Image Capture

Digital Forensics

Evidence Preservation

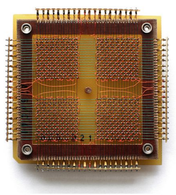
- A forensic investigation is normally part of a larger process, driven by law enforcement or corporate authority
- A forensic examiner gets involved when it is time to gather evidence and analyze it
- Gathering evidence properly is crucial because if it is not done correctly, all derivative works may be unusable
- Properly gathering evidence requires a working knowledge of the various ways information is stored in order to acquire it in a credible way

Non-electronic Information Sources

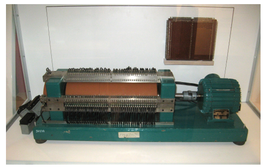
- Storage devices are implemented using one or more of several technologies
- The oldest method of modern information storage is to record it using symbols on paper, whether those symbols are printed text, specially marked pre-formatted cards or paper, or physically punched paper, called paper tape
- Printed text is the only form of paper-based evidence you are likely to encounter anymore
- This type of evidence is captured using cameras and scanners, and may or may not be done by a digital forensic examiner

Mass Storage Devices

- The next significant technologies are based on using magnetism to orient magnetic fields in metals or metal oxides, creating a relatively long-lived record of how those fields are set (polarity of a field indicating zero or one is the idea)
- These types of devices include the original core memory, drum memory, spinning disk platters as found in removable media drives, floppy drives, and hard disks
- It also includes magnetic tapes or cartridges and magstripes such as those found on credit cards and hotel room keycards
- This type of evidence is captured using imaging techniques



https://en.wikipedia.org/wiki/Magnetic-core_memory



https://en.wikipedia.org/wiki/Drum_memory

https://en.wikipedia.org/wiki/Disk_pack

https://en.wikipedia.org/wiki/Floppy_disk

https://en.wikipedia.org/wiki/Hard_disk_drive

https://en.wikipedia.org/wiki/Magnetic_tape_data_storage

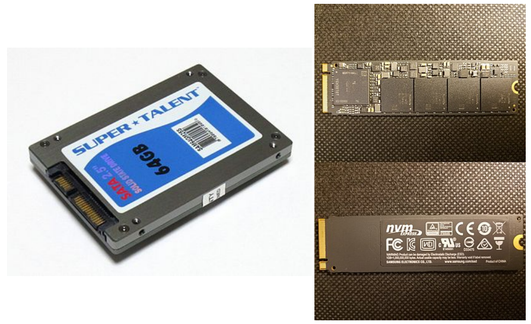
https://en.wikipedia.org/wiki/Linear_Tape-Open

Storage Devices

- More recently information has been imprinted on photo-sensitive media such as CD, DVD, or blu-ray disks using lasers, or physically stamped on those disks so that it can be read back using optical drives, this type of evidence is captured using normal disk imaging techniques
- The newest technologies for storing information implement various types of semiconductor devices, commonly known as flash, NVRAM, or solid-state disks

