Coughlin Associates

Aligning with the Future of Storage

Thomas M. Coughlin
Coughlin Associates, Inc.
www.TomCoughlin.com
June 2011

Table of Contents

Executive Summary

Storage sector misalignment is a serious issue for every industry. Misaligned partitions resulting from a move from 512 byte sectors to 4K byte sectors can cause a high-performance trading system to crash, slow down healthcare providers and generally cause performance and data availability issues for IT organizations, including virtual storage environments, and with many of the most widely used operating systems and storage devices. Clearly, IT providers, CIOs, CTOs and IT managers need to learn about this issue and effective ways to prevent it from bringing their operation to its knees.

This paper explains the issues surrounding sector misalignment with today's 4K sector advance format hard disk drives (HDDs).

- 4K sector storage devices are needed to continue storage capacity increases and to optimize the performance of these devices.

- 4K sector drives will provide more efficient storage utilization, lower energy costs and greater storage economy.

- However, older operating systems and utilities can misalign the logical sectors in the host device and the physical data on the HDD sectors resulting in significant performance degradation.

- Sector misalignment in 4K solid state drives (SSDs) can also result in shortened SSD life expectancy.

- Sector misalignment between RAID partitions and storage device sectors leads to slower SAN and NAS performance.

- Misalignment between the storage clusters in virtual machines, the virtual machine file system blocks and the underlying RAID system blocks can cause significantly more data to be accessed than requested and much slower virtual machine performance.

Special utilities, such as the Paragon Partition Alignment Tool (PAT), can be used to resolve these block alignment issues; resulting in faster storage system performance, longer storage device lifetime and more efficient system utilization and thus lower operating costs.

> "Storage sector misalignment is a serious issue for every industry."
>
> — Tom Coughlin, President, Coughlin Associates

Storage: Transitioning to 4K

For the last 30 years, the size of the sectors on a hard disk drive were limited to 512 bytes for many operating systems including the Windows operating system as well as the popular ATA interface. Just to give this some perspective, large disk drives back then were several megabytes in size. Thus the number of 512 byte blocks in a hard drive 30 years ago was not nearly as large as is required today for hard disk drives over 3 terabyte in size.

Since the number of sectors has increased with storage capacity the total sector header data capacity is larger than in the past, this is data capacity that the user can't record and read from. Thus with a fixed sector size, as the total storage capacity increases the formatting efficiency (the ratio of the user data capacity left on the drive after non-user data is recorded, to the original drive storage capacity) decreases. A larger sector size will result in fewer sector headers and inside of each sector more user data can be stored. This results in a higher formatting efficiency for the disk drive and more of the native hard disk drive storage capacity available to the user.

A longer sector block size also allows creating a more robust ECC (error correction code) for correcting data errors more efficiently, and allows new ECC technologies that will be a critical step to higher disk drive areal densities. The same percentage of overhead for a long block is much more efficient than in a short block, as the number of combinations of correctable error patterns is dependent on the length of the data block. Combining larger format efficiency and more robust ECC to achieve higher drive areal densities means that larger sector size drives provide a boost to the user accessible storage capacity on the disk drive.

IDEMA Advanced Format Drives

IDEMA Advanced Format Drives (AFDs) are the storage device of the future. The improvements that larger hard disk drive sectors enabled in hard disk drive performance and storage capacity led to a standards effort within the trade organization IDEMA (The International Disk Drive Equipment and Materials Association) to create a 4,096 byte sector standard (often referred to as the 4K sector).

> "Under some conditions…newer Windows, Linux and Macintosh computers can suffer from block misalignment. "

This sector size is eight times larger than the older 512 byte sector. 4,096 byte implementation is well underway with Advanced Format (AF) HDD products available in the market from all hard disk drive manufacturers.

A 4K byte sector hard disk drive has a higher storage capacity (about 9 percent higher) than a 512 byte sector hard disk drive with the same native storage capacity. This higher storage capacity is based on the lower system overhead data. Larger sectors also facilitate increased areal density for future drives since additional error correction can be included in a larger sector. With AF disk drives, drive manufacturers can make higher capacity hard drives with increased reliability and a faster time to data. These improvements will allow constructing computers, consumer electronics and storage systems with greater storage capacity and performance than would be possible with the 512 byte sectors.

