



UiO : University of Oslo

FYS3240- 4240

Data acquisition & control

PC-based data acquisition II

Data streaming to a storage device

Spring 2019 – Lecture 9



Data streaming

- Data written to or read from a storage device at a sustained rate is often referred to as **streaming**
- Trends in data storage
 - *Ever-increasing amounts of data (**Big Data**)*
 - *Record “everything” and play it back later*
 - *Hard drives: faster, bigger, and cheaper*
 - *Solid state drives*
 - *RAID hardware*
 - *PCI Express*
 - *PCI Express provides higher, dedicated bandwidth*



Applications requiring high-speed data streaming (examples)

- High speed data acquisition
 - Combined with many DAQ channels
- Radar (Giga-samples/s)
 - RF recording and playback
- High resolution and/or high speed video
 - Digital video recording and playback

Key system components for high-speed streaming

- Hardware Platform with High-Throughput and Low-Latency
 - PXI/PCI Express bus
- High-Speed Data Storage
 - Hard Disk Drives (HDDs) in RAID
 - Solid-State Drives (SSDs)
- Software for Streaming to Disk at High Rates
 - Parallel programming
 - Fast binary file format

BIOS (basic input/output system)

- The BIOS software is built into the [PC](#), and is the first code run by a PC when powered on ('boot firmware'). When the PC starts up, the first job for the BIOS is to initialize and identify system devices such as the [video display card](#), [keyboard](#) and [mouse](#), [hard disk drive](#), [optical disc drive](#) and other [hardware](#). The BIOS then locates [boot loader](#) software held on a [peripheral device](#) (designated as a 'boot device'), such as a [hard disk](#) or a [CD/DVD](#), and loads and executes that software, giving it control of the PC.^[2] This process is known as *booting*, or booting up, which is short for [bootstrapping](#).
- BIOS software is stored on a [non-volatile ROM](#) chip on the [motherboard](#). It is specifically designed to work with each particular model of computer, interfacing with various devices that make up the complementary chipset of the system. In modern computer systems the [BIOS chip's contents can be rewritten](#) without removing it from the motherboard, allowing BIOS software to be upgraded in place.

Drive Partition - 2.2 TB limitation

- A partition is a contiguous space of storage
- Partitions are visible to the system firmware and the installed operating systems.
- This 2.2 TB limitation dates back to the 1980s and the original IBM PC. This introduced the **master boot record (MBR)** partitioning scheme to describe hard disk partitions.
- MBR consists of a sequence of 512 bytes located at the first sector of a data storage device such as a hard disk
- BIOS systems with MBR disks use **32-bit values** to describe the starting offset and length of a partition. Due to this size limit, MBR allows a maximum disk size of approximately 2.2 TB
 - $2^{32} - 1 \text{ sectors} * 512 \text{ bytes/sector} = 2.199\text{TB}$
- This limitation applied for Windows XP 32 bit !



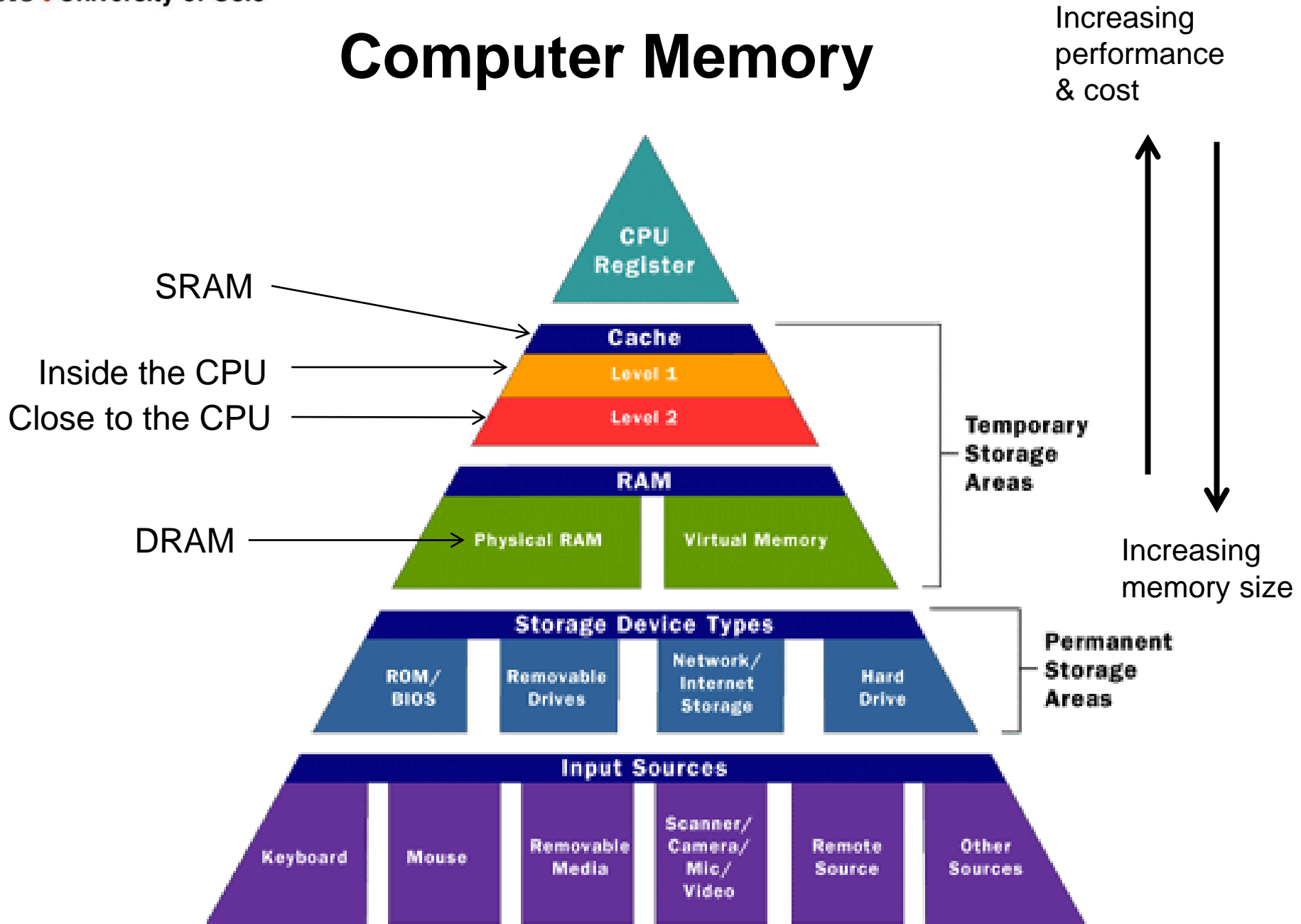
Avoiding 2.2 TB limit : GPT and UEFI

- **GUID Partition Table (GPT)**
 - Provides a more flexible mechanism for partitioning disks than the older Master Boot Record (MBR) partitioning scheme
 - Supported by Windows 7, XP 64-bit, Windows Server 2008 etc.
 - 64-bit values to describe partitions
 - GPT can handle disks of up to 9.4×10^{21} bytes ($2^{64} - 1$ sectors * 512 bytes/sector)
 - Windows file systems currently are limited to 256 terabytes each
 - Windows only support booting to GPT partitions on systems with EFI firmware.
- **Unified Extensible Firmware Interface (UEFI)**
 - A more secure replacement for the older BIOS firmware interface
 - But, BIOS remains in widespread use since EFI booting has only been supported in Microsoft's operating system products supporting GPT and Linux kernels 2.6.1 and greater builds.

32-Bit vs. 64-Bit OS for DAQ applications

- A 32-bit processor can reference 2^{32} bytes, or 4 GB of memory
- A 64-bit processor are theoretically capable of referencing 2^{64} locations in memory. However, all 64-bit versions of Microsoft operating systems currently impose a 16 TB limit on address space
- Note: 64-bit versions of operating systems and application programs are not necessarily faster than their 32-bit counterparts. They can be slower!
- The main benefits of 64-bits for DAQ applications is the large amounts of RAM possible.
- Other applications that can benefit from 64-bits are those working with very large numbers.

