

Nuix Workshop - Introduction to Forensics

eDISCOVERY
INFORMATION GOVERNANCE
INVESTIGATION

nuix
outperforms



Simple. Powerful. Precise.

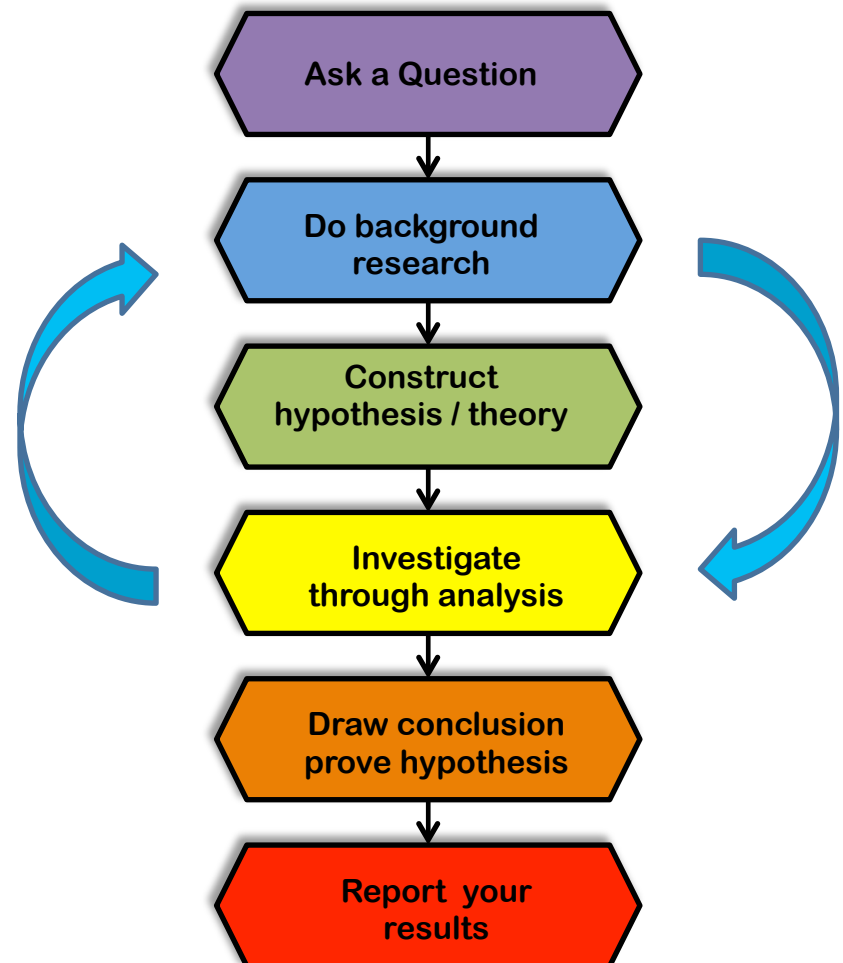
Computer forensics, also called cyber forensics, is the application of scientific method to computer investigation and analysis techniques in order to gather evidence suitable for presentation in a court of law or legal body. The goal of computer forensics is to perform a structured investigation while maintaining a documented chain of evidence to find out exactly what happened on a computer and who was responsible for it.

Computer forensics has become its own area of scientific expertise, with accompanying coursework and certification.

Computer forensics is a science and therefore requires **Scientific Method**....

- The **scientific method** is a recognised body of techniques for investigating incident or occurrence, acquiring new knowledge, or correcting and integrating previous knowledge. To be termed scientific, a method of enquiry must be based on empirical and measurable evidence subject to specific principles of reasoning.
- **Scientific method** is a model applied to all areas of scientific examination. These elements are valuable to computer forensic science

The Scientific Method



Principle 1:

No action taken by law enforcement agencies or their agents should change data held on a computer or storage media which may subsequently be relied upon in court.

Principle 2:

In circumstances where a person finds it necessary to access original data held on a computer or on storage media, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.

Principle 3:

An audit trail or other record of all processes applied to computer-based electronic evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.

Principle 4:

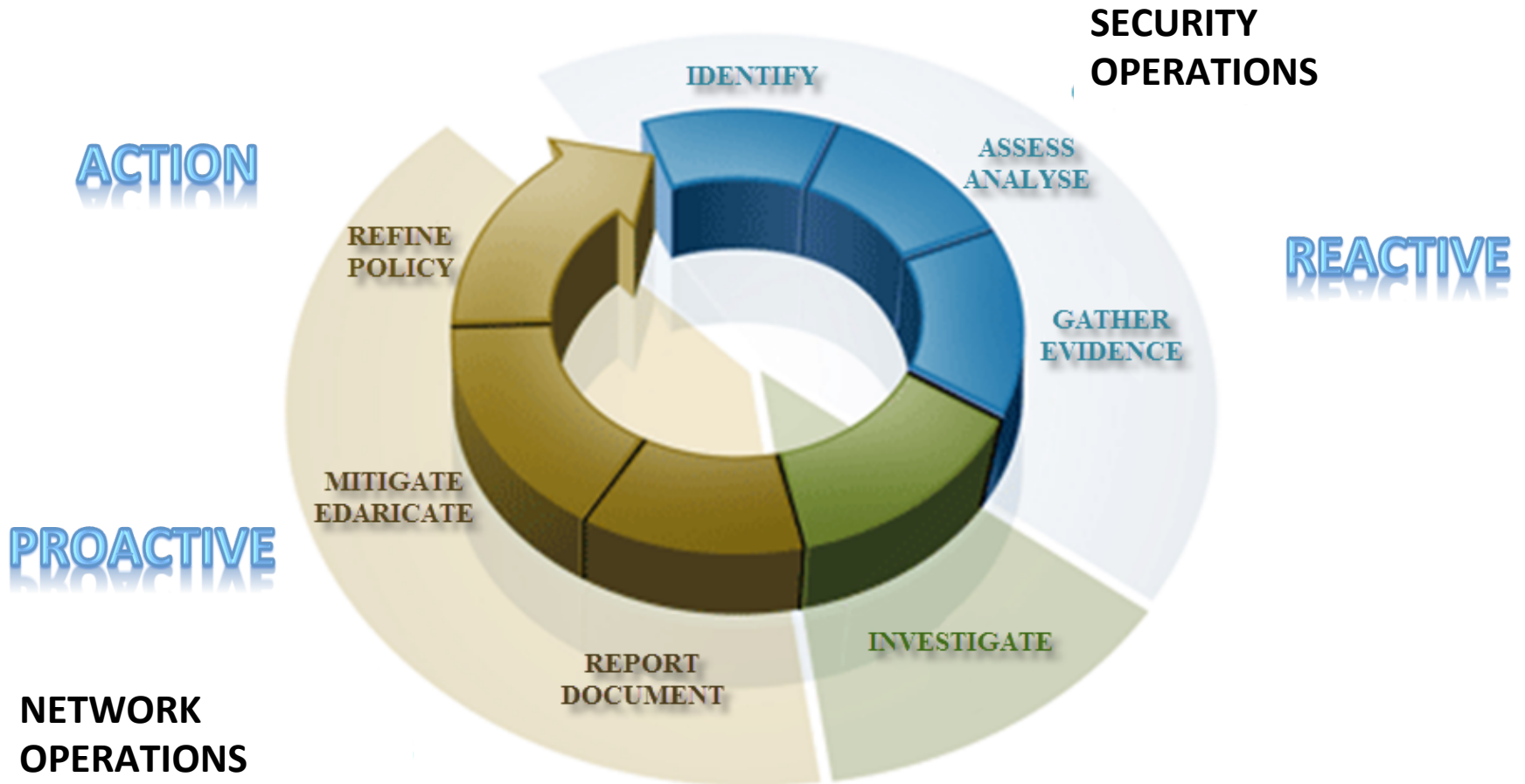
The person in charge of the investigation (the case officer) has overall responsibility for ensuring that the law and these principles are adhered to.

