, and only then update the mapping information. Mapping must always reflect the correct status of the write operation. Only after the success of this sequence should the flash disk update the mapping as "block was successfully transferred".

Since flash disks are designed to operate in mission-critical systems for many years, removing them to perform status checks is unacceptable; hence, remote monitoring is needed. One example of remote monitoring is *SMART (Self-Monitoring, Analysis and Reporting Technology).* When SMART is activated, the disk performs internal monitoring and reports back the test results, indicating the disk status. SMART is commonly used in mechanical ATA/IDE disks to test, among other things, the reliability of the mechanical disk spinning mechanism. Since flash disks have no moving parts but accumulate bad blocks over time, SMART was converted by some flash disk manufacturers to analyze the flash disk bad block status. SMART calculates the total number of bad blocks accumulated after manufacturing relative to the total capacity. This result provides the user with an indication of flash disk reliability and expected life span.

STACKED PC CARDS (PCMCIA ATA CARDS)

Some solid-state flash manufacturers provide hybrid designs that incorporate several units of PC Cards (PCMCIA ATA Cards) to compose a single solid-state flash disk. This hybrid design is less expensive for small capacities that use one or two units of PC Cards, since it uses standard PC Cards available in the marketplace for non-military applications. A solid-state flash disk with higher capacities using three or more units of PC Cards is more expensive than standard solid-state flash disks. With each additional PC Card, there is an additional payment for the PC Card casing, the PC Card internal controller and the adapter mounted within the disk.

Some hybrid PC Cards designs can cause reliability problems under high shock and vibration conditions, due to their Lego-like structure. This issue needs to be address closely, especially in airborne applications.

PC Cards support the ATA/IDE command set. For a SCSI interfaced solid-state flash disk based on PC Cards, there is a need for "on the fly" protocol conversion. ATA/IDE must be converted to SCSI during read operations, and SCSI to ATA/IDE during write operations. The conversion time can lower the disk read/write performance. If the conversion is not smooth, abnormal disk operation may result.

Since PC Cards were originally designed for industrial applications and not specifically for the military, they do not provide Secure Erase and Sanitize. In addition, they must manage bad blocks, an inherent flash limitation, to verify that bad blocks that contain confidential information are completely erased.

Table 2: Stacked PC Card Design vs. Solid-State Flash Disks

Category	Stacked PC Cards (PCMCIA ATA Cards)	Solid-State Flash Disk (not using PC Cards)
1. Fast Secure Erase & Sanitize 2. Partial (Selective) Erase 3. Auto-resume erase during power cycling 4. Erasing confidential data from bad blocks	Does not comply, must be solved externally	Complies, supported internally by some solid-state
5. MIL-STD 810F Shock and vibration	Must to be addressed due to Lego-like design	Complies
6. SCSI commands	Requires on the fly translation from IDE to SCSI and back	Complies, has direct SCSI controller
7. Wear-leveling 8. Bad block mapping out	Does not comply, must verify support	Complies, supported internally
9. Removability	Complies	Complies with: - 80-pin SCA Ultra Wide SCSI - Mounted in removable cartridge
10. Cost	Cost effective for 1-2 units of PC Cards (small capacities)	Cost effective for > 2 PC Cards are needed (higher capacities)

SOLID-STATE FLASH DISKS TRENDS

In the past, cost was a real barrier to solid-state flash disks deployment, but the flash industry has its own version of Moore’s law that overcomes this barrier. According to this law, the density of the flash component doubles within the same silicon footprint size every 12 months, enabling double the capacity every year in the same casing size while sharply reducing flash cost.

Limited capacity and very high flash costs prevented designers in the past from using flash disks. But leading flash manufacturers such as Toshiba and Samsung have improved their processes over the past four years by using less silicon, reducing costs and increasing flash capacity. In 1999, solid-state flash disks were being sold at $5 per MB (1GB for $5,000), while in 2000 the cost came down to $3 per MB (1GB for $3,000). In 2001, it dropped to $2 per MB (1GB for $2,000) and in 2002 it dropped even more to $1.2 per MB (1GB for $1,200). During 2003, the cost has declined yet again to less than $600 per GB. This trend should continue in the future.