Computer Memory



kB vs. KB

Two Different Measurements Systems

Name	Abbreviation	Binary Power	Binary Value (in Decimal)	Decimal Power	Decimal (Equivalent)
kilobyte	KB	2^{10}	1,024	10^3	1,000
megabyte	MB	2^{20}	1,048,576	10^6	1,000,000
gigabyte	GB	2^{30}	1,073,741,824	10^9	1,000,000,000
terabyte	TB	2^{40}	1,099,511,627,776	10^{12}	1,000,000,000,000

How Operating Systems Report Drive Capacity

3TB (E:) Properties

ReadyBoost Previous Versions Quota Customize
General Tools Hardware Sharing Security

3TB

Type: Local Disk
File system: NTFS

Used space: 13,675,864,064 bytes 12.7 GB
Free space: 2,986,906,017,792 bytes 2.71 TB

Capacity: 3,000,581,881,856 bytes 2.72 TB

Total Gigabytes (binary value)

Drive E: Disk Cleanup

Compress this drive to save disk space
 Allow files on this drive to have contents indexed in addition to file properties

OK Cancel Apply

Historically and in common usage, a kilobyte refers to 1024 (2^{10}) bytes in most fields of computer science and information technology. However, hard disk drives measure capacity in strict accordance with the rule of the International System of Units

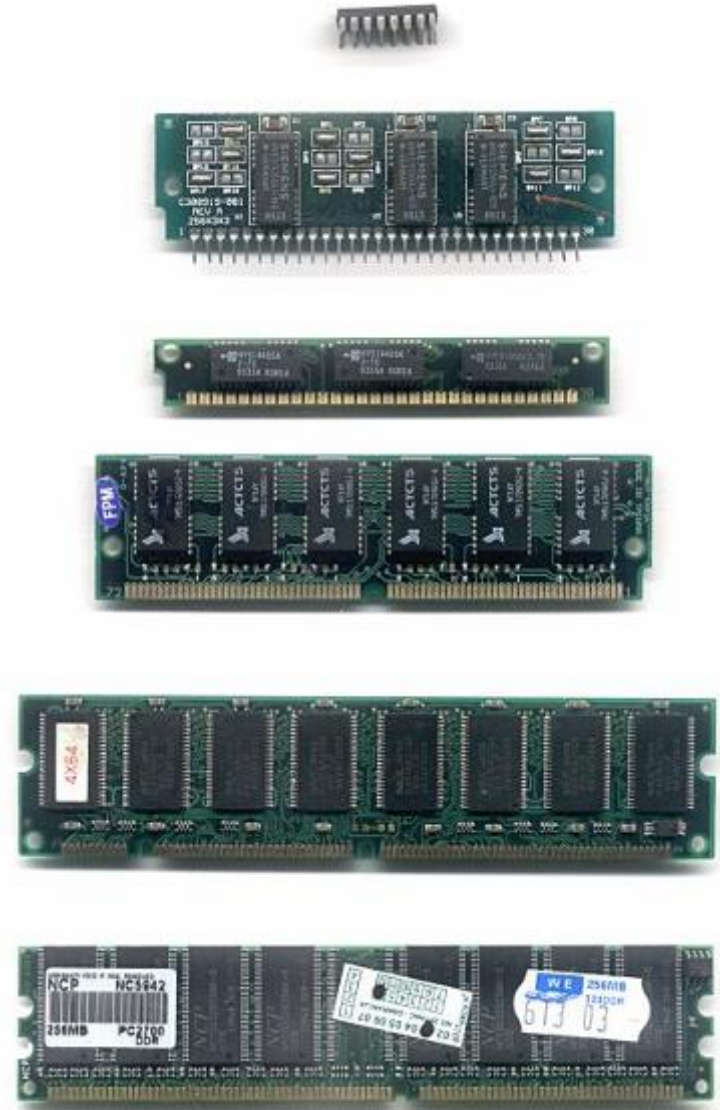
k = 1000 (decimal)

K = 1024 (binary)

TB_{binary} = 1024^4 byte

RAM – Random Access Memory

- **SRAM – Static RAM: Each bit stored in a flip-flop (4-6 transistors)**
- **DRAM – Dynamic RAM: Each bit stored in a capacitor (transistor). Has to be refreshed (e.g. each 15 ms)**
 - EDO DRAM – Extended Data Out DRAM.
Data available while next bit is being set up
 - Dual-Ported DRAM (VRAM – Video RAM).
Two locations can be accessed at the same time
 - SDRAM – Synchronous DRAM.
Latched to the memory bus clock → read/write on one single clock cycle
 - Rambus (RDRAM) - Intel had to use RDRAM 1996-2002.
Much more expensive than SDRAM
 - DDR SDRAM – Double Data Rate SDRAM (2.5 V).
Data transferred on both rising and falling edge of clock.
 - DDR2 SDRAM (1.8 V)
Same, but lower power consumption and higher clock frequency
 - DDR3 SDRAM (1.5 V). Even less power and faster
 - DDR4 SDRAM (1.2 V), available in 2014



Stream to/from disk rates

Peak, not
sustained rate!

Note: old numbers - show the relative data rates

Drive(s)	Write/Read (MB/s)	Rate Types
Laptop	30 (NI PXle-8103 internal drive; 5,400 RPM)	Peak
IDE	57 (Western Digital 160 GB; 7,200 RPM)	Peak
SATA	62 (Western Digital 160 GB; 7,200 RPM)	Peak
SATA	75 (Seagate Barracuda 7,200.10; 250 GB)	Peak

- Important parameters for streaming:
 - Seek times
 - how long it takes the head assembly to travel to the track of the disk that contains data
 - Rotational speed (RPM)
 - 7200, 10000, 15000
 - Buffer size
- Sustained streaming rates are most affected by the **rotational speed**



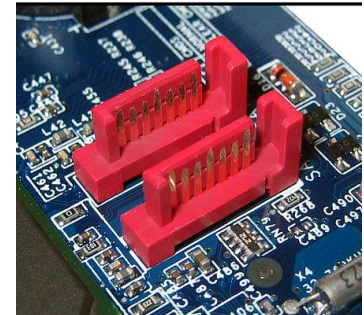
Typical “Memory” speed 2011

- Cache (CPU cache, Disk cache)
 - Caches are normally made from static RAM chips (SRAM), unlike main system memory which is made from dynamic RAM (DRAM)
 - Much faster than DRAM!
- RAM
 - DDR3 SDRAM: **6400 MB/s**
- HDDs
 - Data rate: ~ **100 - 150 MB/s**
 - Note: Peak vs. Sustained data rate
- SATA
 - Serial ATA (SATA or Serial Advanced Technology Attachment)
 - Capacity: **1.5, 3.0, 6.0 Gbit/s** (6 Gbit/s = 750 MB/s * 8/10= 600 MB/s because of the 8b/10b encoding)

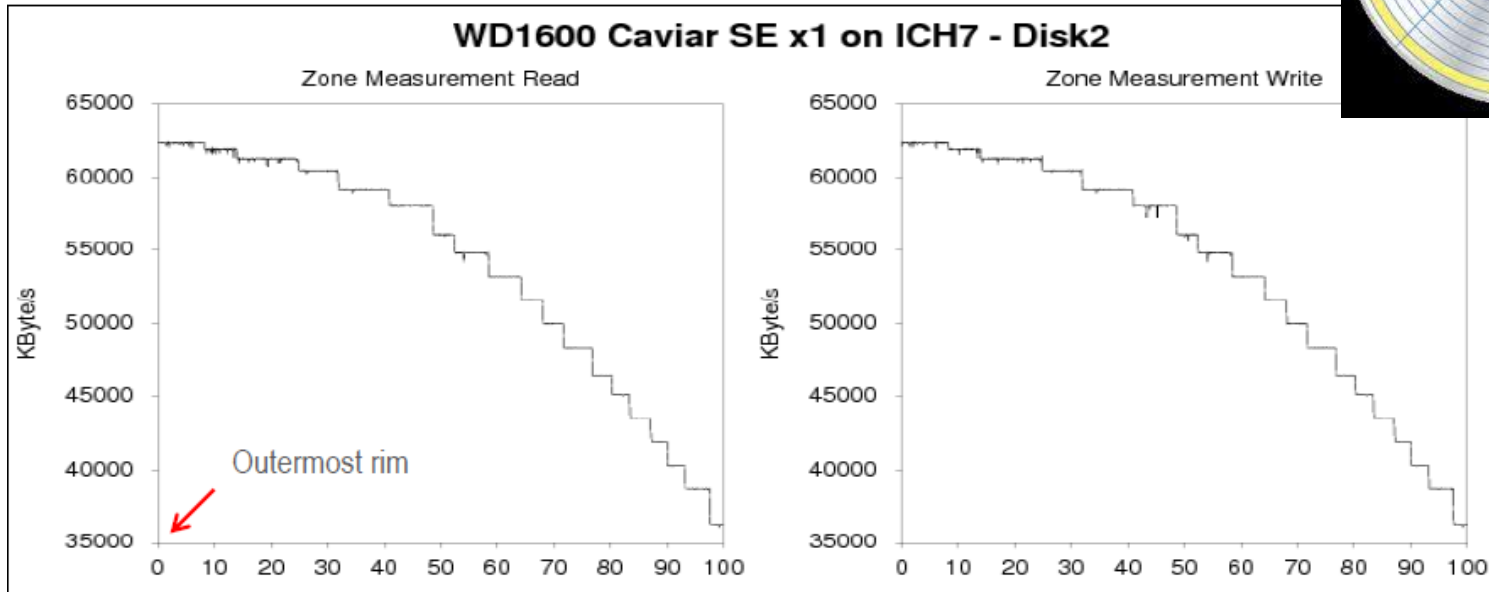
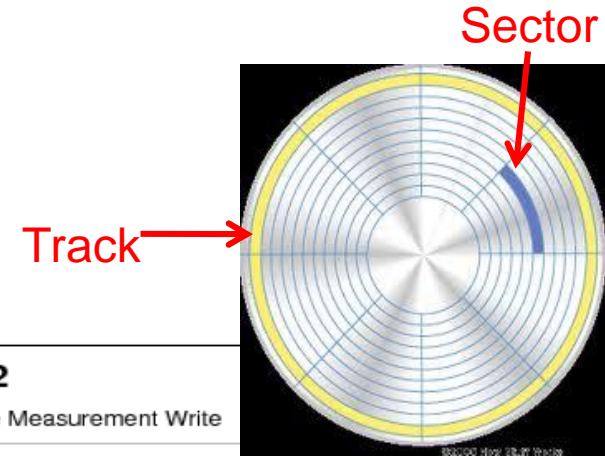
NB! {