Source <http://www.acpo.police.uk/documents/crime/2011/201103CRIECI14.pdf>

The benefits of the application of digital forensics to computer based investigations underpin the following:

- Security of evidence / incident
- Integrity of investigative steps
- Deeper analysis unallocated space / file slack (*The whole story*)
- Auditable response
- Repeatability of action
- Best evidence practice





SECURITY OF INCIDENT - DOING NOTHING EVEN CAUSES CHANGES

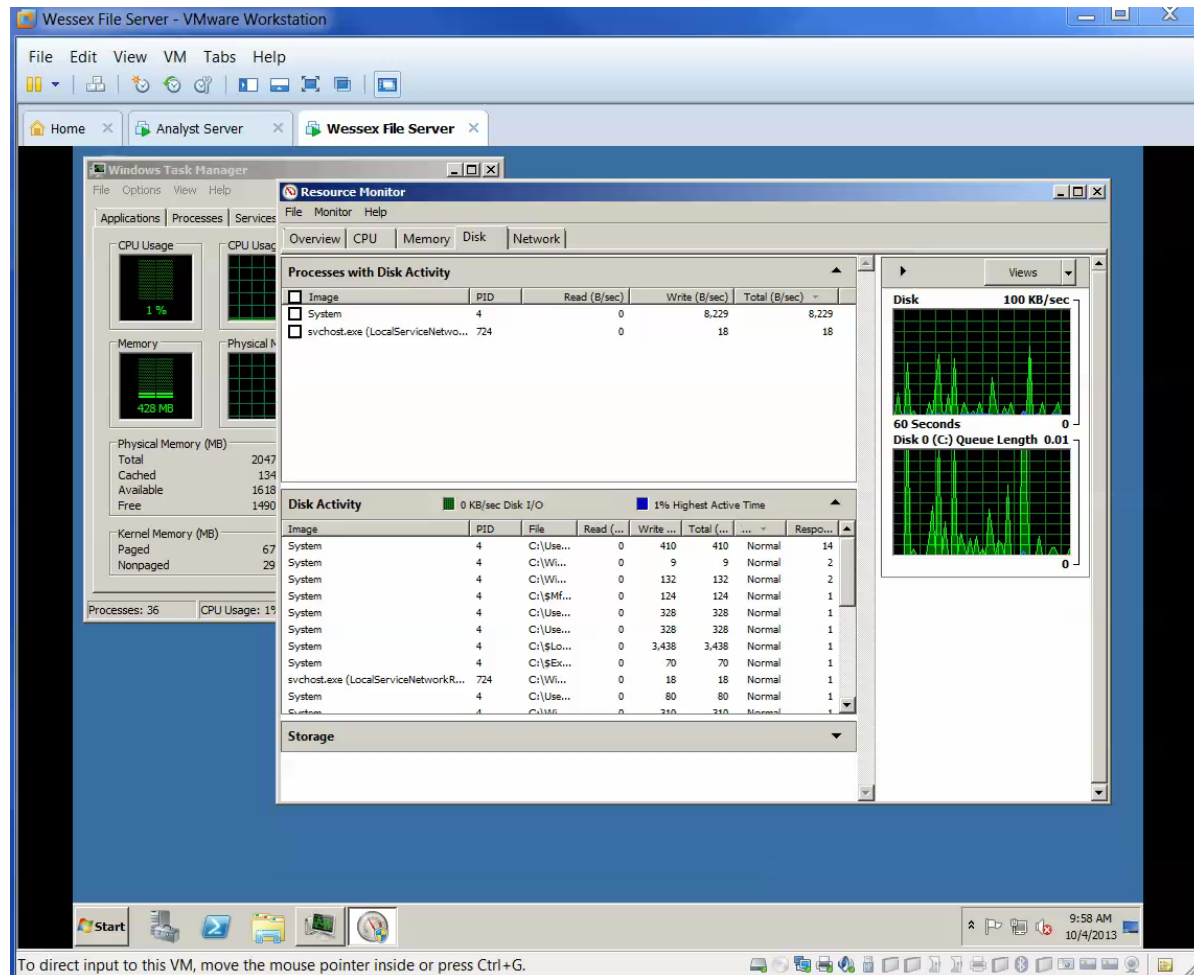


Even at rest a computer is using memory and performing disk writes. This is essential to the operating system.

The capture shows disk activity on a computer with no user activity and no applications running.

Now consider

- Malware
- Anti forensic applications
- Cluster overwrites

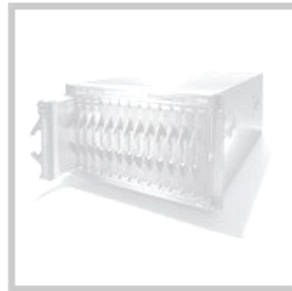
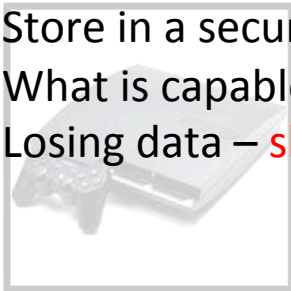
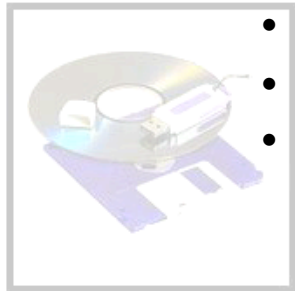


The first step in any investigation is the search & seizure of exhibits which may contain crucial evidence!

Decisions that you, make may result in loss of crucial evidence.

Points to consider

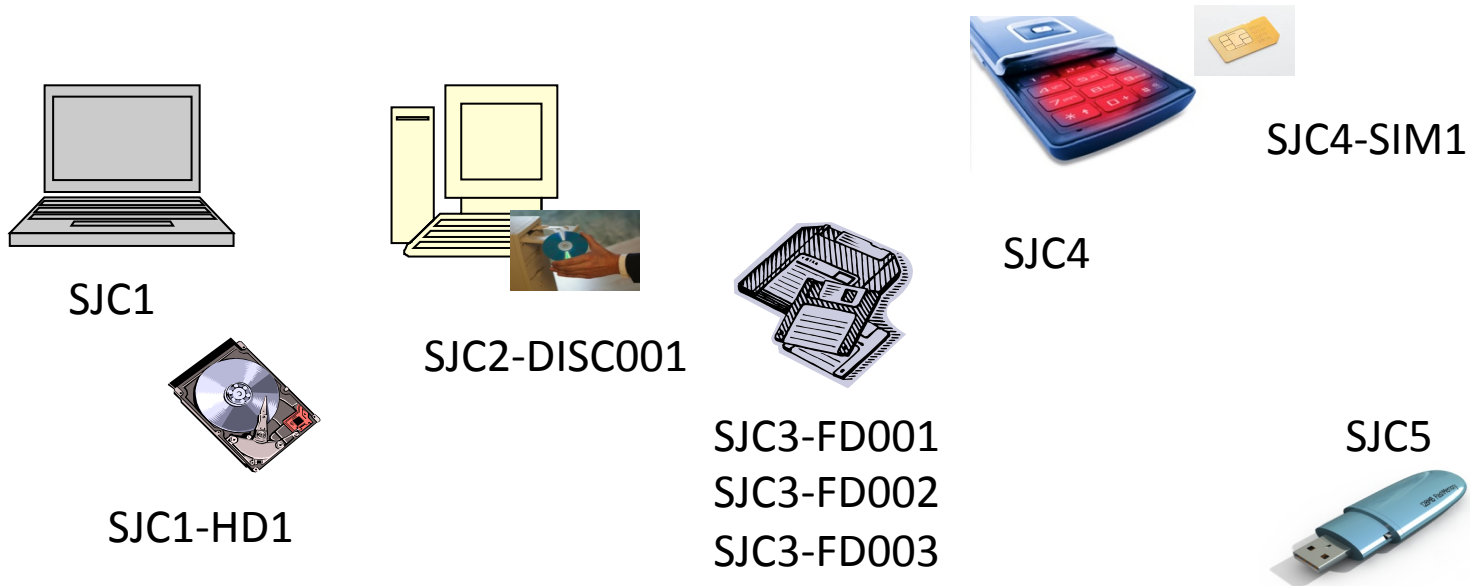
- DNA and/or fingerprints
- Prevent tampering & preserve original condition
- Record details & actions – paperwork!
- Store in a secure location
- What is capable of storing data?
- Losing data – **shutdown or not?**