https://en.wikipedia.org/wiki/Compact_disc

https://en.wikipedia.org/wiki/USB_flash_drive



https://en.wikipedia.org/wiki/Solid-state_drive

Evidence Capture Equipment

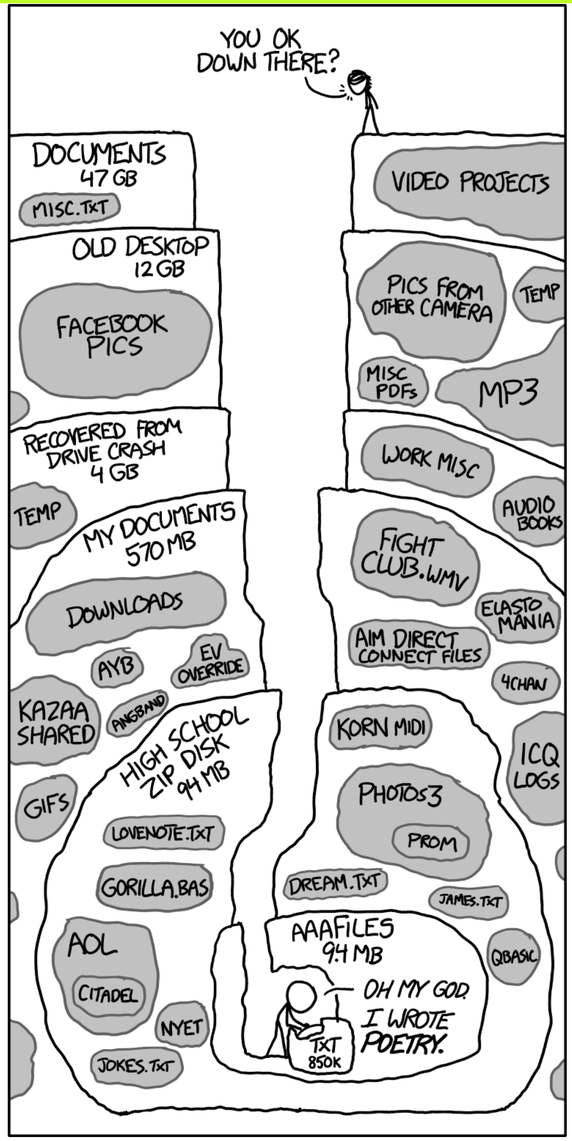
- A digital forensic examiner requires suitable hardware and software to perform data capture
- Seizing suspect hardware to use for this purpose is not always possible or advisable
- An examiner must have access to a system which can be used to capture digital images of physical items, such as scans of paper documents, photos of systems and suspect environments, bitwise images of electronic storage devices, etc. and this equipment may need to be portable

https://en.wikipedia.org/wiki/Sony_a99_II

https://en.wikipedia.org/wiki/Image_scanner

Where To Find Stuff

- Regardless of the technology used to store information, your parent(s) was/were right and your stuff must be organized if you want to be able to find it back
- You were also right in that their storage methods are not the only way you could choose to keep your stuff; people and programs that try to hide information sometimes use that discrepancy to put their data where you do not normally look for it
- Sometimes you are trying to find something even the original author cannot locate for various reasons



https://xkcd.com/1360/

Hiding data on storage devices

- The most common thing a user will do to try to hide information is to rename or delete the file(s) containing the information
- Users may try to transform the information inside a file to make it hard to identify what information in the file is of interest (e.g. compressed, encoded, encrypted, obfuscated, translated, or juxtaposed data doesn't look anything like the original data on normal inspection)
- Strategies to hide data may involve a filesystem that is not the same size as the volume it is stored on, leaving a gap which can hide data, abusing the host protected area of a hard disk, or the use of a partition to hold data without using any standard filesystem (sometimes called a raw disk)
- Sometimes bad actors intentionally use obsolete or obscure file formats or filesystem types to make it harder for evidence to be found or retrieved
- More sophisticated approaches combine one or more of the above, along with adding intentional corruption to the partition, volume, filesystem, or file to further stymie or mislead an investigator

Imaging Process

- In order to have the best chance of recovering information credibly from storage devices, we must begin by obtaining a copy of the storage device, using various imaging techniques
- If possible and permitted, remove the storage device physically from the computer under investigation and install it as an add-on drive on a forensic workstation, record this activity for your final report
- When forced to do imaging with Windows, always use a physical write-blocking device because just accessing the disk can update timestamps on it, subtly altering the disk image, software-based write blocking may be an option for USB devices, record this activity for your final report
- Once an image has been made of the device, disconnect it and store it securely for future validation if required, record this activity for your final report
- Always do investigative work on a read-only image, or on a copy of the image

Storage Device Imaging

- Imaging a storage device is done to preserve all the current information stored on it and includes all the bits it is possible for the storage medium to store, not just the ones that look like normal files
- To use an image for legal evidence, there are processes and rules for what can be copied, how to perform the copying, who is involved, and how the copy is handled
- An image is simply a copy of every bit stored on the device, so it is important to give consideration to where the image will be stored since it can be a large amount of data (the size of the partition, volume, or device you are capturing)
- Compression can be used when storing images, but it must be lossless compression
- There are techniques to create images of RAM if you have physical access to the computer; this allows attempts at recovery of things like encryption keys and passwords (e.g. https://en.wikipedia.org/wiki/Cold_boot_attack)