Combined with the growing use of virtualization, de-duplication and other storage technology, 4K sector drives will provide more efficient storage utilization, lower energy costs and greater storage economy. While virtualization and de-duplication are mostly software improvements, 4K sectors offer a true hardware enhancement.

Problems for the Misaligned

Hard disk drives hide the actual physical location of where data is written to or read from along the tracks. The file system in a host device accesses logical sectors of information. The actual location of the physical sectors of data on the hard disk drive surfaces is controlled by the electronics on the hard disk drive.

Newer versions of Windows as well as Linux and Macintosh operating systems will transfer 4K data blocks that match the native 4K hard disk drive sector size. Windows operating systems before Vista experience a problem with 4K sectors called misalignment. Under some conditions, such as system migration, even newer Windows, Linux and Macintosh computers can suffer from block misalignment.

In order to be compatible with older operating systems, 4K sector hard disk drives remap 512 byte sectors requested by the operating system to the 4K sector within the hard disk drive, this is called 512 byte sector emulation. The 4,096 byte sector appears to the operating system as eight 512 byte sectors. With emulation, write and read is done in eight 512 byte sectors.

Windows XP and earlier operating systems may not align the first block in an eight 512 block group with the beginning of the 4K sector. As a result the drive will not be able to write the eight 512 byte blocks directly to a 4K sector. Another way of saying this is, the eight 512 blocks are spanning parts of two separate 4K byte physical sectors. This is called sector misalignment. **Figure 1** shows how the eight 512 byte logical sectors (0 through 7) may be off-set by one block from the 4K physical sector they are trying to map to. In fact a one block off-set might occur with any version of Windows with data migration.

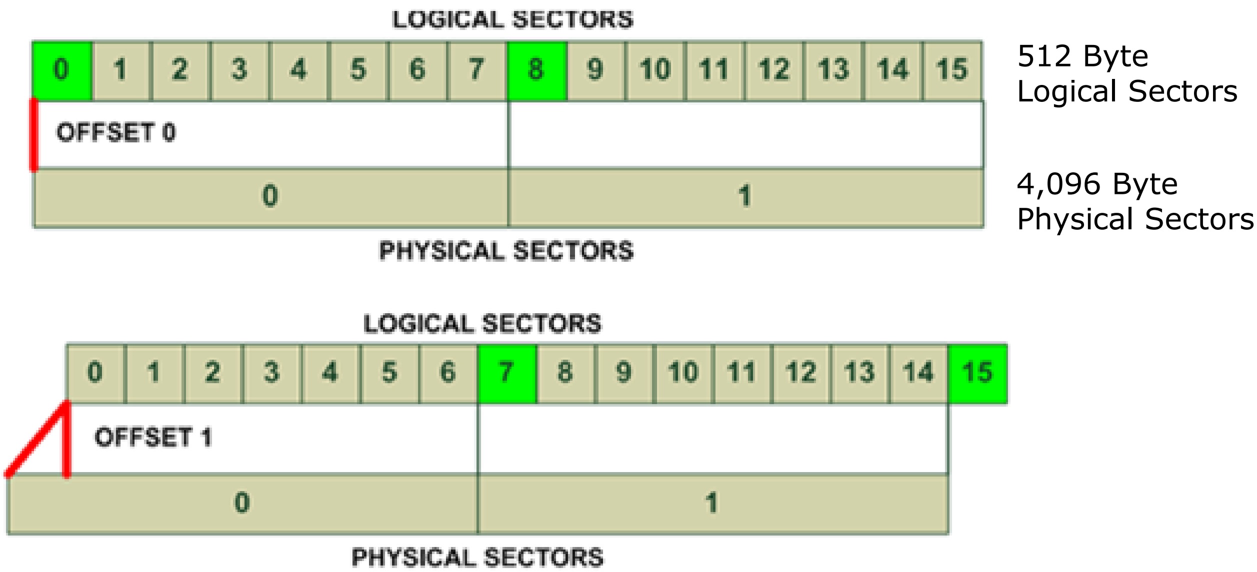


Figure 1. 512 byte Sector Off-Set to a 4096 byte Sector

If there is no sector misalignment the eight 512 byte sectors can be directly written on the 4K physical sector. However if there is misalignment of the 512 byte logical sectors to the 4K byte physical sectors, it forces the hard disk drive to perform an additional read operation to read the two physical 4K sectors where the eight 512 byte sectors are to be written (since the eight 512 byte blocks span the two 4K sectors). The data from these sectors is put into cache memory.

