ARIZONA STATE UNIVERSITY

CSE 469: Computer and Network Forensics

Topic 9: Semester Review

Dr. Mike Mabey | Spring 2019

ARIZONA STATE UNIVERSITY

Review:

Topic 1: Forensics Intro

Digital Forensics: Basics

What is Computer Crime?

- A crime in which technology plays an important, and often a necessary, part.
- What about the computer?
  - the tool used in an attack
  - the target of an attack
  - used to store data related to criminal activity
- **3 generic categories**
  - Computer assisted
    - e.g., fraud, child pornography
  - Computer specific or targeted
    - e.g., denial of service, sniffers, unauthorized access
  - Computer incidental
    - e.g., customer lists for traffickers

Digital Forensics: Objectives (1)

- Digital forensics involves <u>data</u> retrieved from a suspect's:
  - Hard drive
  - Other storage media also:
    - Cell phones
    - Flash drives
    - Cloud services
    - Cars
    - Thermostats
    - Smart speakers

NOTE: The data might be

- Hidden
- Encrypted
- Fragmented
- Deleted
- Outside the normal file structure

Digital Forensics: Objectives (2)

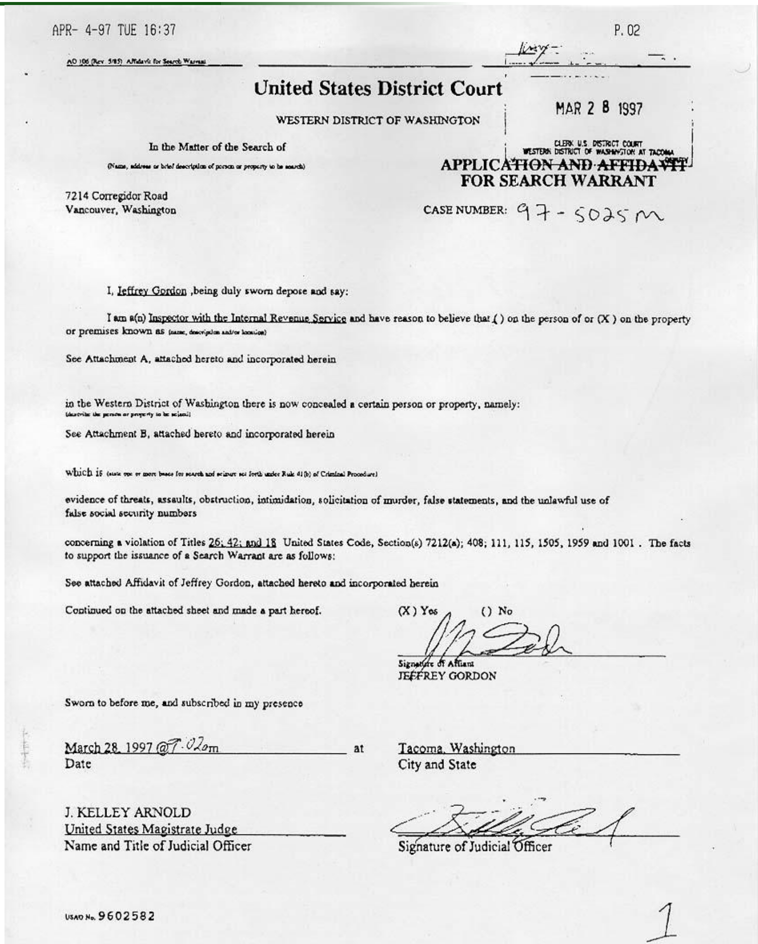
- Figure out *what* happened, *when*, and *who* was responsible.
- Computer forensics is a discipline dedicated to the collection of computer evidence for judicial purposes.
  - Source: EnCase Legal Journal
- Computer forensics involves the preservation, identification, extraction, documentation and interpretation of computer data.
  - Source: Kruse and Heiser, Computer Forensics Incident Response Essentials
- Must be able to show proof

Understanding Digital Forensics

- Digital forensics involves:
  a. Obtaining and analyzing
  b. digital information
  c. for use as evidence
  d. in civil, criminal, or administrative cases.
- Critical condition:
  a. Obtaining evidence covered by the **Fourth Amendment to the U.S. Constitution**
  b. **Protects everyone's rights** to be secure in their person, residence, and property **from search and seizure**.

Fourth Amendment

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.


APR- 4-97 TUE 16:37 P. 02

AO 106 (Rev. 5/85) Affidavit for Search Warrant

United States District Court

WESTERN DISTRICT OF WASHINGTON

MAR 2 8 1997

CLERK U.S. DISTRICT COURT
WESTERN DISTRICT OF WASHINGTON AT TACOMA

In the Matter of the Search of
(Name, address or brief description of person or property to be searched)

7214 Corregidor Road
Vancouver, Washington

APPLICATION AND AFFIDAVIT
FOR SEARCH WARRANT

CASE NUMBER: 97-5025M

I, Jeffrey Gordon ,being duly sworn depose and say:

I am a(n) Inspector with the Internal Revenue Service and have reason to believe that ( ) on the person of or (X) on the property or premises known as (name, description and/or location)

See Attachment A, attached hereto and incorporated herein

in the Western District of Washington there is now concealed a certain person or property, namely:
(describe the person or property to be seized)

See Attachment B, attached hereto and incorporated herein

which is (state one or more bases for search and seizure set forth under Rule 41(b) of Criminal Procedure)

evidence of threats, assaults, obstruction, intimidation, solicitation of murder, false statements, and the unlawful use of false social security numbers

concerning a violation of Titles 26; 42; and 18 United States Code, Section(s) 7212(a); 408; 111, 115, 1505, 1959 and 1001 . The facts to support the issuance of a Search Warrant are as follows:

See attached Affidavit of Jeffrey Gordon, attached hereto and incorporated herein

Continued on the attached sheet and made a part hereof. (X) Yes ( ) No

Signature of Affiant
JEFFREY GORDON

Sworn to before me, and subscribed in my presence

March 28, 1997 @7:02pm at Tacoma, Washington
Date City and State

J. KELLEY ARNOLD
United States Magistrate Judge
Name and Title of Judicial Officer

Signature of Judicial Officer

USAO No. 9602582

1

Digital Forensics vs Data Recovery

- Data recovery
  - Retrieving data accidentally deleted
  - Damaged or destroyed (fire, power failure, etc.)
  - User WANTS it back

- Digital forensics
  - Retrieving data the user *deliberately obscured*
  - User DOESN'T want it back

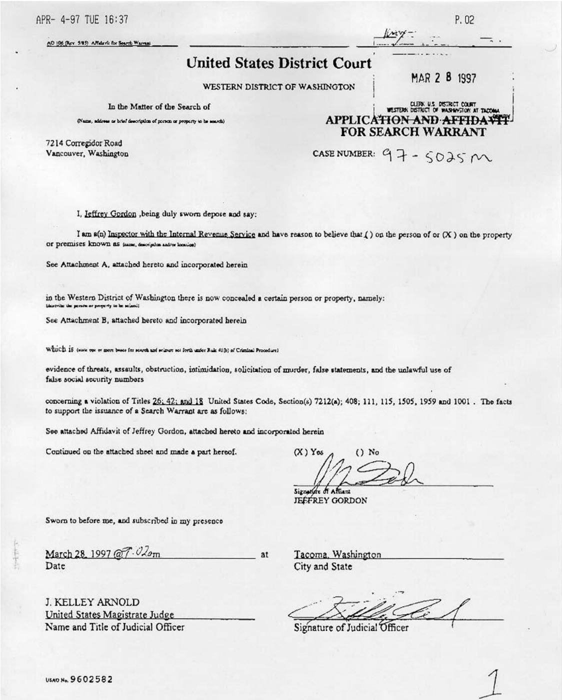
Public vs Private Sector Investigations

Need to Know

- File system and operating system
  - How a PC saves a file to disk
  - What happens when you delete a file?
    - Data is not changed
    - OS indicates that clusters used by the file are available for reuse
- Understanding Data
  - Hex editor
  - Binary analysis
- Basic OS-level commands are useful and critical

Public Investigations

- **Government agencies** are responsible for criminal investigations and prosecution.
- The law of search and seizure protects the rights of all people, including people suspected of crimes.



APR- 4-97 TUE 16:37 P. 02

AO 106 (Rev. 5/85) Affidavit for Search Warrant

United States District Court

WESTERN DISTRICT OF WASHINGTON

MAR 2 8 1997

CLERK U.S. DISTRICT COURT
WESTERN DISTRICT OF WASHINGTON AT TACOMA
DEPUTY

In the Matter of the Search of
(Name, address or brief description of person or property to be search)

7214 Corregidor Road
Vancouver, Washington

APPLICATION AND AFFIDAVIT FOR SEARCH WARRANT

CASE NUMBER: 97-5025M

I, Jeffrey Gordon ,being duly sworn depose and say:

I am a(n) Inspector with the Internal Revenue Service and have reason to believe that ( ) on the person of or (X) on the property or premises known as (name, description and/or location)

See Attachment A, attached hereto and incorporated herein

in the Western District of Washington there is now concealed a certain person or property, namely:
(describe the person or property to be seized)

See Attachment B, attached hereto and incorporated herein

which is (state one or more bases for search and seizure set forth under Rule 41(b) of Criminal Procedure)

evidence of threats, assaults, obstruction, intimidation, solicitation of murder, false statements, and the unlawful use of false social security numbers

concerning a violation of Titles 26; 42; and 18 United States Code, Section(s) 7212(a); 408; 111, 115, 1505, 1959 and 1001 . The facts to support the issuance of a Search Warrant are as follows:

See attached Affidavit of Jeffrey Gordon, attached hereto and incorporated herein

Continued on the attached sheet and made a part hereof. (X) Yes ( ) No

Signature of Affiant
JEFFREY GORDON

Sworn to before me, and subscribed in my presence

March 28, 1997 @9:02am at Tacoma, Washington
Date City and State

J. KELLEY ARNOLD
United States Magistrate Judge
Name and Title of Judicial Officer

Signature of Judicial Officer

USAO No. 9602582

1

Public Investigations

- Public investigation == Law enforcement agency investigation
  - Need to understand laws on computer-related crimes: local city, county, tribal, state/province, and federal.
  - Understand the standard legal process.
  - How to build a criminal case.