Evidence Capture Equipment

- An examiner's system must have workstation-class capabilities (CPU, RAM, USB3 or better external ports, SSD operating drive(s), high capacity bulk storage, high speed networking, etc.) because examiners work with unpredictably large volumes of data
- Don't forget to include things like cameras, scanners, external drive bays, high-speed USB, thunderbolt ports, write-blockers, USB docking devices, and more mundane items like flashlights, screwdrivers, and anti-static evidence bags

Image Verification

- To use an image and any information drawn from it in legal proceedings, the image must be verifiable as authentic and unaltered
- For the image to be verified as unaltered, it has to be compared to the original
- Bit by bit comparison is very slow and resource intensive, so we use a method known as hashing to derive a relatively distinctive sequence of bits to represent the contents of the image which is much shorter than the actual image so it is easier to store and transfer
- Forensic grade tools create the hash while making an image
- The hash itself can be stored in a separate file, or may be embedded into the forensic image file in one of several formats

Image File Formats

- The Unix/Linux family of imaging tools are usually based on the dd program, and they capture raw image files (often named with a .dd suffix); additional files are often kept with them containing information relevant to an investigation such as digital copies of signed documents, hashes, notes, etc.
- Advanced Forensic Format (AFF) is an open source format that supports encryption and storing metadata for simplicity, consistency checking, and self-authentication (not always a good thing)
- Many commercial software packages have their own proprietary formats for image files, Expert Witness Format (EWF) and SMART are used by some popular commercial software packages
- Most formats are fairly easily converted to other formats using the standard tools that read those formats, raw format is more or less universally usable

Large Image Files

- Imaging storage devices can result in very large files
- Compressing the files is helpful, but they need to be uncompressed to work with them
- An image is stored in a single file if possible, sparse storage is a good option for imaging devices that have significant unused space
- Images may be split up into a sequence of files of a maximum size but this can make them more time-consuming to work with
- The maximum size can be driven by the media used to store them, the media used to archive them, or the limitations of the system doing the investigation
- Images may be written to regular files, other drives, or external storage media such as optical disks or magnetic tape

Bootable Forensics Software

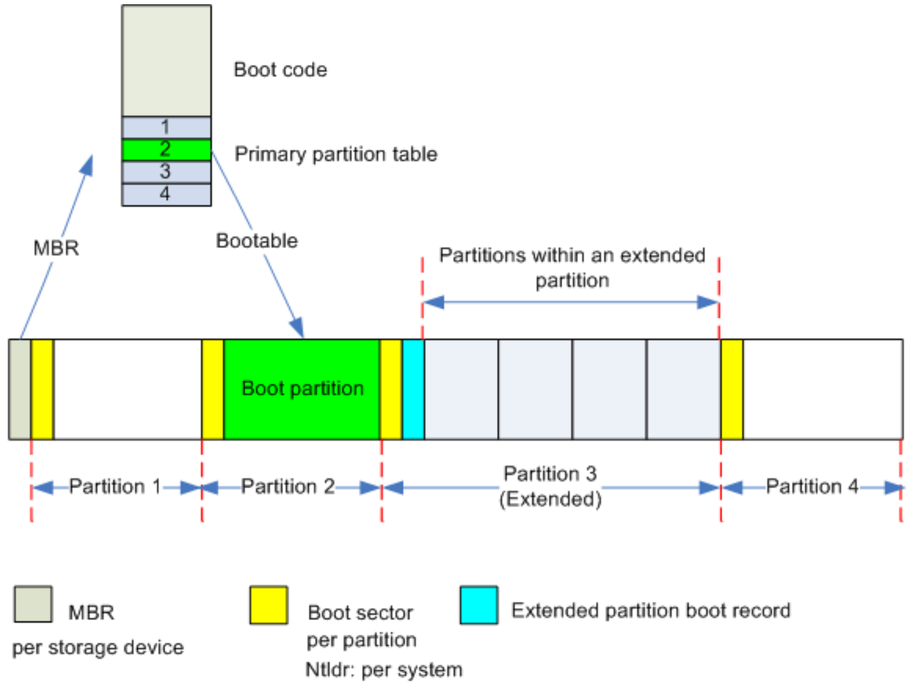
- Several toolsets have been crafted into pre-configured operating systems, and some operating environments have been built and preconfigured to run security software including forensic tools
- SIFT Workstation, set of tools pre-built for a Ubuntu system
- Kali Linux, everything including the kitchen sink system
- Parrot, streamlined Linux distro for pen testing, forensics, reverse engineering, cryptography, and privacy
- CAINE, a distro focused exclusively on forensics
- See https://distrowatch.com/search.php?category=Forensics#simple for a list of Forensics-oriented Linux distros
- Whatever you choose, consider running it as a live boot OS instead of an installed OS in order to maintain credibility of your toolset