> "Windows XP and earlier operating systems may not align…"

In a separate operation the hard drive then writes both the 4K sectors along with ECC data recalculated for the new data in the sectors. These extra operations are called Read-Modify-Write and they can result in slow write times and generally reduced drive performance. **Figure 2** shows how misaligned blocks increase the effective write time by at least 30 percent for a 7200 RPM hard disk drive (due to the additional read and data organization steps).

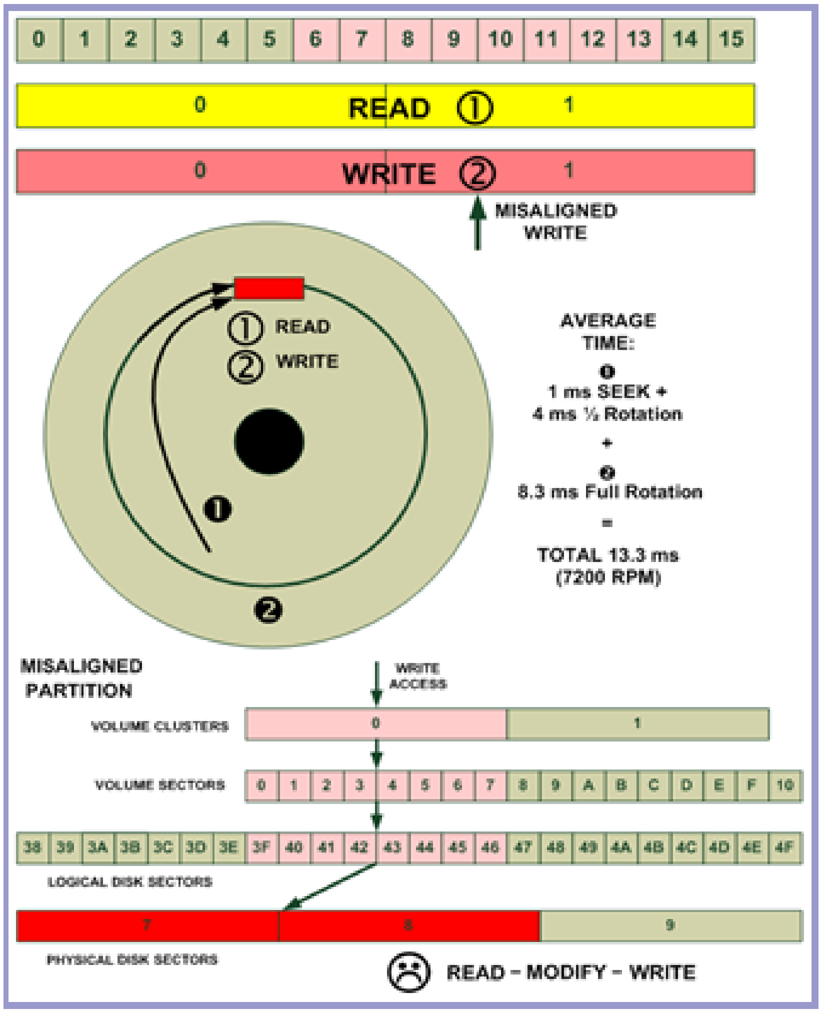


Figure 2. Additional Write Time for Misaligned Sectors

Block misalignment can occur with Vista and Windows 7 computers if the OS installation was performed from an XP or earlier Windows OS where the pre-existing partition was misaligned. Also non-4K aware utilities can create block misalignment. For instance a partitioning utility might resize a Vista/Win7 partition creating an alignment issue. Also for Windows, Linux and Mac computers a non-4K aware application that writes directly to the disk drive (bypassing the OS) can create misaligned data within an aligned partition.