Private Sector Investigations

- Deals with private organizations are not governed directly by criminal law or the Fourth Amendment...
- But by **internal policies** that define expected employee behavior and conduct in the workplace.

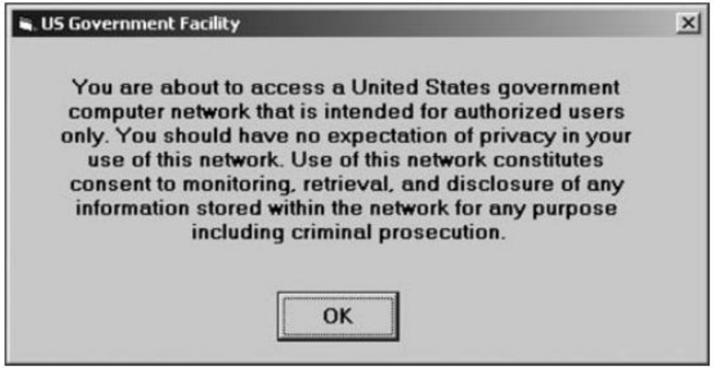
- Private investigations are usually conducted in civil cases...
- However, a civil case can escalate into a criminal case...
- And a criminal case can be reduced to a civil case.

Private Sector Investigations

- Guiding principle:
  - Business must continue with minimal interruption from the investigation.
- Corporate computer crime examples:
  - Email-harassment
  - Falsification of data
  - Gender/age/… discrimination
  - Embezzlement
  - Industrial espionage

Organizations' Responsibilities

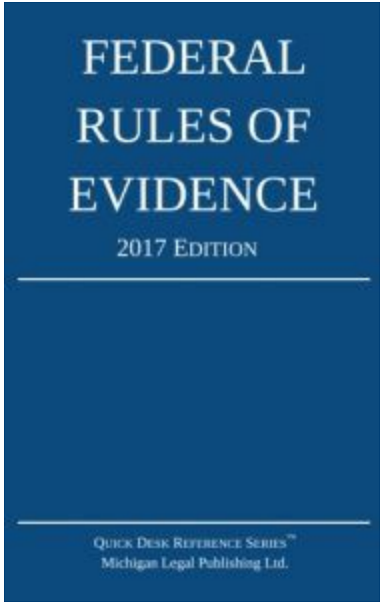
- Organizations must help prevent and address computer crime by:
  - Establishing company policies for acceptable use of systems.
    - Bring your own device (BYOD)
  - Clearly defining what distinguishes private property and company property.
  - Display warning banners.



Rules of Evidence

Rules of Evidence

- Authenticity
- Admissibility
- Completeness
- Reliability / Accuracy



Rules of Evidence: Authenticity

- Can we explicitly link files, data to specific individuals and events?
- Typically uses:
  - Access control
  - Logging, audit logs
  - Collateral evidence
  - Crypto-based authentication
    - Non-repudiation

Rules of Evidence: Admissibility

- Legal rules which determine whether potential evidence can be considered by a court.
  - Common / civil code traditions
  - Adversarial / inquisitorial trials
  - “Proving” documents, copies
- US: 4th amendment rights / Federal Rules of Evidence
- UK: PACE, 1984; “business records” (s 24 CJA, 1988) etc

Rules of Evidence: Completeness

- Evidence must tell a complete narrative of a set of particular circumstances, setting the context for the events being examined so as to avoid "any confusion or wrongful impression."
- If an adverse party feels evidence lacks completeness, they may require introduction of additional evidence "to be considered contemporaneously with the [evidence] originally introduced."
  - Wex Legal Dictionary / Encyclopedia. Doctrine of Completeness. Legal Information Institute at Cornell University Law School. URL: https://www.law.cornell.edu/wex/doctrine_of_completeness.

Rules of Evidence: Accuracy

- Reliability of the *computer process* that created the content **not** the data content itself.
- Can we explain how an exhibit came into being?
  - What does the computer system do?
  - What are its inputs?
  - What are the internal processes?
  - What are the controls?

Chain of Custody

- When you are given an original copy of media to deal with, you need to document the handling:
  - *Where* it was stored
  - *Who* had access to it and *when*
  - *What* was done to it
- Shows that the **integrity** of evidence/data was preserved and not open to compromise.
- Route the evidence takes from the time you find it until the case is closed or goes to court.

Time Attributes

- Allow an investigator to develop a timeline of the incident
- M-A-C
  - **m**time: Modified time
    - Changed by modifying a file's content.
  - **a**time: Accessed time
    - Changed by reading a file or running a program.
  - **c**time : changed time
    - Keeps track of when the meta-information about the file was changed (e.g., owner, group, file permission, or access privilege settings).
    - Can be used as approximate ***dtime*** (deleted time).

The Forensic Process

Forensics Process/Flow (AAA)

- **A**cquisition/Preparation/Preservation
  - Copy the evidence/data without altering or damaging the original data or scene.
- **A**uthentication/Identification
  - Prove that the recovered evidence/data is the same as the original data.
- **A**nalysis/Examination/Evaluation
  - Analyze the evidence/data without modifying it.
- Reporting/ Presentation/ Documentation/ Interpretation

Review:

Topic 2: Evidence Acquisition

Acquisition

- First step in the forensic process:
  - Copy the evidence/data without altering or damaging the original data or scene.
  - Can you think of a circumstance where analyzing the original would be impossible?
- Must be done concurrently with Authentication:
  - Prove that the recovered evidence/data is the same as the original data.
  - Why?

Purpose of Authentication

- Acquired copy of evidence provides protection for the original.
- Authentication proves the copy is ***exactly the same*** as the original.
- How can you prove two digital things are exactly the same?
  - Compare every single bit.
  - OR...
  - Compute a cryptographic hash of both.

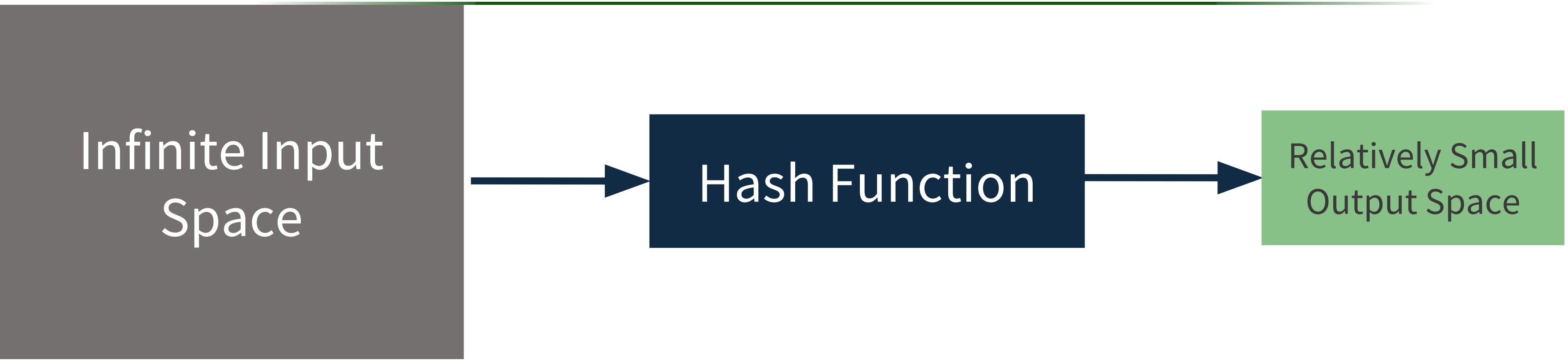
Message Digests

- Also called *cryptographic hash functions*
- Purposes:
  1. Uniquely identify data using the data itself as the source
     - Better than an index or a random number because others can generate the same identification using just the data
     - Should be easy to generate for any input (message)
  2. Infeasible to find data that will generate a specific digest
     - Can't process the hash in reverse
  3. Infeasible to find two messages that will generate the same digest
  4. The digest changes if the data changes

Called a "collision"

- Usually based on "lossy" computations

Hash Function: One-Way



- One-way function: It is impossible to calculate m from $H(m)$

1TB Hard Disk

128 bit digest

Acquisition Types and Methods

Acquisition Types

- Live acquisitions
  - System is still running
  - Data still available in RAM
  - Crucial if the storage is ***encrypted*** - only way to recover the key to decrypt the data
  - Inherently trusts the system to get the data...
- Static (or dead) acquisitions
  - System is turned off
  - ***Preferred method*** of acquisition
  - Limits the data available
    - No RAM data
    - No way to decrypt