Seagate Barracuda 7200.12 Specifications

Available capacities	250, 500, 750GB, 1TB
Maximum external transfer rate	300MB/s
Maximum sustained data rate	125MB/s
Cache size	32MB (750GB, 1TB) 16MB (500GB) 8MB (250GB)
Spindle speed	7,200 RPM



HDD Performance



- HDD's Internal Data Rate (IDR) = density * RPM * disk diameter
- **Outer HDD track is faster, inner track is slower**
 - more data sectors on outer tracks, fewer data sectors on inner tracks
- Example above: 62 MB/s at outer track, 36 MB/s at inner track
- **Windows OS allocates file space from outer track and inward**

SSD (Solid-State drive)

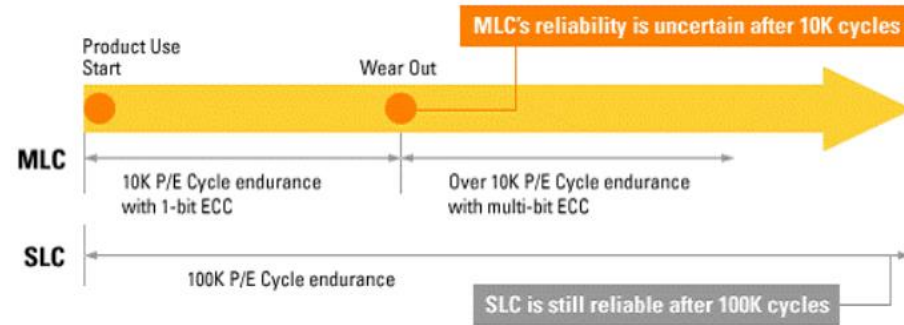
- A data storage device that uses solid-state (Flash) memory to store data.
- SSDs are distinguished from traditional hard disk drives (HDDs), which are electromechanical devices containing spinning disks and movable read/write heads.
- SSDs, in contrast, use microchips which retain data in non-volatile memory chips and contain no moving parts.
- SSDs use the same interface as HDDs, thus easily replacing them in most applications



SSD (Flash Chip) Types (TLC, MLC, SLC)

SLC- Single Layer Cell

- High performance
- Lower power consumption
- Faster write speeds
- 100,000 program/erase cycles per cell
- Higher cost
- A good fit for industrial grade devices, embedded systems, critical applications.

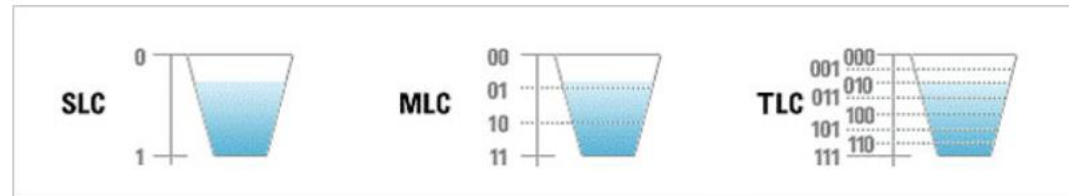


MLC- Multi Layer Cell

- Lower endurance limit than SLC
- 10,000 program/erase cycles per cell
- Lower cost
- A good fit for consumer products. Not suggested for applications which require frequent update of data.

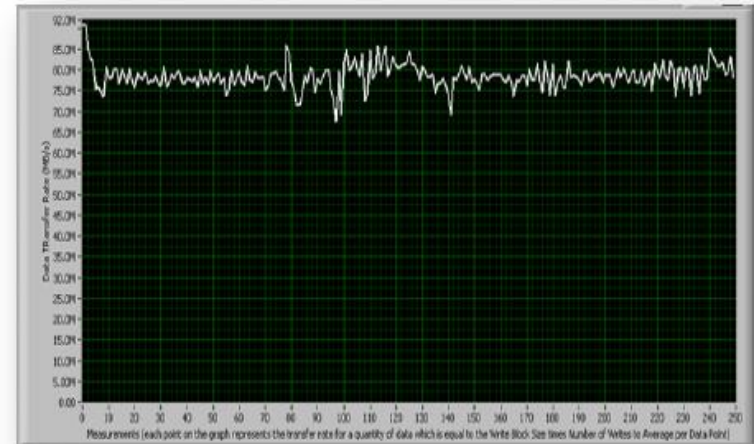
TLC- Three Layer Cell

- Higher density
- Lower endurance limit than MLC and SLC
- TLC has slower read and write speeds than conventional MLC
- 5,000 program/erase cycles per cell
- Best price point
- A good fit for low-end basic products. Not suggested for critical or important applications at this time which require frequent updating of data.



SSDs pros and cons

- Pros
 - Robustness (Less susceptible to vibrations and shock)
 - Increased write/read speeds (low access time and latency)
 - Not a drop in write speed when the memory fills up (as for HDDs)
 - Low power consumption (reduced heat generation)
 - Low boot-up time (for OS) and quicker application-launches
- Cons
 - High cost (in price/GB)
 - Low capacity (in # GB)
 - Great quality variations have been experienced
 - Reduced write speed experienced over time (for some suppliers)



[SSD - NI tutorial](#)

Selecting hard drives for DAQ systems

- A standard HDD is a 8/5 drive
 - Designed for power on for 8 hours, 5 days a week
- **Select Enterprise/Extended operations/ES version HDDs (when available)**
 - They are 24/7 drives, meaning that they are designed for continuous operation and high sustained throughput
 - Used in servers!

Key Care-Abouts



	Price	Temperature	Shock & Vibration	Power on Hours	Duty Cycle
Standard	Best	Average	Average	Worst	Worst
Extended Temperature ¹	Average	Best	Average	Worst	Worst
Extended Operation (24/7) ¹	Average	Average	Average	Best	Average
Solid State (Flash) ²	Worst	Best	Best	Best	Best

Best
 Average
 Worst

¹ Hard drives are available that are both extended-temperature and extended-operation.

² Care must be taken when selecting a solid-state hard drive in order to ensure that it will meet the requirements of your application. Additional details are provided in this white paper.

Example – Desktop HDD from WD

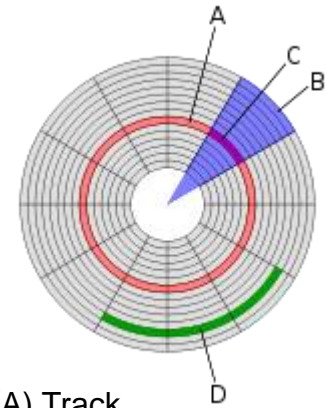
Desktop/Workstation Hard Drives	 WD Blue	 WD Green	 WD Black	 WD Red	 WD VelociRaptor
Product Selector >>	Solid performance and reliability for everyday computing.	Cool, quiet operation with massive capacity.	Maximum performance for power computing.	Designed and tested for RAID environments.	High-performance storage for power users.
Designed For	Desktop	Desktop	Desktop	Desktop RAID	Workstation
Capacity	80 GB – 1 TB	500 GB – 4 TB	500 GB – 4 TB	750 GB – 4 TB	250 GB – 1 TB
Interface	SATA 6 Gb/s PATA 100 MB/s	SATA 6 Gb/s	SATA 6 Gb/s	SATA 6 Gb/s	SATA 6 Gb/s
Form Factor	3.5-inch	3.5-inch	3.5-inch	3.5-inch 2.5-inch	3.5-inch 2.5-inch
Cache	8 MB – 64 MB	64 MB	64 MB	16 MB – 64 MB	64 MB
Limited Warranty	2-year	2-year	5-year	3-year	5-year