Solid-state flash disk performance has also improved dramatically. Until 2000, Narrow SCSI sustained read/write rates were only up to 3MB/sec. In 2001, Ultra Wide SCSI solid-state flash disks were introduced, increasing sustained read/write rates to more than 20MB/sec. This performance enables the storage of video applications and high resolution images.

SECURITY IN EMERGENCY CONDITIONS

On April 1, 2001, a US Navy surveillance plane (EP-3E ARIES II 156511/PR-32) was forced to make an emergency landing in China after what officials described as a "minor" mid-air collision with a Chinese Navy Shenyang F-8-II "Finback" fighter. The incident occurred over the South China Sea when two Finbacks intercepted the EP-3E ARIES during what the U.S. Navy described as a “routine” patrol flight.

The EP-3E crew's procedures for handling classified materials would have involved erasing data and software from computer hard drives and destroying CD-ROMs, floppy disks and key pieces of equipment, including cryptographic systems that encode the electronic signals gathered by the aircraft. The crew had between 12 and 20 minutes in the air to destroy all classified material before making the emergency landing. They had approximately another 12 minutes after landing and before emerging from the aircraft to complete what they probably began while airborne, according to reports. U.S. officials said that in the final moments before the spy plane landed, the crew may have been trying to destroy the hardware with hammers and axes.

Just how much the crew was able to destroy is not clear. However, the potential damage to the collaborative intelligence effort is clear. If the Chinese were able to access the aircraft's top secret equipment, they were able to discover what the Pentagon knows and doesn't know about their communications and operations. This information would enable them to change their methods and develop countermeasures, thus depriving Americans of a powerful advantage in wartime.

Cryptographic keys, databases that maintain U.S. intelligence information on Chinese systems and classified computer codes are of far greater value to the Chinese than the hardware systems alone. However, electricity is required to degauss, or erase, crypto systems and other hard drives. Damage to the aircraft could have hampered the crew's ability to take such actions.

CLEANING, SANITIZING AND DESTROYING PROCEDURES

Security agencies in the US define several levels of “erasing” sensitive data for various storage media type, such as tapes, magnetic disks and optical disks. These levels were originally set by the DoD (Department of Defense) 5200.28 and by the NSA (National Security Agency) CSS 130-2, Media Declassification and Destruction Manual. In 1995, the DoD published the 5220.22 NISPOM (National Industrial Security Program Operating Manual). This manual was issued in accordance with the National Industrial Security Program (NISP) and was developed in close cooperation with the industry. It prescribes requirements, restrictions, and other safeguards to prevent unauthorized disclosure of classified information and to control authorized disclosure of classified information released by the US.

Each US military force has compiled its own internal version, drawn from the DoD/NSA instructions as described by the US Air Force AFSSI (Air Force System Security Instruction) 5020, the US Army 380-19 Information Systems Security (ISS) and the US Navy NAVSO P-5239-26 INFOSEC (Information Systems Security).

Since “erasing” is an ambiguous term, several others terms are used. These terms are applied to magnetic tapes, magnetic disks, optical disks and various other types of memory devices (such as DRAMs, EEPROMs and SRAMs):

- Clearing: Clearing is the process of eradicating data on the media before it is reused in an environment that provides an acceptable level of protection for the data previously stored on the media before clearing.
- Sanitizing (also called Purging): Sanitizing is the process of removing data on the media before it is reused in an environment that does *not* provide an acceptable level of protection for the data previously on the media before sanitizing. Sanitizing mechanical disks and magnetic tapes requires the use of a degausser or otherwise destroying by means of disintegrating, incinerating, pulverizing, shredding or smelting the disks and tapes.
- Destroying: Destroying is the process of physically damaging the media to make it totally unusable as a media, thereby making it impossible to retrieve data.

Declassification is a separate administrative process, whereby classified data stored on a media is deemed to no longer require protection as classified information.

