MODULE II DATA ACQUISITION

Storage Formats for Digital Evidence – Best Acquisition Method – Contingency Planning for Image Acquisitions – Acquisition Tools – Validating Data-Acquisition – RAID Data Acquisitions.

WHAT IS DATA ACQUISITION?

The gathering and recovery of sensitive data during a digital forensic investigation is known as data acquisition. Cybercrimes often involve the hacking or corruption of data. Digital forensic analysts need to know how to access, recover, and restore that data as well as how to protect it for future management. This involves producing a forensic image from digital devices and other computer technologies.

TYPES OF DATA SOURCES

In cyber forensic investigations, digital evidence can be sourced from various types of devices and platforms. These sources can be broadly categorized into three main types

1. Primary Sources

- Primary sources are the devices that are directly involved in the incident or contain the original data.
- Examples:

Computers (desktops, laptops),Servers (file servers, web servers, email servers),Mobile devices (smartphones, tablets)

- These devices typically store a wide range of data, including files, system logs, internet history, emails, and application data.

2. Secondary Sources

- Secondary sources are additional devices or platforms that may indirectly contain relevant digital evidence related to the incident.
- Examples:

Cloud storage services (Google Drive, Dropbox, iCloud), Network logs (firewall logs, DHCP logs, DNS logs), IoT (Internet of Things) devices (smart home devices, wearables)

- Secondary sources often provide contextual information, communication logs, or metadata that complement primary sources.

3. Tertiary Sources

- Tertiary sources include external sources beyond the immediate control of the investigator, yet they may contain pertinent data.
- Examples:

Social media platforms (Facebook, Twitter, Instagram), External databases (public records, online forums)

STORAGE FORMATS FOR DIGITAL EVIDENCE

1.Raw format

Raw format method is make it possible to write bit-stream data to files

The **advantages** of raw format are as follows −

- Fast data transfers
- Ignores minor data read errors on source drive
- Most computer forensics tools can read raw format.

The **disadvantages** of raw format are as follows −

- It requires as much storage as the original disk or data.
- Tools might not collect marginal (bad) sectors.

2.Proprietary formats

Most forensics tools have their own formats

The **features** offered in proprietary format are as follows −

- Option to compress or not compress image files.
- Can split an image into smaller segmented files.
- Can integrate metadata into the image file.

The **disadvantages** of proprietary format are as follows −

- This format has an inability to share an image between different tools.
- File size limitation for each segmented volume.
- The Expert Witness format is an unofficial standard.
- FTK uses and Encases USES.

3.Advanced forensics format

This format was developed by Dr. Simson L. Garfinkel as an open-source acquisition format. It designs goals and provides compressed or uncompressed image files.

There is No size restriction for disk-to-image files and it can provide space in the image file or segmented files for metadata.

It has simple design with extensibility and is an open-source for multiple platforms and OS. Moreover, its internal consistency checks for self-authentication.

File extensions in advanced forensics format include the following −

- .aff – variation that stores all data and metadata in a single file
- .afm – variation stores all the data and metadata in separate files
- .afd – variation stores all the data and metadata in multiple small files.
- AFF is open source

DETERMINING THE BEST ACQUISITION METHOD

In digital Forensics, there are 2 types of acquisitions:

1. **Static Acquisition:** which is the preferred way to collect a digital evidence when a computer seized during police raid.
2. **Live Acquisition:** is the way to collect digital evidence when a computer is powered on and the suspect has been logged on to. This type is preferred when the hard disk is encrypted with a password.

For both types, there are 4 methods of collecting data:

1. **Creating a disk-to-image file:** the most common method to collect data. It allows the investigator to create on or many bit-for-bit replications of the original drive. By using this method, we can use any of the forensics tools such as ProDiscover, EnCase, FTK, X-ways, ILook, SMART, and Sleuth Kit to read the different types of disk-to-image files.
2. **Creating a disk-to-disk copy:** is used when disk-to-image faces hardware of software errors due to incompatibilities. It copies the entire disk to a newer disk by using any of the forensics tools such as EnCase and SafeBack. These tools can adjust the target disk's geometry to match the original drive.
3. **Creating a logical disk-to-disk or disk-to-data file:** this is the preferred method with large data storage such as RAID servers. This method captures only specific files or file types of interest to the case. It is used when time is limited.
4. **Creating a sparse copy of a folder or file:** this method is similar to creating a logical acquisition but it also collects deleted data (unallocated). Also this method is used when an investigator doesn't need to examine the whole drive.