Factors that affect streaming performance

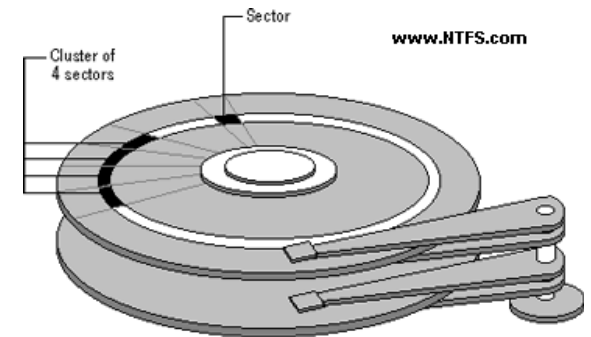
- Beyond overall application architecture, stream-to-disk or stream-from-disk rates can be affected by some of the following factors:
 - Running background programs such as virus scan
 - Recommended to disable the scheduled scans and updates for the entire duration of data streaming
 - How the hard drive is formatted to group data
 - Location of the file on the hard drive(s)
 - Locate the OS on a separate HDD (to free the fastest outer tracks for data storage)

HDD Sectors and Clusters

- The smallest unit of space on the hard disk that any software program can access is a **sector**, which usually consists of 512 bytes
- Traditional formatting provides space for 512 bytes per sector. Newer hard drives use 4096 byte (4K)
- Sectors are grouped into larger blocks that are called **clusters** (or *allocation units*)

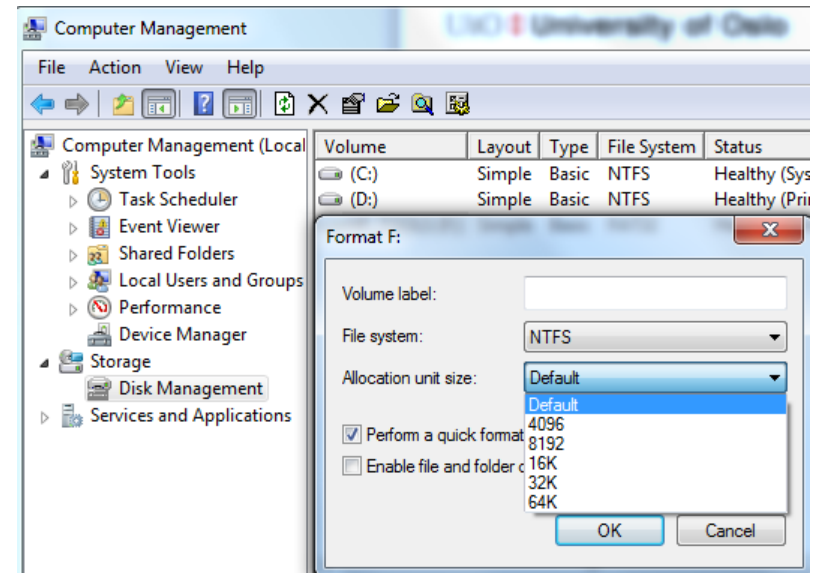


(A) Track
(B) Geometrical sector
(C) Track sector
(D) Cluster

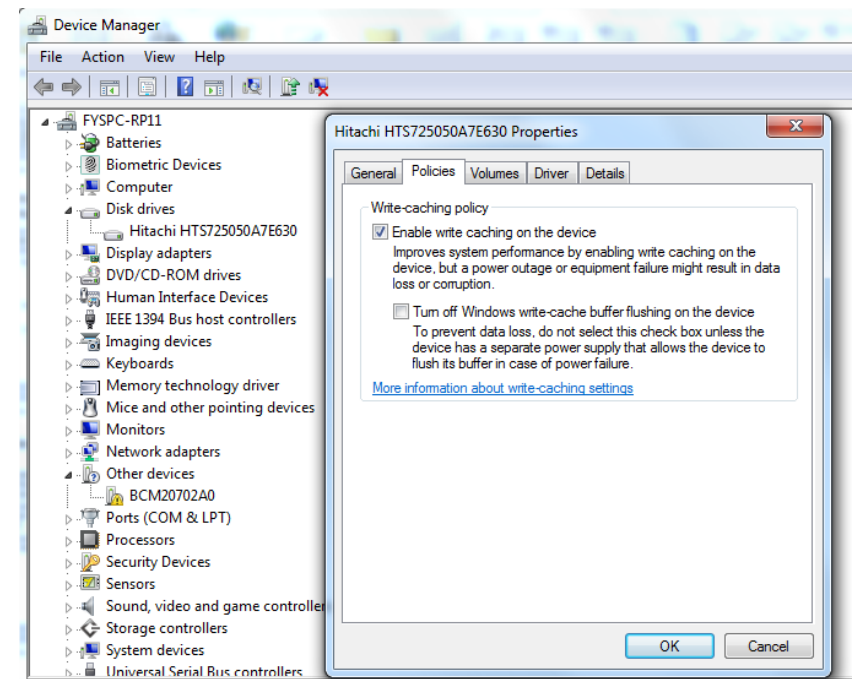


HDD optimization

- Larger cluster size (allocation unit size) can provide better streaming performance



- Enable the HDD onboard cache memory



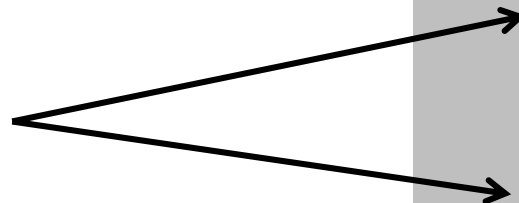
Determining Storage Format

When determining the appropriate storage format for the data, consider the following:

- At what sample rate will you acquire data?
- How much data will you acquire?
- Will you need to exchange data with another program?
- Will you need to search your data files?

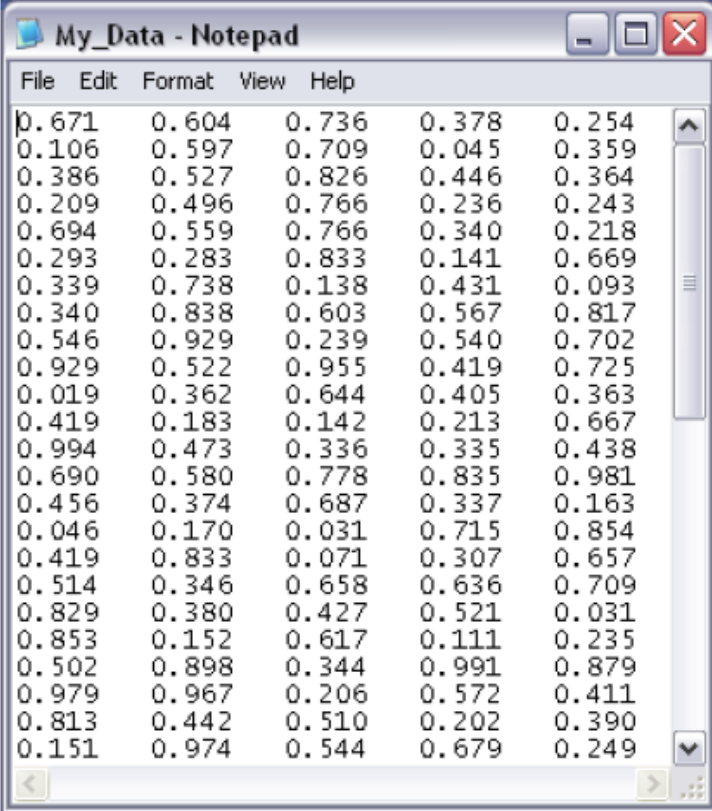
Configuration files

- ASCII (text)
- CSV
- Binary
- TDMS
- INI
- Spreadsheet
- AVI
- XML



ASCII Files

- **Pros**
 - Human-readable
 - Easily portable to other applications such as Microsoft Excel
 - Can easily add text information (first line) for each data column
- **Cons**
 - Large file size
 - Slow read and write

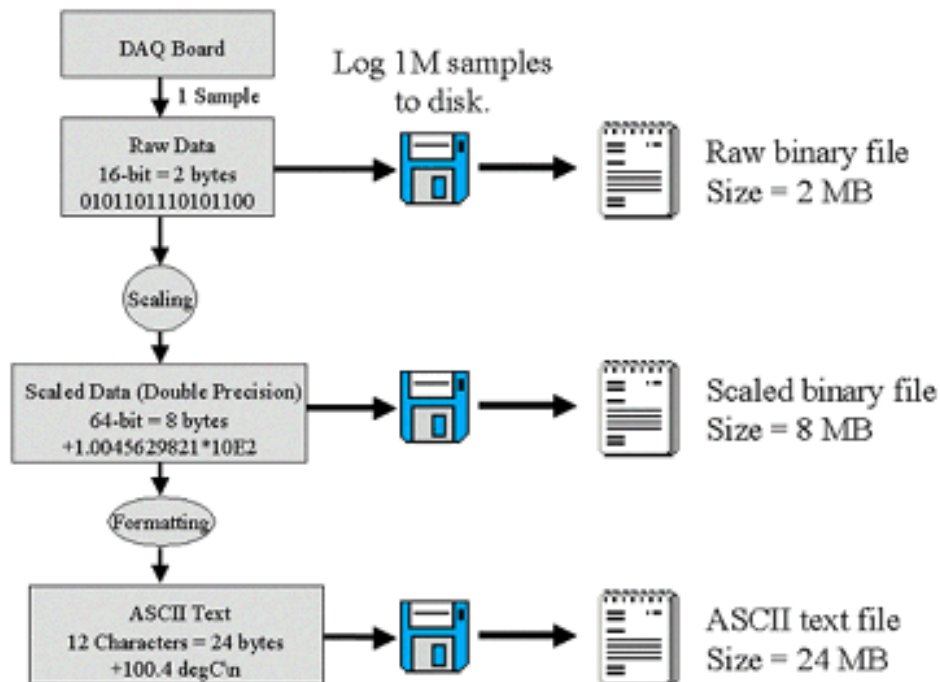


The screenshot shows a Notepad window titled "My_Data - Notepad" with a menu bar containing "File", "Edit", "Format", "View", and "Help". The window contains a table of numerical data with 5 columns and 20 rows. The data is as follows:

0.671	0.604	0.736	0.378	0.254
0.106	0.597	0.709	0.045	0.359
0.386	0.527	0.826	0.446	0.364
0.209	0.496	0.766	0.236	0.243
0.694	0.559	0.766	0.340	0.218
0.293	0.283	0.833	0.141	0.669
0.339	0.738	0.138	0.431	0.093
0.340	0.838	0.603	0.567	0.817
0.546	0.929	0.239	0.540	0.702
0.929	0.522	0.955	0.419	0.725
0.019	0.362	0.644	0.405	0.363
0.419	0.183	0.142	0.213	0.667
0.994	0.473	0.336	0.335	0.438
0.690	0.580	0.778	0.835	0.981
0.456	0.374	0.687	0.337	0.163
0.046	0.170	0.031	0.715	0.854
0.419	0.833	0.071	0.307	0.657
0.514	0.346	0.658	0.636	0.709
0.829	0.380	0.427	0.521	0.031
0.853	0.152	0.617	0.111	0.235
0.502	0.898	0.344	0.991	0.879
0.979	0.967	0.206	0.572	0.411
0.813	0.442	0.510	0.202	0.390
0.151	0.974	0.544	0.679	0.249

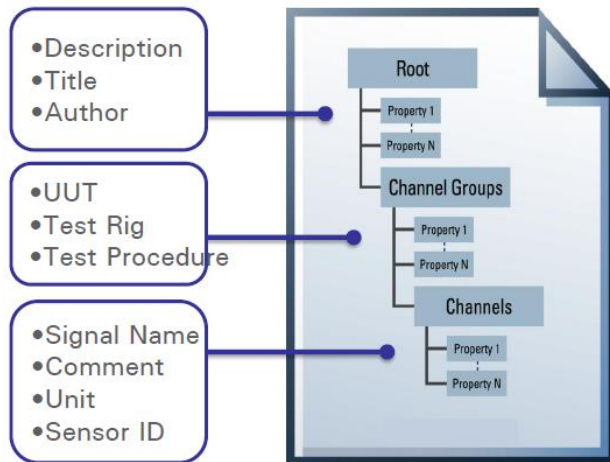
Data Types – file size

- Reduce file size
 - 1 16 sample = 16 bits = 2 bytes
 - 1 DBL sample = **64 bits = 8 bytes = 4X increase in bandwidth**



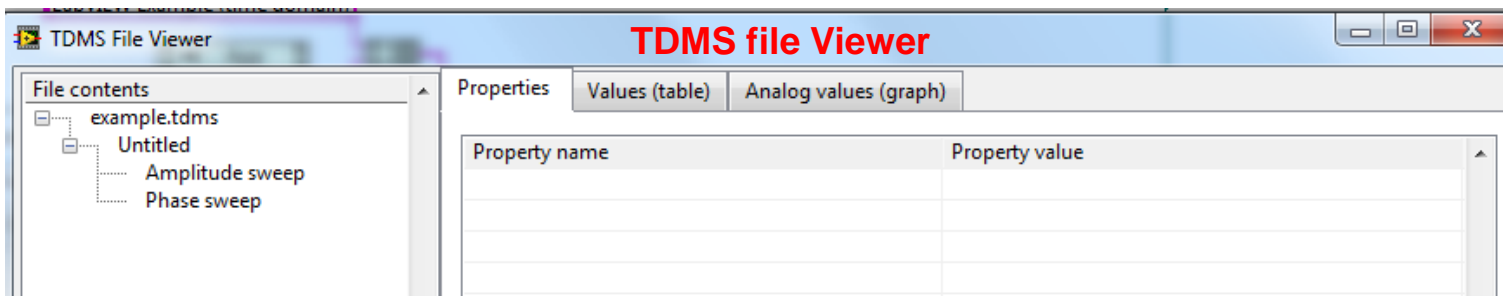
TDMS

- A file format from National Instruments – **use with LabVIEW!**
- TDMS = Technical Data Management Streaming
- Three levels of hierarchy
- Microsoft Excel add-in
- C-DLL
- Can download reader for Matlab

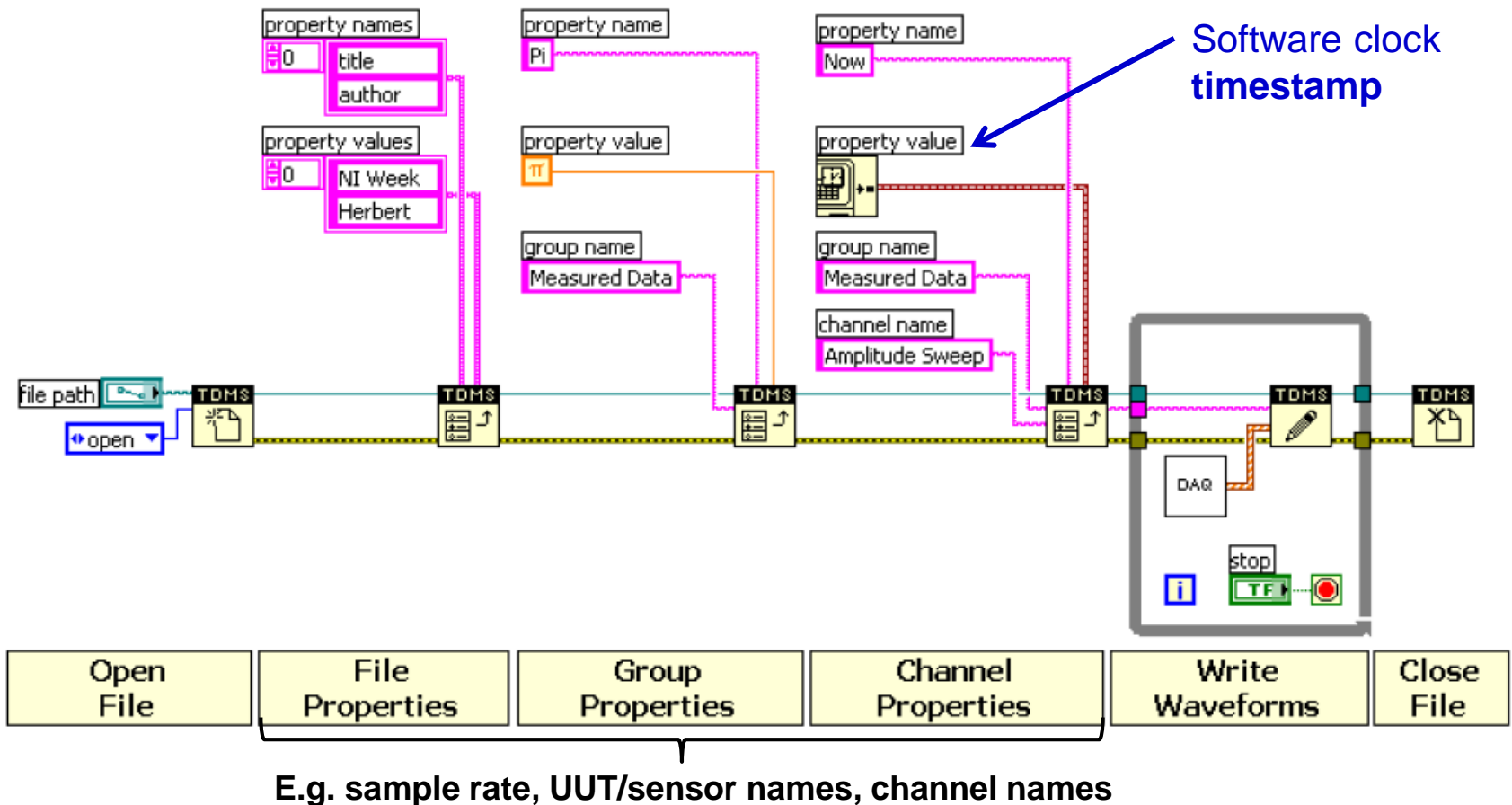


- ***Optimized for high-speed streaming***
- ***Single file***
 - ***Binary header***
 - ***Binary data***