Three Acquisition Methods

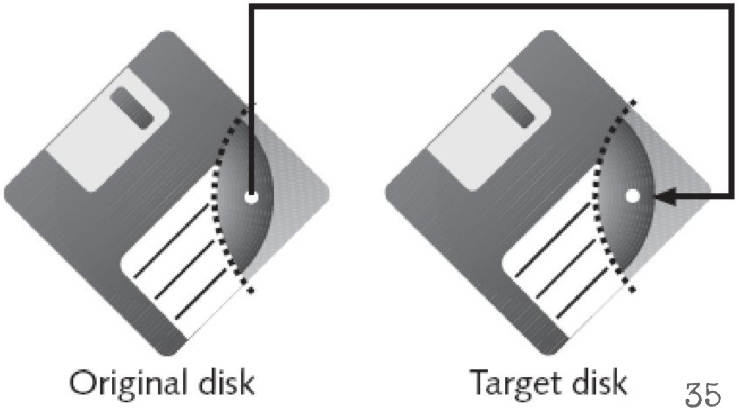
Ordered from the least amount of data collected to the most:

1. Logical Acquisition
   - Captures only ***specific files*** of interest to the case or specific ***types of files***.
   - Example: Email investigation - `.pst` and `.ost` files.
   - Focus: Filesystem (relies on filesystem to list files correctly)
2. Sparse Acquisition
   - Same as logical, but includes fragments of ***unallocated*** (deleted) data.
   - Focus: Partition or Volume
3. Bit-stream Copy or Acquisition
   - ***Exact copy*** (bit for bit) of the entire device; also called a ***forensic copy***.
   - Includes deleted files, fragments, etc.
   - Focus: Disk or other storage medium.

NOTE: A logical or sparse acquisition may be more appropriate if ***time is limited*** or if the ***original storage isn't accessible***, such as in web or cloud forensic cases.

More on Bit-Stream Acquisitions (1)

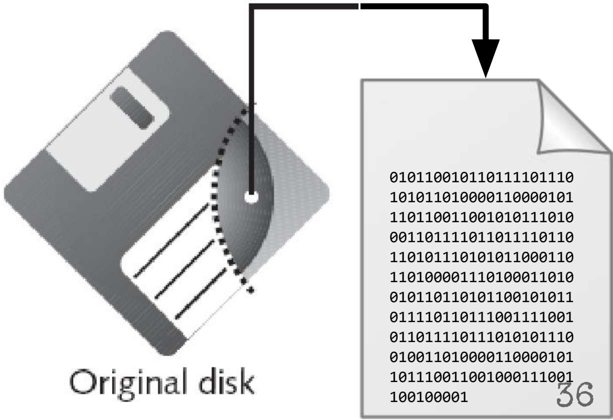
- Two types of bit-stream copies:
  1. Bit-stream disk-to-disk
     - Contents of evidence written to a storage device that exactly matches the make and model of the original: a ***literal duplicate*** of the original.
     - Only used when something about the storage device itself is important.



More on Bit-Stream Acquisitions (2)

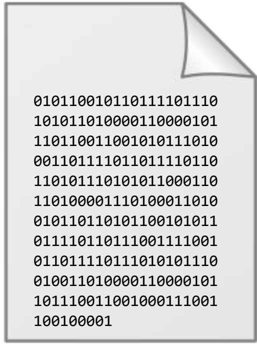
- Two types of bit-stream copies:

  2. Bit-stream disk-to-image file
     - All bits from the evidence are copied to a file: a ***virtual duplicate*** of the original.
     - More common method than disk-to-disk.
     - Referred to as an “image” or “image file”.
     - File is the exact size of the original evidence.



Evidence Formats

Raw



- Bit-stream image file
- Advantages
  - Fast (but uncompressed) data transfers.
  - Can ignore minor data read errors on source drive.
  - “Universal” format - not specific to any tool.
- Disadvantages
  - Requires as much storage as original disk or data.
  - Tools might not collect marginal (bad) sectors.

Proprietary Formats

- Features:
  - Compressed image files.
  - Split an image into smaller segments.
  - Integrate metadata into the image file.
- Disadvantages:
  - Inability to share an image between different tools.
  - File size limitation for each segmented volume.
- Unofficial standard: Expert Witness
  - Files end in `.e01`, `.e02`, `.e03`, etc.

Advanced Forensics Format

- Developed by Dr. Simson L. Garfinkel
- Design goals
  - Provide compressed or uncompressed image files.
  - No size restriction for disk-to-image files.
  - Provide space in the image file or segmented files for metadata.
  - Simple design with extensibility.
  - ***Open source*** for multiple platforms and OSs - no vendor lock-in.
  - Internal consistency checks for self-authentication.
- File extensions
  - *.afd for segmented image files.
  - *.afm for AFF metadata.

Review:

Topic 3: Drives, Volumes, and Files

Big- and Little-Endian

- Big-endian ordering:
  - Puts the **most significant byte** of the number in the **first** storage byte.
  - Sun SPARC, Motorola Power PC, ARM, MISP.
- Little-endian ordering:
  - Puts the **least significant byte** of the number in the **first** storage byte.
  - IA32-based systems.

Endianness: Example

Actual Value: 0x12345678 (4 Bytes)

- Big-endian ordering

23	24	25	26	27	28
00	12	34	56	78	00

- Little-endian ordering

23	24	25	26	27	28
00	78	56	34	12	00

Data Structure: Example

Byte Range	Description
0-1	2-byte house number
2-31	30-byte ASCII street name

```
0000000: 0100 4d61 696e 2053 742e 0000 0000 0000   ..Main St....
0000016: 0000 0000 0000 0000 0000 0000 0000 0000   ..............
0000032: bb02 536f 7574 6820 4d69 6c6c 4176 652e   ??
0000048: 0000 0000 0000 0000 0000 0000 0000 0000
```

The byte offset in decimal

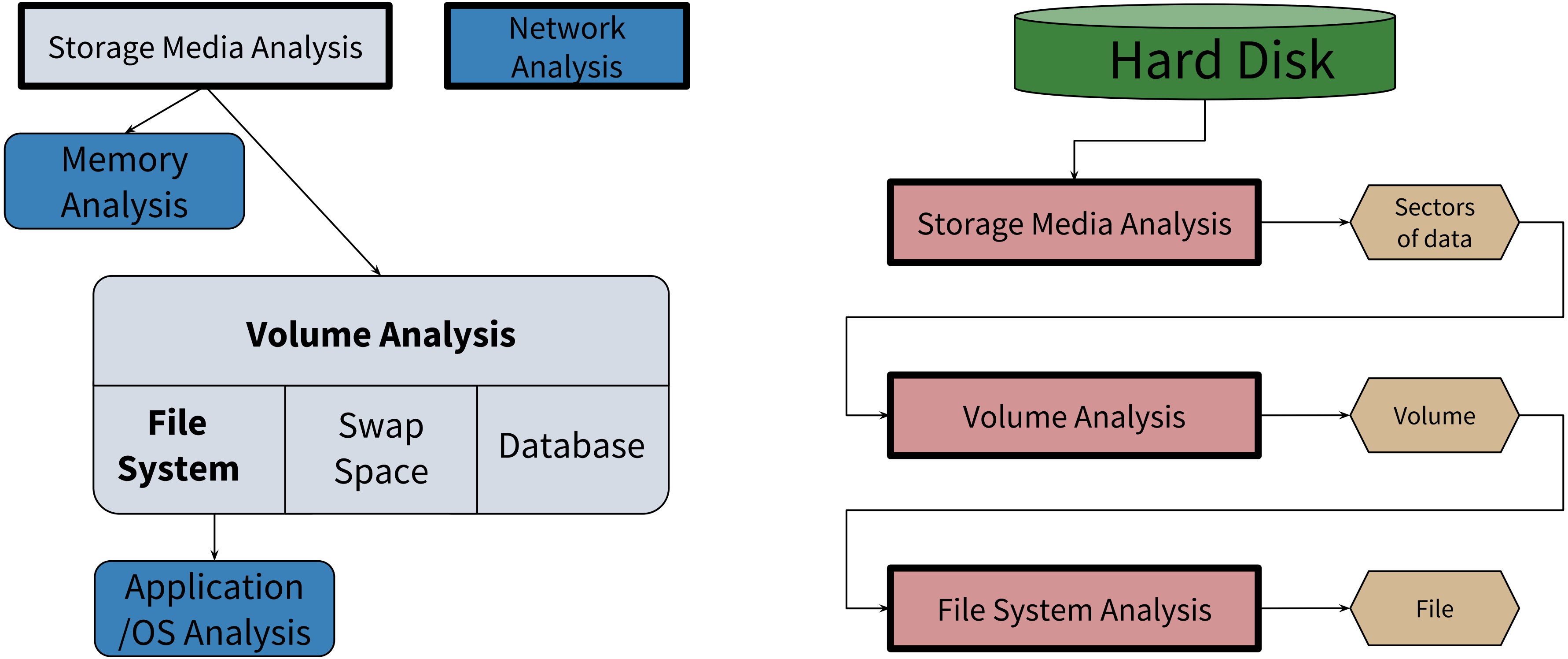
16 bytes of the data in hexadecimal

ASCII equivalent

Data structures are important!!