Image Capture Tools

- UNIX/Linux provides an excellent operating environment for image capture when windows-specific drivers or encryption are not required to access a storage device
- UNIX/Linux provides the dd command, and variations based on dd, that do not require the target device to be mounted, and will not alter the target device unless explicitly commanded to do so
- dd can save raw image files, runs very fast, supports saving as sparse files, handles drive errors, and can be used in a pipeline to allow for encryption/decryption, hashing, etc.
- dc3dd is a forensics-enhanced form of dd that can save more information as metadata, generate hashes on the fly to save imaging time and resources, and has more sophisticated drive error handling
- There are other tools both for the command line and with GUIs, but these are all that are needed for typical storage devices
- Using any of these tools requires you to know what devices, partitions, volumes, and filesystems might be present and of interest

Storage Device Partitioning

- A single storage device is usually logically divided into pieces called slices or partitions to allow multiple independent storage spaces to exist on a single device
- Storage devices that can be partitioned have a table stored in the first block of data on them that describes the partitioning in use; that block contains a disk label, partition table, or partition map
- A single partition is often used for simple systems with that partition being defined as including the entire storage device
- The maximum number of partitions, their maximum size, and the maximum size of disk which the partition table is usable on is dependent on what partition table type is in use
- Commonly used partition table types are MBR (https://en.wikipedia.org/wiki/Master_boot_record) and GPT (https://en.wikipedia.org/wiki/GUID_Partition_Table)



Partitioning Investigation

- The easiest way to view a partition table is to use fdisk (fdisk -l will display the partition table) which is available in Windows and Linux, gparted is an example of a program that will handle a greater variety of partition table types
- The table shows each partition that is defined, where it starts on the drive, how long it is, and has a type code indicating what might be stored in the partition
- The table can be viewed without altering it
- The Linux lsblk command can be very helpful in providing a concise list of disk devices, their partitions, and their mount points
- The entire drive can be used for imaging, or the partitions can be individually imaged
- When looking at a table, be sure to watch for gaps between partitions, before partitions, or at the end of the drive; bad actors can use these normally inaccessible chunks of disk space to store information they want to hide (this is a non-trivial act)
- hdparm can be used to view the host protected area (HPA) which may be present on a drive and make it accessible, as well as show the DCO area

Volumes

- In order to allow a single collection of file data to be larger than any disk partition, the volume concept provides the ability to group partitions from one or more storage devices into a single logical space called a volume
- If you have a multi-device volume, imaging it may be a non-trivial task requiring additional hardware and software both to read the target volume, and to save the image obtained
- Single partition volumes are the most common type of volume for desktop and small server systems
- RAID and other complex forms of disk management may not use a disk partition map in the same way or at all, and use volume management software to organize the contents of their volumes; partition tables on such drives may only be there by convention
- http://www.ntfs.com/ldm.htm has a very detailed discussion of PC disk organization and partition structures, but it is written as if nothing has ever existed but PCs

Volume Identification

- A partition or whole drive may not hold a complete filesystem, spanned disks and RAID configurations can complicate the picture
- These types of drives have the partition type codes set to indicate when a partition belongs to a group comprising a RAID volume
- Some software packages can put the RAID volume together from images of the raw devices, other times, the entire volume might be possible to image using the metadisk drive name that the operating system uses to refer to the volume (mdadm is a Linux tool to manage and monitor RAID volumes)
- RAID volumes can also be stored on drives attached to hardware-based RAID controllers and that specific hardware configuration is usually required to be intact to retrieve the data on the volume
- Linux commands like lsblk, df, mount, mdadm, and simply using ls /dev can help identify volume names

Filesystems

- A filesystem is a data structure applied to storage spaces to enable storage, retrieval, and management of the data stored therein using the file concept
- On a simple desktop type of computer, a single filesystem is stored on a single volume and occupies the entire volume, and that volume is stored on a single partition and occupies the entire partition, and that partition is stored on a single drive and occupies the entire drive
- There have been many distinct filesystem formats, or types (structure definitions), since computers were invented; most are no longer in regular use
- Awareness of many different types of filesystems is required do be able to do forensic work in varying environments and situations