To determine the appropriate acquisition method, the investigator must consider the following:

1. The size of the source disk.
2. Can you retain the source disk as an evident or must you return it to the owner?
3. Time to do perform the acquisition.
4. Location of the evidence

CONTINGENCY PLANNING FOR IMAGE ACQUISITIONS

Contingency planning for data acquisitions in digital forensics involves preparing for unexpected events or challenges that may arise during the acquisition process. These contingencies can include technical failures, legal obstacles, resource constraints, or other unforeseen circumstances that may disrupt or complicate the acquisition of digital evidence. A robust contingency plan helps forensic examiners mitigate risks, maintain the integrity of the evidence, and ensure the continuity of the investigation. Here's a detailed discussion of the key components of contingency planning for data acquisitions:

1.Risk Assessment:

- Identify potential risks and challenges that may impact the data acquisition process. This includes technical risks such as hardware or software failures, legal risks related to admissibility of evidence, logistical risks such as resource constraints or scheduling conflicts, and procedural risks such as chain of custody issues.
- Evaluate the likelihood and potential impact of each risk to prioritize contingency planning efforts. High-risk scenarios may require more comprehensive mitigation strategies.

2.Backup Plans:

- Develop backup plans to address common technical failures or disruptions that may occur during data acquisition. This may include having redundant hardware or software tools available, maintaining spare equipment, and establishing alternative acquisition methods or procedures.

• Ensure that forensic examiners are trained and equipped to quickly switch to backup plans in the event of a failure, minimizing downtime and mitigating the risk of data loss or corruption.

3.Contingency Procedures:

• Define clear procedures and protocols for responding to unexpected events or challenges during data acquisition. This includes steps for troubleshooting technical issues, escalating problems to appropriate personnel or authorities, and documenting deviations from the original acquisition plan.

• Establish communication channels and escalation paths to coordinate response efforts and seek assistance from relevant stakeholders, such as IT support teams, legal advisors, or law enforcement agencies, as needed.

4.Redundancy and Resilience:

• Implement redundancy measures to ensure the resilience of data acquisition systems and processes. This may include redundant hardware components, backup power supplies, off-site storage solutions, and failover mechanisms to switch to alternative acquisition methods or tools if necessary.

ACQUISITION TOOLS

Acquisition Tools refer to software and hardware solutions designed to acquire, extract, and preserve digital evidence from computers, mobile devices, cloud storage, and other digital sources. These tools utilize specialized techniques to ensure the integrity and admissibility of collected evidence in legal proceedings.

The Importance of Forensic Acquisition Tools for Digital Investigators

1. Data Collection and Preservation

Forensic Acquisition Tools enable digital investigators to collect and preserve digital evidence in a forensically sound manner. They employ write-blocking and hashing techniques to ensure data integrity during the acquisition process.

2. Comprehensive Device Support

Digital investigations involve a diverse range of devices and operating systems. Forensic Acquisition Tools offer support for various devices, including computers, smartphones, tablets, IoT devices, and more, ensuring investigators can access evidence from different sources.

3. Efficient Data Extraction

Time is of the essence in digital investigations. Forensic Acquisition Tools streamline the data extraction process, enabling investigators to quickly access and analyze crucial evidence.

4. Acquiring Hidden and Deleted Data

Digital evidence may be hidden or deleted by individuals seeking to conceal their activities. Forensic Acquisition Tools employ advanced techniques to access hidden data and recover deleted files, revealing valuable evidence.

5. Forensically Sound Practices
Adhering to forensically sound practices is essential for digital investigators to maintain the integrity and admissibility of evidence. Forensic Acquisition Tools ensure proper handling of data, preserving its evidentiary value.

6. Preservation of Metadata
Metadata, such as timestamps and file attributes, can provide critical context in digital investigations. Forensic Acquisition Tools preserve metadata during the acquisition process, enhancing the reliability of evidence.

7. Chain of Custody Documentation
Maintaining a clear and documented chain of custody is essential in digital investigations. Forensic Acquisition Tools generate detailed reports, documenting each step of the acquisition process for legal purposes.

8. Cloud Data Extraction
As data storage shifts to the cloud, digital investigators must access evidence from cloud-based sources. Forensic Acquisition Tools offer capabilities to acquire data from cloud storage services securely.