Layers of Forensic Analysis

Layers of Forensic Analysis



Layers of Analysis (1)

- Storage media analysis:
  - Non volatile storage such as hard disks and flash cards.
  - Organized into partitions / volumes:
    - Collection of ***storage locations*** that a user or application can write to and read from.
    - Contents are file system, a database, or a temporary swap space.

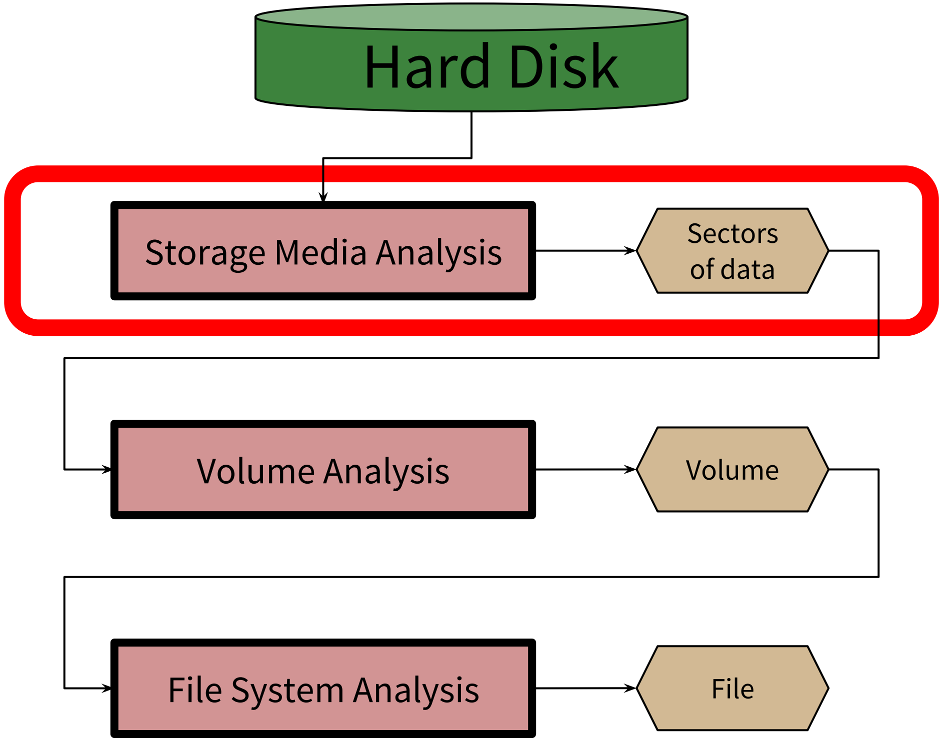
- Volume analysis:
  - Analyze data at the volume level.
  - Determine ***where*** the file system or other data are located.
  - Determine ***where*** we may find hidden data.

Layers of Analysis (2)

- File system analysis:
  - A collection of ***data structures*** that allow an application to create, read, and write files.
  - Purpose: To find files, to recover deleted files, and to find hidden data.
  - The result could be ***file content***, ***data fragments***, and ***metadata*** associated with files.

- Application layer analysis:
  - The structure of each file is based on the application or OS that created the file.
  - Purpose: To ***analyze files*** and to determine ***what program we should use***.

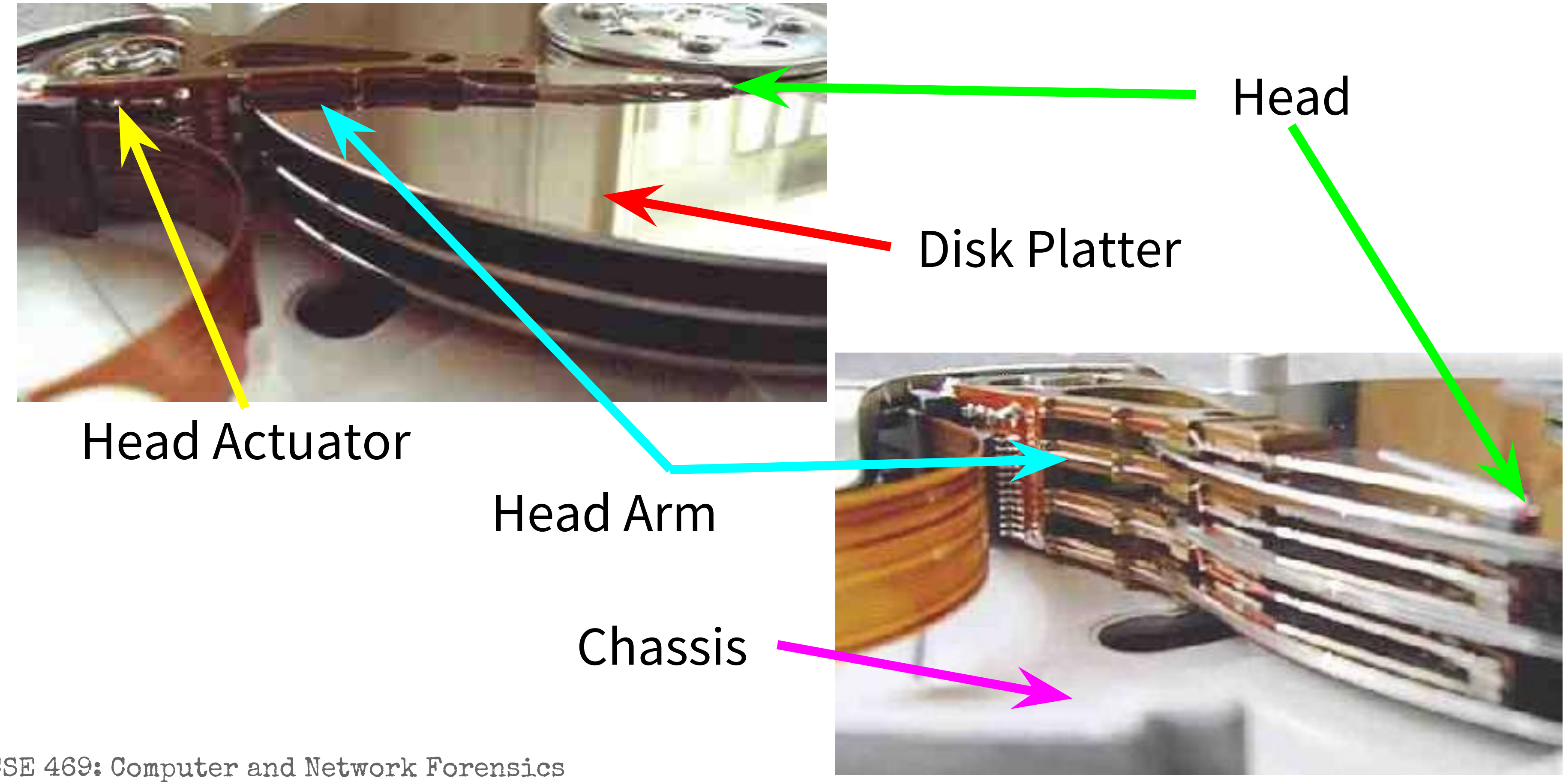
Disk Drive Geometry


Hard Disk
Storage Media Analysis
Sectors of data
Volume Analysis
Volume
File System Analysis
File

Storage Media Analysis

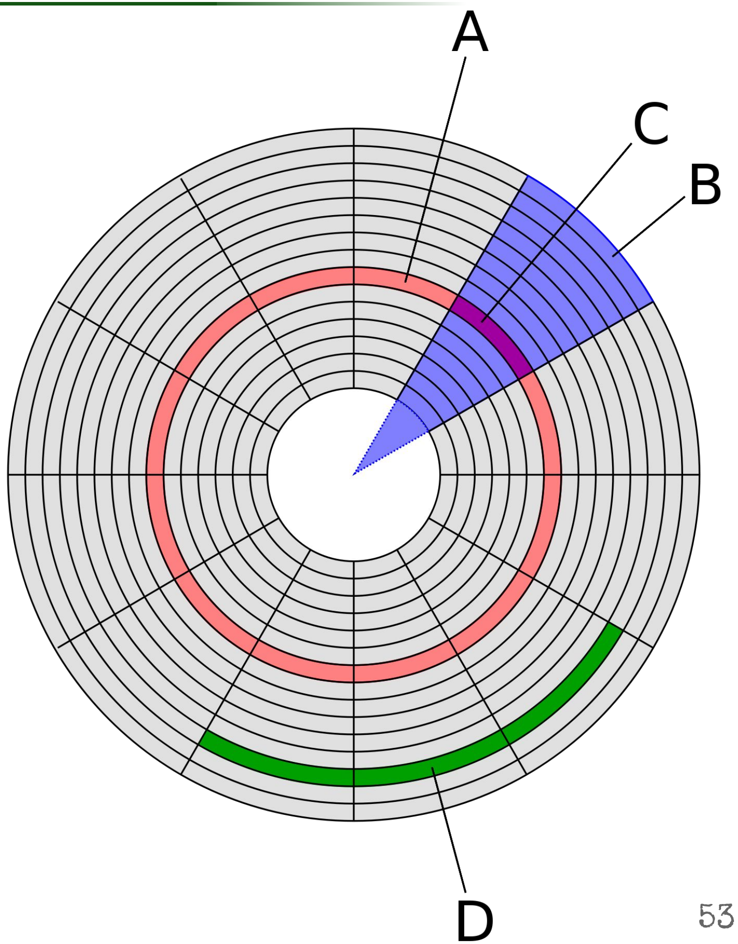
- Hard Disk Geometry
  - Head: The device that reads and writes data to a drive.
  - Track: Concentric circles on a disk platter.
  - Cylinder: A column of tracks on disk platters.
  - Sector: A section on a track.