ni.com/tdm

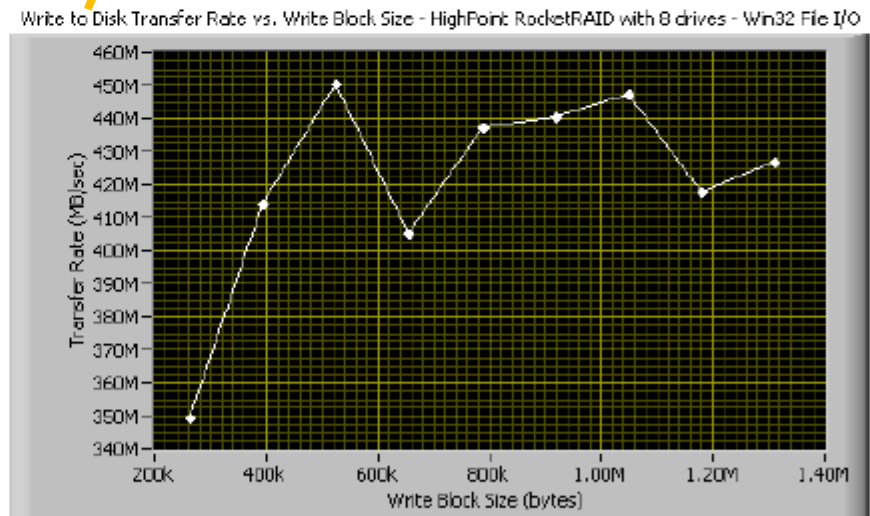
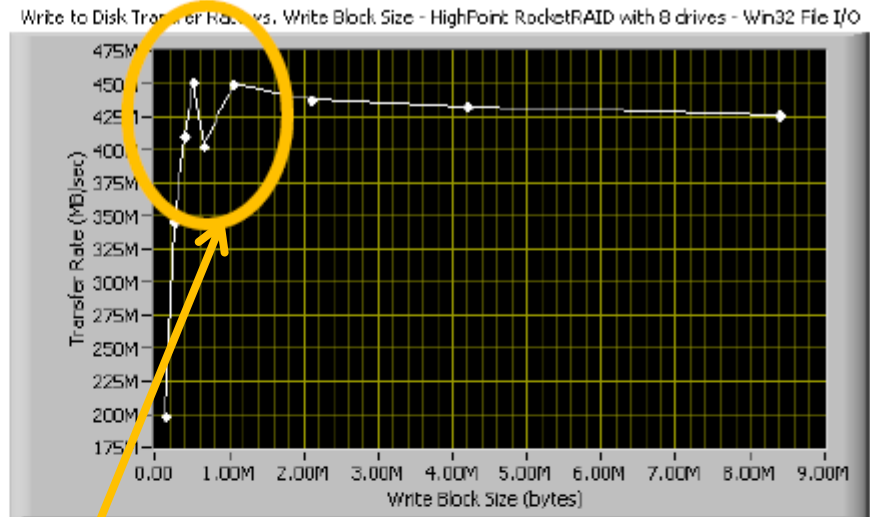


TDMS – Write Data and Set Properties



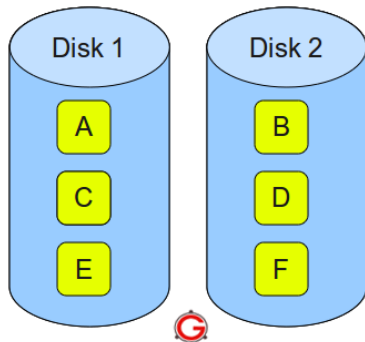
File I/O Block Size

- Different file I/O block sizes can impact performance
- Specifically relevant to RAID controllers
- Ideal file I/O block size may not line up with ideal measurement I/O block size

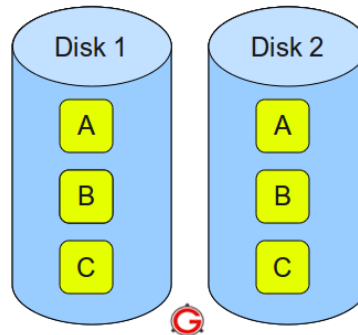


RAID introduction

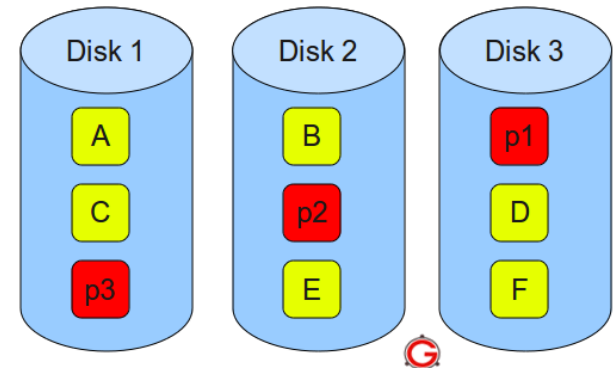
- RAID = Redundant Array of Independent Drives.
- RAID is a general term for mass storage schemes that split or replicate data across multiple hard drives.
 - To increase write/read performance and/or increase safety (redundancy)



RAID 0 – Blocks Striped. No Mirror. No Parity.



RAID 1 – Blocks Mirrored. No Stripe. No parity.



RAID 5 – Blocks Striped. Distributed Parity.

RAID examples

- Internal RAID of the workstation/PC
- Network attached storage (NAS) with RAID
- Server RAID
- Externally connected RAID (e.g. to the PXIe chassis)
- In-chassis PXI RAID
- SSDs (FLASH circuits) on a PCIe card