Later Mac OSX versions can use the GUID partitioning table scheme, which aligns AF drives properly. However Apple's Partition Manager might create non-aligned partitions. Computers using the latest versions of Linux will provide good support for aligned partitions on AF drives, although the partitioning tool must be used properly for the best results.

Use Case: Making Money Faster at a Credit Card Processing Company

A credit-card processing company upgraded 350 servers to high-capacity drives (4K AFDs) via disk imaging (all machines were running Windows Server 2003 and each host server was configured with either RAID 0, RAID 1, RAID 5, or RAID arrays). The migrated machines ended up misaligned, causing significant performance degradation of the drives. Paragon PAT was used to realign partitions across all the drives to regain overall system performance.

Misaligned Sectors Hinder Storage Devices and Storage Systems

Besides causing performance problems for AF HDDs with older Windows operating systems, sector misalignment with 4K sector solid state drives (SSDs) can lead to endurance and lifespan issues. This is because SSDs using flash memory have to perform a separate erase cycle before writing over a used memory cell. SSDs can only erase each memory cell a certain number of times before the cell "wears" and can no longer retain the charges in the floating gate that make the cell a non-volatile memory device. Thus SSD controllers spread the "wear" of the memory cells by writing new data on less used cells and erasing previously used cells in bulk erase cycles.

When the extra complications of the read, modify, and then write for misaligned sectors is applied to SSDs, it can lead to additional erase/write cycles during normal system use. This accelerates the "wear" of the memory cells and shortens the useful life of the SSD. Thus misaligned sectors can result in shorter life for SSDs in addition to performance degradation.

Use Case: Recovering SSD Value at a Start-Up Company

A start-up technology company invested over $20,000 upgrading 14 servers to ultra-fast SSDs (all machines were running Windows Server 2003). After the imaging process, the drives were not operating at the expected performance level, rendering the entire investment a total waste of money. Paragon's PAT software was used to realign the drives to fully capitalize on the drives' potential performance and also to extend the life expectancy of the SSDs.

Storage arrays using RAID (Redundant Array of Independent Disks), or other methods to spread data over a group of storage devices, can suffer performance degradation when sector misalignment occurs. This is because the starting location of a storage partition in the array (from the logical level) may not align properly with a stripe unit boundary in a partition on a disk in the array (at the physical level). This results in slower performance since operations that are intended to occur on only one storage device may in fact occur over several storage devices in the array.

Block alignment in a virtual machine environment is critical to performance, hardware life, and storage efficiencies. Misalignment results in retrieving more data from an underlying array than what the virtual machine is requesting. The resulting inefficiencies require more storage hardware resources to serve a given workload.

In **Figure 3** a VMware Server configuration uses RAID/SAN storage. Here there are two levels of misalignment that cause a severe decline in performance. For example, to read data from the first cluster, the system has to read three VMFS blocks and four RAID blocks (and consequently more stripe-units).

If we could align partitions on the virtual drives we would get the results shown in **Figure 4**. In this configuration all volumes and partitions are properly aligned and overall data operations are two times faster. In order to get data from one cluster the system only needs to read two VMFS blocks and two RAID blocks.

Misalignment can also occur in Linux-based NAS systems when embedded Linux devices create partitions that are not aligned with the underlying storage device blocks. As in other misalignment cases block misalignment results in slower system performance.