Inside a Hard Drive



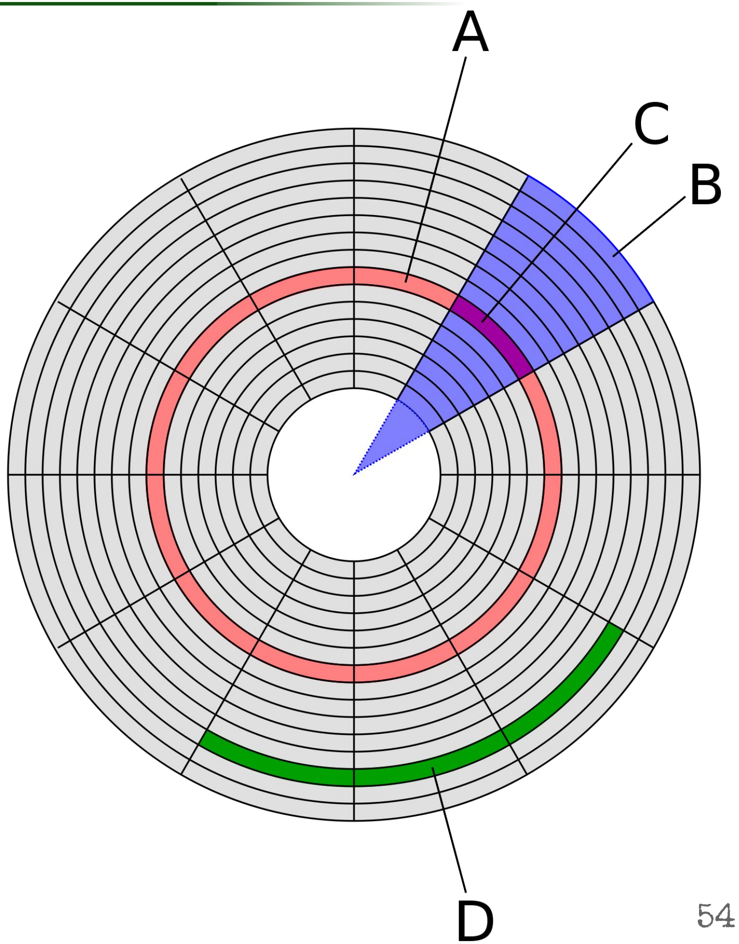
Tracks, Sectors, and Clusters

- Platters are divided into concentric rings called ***tracks*** (A).
- Tracks are divided into wedge-shaped areas called ***sectors*** (C).
  - A sector typically holds 512 bytes of data.
  - A collection of sectors is called a ***cluster*** or ***block*** (D).
- (B) is apparently called a *geometrical sector* (uncommon).



CHS Addresses

- ***Tracks/Cylinders***: Numbered from the outside in, **starting at 0**.
  - All sectors of all tracks in cylinder 0 will be filled up before using cylinder 1.
- ***Heads***: Numbered from the bottom up, **starting at 0**.
  - All platters are double-sided, one head per side.
- ***Sectors***: Each sector is numbered, **starting at 1**.
  - Typically holds 512 bytes of data.
- First sector has CHS address: **0,0,1**



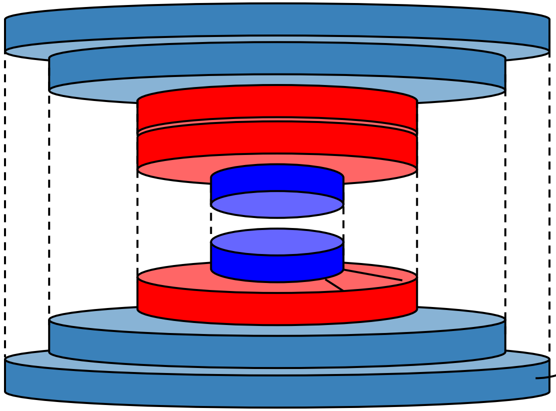
Logical Block Address (LBA)

- CHS addresses have a limit of 8.1 GB.
  - Not enough bits allocated to store values in the Master Boot Record of disks.
- Logical Block Addresses (LBA) overcome this:
  - Singe address instead of three.
  - **Starts at 0**, so LBA 0 == CHS 0,0,1.
  - To convert from CHS, need to know:
    - CHS address.
    - Number of heads per cylinder.
    - Number of sectors per track.

CHS to LBA Conversion

- LBA = (((**CYLINDER** * heads_per_cylinder) + **HEAD**) * sectors_per_track) + **SECTOR** -1

== num_platters * 2



- CHS (***x***,***y***,***z***)
  - Locate the ***x***-th cylinder and calculate the number of sectors
  - Locate the ***y***-th head and calculate the number of sectors
  - Add (***z***-1) sectors

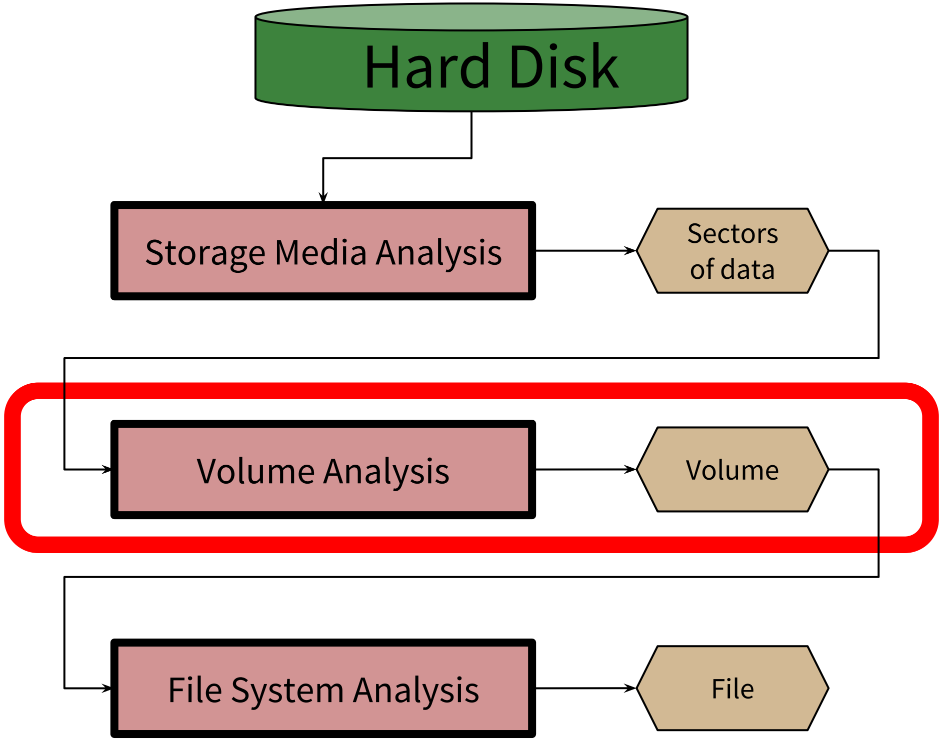
Address Conversion: Practice

- Given a disk with **16 heads** per cylinder and **63 sectors** per track, if we had a CHS address of **cylinder 2**, **head 3**, and **sector 4**, what would be the LBA (a.k.a CHS (2,3,4))?

LBA = (((**CYLINDER** * heads_per_cylinder) + **HEAD**) * sectors_per_track) + **SECTOR** -1

(((2*16)+3)*63)+4-1=2208

Volumes and Partitions


Hard Disk
Storage Media Analysis
Sectors
of data
Volume Analysis
Volume
File System Analysis
File

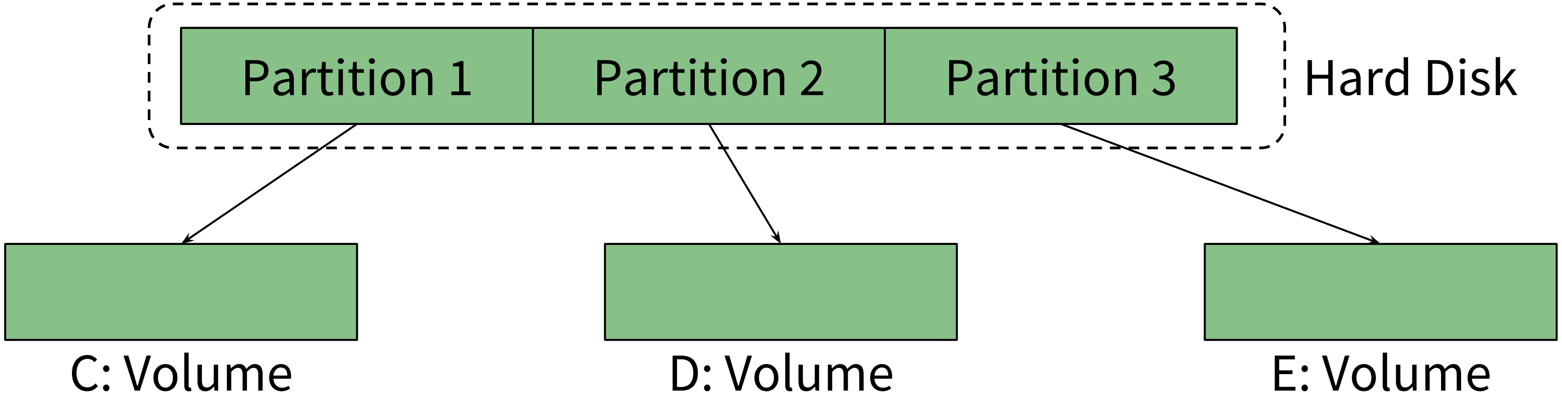
Volume Analysis

- Volume/Partition:
  - Collection of ***addressable sectors*** that an OS or application can use for data storage.
  - Used to store file system and other structured data.