Microsoft Filesystems

- The Microsoft File Allocation Table (FAT) filesystem format is the oldest one you are likely to encounter on any machine running a Microsoft operating system
- exFAT was the replacement for FAT to get past a number of limitations of the FAT design
- NTFS was the replacement for the FAT family of filesystems to redesign how files can be stored and managed in Windows and is the most commonly used filesystem for internal drives for Microsoft operating systems
- http://www.ntfs.com/ntfs_vs_fat.htm

UNIX Filesystems

- UNIX and its derivatives have always been implemented on a considerably wider variety of system hardware than Microsoft operating systems
- The wide variety of systems and applications for UNIX and UNIX-derived systems and the open source nature of the operating systems have led to many filesystems being developed, each of which is intended to be well matched to the task that system is deployed to perform
- Linux systems typically use EXT2 and EXT4 currently
- MacOSX systems use the older HFS+ and the newer APFS, depending on the drive type and OS version
- UNIX systems may use ufs, although the newer zfs is gaining market share

Cross-platform Filesystems

- Optical media such as CDROM, DVD, and Blu-ray disks store digital data and are sometimes used for saving or shipping information
- They commonly use ISO 9660 or CDFS format, or ISO 13346 format and can be imaged just like the previously discussed filesystems
- Other storage media you may need to capture might include magnetic tape or other serial data storage devices, these are imaged by simply reading the bit streams from them and saving those streams in files
- It is a common practice to use portable USB drives or external hard drives to transfer data between people and systems and these are often formatted as FAT, FAT32, or NTFS
- Any filesystem files are copied to will silently discard metadata if the file source is a different filesystem type
- Always transfer evidence files as raw files, forensic format files, or archives produced using the source OS to preserve metadata

Working With Filesystems Not Stored On Volumes

- A volume is the container normally used to store a filesystem
- A filesystem is just a data structure and can be stored in anything that will hold data
- A not uncommon technique to create portable filesystems is to store them in files instead of hardware-backed volumes which makes them easy to move between systems, but also gives rise to some opportunities to make them hard to read unless you have the necessary knowledge (the keys and tools) to access the actual data
- The filesystem can be encrypted, and the file it is stored on can be encrypted, and you can even store the encrypted filesystem file on an encrypting filesystem which can be kept on a device that does hardware encryption - this would make it cpu-intensive to use, but impossible to crack open in any practical way

File-based Filesystems

- There are filesystems that were designed to be stored in a file that lives on another filesystem, although it is only the filesystem management tools that typically prevent storing all filesystem formats in ordinary files
- An example is TrueCrypt, which was discontinued in 2014, but forks are active as well as at least one complete re-implementation (Veracrypt)
- These typically provide some capability not provided by the host filesystem, and sometimes are there purely for portability
- TrueCrypt can give a suspect deniability, because the contents are not something you can examine directly and it can even be used to run a hidden operating system
- https://en.wikipedia.org/wiki/TrueCrypt has good information on this type of filesystem, as well as several examples of how TrueCrypt was relevant in a number of legal proceedings

Filesystem Identification

- Identifying a filesystem type may not be as simple as examining partition table type codes
- mount and /etc/fstab or /etc/vfstab in UNIX/Linux
- Disk Management app in Windows
- lsblk -f, blkid -p, disktype, and file -s commands in Linux
- Once the filesystem has been identified in an image, analysis on the contents can begin

Imaging Caveats for Windows

- There are several Windows applications which can be used to capture FAT and NTFS filesystem images
- In addition to image capture, most can use the files in an image for deeper analysis, such as viewing the registry stored on the image, or searching for specific data or files
- Windows does not provide the capability to create an image of a drive without mounting it, which alters the drive if hardware-based write-blocking is not in use (strongly recommended when using Windows for capture, requires special hardware)
- Sometimes software is needed that only runs on the target system in order to access the drive data because it is encoded, encrypted, or is otherwise under access control mechanisms that require the OS to be running

Image Capture Lab

- Simulate an incident
- Create investigation report
- Seize drives
- Image drives
- Compare imaging tools
- Produce image files for next lab