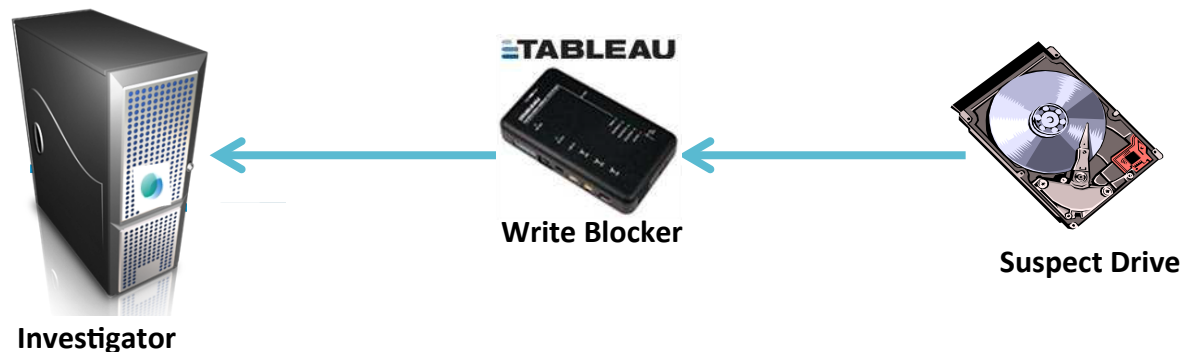
Whether we are considering logical or physical collection we must ensure that we collect data in accordance to industry guidelines and take every step to protect the data from any change due to our action. In accordance to guidelines if this is impractical we must ensure we understand the implications of our actions.

A write blocker is a hardware or software device that prevents ANY write activity to a connected device. We can then use a forensic application or DD command to collect the data into a forensic container.

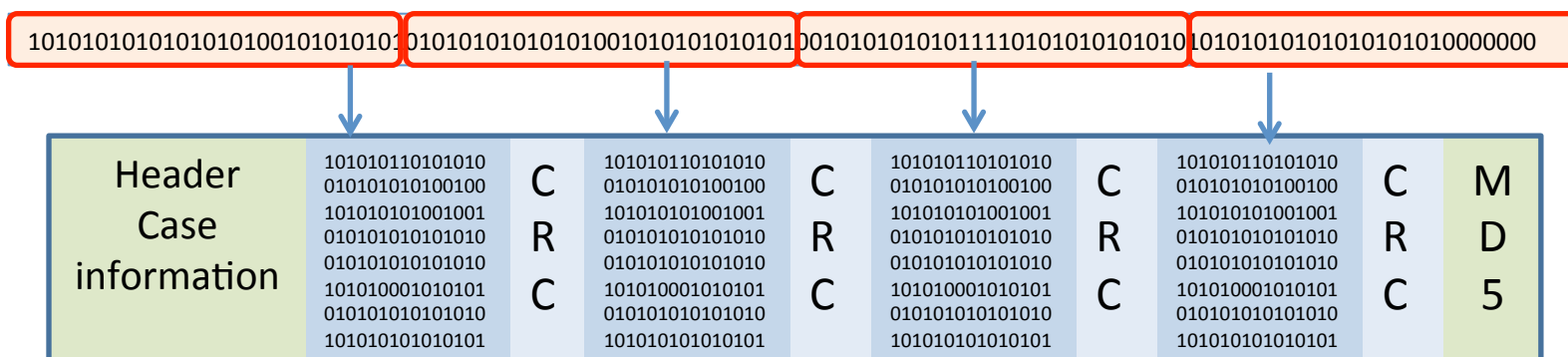




- Forensic image files are generated with specialist tools
- Are an exact 'bit for bit' acquisition of the data
- All devices should be unique referenced

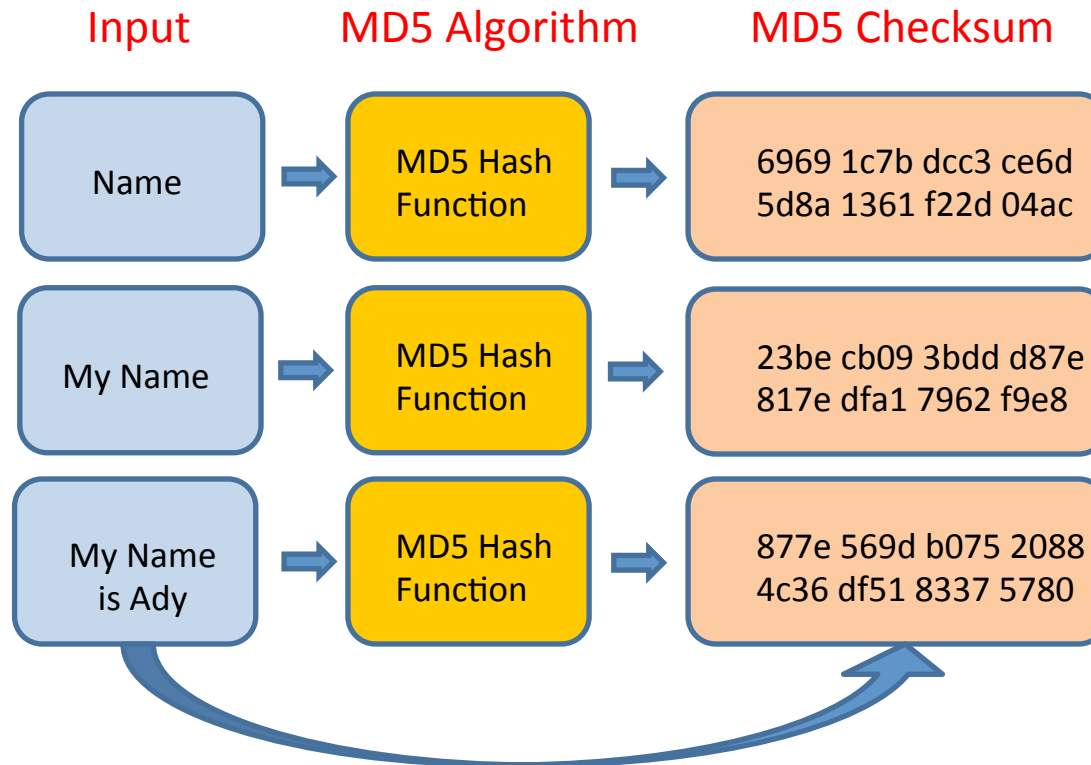


Data is collected from the source device at binary/disk level by pre defined size and each section is checked with a CRC checksum. The whole image is then verified with an MD5 checksum



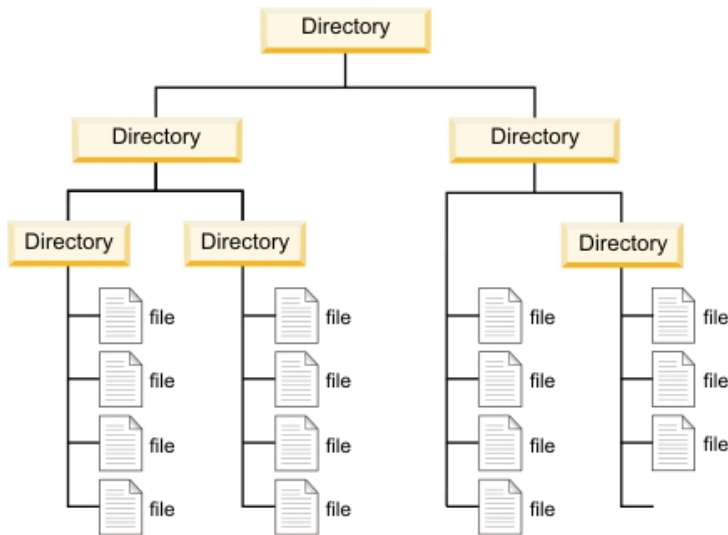
Should any single value be change then the CRC would fail and the MD5 checksum would present a different value. Therefore verification would fail and the collection process would be undermined.