- Purpose of Volume Analysis:
  - Involves looking at the data structures that are involved with partitioning and assembling the bytes in storage devices.

Partitions

- Collection of ***consecutive*** sectors in a volume.
- Each OS and hardware platform use a different partitioning method.



Partitions: Purpose

- Partitions organize the layout of a volume.
- Essential data are the ***starting*** and ***ending*** location for each partition.
- Common partition systems have one or more tables and each table describes a partition:
  - Starting sector of the partition.
  - Ending sector of the partition (or the length).
  - Type of partition.

Master Boot Record (MBR)

- First sector (CHS 0,0,1) stores the disk layout.
- Each **partition entry** has the structure shown on the next slide.

Offset	**Description**	**Size**
0x0000	Executable Code (Boots Computer)	446 Bytes
0x01BE	1st Partition Entry	16 Bytes
0x01CE	2nd Partition Entry	16 Bytes
0x01DE	3rd Partition Entry	16 Bytes
0x01EE	4th Partition Entry	16 Bytes
0x01FE	**Boot Record Signature (0x55 0xAA)**	2 Bytes

MBR Partition Entry

Offset	Description	Size
0x00	Current State of Partition (0x00=Inactive, 0x80=Active)	1 byte
0x01	Beginning of Partition - Head	1 byte
0x02	Beginning of Partition - Cylinder/Sector	1 word (2 bytes)
0x04	Type of Partition	1 byte
0x05	End of Partition - Head	1 byte
0x06	End of Partition - Cylinder/Sector	1 word (2 bytes)
0x08	LBA of First Sector in the Partition	1 double word (4 bytes)
0x0C	Number of Sectors in the Partition	1 double word

Volume Analysis (MBR)

```
0000432: 0000 0000 0000 0000 0000 0000 0000 0001
0000448: 0100 07fe 3f7f 3f00 0000 4160 1f00 8000
0000464: 0180 0bfe 3f8c 8060 1f00 cd2f 0300 0000
```

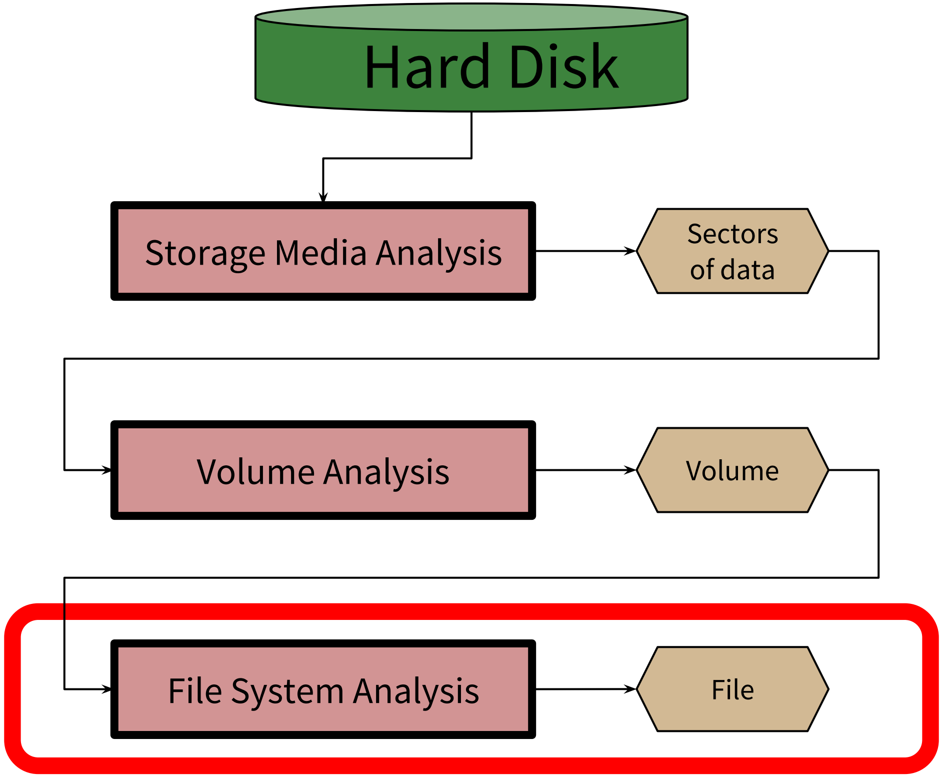
The first 446 bytes contain boot code

The byte offset in decimal

16 bytes of the data in hexadecimal

#	Flag	Type	Starting Sector	Size
1	0x00	0x07	0x0000003f (63)	0x001f6041 (2,056,257)
2	?	?	?	?

Files and Directories


Hard Disk
Storage Media Analysis
Sectors of data
Volume Analysis
Volume
File System Analysis
File

File Systems and Disks

- User view:
  - File is a ***named***, ***persistent*** collection of data.

- OS & file system view:
  - File is collection of disk blocks — i.e., a ***container***.
  - File System ***maps*** file names and offsets to disk blocks.

File Attributes

- Name:
  - Although the name is not always what you think it is!
- Type:
  - May be encoded in the name (e.g., `.cpp`, `.txt`)
- Dates:
  - Creation, updated, last accessed, etc.
  - (Usually) associated with container.
  - Better if associated with content.
- Size:
  - Length in number of bytes; occasionally rounded up.
- Protection:
  - Owner, group, etc.
  - Authority to read, update, extend, etc.
- Locks:
  - For managing concurrent access.
- …

File Metadata

- Definition:
  - Information ***about*** a file. Data ***about*** the data.
- Maintained by the file system.
- Separate from file itself.
- Usually attached or connected to the file.
- Some information visible to user/application:
  - Dates, permissions, type, name, etc.
- Some information primarily for OS:
  - Location on disk, locks, cached attributes

Directory – A Special Kind of File

- A tool for users and applications to organize and find files.
  - User-friendly names.
  - Names that are meaningful over long periods of time.

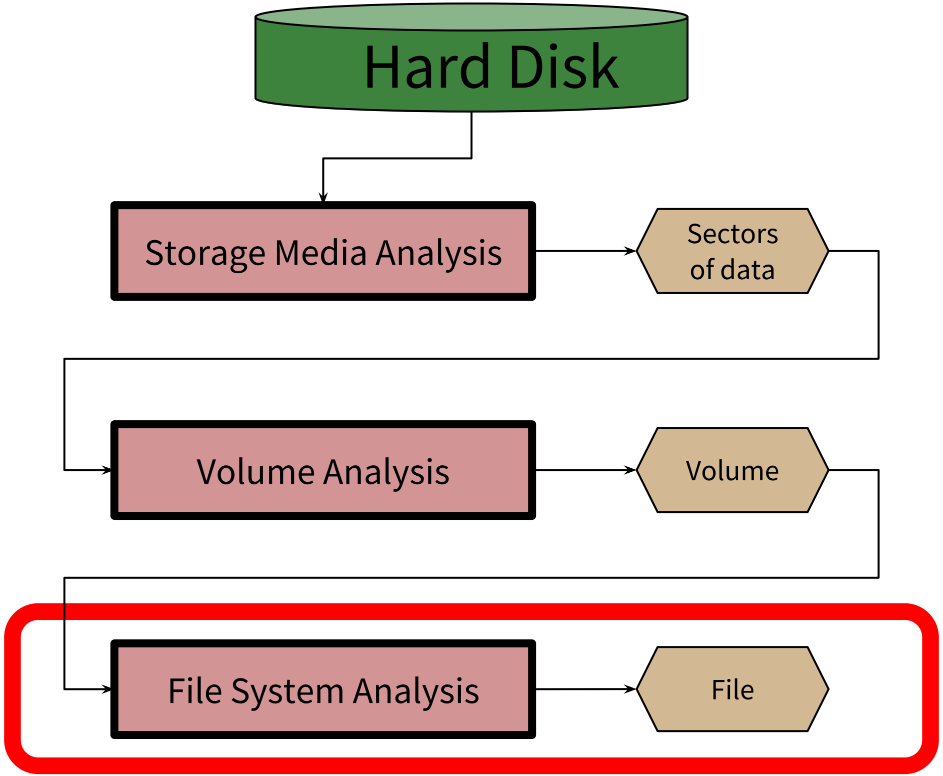
- The data structure for OS to locate files (i.e., containers) on disk.