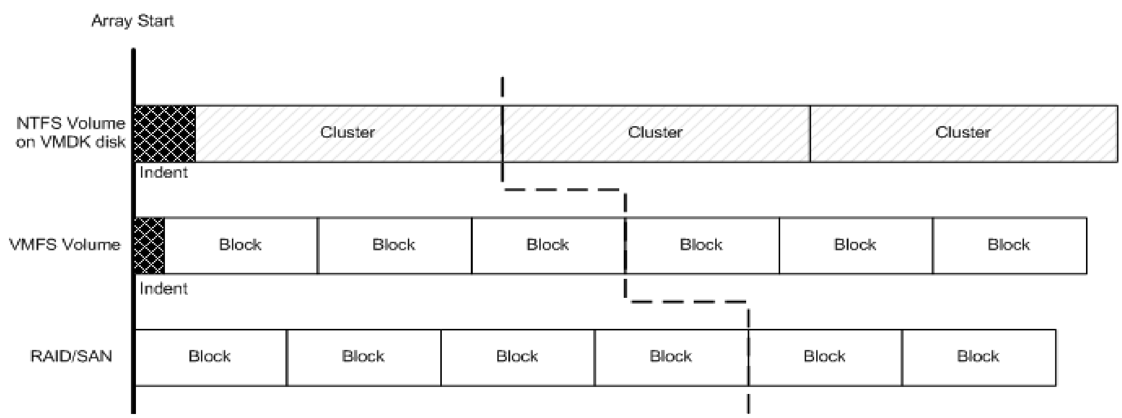


Figure 3. Virtualization Environment with Partition Alignment Issues

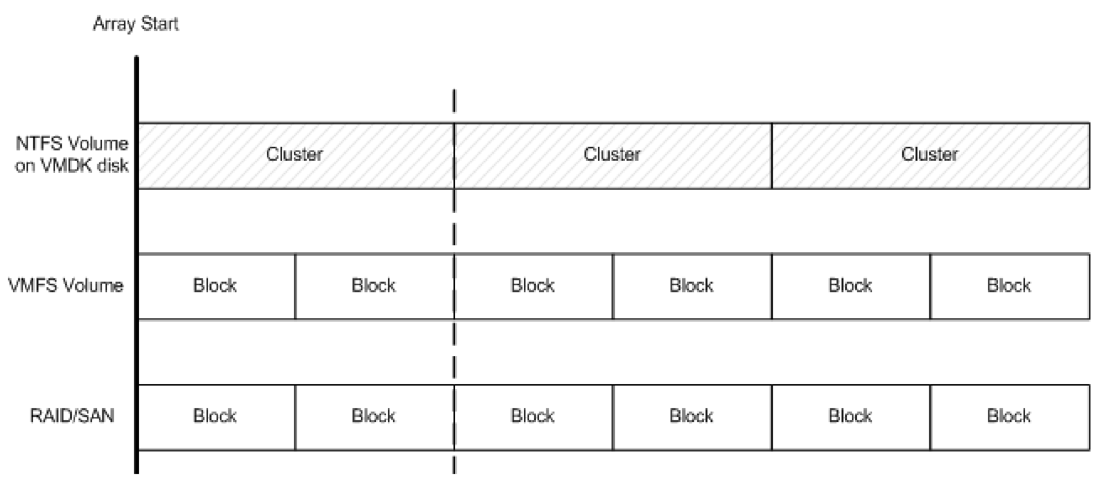


Figure 4. Virtualization Environment after Partition Alignment

It is clear that users and system integrators need good tools to help them align their computers and storage systems with the 4K blocks in AF storage devices.