MD5 is an algorithm that is used to verify data integrity through the creation of a 128-bit message digest from data input (which may be a message of any length). The result is as unique to that specific data as a fingerprint is to the specific individual.

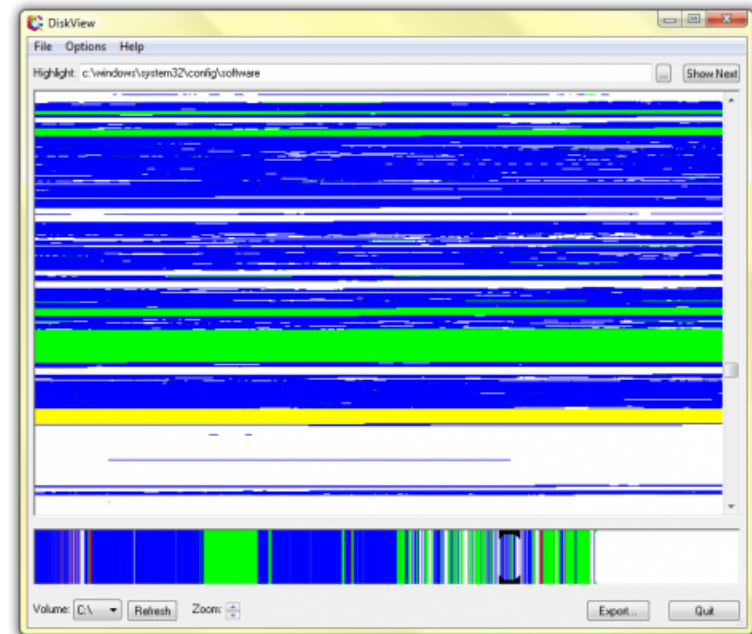


Identical data will provide identical MD5

- Investigations frequently involve large numbers of devices including multiple computers, mobile devices and a variety of digital storage media.
- Traditional methods of analysing each data repository individually are immensely time consuming and often ineffective.
- Typical collection of devices for investigation analysis
 - Suspect's personal possessions
 - Apple Mac book Laptop (HFS+)
 - Apple iPhone (iOS)
 - External Hard Drive
 - Company/Employer data relating to suspect
 - Microsoft Windows Desktop PC
 - Microsoft Exchange Mailbox
 - Folder and files stored on a Windows Network share
 - RIM Blackberry mobile phone
- Nuix is engineered to triage, process, analyze and bring to the surface critical evidence from entire data sets.
- This saves time and effort, freeing investigators to test hypotheses, follow evidence trails and find links between suspects.

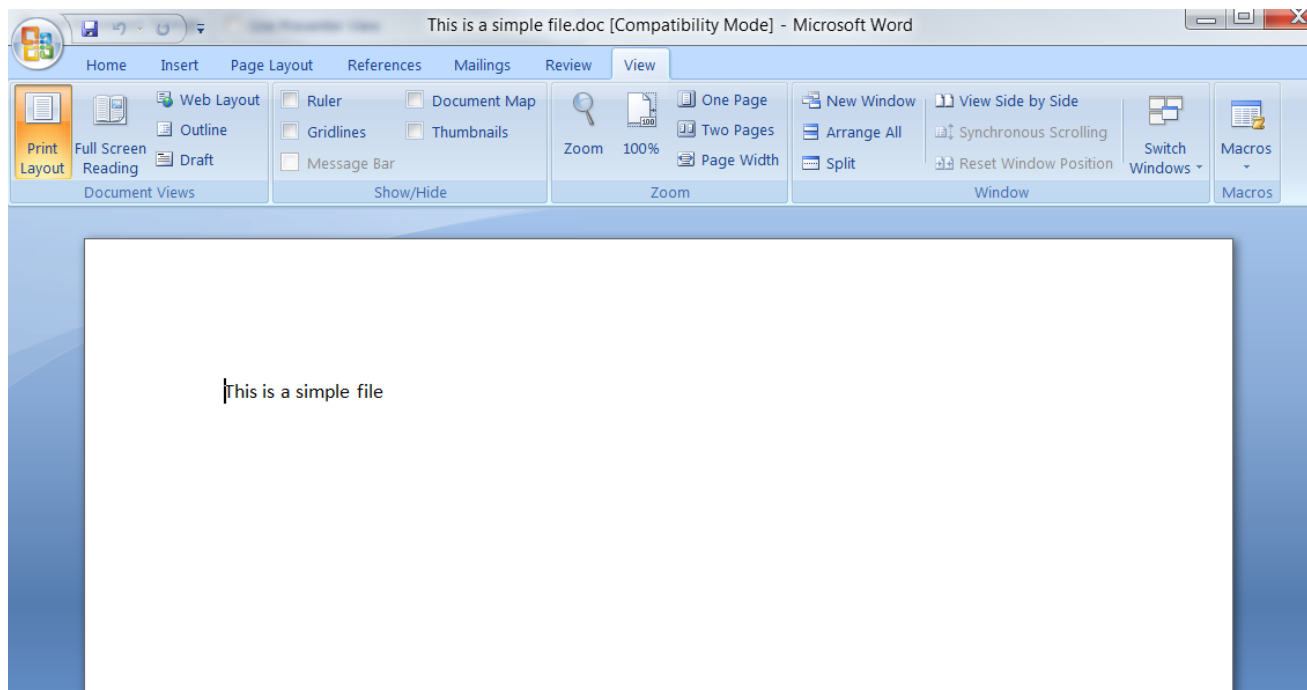


Logical application only allows an analyst to investigate live files and folders and whilst investigation is undertaken important attributes are changing



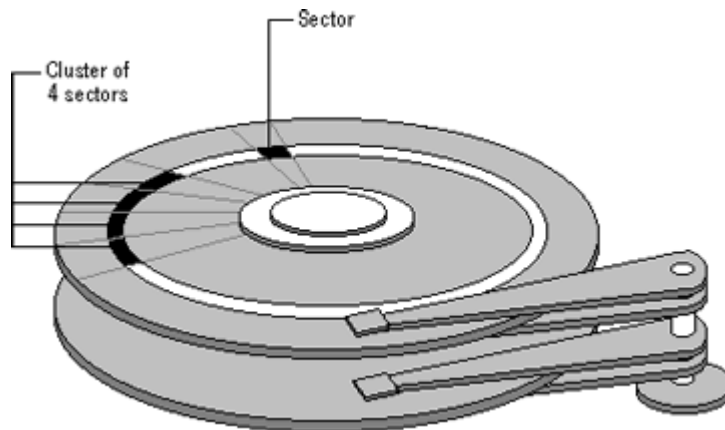
This Sysinternals utility will make a graphical map of a hard drive - we can see all clusters and view information about every single one of them. Capturing the data at disk level in a forensic container ensures no changes can be made to the data

Lets take a look at a simple word document. From a logical view we can see the content and some simple meta data

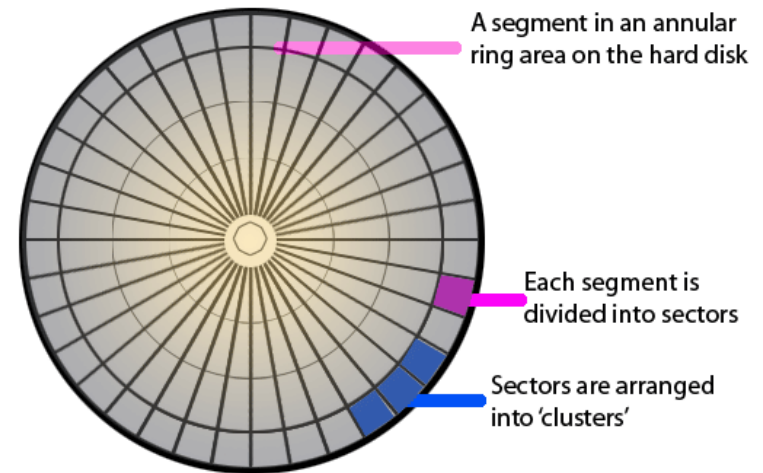


Now lets take a look at the document from a forensic image

To gain a better understanding of how data is recovered we must appreciate how data is stored and managed by an operating system.