9. Reduced Risk of Data Contamination
Forensic Acquisition Tools employ write-blocking mechanisms to prevent accidental data contamination during acquisition. This feature ensures that investigators can collect evidence without altering the original data.

10. Support for Encrypted Devices
Encrypted devices pose a challenge in digital investigations. Forensic Acquisition Tools are equipped with decryption capabilities, allowing investigators to access data on encrypted devices with proper authorization.

VALIDATING DATA-ACQUISITION

Validating evidence may be the most critical aspect of computer forensics

- Requires using a hashing algorithm utility
- Validation techniques – CRC-32, MD5, and SHA-1 to SHA-512

Linux Validation Methods

- **Validating dd acquired data**
  - – You can use md5sum or sha1sum utilities
  - – md5sum or sha1sum utilities should be run on all suspect disks and volumes or segmented volumes
- **Validating dcfldd acquired data**
  - – Use the hash option to designate a hashing algorithm of md5, sha1, sha256, sha384, or sha512

– hashlog option outputs hash results to a text file that can be stored with the image files – vf (verify file) option compares the image file to the original medium

Windows Validation Methods

- Windows has no built-in hashing algorithm tools for computer forensics
  - Third-party utilities can be used
- Commercial computer forensics programs also have built-in validation features
  - Each program has its own validation technique
- Raw format image files don't contain metadata

Separate manual validation is recommended for all raw acquisitions

RAID DATA ACQUISITIONS.

Acquisitions of RAID drives can be challenging and frustrating for digital forensics examiners because of how RAID systems are designed, configured, and sized.
Size is the biggest concern because many RAID systems are now pushing into exabytes or more of data.

Understanding RAID

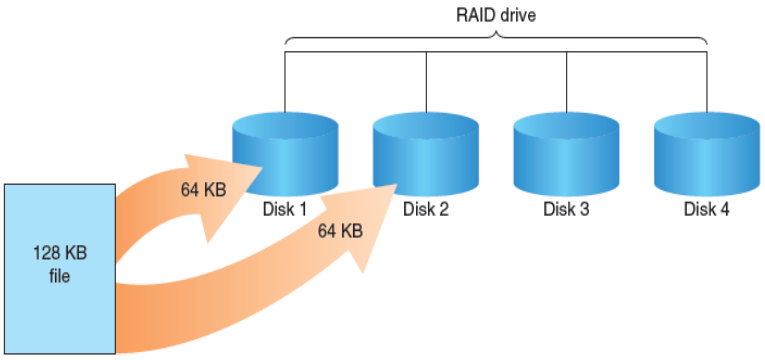
Redundant array of independent disks (RAID) is a computer configuration involving two or more physical disks. Originally, RAID was developed as a data-redundancy measure to minimize data loss caused by a disk failure. As technology improved, RAID also provided increased storage capabilities.

Several levels of RAID can be implemented through software (known as "software RAID") or special hardware controllers (known as "hardware RAID").
Software RAID is typically implemented from the host computer's OS.
Hardware RAID uses its own controller as well as a processor and memory connected to the host computer.

For Windows XP, 2000, and NT servers and workstations, RAID 0 or 1 is available.

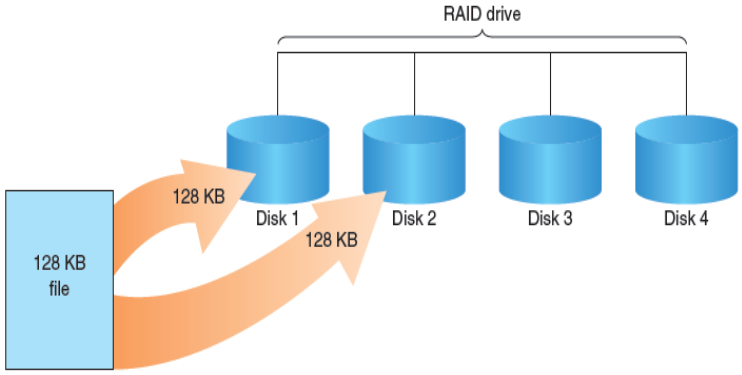
For a high-end data-processing environment, RAID 5 is common and is often based on special RAID towers.
These high-end RAID systems usually have integrated controllers that connect to high-end servers or mainframes.
These systems provide redundancy and high-speed data access and can make many small disks appear as one very large drive.