Paragon Alignment Tool

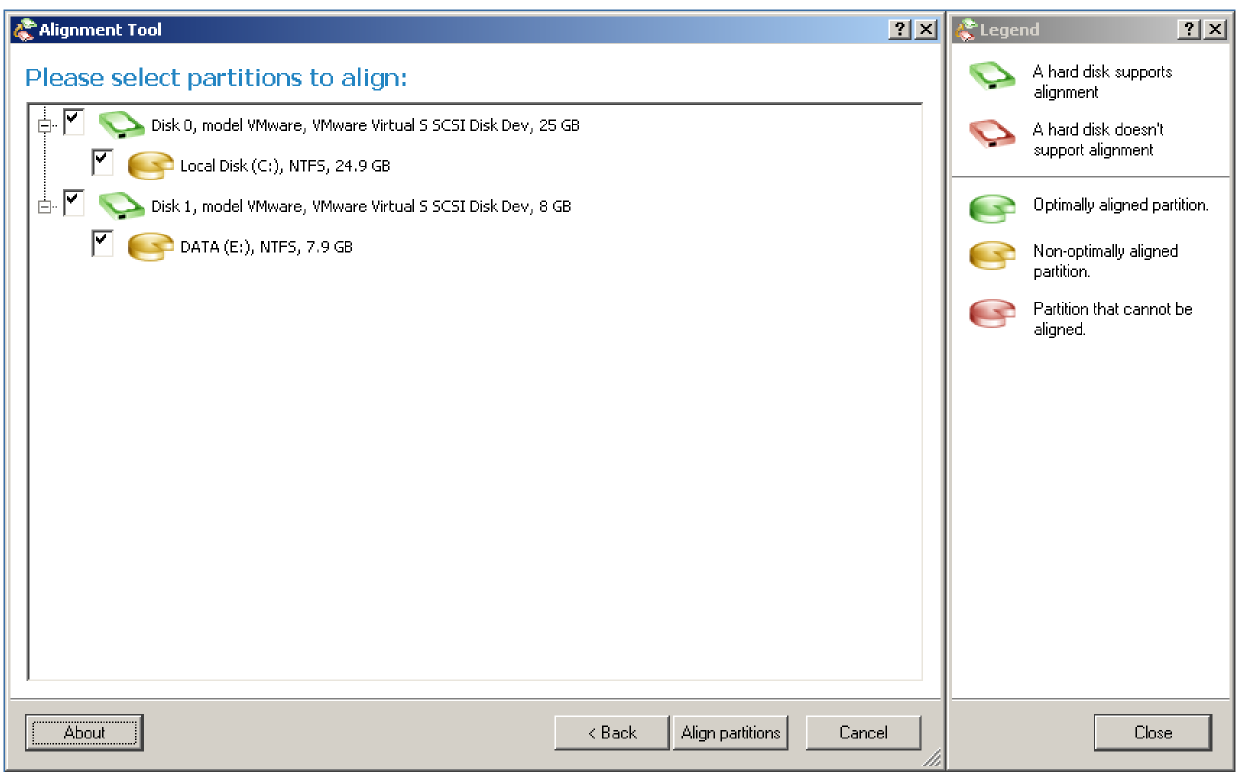
Paragon Software offers powerful products that allow their customers to detect and correct issues with block alignment on 4K byte block storage devices, resulting in faster system performance, higher efficiency and longer storage device and system life.

The Paragon Alignment Tool (PAT) can resolve partition alignment problems associated with 4K advanced format HDDs, SSDs and RAID systems built with these devices. This utility can also reduce alignment issues in virtual machine environments running on VMware, Hyper V or Citrix.

PAT is a powerful and easy-to-use software utility designed to automatically determine if a drive's partitions are misaligned. If there is misalignment the utility then properly realigns all the existing partitions, including boot partitions, to 4K sector boundaries.

PAT has proven useful to relieve misalignment issues for many organizations, from large corporations to small- to medium-sized businesses. Following are some use cases showing customer applications that have benefited from the use of PAT.

Paragon's PAT technology can be found integrated into many of Paragon's mainstream products such as Hard Disk Manager and Partition Manager. It is also available as an add-on feature for Virtualization Manager.



Paragon Alignment Tool Screen Shot

Use Cases in Virtual and Enterprise Environments

U.S. Airline Takes Off after Virtual Environment Realignment

A U.S.-based airline virtualized 3,000 employee desktops only to discover soon after that all of the client virtual machines were underperforming due to misalignment issues. As a result, they could not garner acceptable performance while running the virtual desktops on the available number of servers. Purchasing an additional 30 high-end servers was considered, but quickly dismissed as an alternative solution to the problem. Looking for other options, they decided on Paragon's PAT software. After realigning all of the virtual desktops' partitions, performance was so greatly improved that the expected virtual desktop-to-server ratio was exceeded.

PAT Staged Alignment for Mass System Deployment in the Enterprise

The enterprise IT department of a Fortune 500 corporation was tasked with implementing images for 30,000 newly purchased desktops, laptops and servers that incorporated 4K AFDs and the associated alignment challenges. This project required a flexible alignment solution that could provide post-configuration image alignment. After the image was deployed in their staging operations, the IT team used a scriptable command line version of PAT to efficiently scan disks, determine alignment status and automatically correct any misaligned partitions quickly and efficiently. The PAT System Staging command line version streamlined the entire process of the project and ensured on-time delivery of the assignment.