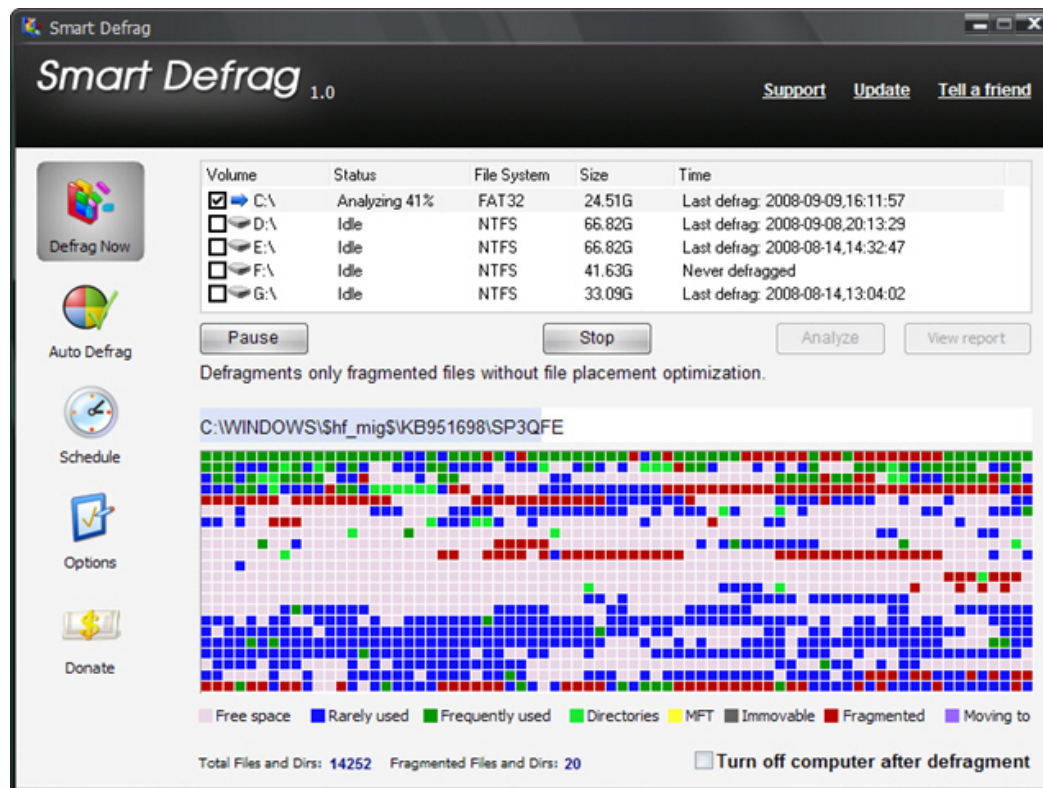


Hard disk format



The above example uses a traditional single HDD however the same principle applies to other data medium e.g. USB, Solid State and RAID configuration.

This is sometimes easier to represent and more familiar when we use an application to show the fragmentation of files across the hard drives

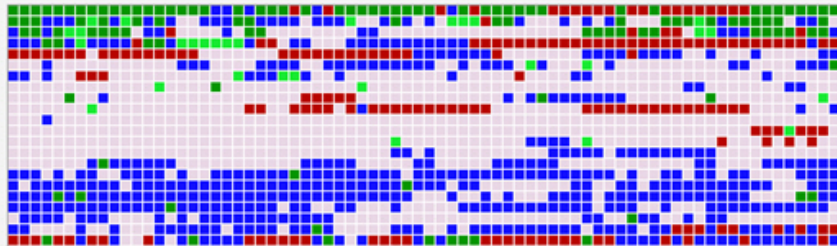


Smart Defrag 1.0 [Support](#) [Update](#) [Tell a friend](#)

Volume	Status	File System	Size	Time
<input checked="" type="checkbox"/> C:\	Analyzing 41%	FAT32	24.51G	Last defrag: 2008-09-09,16:11:57
<input type="checkbox"/> D:\	Idle	NTFS	66.82G	Last defrag: 2008-09-08,20:13:29
<input type="checkbox"/> E:\	Idle	NTFS	66.82G	Last defrag: 2008-08-14,14:32:47
<input type="checkbox"/> F:\	Idle	NTFS	41.63G	Never defragged
<input type="checkbox"/> G:\	Idle	NTFS	33.09G	Last defrag: 2008-08-14,13:04:02

Defragments only fragmented files without file placement optimization.

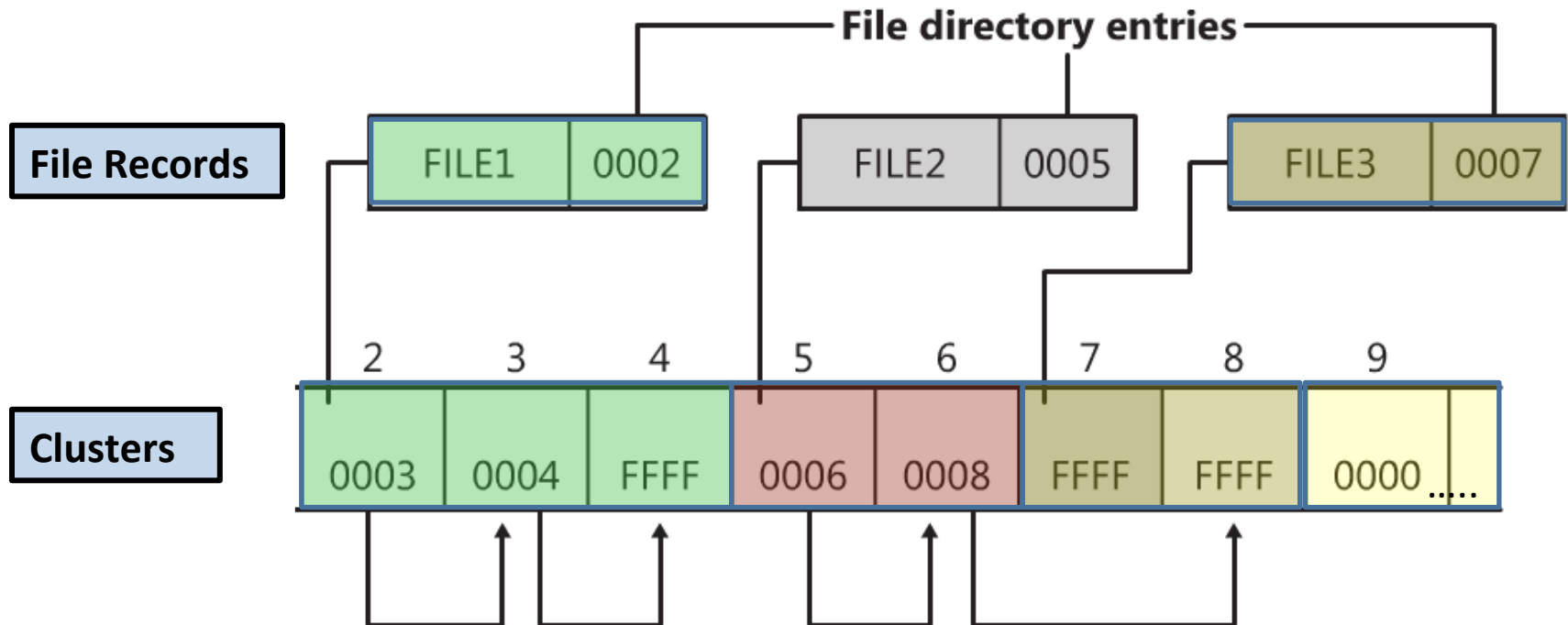
C:\WINDOWS\Shf_mig\$KB951698\SP3QFE



Free space Rarely used Frequently used Directories MFT Immovable Fragmented Moving to

Total Files and Dirs: 14252 Fragmented Files and Dirs: 20 Turn off computer after defragment

The allocation of the sectors and cluster is managed by the operating system. On FAT it is the File Allocation Table and NTFS is the Master File Table. The system records much information about the files it is storing in these tables and this is referred to as Meta Data. The table records the whereabouts of all files on a system and also which clusters are available for future use.