Links

- Symbolic (soft) links:
  - Unidirectional relationship between a filename and the file.
  - Directory entry contains *text* describing *absolute* or *relative* path name of original file.
  - If the source file is deleted, the link exists but pointer is invalid.
- Hard links:
  - Bidirectional relationship between file names and file.
  - A hard link is directory entry that points to a source file's metadata.
  - Metadata maintains *reference count* of the number of hard links pointing to it – *link reference count*.
  - Link reference count is decremented when a hard link is deleted.
  - File data is deleted and space freed when the link reference count goes to zero.

Review:

Topic 4: File Systems

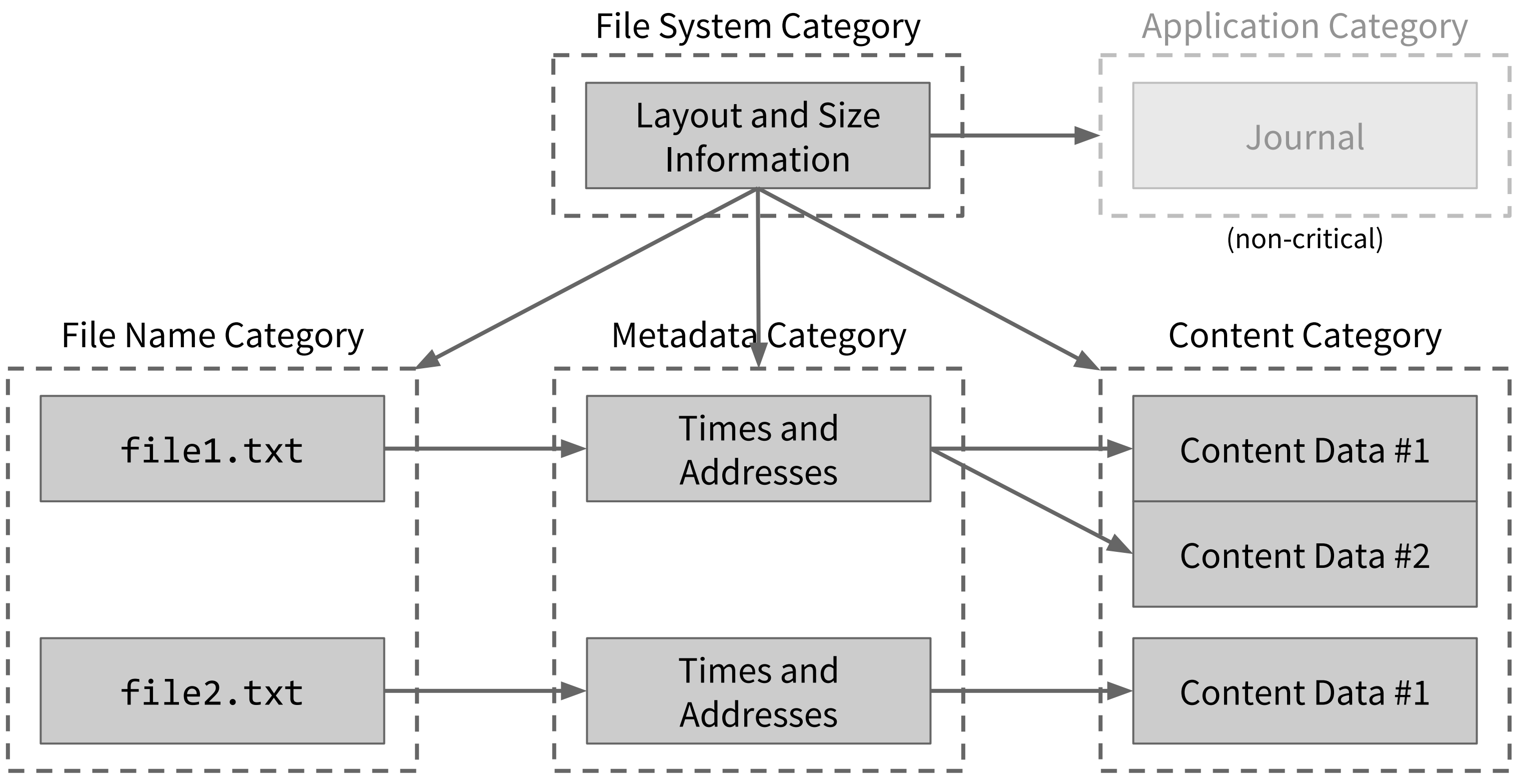

Hard Disk
Storage Media Analysis
Sectors of data
Volume Analysis
Volume
File System Analysis
File

File System Reference Model

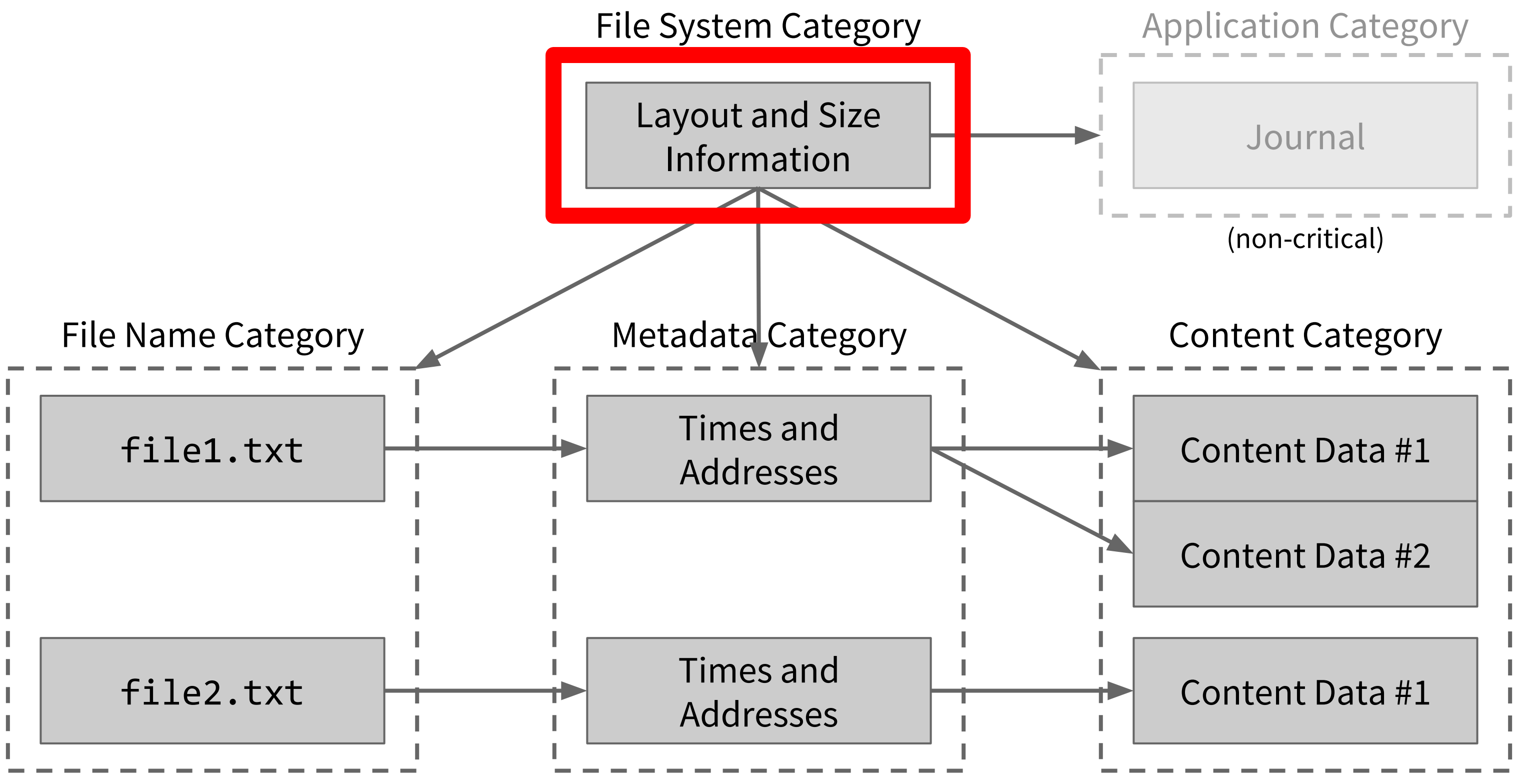
Reference Model Categories

1. File system category:
   - General info about the file system.
   - Size and layout, location of data structures, size of data units.
2. Content category:
   - Data of the actual files - the reason file systems exist.
   - Organized into collections of standard-sized containers.
3. Metadata category:
   - Data that describes a file (except for the name of the file!).
   - Size, locations of content, times modified, access control info.
4. File name category:
   - a.k.a Human interface category.
   - Name of the file.
   - Normally stored in contents of a directory along with location of the file’s metadata.
5. Application category:
   - Not essential to file system operations.
   - Journal.