Use Case: Storage Resuscitation at a National Healthcare Company

A national healthcare company virtualized 461 of their roughly 800 physical servers (all hosts and guests were running Windows Server 2003 and each host server was configured with RAID 5 or RAID 1+0 arrays). After the migration process, all virtual servers ended up misaligned, causing a significant decrease in performance. A large portion of the servers were using high-capacity 4K Advanced Format Drives where the performance loss was compounded by the larger sector sizes. Paragon PAT was used to realign all partitions across the affected drives to regain optimum performance.

<u>Addendum</u>
The Anatomy of an HDD

Tracks and Sectors: Architecture of an HDD

Hard disk drives provide cost effective digital storage technology. These useful devices enable a vast number of applications from computers to enterprise storage systems and consumer electronics. The information in a hard disk drive is stored on the magnetic surface of rotating rigid disks. This information is written to and read from the rotating disks using magnetic heads that move across the radius of the disks flying close to the surface. These heads are moved across the disk surface by a rotary motor. **Figure 5** shows key components within a hard disk drive.

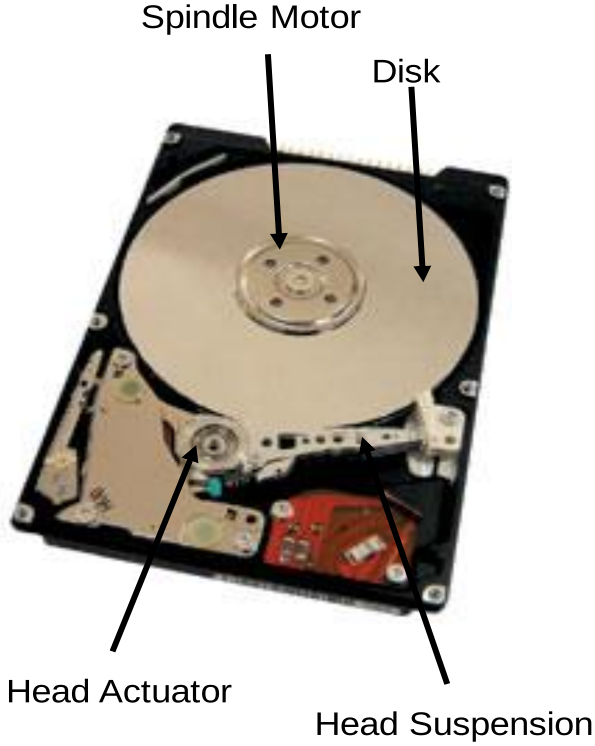


Figure 5. Illustration of hard drive components.

Hard disk drives (and other storage devices) organize their data into blocks. In a hard disk drive data blocks are regions of the magnetic media that contain user data. These blocks are located in regions of the disk called a sector. Hard disk drive sectors are located along the concentric tracks on the surface of the disks. Within each sector are also located servo and error correction information that allow the magnetic heads to stay on track so they can accurately read and write data and for the electronic channel to recover the data without errors. The sectors also contain information that allows the file system to reassemble the data blocks into computer files. The file system, usually located in a computer or server, creates files with useful user data such as documents, spreadsheets, presentations, etc.

Figure 6 shows a sector taken from a data track on a hard disk drive. **Figure 7** shows a typical sector format.

The header information in a hard disk drive sector contains information on the position of the sector on the disk to help the drive find the proper sector where a piece of data is located. It also contains error correction information that allows finding and fixing bytes in the sector that may contain errors. *

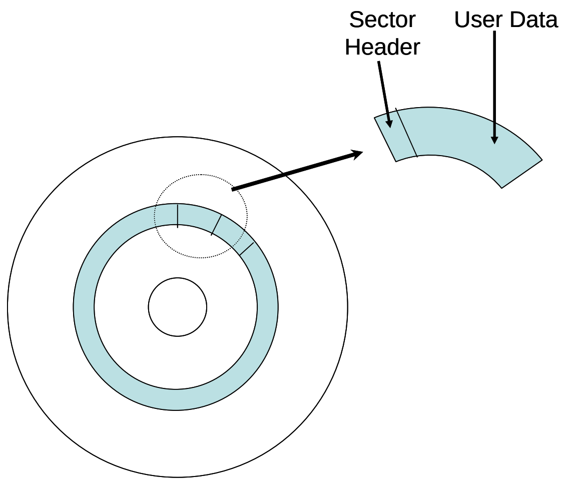


Figure 6. Hard Disk Drive Tracks are Organized into Sectors.