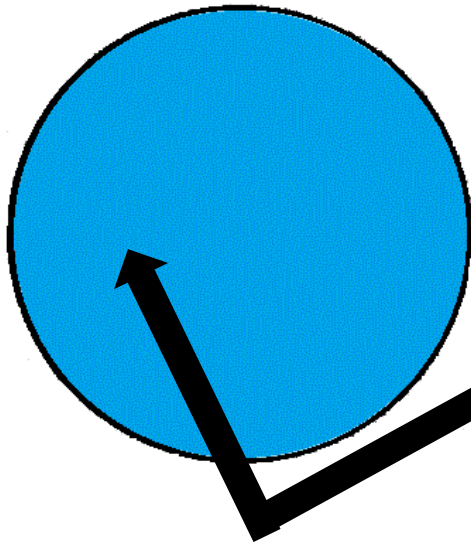
When files are erased or deleted the content of the file is not actually erased. Unless security grade file deletion software is used data from the 'erased file' remains behind in an area called unallocated storage space. The same is true concerning file slack that may have been attached to the file before it was deleted. As a result, the data remains behind for discovery through the use of data recovery and/or computer forensics software utilities.

Unallocated file space and file slack are both important sources of leads for the computer forensics investigator.

Until the first file is written to the data storage area of a computer storage device, the clusters are unallocated by the operating. As files are created by the computer user, clusters are allocated in the file table to store the data. When the file is 'deleted' by the computer user, the clusters allocated to the file are released by the operating system so new files and data can be stored in the clusters when needed. However, the data associated with the 'deleted' file remains behind. This data storage area is referred to as unallocated storage space and it is fragile from an evidence preservation standpoint. However, until the unallocated storage space is reassigned by the operating system, the data remains behind for discovery and extraction by the computer forensics specialist.

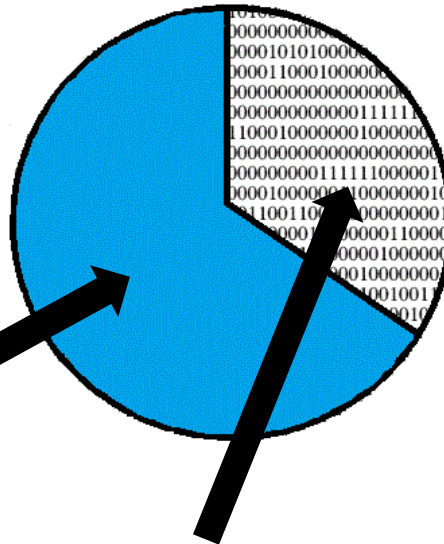
As data is deleted through system or user activity then more and more data becomes recoverable from unallocated sectors

Unused Drive



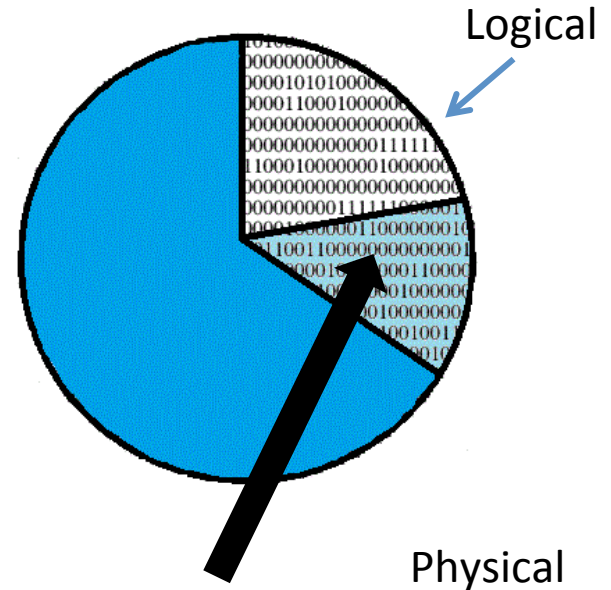
Unused Disk Space (Zero Sectors)

Data Stored



Allocated Sectors

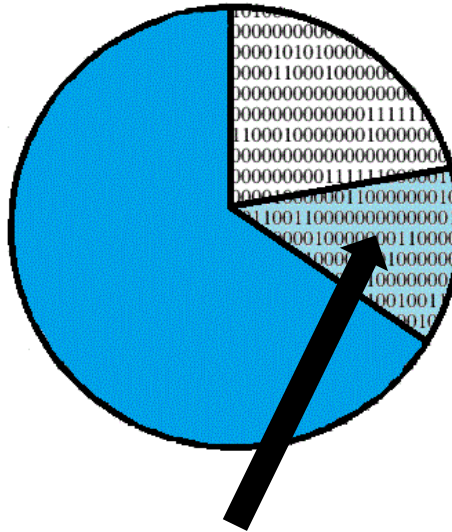
Residual Data



Residual Data in Unallocated Sectors

Carve file system unallocated space

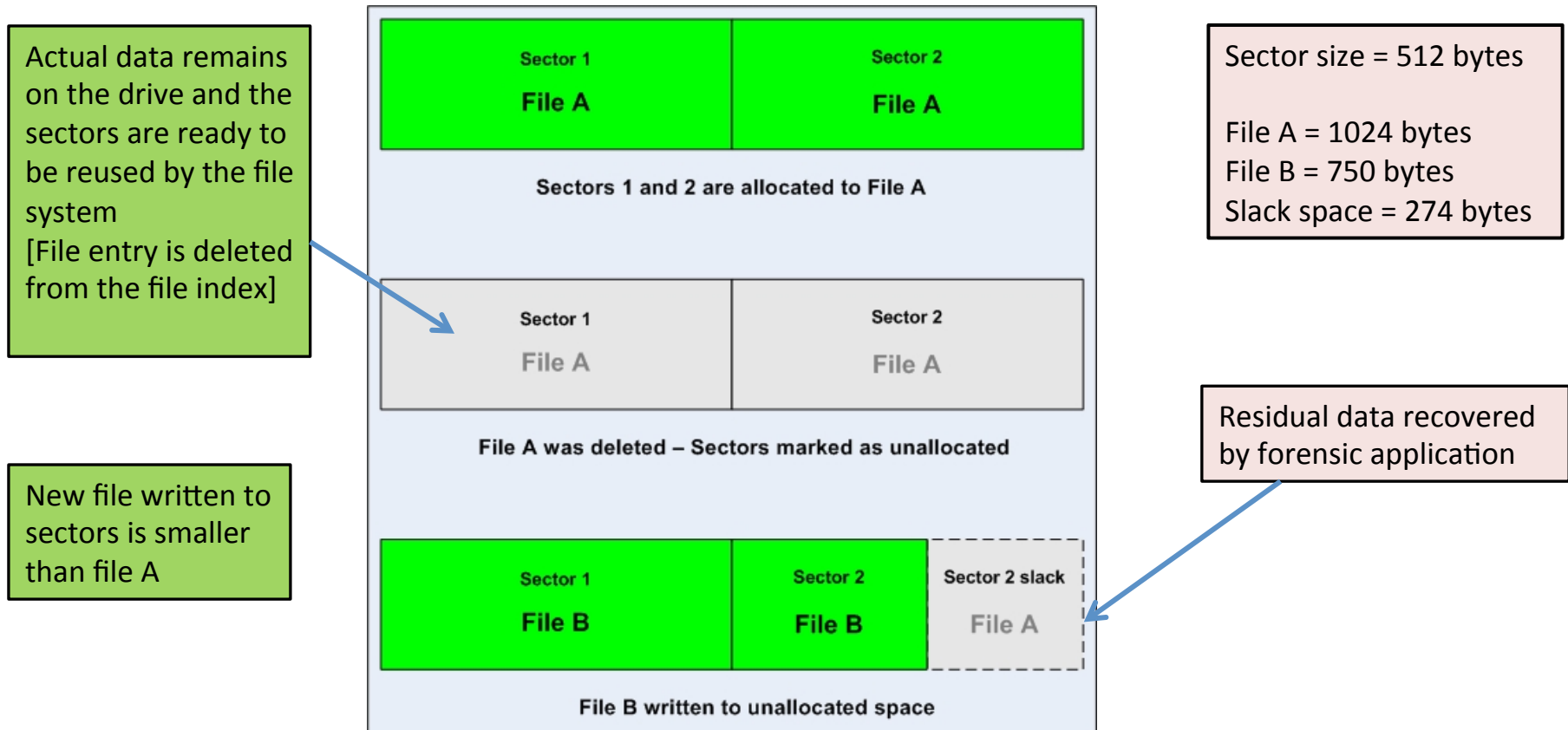
Data carving, or file carving is a process of reading files without reference to a file system. The technique can be applied to any type of disk that stores data on sector boundaries which includes camera memory, USB devices as well as hard drives. It is based on the fact that most files start with a recognisable data signature