RAID 0



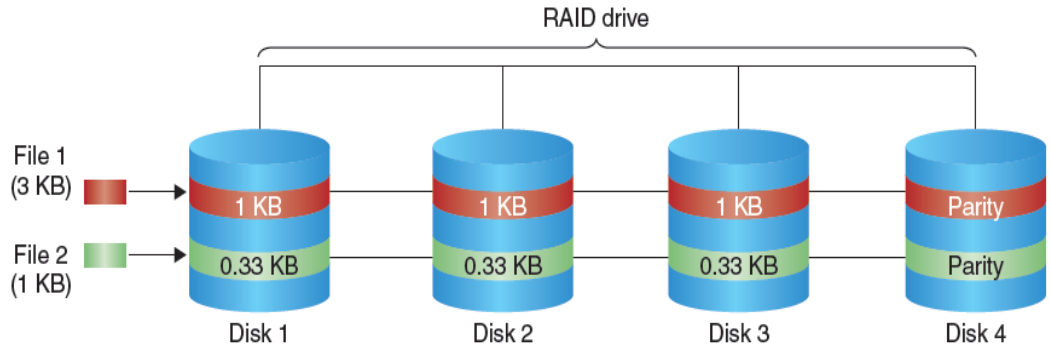
RAID 0 provides rapid access and increased data storage.
In RAID 0, two or more disk drives become one large volume, so the computer views the disks as a single disk.
The tracks of data on this mode of storage cross over to each disk.
The logical addressing scheme makes it seem as though each track of data is continuous throughout all disks.
If you have two disks configured as RAID 0, track one starts on the first physical disk and continues to the second physical disk.
When viewed from a booted OS, such as Windows XP or later, the two disks appear as one large disk.
The advantage of RAID 0 is increased speed and data storage capability spread over two or more disks that can be one large disk partition.
Its biggest disadvantage is lack of redundancy; if a disk fails, data isn't continuously available.

RAID 1



RAID 1 is made up of two disks for each volume and is designed for data recovery in the event of a disk failure.
The contents of the two disks in RAID 1 are identical.
When data is written to a volume, the OS writes the data twice—once to each disk at the same time.
If one drive fails, the OS switches to the other disk.

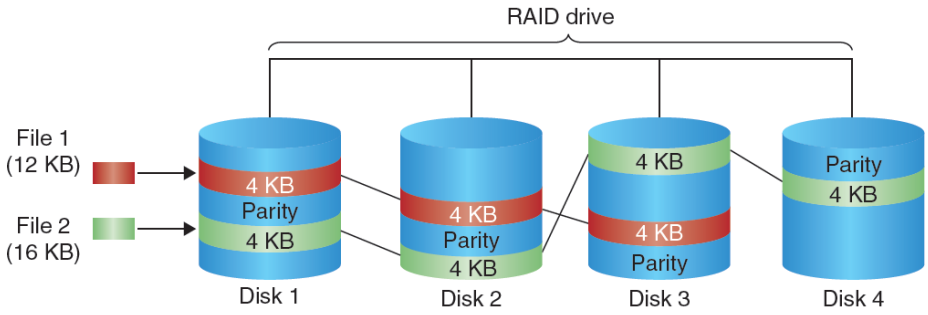
RAID 1 ensures that data isn't lost and helps prevent computer downtime.
The main disadvantage of RAID 1 is that it takes two disks for each volume, which doubles the cost of disk storage.



Like RAID 1, RAID 2 (see Figure 3-10) provides rapid access and increased storage by configuring two or more disks as one large volume.
The difference with RAID 2 is that data is written to disks on a bit level.
An error-correcting code (ECC) is used to verify whether the write is successful. RAID 2, therefore, has better data integrity checking than RAID 0.
Because of the bit-level writes and the ECC, however, RAID 2 is slower than RAID 0.

RAID 3 uses data striping and dedicated parity and requires at least three disks.
Similar to RAID 0, RAID 3 stripes tracks across all disks that make up one volume.
RAID 3 also implements dedicated parity of data to ensure recovery if data is corrupted.
Dedicated parity is stored on one disk in the RAID 3 array.
Like RAID 3, RAID 4 uses data striping and dedicated parity (block writing), except data is written in blocks rather
than bytes.