Various procedures have been set forth by the DoD to meet these “erasing” and declassification processes. For the disposition of unclassified DoD mechanical hard drives, the DoD issued special instructions. Software packages can be disposed of after undergoing six passes of special overwriting procedures. These procedures meet the minimal security standard, but are not authorized for use to sanitize classified data. Damaged mechanical hard disks with sensitive data that cannot be sanitized must be degaussed or destroyed. For declassification, degaussing procedures are used to reduce the magnetic flux on the media virtually to zero by applying a reverse magnetic field.

NSA approved special degaussing equipment to declassify magnetic tapes, mechanical disks and optical disks (as within the NSA L1-MTC-4A testing procedure). This equipment is available in several levels indicative of its magnetic field strength, each with a different Oersted (Oe) rating. Each type of magnetic media is distinguished by the rate of coercivity required to return it to its zero state.

Some solid-state flash disks comply with the NSA, DoD, Air Force, Army and Navy sanitization processes for flash NAND EEPROM, which require disk erases, character fill and random data fill in a specific sequence. Therefore, such solid-state products do not require degaussing or destruction.

ENHANCED DATA SECURITY IN SOLID-STATE FLASH DISKS

Solid-state flash disks use NAND flash technology (non-volatile EEPROM) as opposed to NOR flash technology for mass data storage. NAND’s high density and capacity (256 and 512MByte memory chips) and lower price per MB make it more attractive as a mass memory solution. NAND flash memory writes zeroes in page size (512 to 2048 Bytes) and erases them to ones in block size (one block equals 32 pages).

Some flash disk vendors provide attractive features for data security in emergency conditions. None of these features requires degaussing or disk destruction to ensure that what has been erased remains permanently erased. This makes the procedures quicker, easier and more cost-effective. Some solid-state disks provide features described below.

Fast Secure Erase enables the user to erase the entire disk contents in a matter of seconds, between 10 to 60 seconds, depending on the disk capacity. The very nature of NAND flash technology, whereby data is "burned" into the silicon to perform both read/write and erase operations, ensures that erased data cannot be retrieved. In addition, there is no indication of the number of erase cycles that have been performed since a cell was programmed to zero. In contrast, magnetic technology continues to retain some of the original level of magnetic data, even after being overwritten more than 20 times. Fast Secure Erase can be activated by a software or hardware interrupt locally or remotely.

Selective Fast Secure Erase enables the user to erase only part of the disk, for example, only sensitive data but not the operating system. This ensures disk readiness for the next mission without reinstallation, saving valuable time. The user can partition the disk into up to 8 different areas, and decide which to designate for fast erasure.

Usually during emergency conditions, the power supply is unstable, preventing the completion of special security processes to erase sensitive data. In such cases, ***Auto-Resume Security Erase*** ensures that a Fast Secure Erase that was halted before completion, due to lack of power, will automatically be completed as soon as power is restored. Auto-Resume Security Erase is supported by the flash disk internal processor, which guarantees Fast Secure Erase completion, regardless of the host.

The ***Sanitize (Purge)*** process for NAND flash (in the NSA/CCS 130-2, DoD 5220.22 NISPOM, Air Force AFSSI 5020, Army 380-19 and Navy NAVSO P-5239-26) require sanitizing NAND flash with disk erase, same character fill and random data fill, using the specific sequence specified in each of the above procedures. Bad blocks must be repeatedly erased ("strong erasure") to guarantee total disk sanitization.

CONCLUSION

Mechanical disks continue to be the weak link in total system reliability under harsh environmental conditions. As such, solid-state flash disks are being designed as true "drop-in replacements" for data storage in mission-critical applications. Solid-state flash disks provide top data integrity under harsh environmental conditions of extreme shock, vibration and humidity, and meet industrial temperature requirements. Cost is no longer an insurmountable barrier in using solid-state flash disks in mission-critical applications, since flash prices have declined dramatically over the past two years, a trend that experts expect to continue.

Securing confidential data in emergency situations is essential. Sanitizing mechanical disks and magnetic tapes is an arduous process, requiring special degaussers, stable power conditions during the process, and ample time, all of which may be lacking during an emergency. Solid-state flash disks can sanitize the disk in seconds with Fast Secure Erase and Sanitize procedures. In addition, once Fast Secure Erase has been activated, Auto-Resume Secure Erase guarantees successful completion of the process.