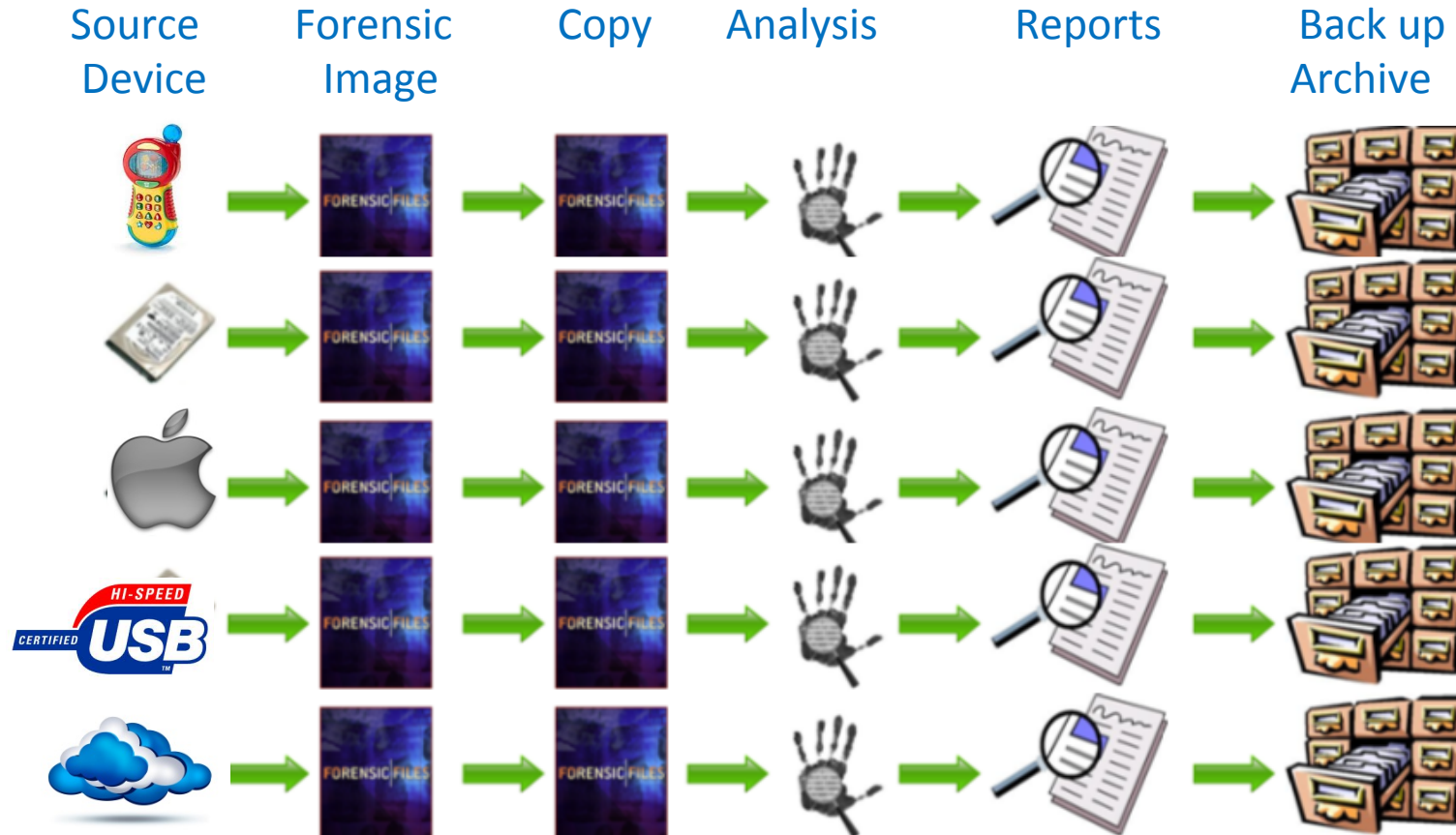
Unallocated Sectors with residual data

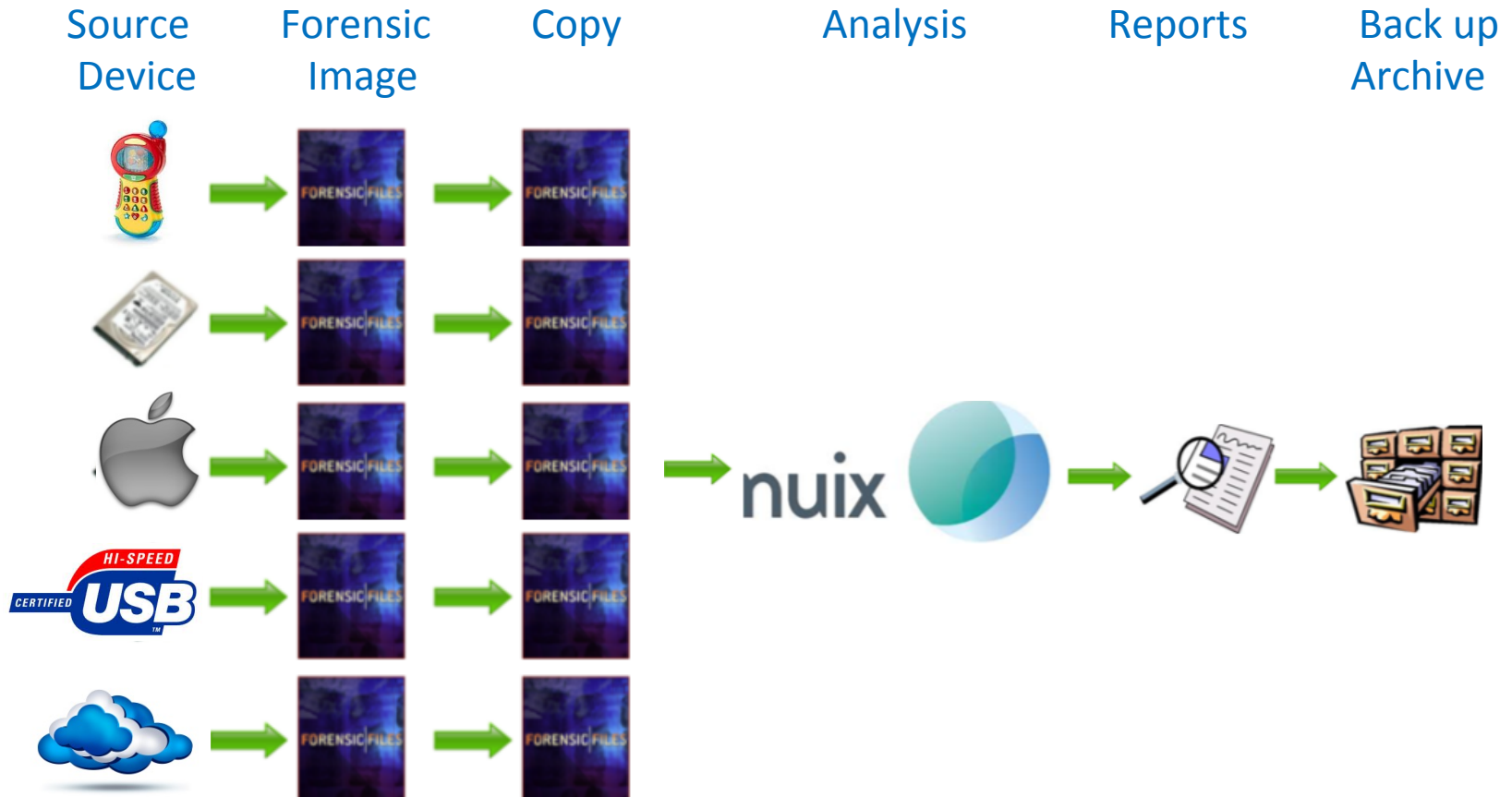
Extract end of-file slack space from disk images