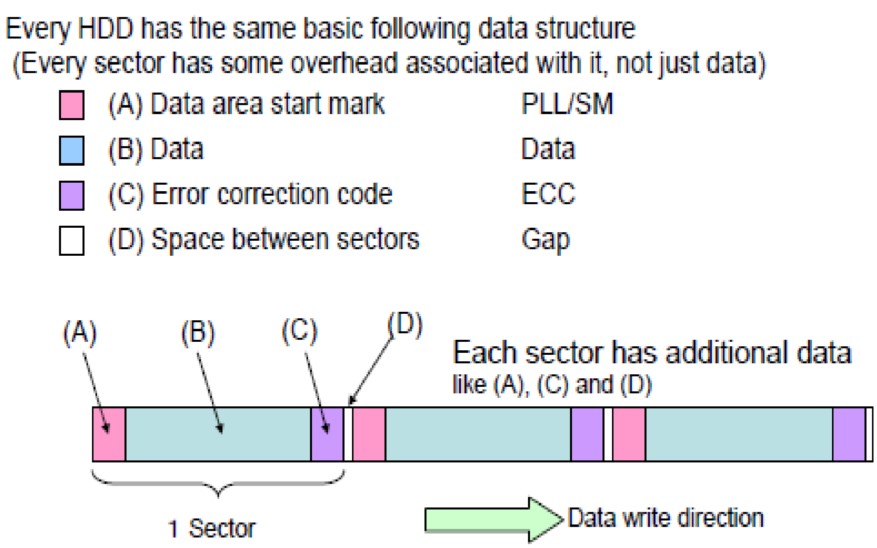


Figure 7. A Typical Hard Disk Drive Sector Format

The digital storage capacity of one surface of a hard disk drive platter is a product of the number of average bytes of information along the disk tracks (this is not constant but varies with the track radius) times the number of tracks along the radius of the disk. Since disk drives were introduced in 1956 storage capacities of disk drives have increased from 5 MB to over 3 TB (an increase of about a million times).

* *Michael E. Fitzpatrick, **4K Sector Disk Drives: Transitioning to the Future with Advanced Format Technologies,** Toshiba America Information Systems, 2011.*

About the Author

Tom Coughlin, President, Coughlin Associates is a widely respected storage analyst and consultant. He has over 30 years in the data storage industry with multiple engineering and management positions at high profile companies.

Dr. Coughlin has many publications and six patents to his credit. Tom is also the author of Digital Storage in Consumer Electronics: The Essential Guide, which was published by Newnes Press. Coughlin Associates provides market and technology analysis (including reports on several digital storage technologies and applications and a newsletter) as well as Data Storage Technical Consulting services. Tom publishes a *Digital Storage in Consumer Electronics Report,* a *Media and Entertainment Storage Report*, and a *Capital Equipment and Technology Report for the Hard Disk Drive Industry*.

Tom is active with SMPTE, SNIA, IDEMA, the IEEE Magnetics Society, IEEE CE Society, and other professional organizations. Tom is the founder and organizer of the Annual Storage Visions Conference (www.storagevisions.com), a partner to the International Consumer Electronics Show, as well as the Creative Storage Conference (www.creativestorage.org). He is also a Senior member of the IEEE, Leader in the Gerson Lehrman Group Councils of Advisors and a member of the Consultants Network of Silicon Valley (CNSV). For more information on Tom Coughlin and his publications. go to www.tomcoughlin.com.

Aligning with the Future of Storage

Thomas M. Coughlin
Coughlin Associates, Inc.
June 2011

Coughlin Associates