RAID 5
RAID 5: Block-level striping with distributed parity



RAID 5is similar to RAID 0 and RAID 3 in that it uses distributed data and distributed parity and stripes data tracks across all disks in the RAID array.
Unlike RAID 3, however, RAID 5 places parity data on each disk. If a disk in a RAID array has a data failure, the parity on other disks rebuilds the corrupt data automatically when the failed drive is replaced.

Acquiring RAID Disks

There's no simple method for getting an image of a RAID server's disks.
We need to address the following concerns:

- ✓ How much data storage is needed to acquire all data for a forensics image?

- ✓ What type of RAID is used? Is it Windows RAID 0 or 1 or an integrated hardwarefirmware vendor's RAID 5, 10, or 15? Is it another unknown configuration or OS?
- ✓ If it's a RAID 1, 10, or 15 server, do you need to have all drives connected so that the OS sees their contents? Some older RAID 1 systems required connecting both drives to make the data readable, which might also apply to RAID 10 and 15.
- ✓ Do you have an acquisition tool capable of copying the data correctly?
- ✓ Can the tool read a forensic copy of a RAID image?

Can the tool read split data saves of each RAID disk, and then combine all images of each disk into one RAID virtual drive for analysis?

With the larger disks now available, copying small RAID systems to one large disk is possible, similar to the way non-RAID suspect drives are copied.
For example, a small server running eight 36 GB SCSI drives in a RAID 0 tower requires about a 300 GB SATA or IDE (PATA) drive.
Less data storage is needed if a proprietary format acquisition is used with compression applied.
All forensics analysis tools can analyze an image because they see the acquired data as one large drive, not eight separate drives.
Several forensics vendors have added RAID recovery features.
These vendors typically specialize in one or two types of RAID formats.

The following are some vendors offering RAID acquisition functions:
• Guidance Software EnCase
• X-Ways Forensics
• AccessData FTK
• Runtime Software
• R-Tools Technologies
You should know which vendor supports which RAID format and keep up to date on the latest improvements in these products.

Being able to separate each physical disk into smaller save sets eliminates the need to have one large drive for storing acquired data. Acquiring RAID data requires only similar-size drives that match each disk in the RAID array.

For example, with a RAID 0 array of three 2 TB disks, all you need are three target drives of the same size.

If each acquisition is compressed, you might be able to get by with slightly smaller target drives.
Tools such as Runtime Software (*www.runtime.org)* and R-Tools Technologies (*www.r-tt.com*) are designed as data recovery tools.
Although not intended as forensics acquisition tools, they have unique features that can aid in recovering corrupted RAID
data and can perform raw format acquisitions and repair broken RAID 0 and 5 systems.

The Runtime RAID Reconstructor tool copies the original RAID to a raw format file, which must then be restored on another RAID-configured system where repairs can be performed. It also scans and corrects errors on the newly copied RAID. R-Tools R-Studio creates a virtual volume of the RAID image file.
All repairs are made on the virtual volume, which can then be restored to the original RAID.

Acquiring Data with a Linux Boot CD

The Linux OS has many features that are applicable to digital forensics, especially data acquisitions. One unique feature of older Linux versions is that it can access a drive that isn't mounted. Physical access for the purpose of reading data can be done on a connected media device, such as a disk drive, a USB drive, or other storage devices.

In Windows OSs and newer Linux kernels, when you connect a drive via USB, FireWire, external SATA, or even internal PATA or SATA controllers, both OSs automatically mount and access the drive.

On Windows drives, an acquisition workstation can access and alter data in the Recycle Bin; on Linux drives, the workstation most likely alters metadata, such as mount point configurations for an Ext3 or later drive. If you need to acquire a USB drive that doesn't have a write-lock switch, use one of the forensic Linux Live CDs (discussed in the next section) to access the device.

Using Linux Live CD Distributions

Several Linux distributions, such as Ubuntu, openSUSE, Arch Linux, Fedora, and Slackware, provide ISO images that can be burned to a CD or DVD.

A few Linux ISO images are designed specifically for digital forensics. These images contain additional utilities that aren't typically installed in normal Linux distributions. They're also configured not to mount, or to mount as read-only, any connected storage media, such as USB drives.

This feature protects the media's integrity for the purpose of acquiring and analyzing data. To access media, you have to give specific instructions to the Live CD boot session through a GUI utility or a shell command prompt.

Linux can read data from a physical device without having to mount it. As a usual practice, don't mount a suspect media device as a precaution against any writes to it.