The unused space in a disk cluster. The DOS and Windows file systems use fixed-size clusters. Even if the actual data being stored requires less storage than the cluster size, an entire cluster is reserved for the file. The unused space is called the *slack space*.



CURRENT APPROACH TO FORENSIC INVESTIGATION





Offering a scalable collaborative analysis solution saving time, storage and money

More power, more precision, more speed

- ❑ Speed – The ability to react
 - Upwards of 1 TB of data on a single server in 24 hours

- ❑ Breadth of supported file types – Save the sorting for later
 - As data quantities rise so does the range of file formats encountered for analysis.

- ❑ Indexing – Get to the detail
 - Fully double-byte Unicode compliant – Search in any language, search both file content and metadata, search for special characters

- ❑ Ease of deployment – Install and go
 - Download and install in under 5 minutes

- ❑ Ease of use – Quickly realise the value
 - Within minutes, identify IP leaks, transmission of data outside the organization, or run 1000's of queries in an automated fashion

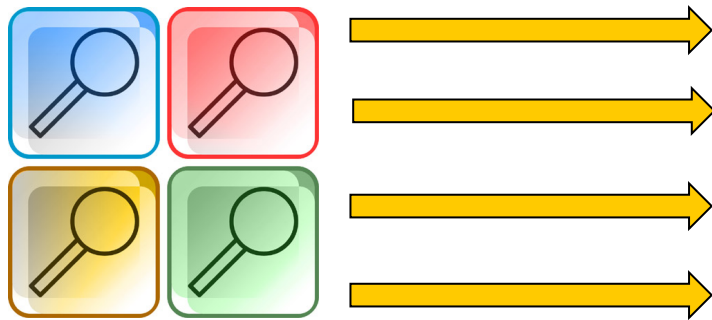
NUIX SEARCHABLE INDEXES

As Nunix processes data it extracts valuable information and places the data into separate searchable indexes which can be searched against in whole or individually.

Powerful dynamic component that allows investigators to be flexible and intuitive in the approach to data management from ECA through analysis.

Allows for the application of specific function to relevant data throughout workflow.

Enable investigators to quickly target and hydrate function to relevant material through ECA and NVA.




The database architecture of Nuix offers the investigator powerful options in order to get to relevant data very quickly and decreasing false hits from search criteria. We can draw comparison to a well known entity that we use every day - Google

QuickTime Player File Edit View Window Help

Google

https://www.google.co.uk

+You Gmail Images  [Sign In](#)



[Google Search](#) [I'm Feeling Lucky](#)

New! Wander through the streets and canals of Venice with [Google Maps](#)

As previously discussed Nuix extracts values of credit cards, money values, IP addresses, emails, company names and countries – these are referred to Extracted Entities. Nuix uses a method of searching for these values by pre defined regular expression.

In computing, a **regular expression** (abbreviated **regex** or **regexp**) is a sequence of characters that form a search pattern.

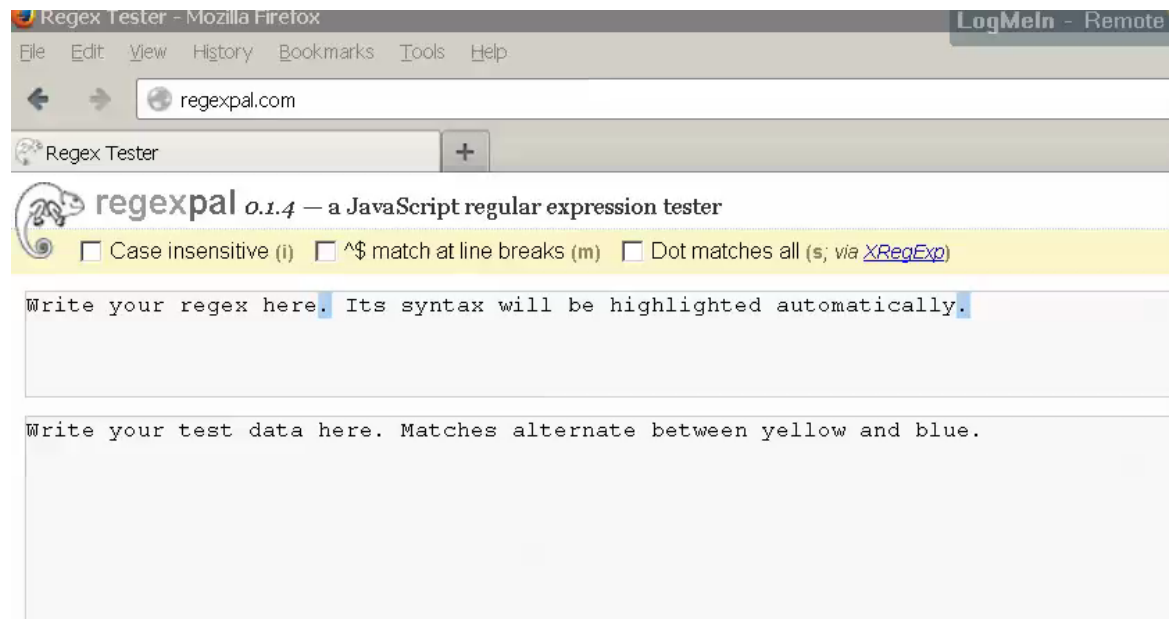
Unlike MD5 which requires exact matching regular expression allows for rules of sequence to be applied into a search string that can be used to match any entities that meet that rule.

As long as the entity meets a predefined rule then we can use this method to find data

This is a simple example of how the use of Regular expression can locate data. All Mastercard numbers start with a 51,52,53,54, or 55 followed by 14 numbers. This rule can be applied to a regular expression as follows:

```
^([51|52|53|54|55]{2})([0-9]{14})$
```

The number must start with 2 digits that are defined in the expression followed by 14 digits that are between 0 and 9.



- **Typical collection of devices for investigation analysis**

- Suspect's personal possessions

- Apple Macbook / Laptop (PC/HFS+)
 - Mobile device Apple iPhone (iOS)
 - External Hard Drive

- Company/Employer data relating to suspect

- Microsoft Windows Desktop PC
 - Microsoft Exchange Mailbox
 - Folder and files stored on a Windows Network share
 - RIM Blackberry mobile phone
 - USB Devices - Media

Extended range of supported file formats

- Forensic Images
- Cloud based email
- HFS, HFS+ Filesystems
- Cellebrite mobile phone images
- XRY Mobile device files
- Mobile phone backup files
- Apple plist files
- Apple iWork files types
- Apple iPhone/iPad images (iOS)
- SQLite files
- Entire Exchange databases
- Outlook mailstore containers
- Lotus Notes mailstore container



Nux addresses all of the topics we have discussed along with many more and automates them into its process.

Allows the user to feel assured that forensic integrity is maintained and data is presented in a format that is ready to be immediately searched and investigated.

Lets take a look!