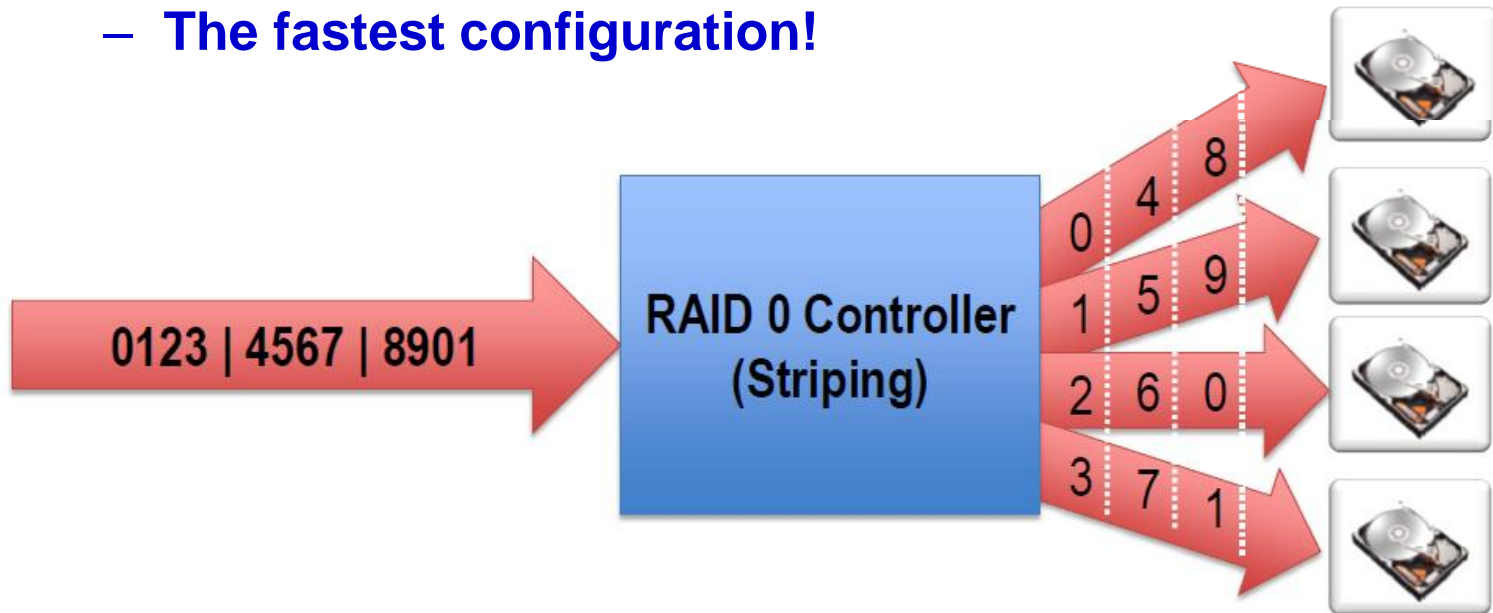


NI 8264
12-drive RAID



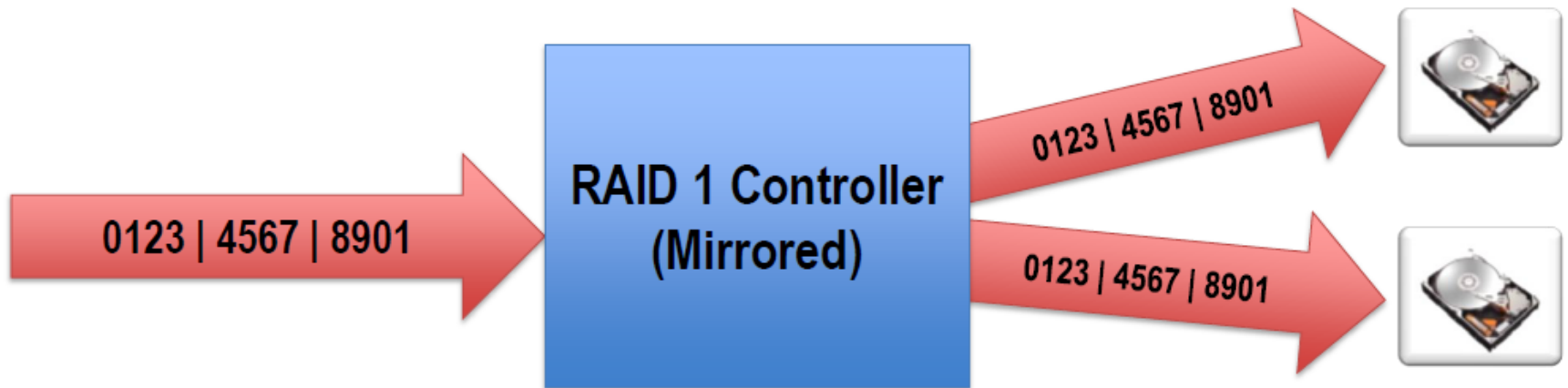
RAID-0

- **Striping without redundancy**
 - Improved speed over streaming to a single hard drive
 - Unimproved system reliability
 - Transparently supported by Windows OS
 - **The fastest configuration!**



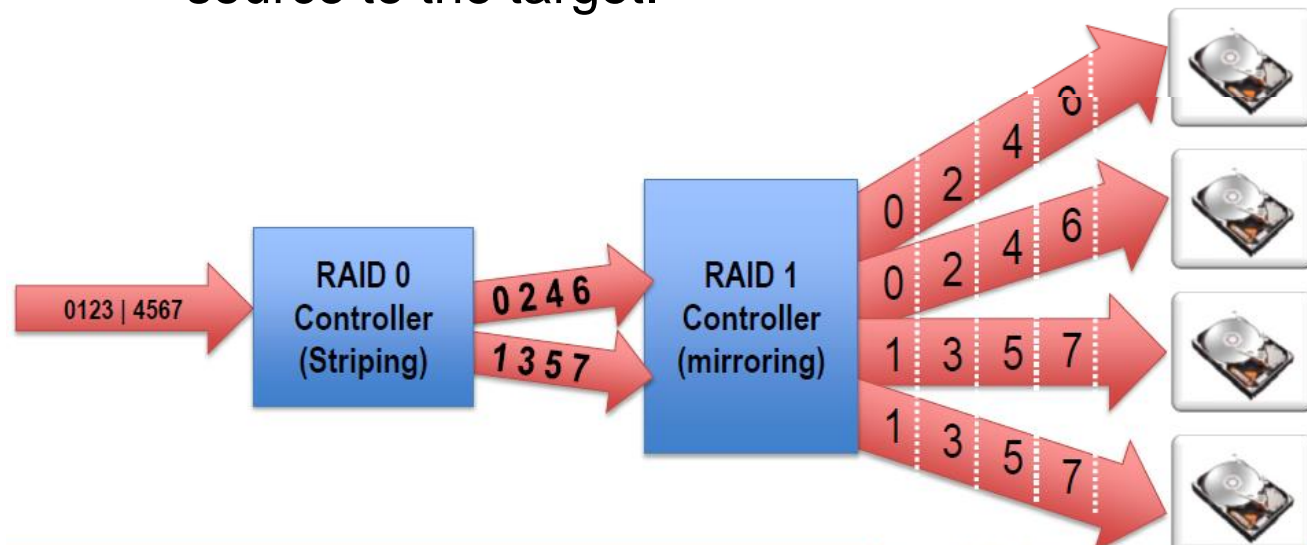
RAID-1

- **Mirrored (redundancy)**
 - 100% data redundancy
 - Each piece of data is written to **two** (or more) hard drives
 - No write speed increase over single disk
 - Highest overhead of all raid configurations
 - Often used for the operating system (OS) disks



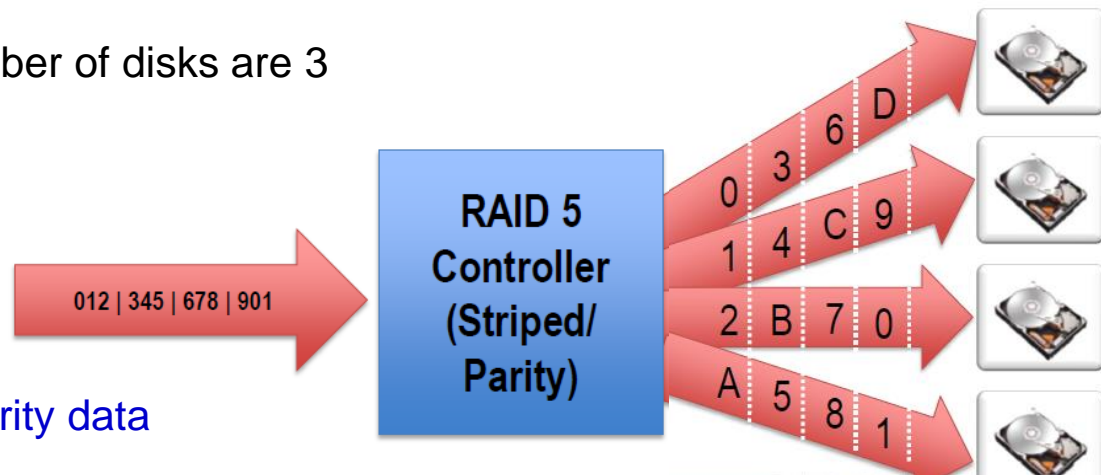
RAID-10 (1+0)

- **Striping and mirroring**
 - Both increased speed and redundancy compared to single drive
 - Can sustain multiple drive failures
 - Configuration requires twice the number of hard drives
 - Fast **rebuild** as data is copied block for block from the source to the target.



RAID-5

- **Distributed parity (single parity)**
 - Parity data distributed on all disks
 - Can only tolerate one drive failure (array continues to operate with one failed drive)
 - Single-parity RAID levels are as vulnerable to data loss as a RAID-0 array until the failed drive is replaced and its data rebuilt
 - More write overhead than RAID-1 because of the additional parity data that has to be created and written to the disk array
 - Poor performance with small files
 - Gives less space for measurement data (due to parity)
 - Reduces the amount of storage space available due to parity information
 - Minimum number of disks are 3



```

0 XOR 0 = 0
0 XOR 1 = 1
1 XOR 0 = 1
1 XOR 1 = 0

```

Parity: XOR method

- Parity is used both for the protection of data, as well as for the recovery of missing data.
- To calculate the parity for a RAID the [bitwise XOR](#) of each drive's data is calculated.
- The parity data is written to the dedicated parity drive.
 - **Note: distributed parity is used in the common RAID configurations!**
- In order to restore the contents of a failed drive, the same bitwise XOR calculation is performed against all the remaining drives, substituting the parity value (here 11100110) in place of the missing/dead drive:

```

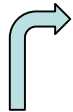
Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: 11110111 (Data)
Drive #4: 10110101 (Data)
Drive #5: ----- (Hot Spare)
Drive #6: 11100110 (Parity)

```

```

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: --Dead-- (Data)
Drive #4: 10110101 (Data)
Drive #5: 11110111 (Hot Spare)
Drive #6: 11100110 (Parity)

```



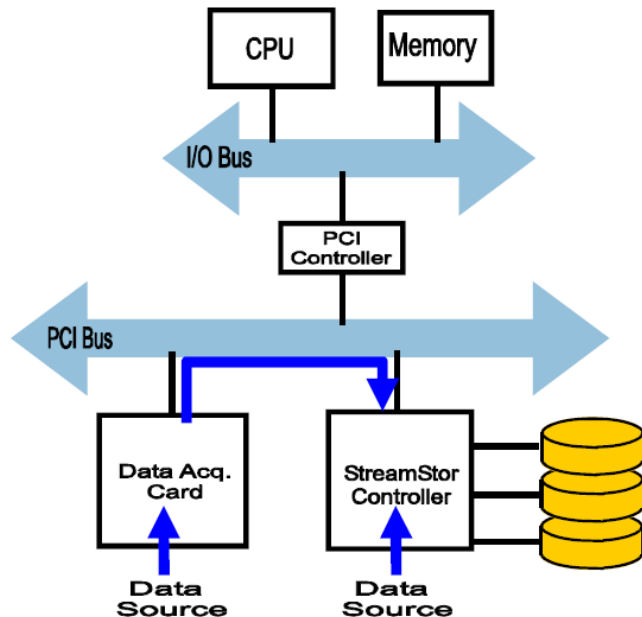
00101010 XOR 10001110 XOR 11110111 XOR 10110101 = 11100110

RAID-6

- **Double distributed parity**
 - Extends RAID 5 by adding an additional parity block
 - Provides fault tolerance from two drive failures (array continues to operate with up to two failed drives)
 - This becomes increasingly important as large-capacity drives lengthen the time needed to recover from the failure of a single drive
 - Double parity gives time to rebuild the array without the data being at risk if a single additional drive fails before the rebuild is complete.
 - Very slow write
 - because of the overhead associated with parity calculations

Direct-to-Disk controller

- With a PCI Express direct-to-disk controller module, data is streamed directly from the FIFO memory onboard a DAQ-card, across the PCI or PCI Express bus, to the direct-to-disk controller for acquisition
- This gives a minimum use of the PCI-buss and the CPU



<http://www.conduant.com/>

Streamstor Amazon Express Controller
(using PCI Express)

Note: Only a few DAQ cards supported!

Remote DMA

- In computing, remote direct memory access (RDMA) is a direct memory access from the memory of one computer into that of another **without involving either one's operating system**. This permits high-throughput, low-latency networking, which is especially useful in massively parallel computer clusters.
- RDMA supports zero-copy networking by enabling the network adapter to transfer data directly to or from application memory, eliminating the need to copy data between application memory and the data buffers in the operating system. Such transfers require **no work to be done by CPUs**, caches, or context switches, and transfers continue in parallel with other system operations. When an application performs an RDMA Read or Write request, the application data is delivered directly to the network, reducing latency and enabling fast message transfer.
- However, this strategy presents several problems related to the fact that the target node is not notified of the completion of the request.
- Hardware vendors have started working on higher-capacity RDMA-based network adapters, with rates of 100 Gbit/s reported.

InfiniBand

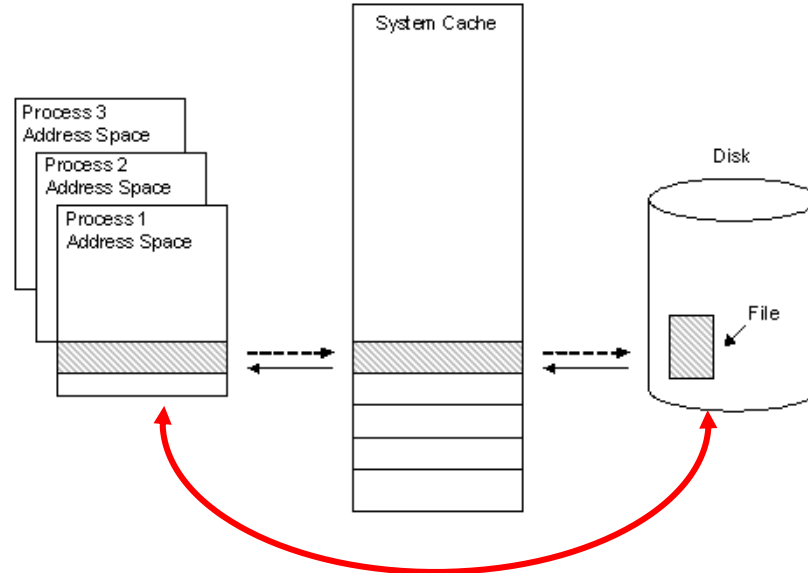
- InfiniBand (IB) is a computer-networking communications standard **used in high-performance** computing that features **very high throughput and very low latency**. It is used for data interconnect both among and within computers. InfiniBand is also used as either a direct or switched interconnect between servers and storage systems, as well as an interconnect between storage systems. **As of 2014, it was the most commonly used interconnect in supercomputers**
- InfiniBand is just a specific network architecture offering RDMA.

File Buffering - flushing

- File buffering is usually handled by the system behind the scenes and is considered part of file caching within the Windows operating system
- If your application continuously saves data to a hard drive, the operating system usually buffers these data in memory, without telling you, until it decides to physically write to disk.
- LabVIEW's **Flush file function** writes all the data in buffers to the disk
- When a file is closed it is automatically flushed

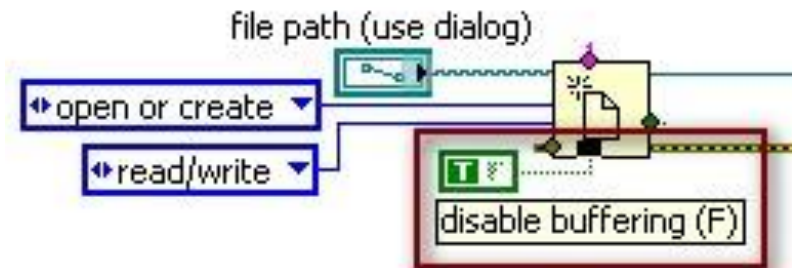
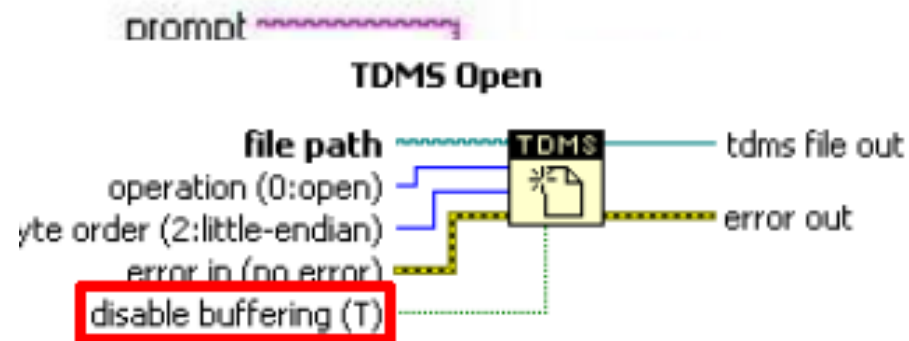
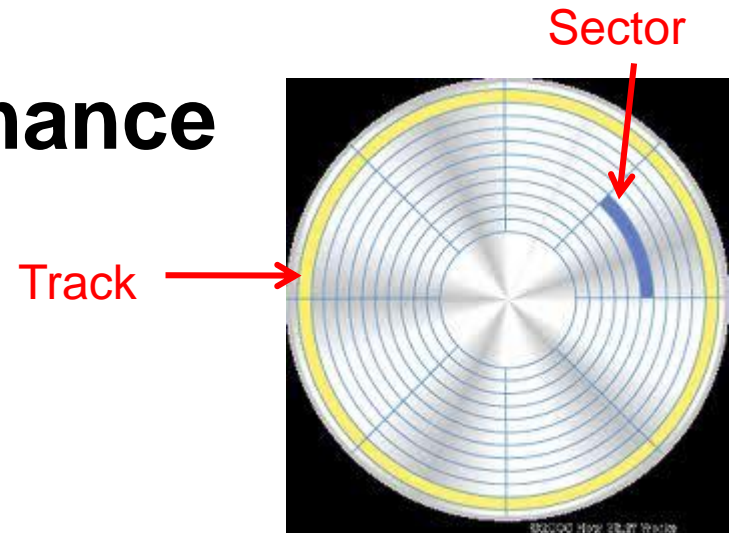
Unbuffered File I/O

- Possible to bypass file caching within the Windows operating system
- Can use unbuffered file I/O to increase write/read speed for RAID systems



Increase DAQ I/O Performance

- Use the option to **disable buffering (unbuffered file I/O)** in **Windows API**
 - Optimizes streaming applications
 - Important for RAID systems
 - Supported in LabVIEW
- Note that you must read from or write to the file in integer multiples of the disk sector size (usually 512 bytes)
- The data can span multiple sectors but must fill each sector completely
- If the data is not a multiple of the sector size, you must pad the data with filler data and delete the filler data before the data is used



New high performance sensors and buses

- Examples:
 - 4k cameras
 - A 4096 x 3000 pixel colour camera operated at 50 Hz will generate about 615 MB/s with 8 bit Bayer coding, and about 2500 MB/s with RGB coding.
 - 10 Gbit/s (1250 MB/s) Ethernet
- Problem:
 - SATA III gives only a 6 Gbit/s = 750 MB/s connection for a hard drive or SSD.
 - Effectively only 600 MB/s due to 8/10 decoding.
 - Need a faster interface to SSD for high speed storage!
- Solutions:
 - SATA Express gives 1969 MB/s
 - **The M.2 interface for SSD use PCI Express x4.**
 - For instance, the Samsung 970 EVO 1 TB M.2 **NVMe** SSD provide write speed up to 2500 MB/s, and the price is about 2500 NOK (feb. 2019).