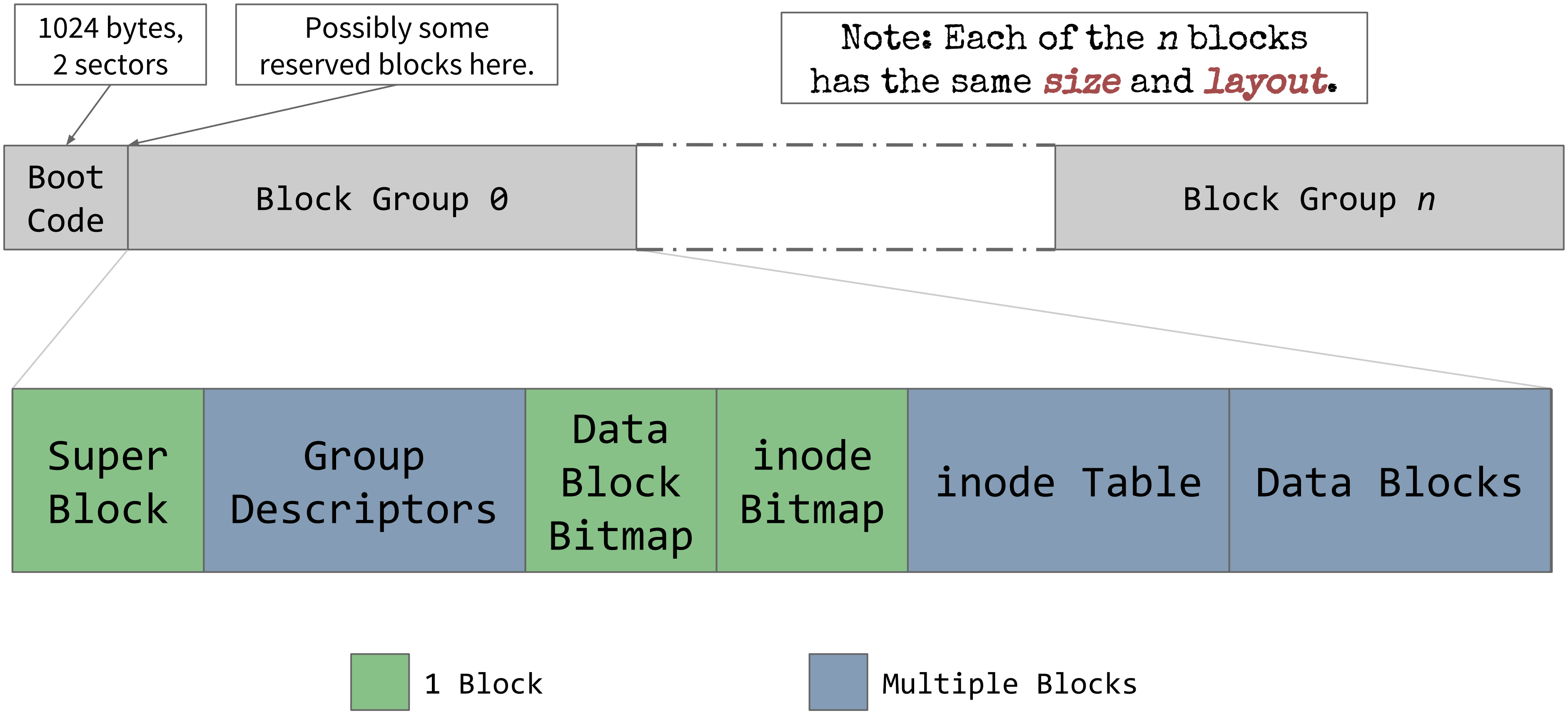
Reference Model Illustrated



ext4


File System Category
Layout and Size Information
Application Category
Journal
(non-critical)
File Name Category
Metadata Category
Content Category
file1.txt
Times and Addresses
Content Data #1
Content Data #2
file2.txt
Times and Addresses
Content Data #1

ext4 Layout


1024 bytes,
2 sectors
Possibly some
reserved blocks here.
Note: Each of the n blocks
has the same size and layout.
Boot
Code
Block Group 0
Block Group n
Super
Block
Group
Descriptors
Data
Block
Bitmap
inode
Bitmap
inode Table
Data Blocks
1 Block
Multiple Blocks

Superblock

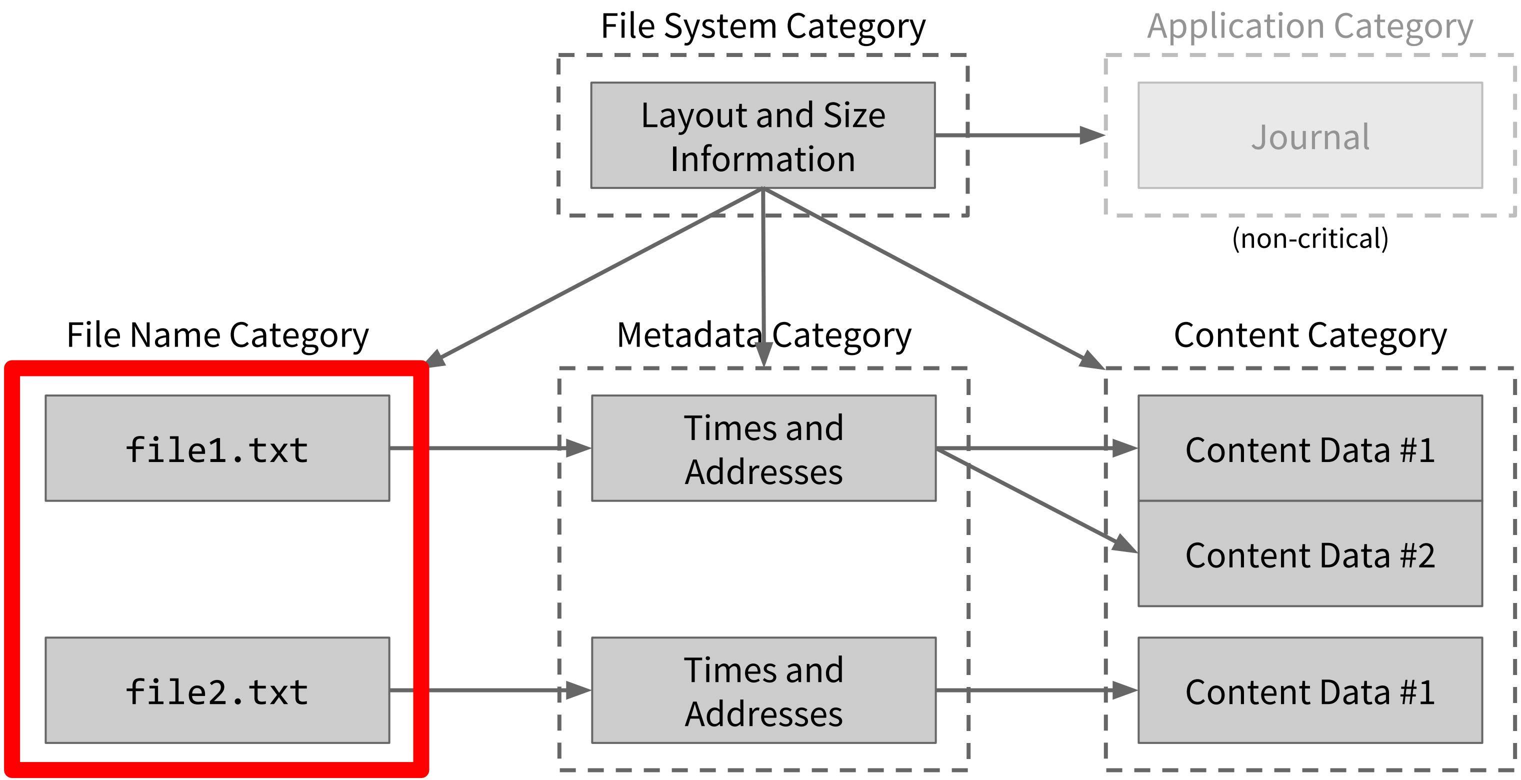
- Stores layout information for the file system.
- Duplicated in *every block group* in the file system.
  - Kernel only reads the superblock in group 0. The others are backup copies.
- Stores:
  - Block size
  - Total # of blocks
  - # blocks per group
  - # reserved blocks before group 0
  - # of inodes (total)
  - # of inodes per block group

Super Block	Group Descriptors	Data Block Bitmap	inode Bitmap	inode Table	Data Blocks

Group Descriptor

- Has the following fields:
  - Block numbers of the block bitmap and inode bitmap.
  - Block number of the first inode table block.
  - Number of free blocks, free inodes, and directories in the group.
- The descriptor table contains **all** the descriptors for the whole file system.
- Duplicated in ***every block group***, just like the superblock.

Super Block	Group Descriptors	Data Block Bitmap	inode Bitmap	inode Table	Data Blocks


File System Category
Layout and Size Information
Application Category
Journal
(non-critical)
File Name Category
Metadata Category
Content Category
file1.txt
file2.txt
Times and Addresses
Times and Addresses
Content Data #1
Content Data #2
Content Data #1

Directory

- Just another file, but with a simple structure that identifies the files it contains.
- Always includes '.' (self) and '..' (parent) entries (even for the root directory!).
- Directory entry fields:
  - inode number
  - File name
  - File type number →

	File Type
0	Unknown
1	Regular file
2	Directory
3	Character device
4	Block device
5	Named pipe
6	Socket
7	Symbolic link

Directory Entry Example

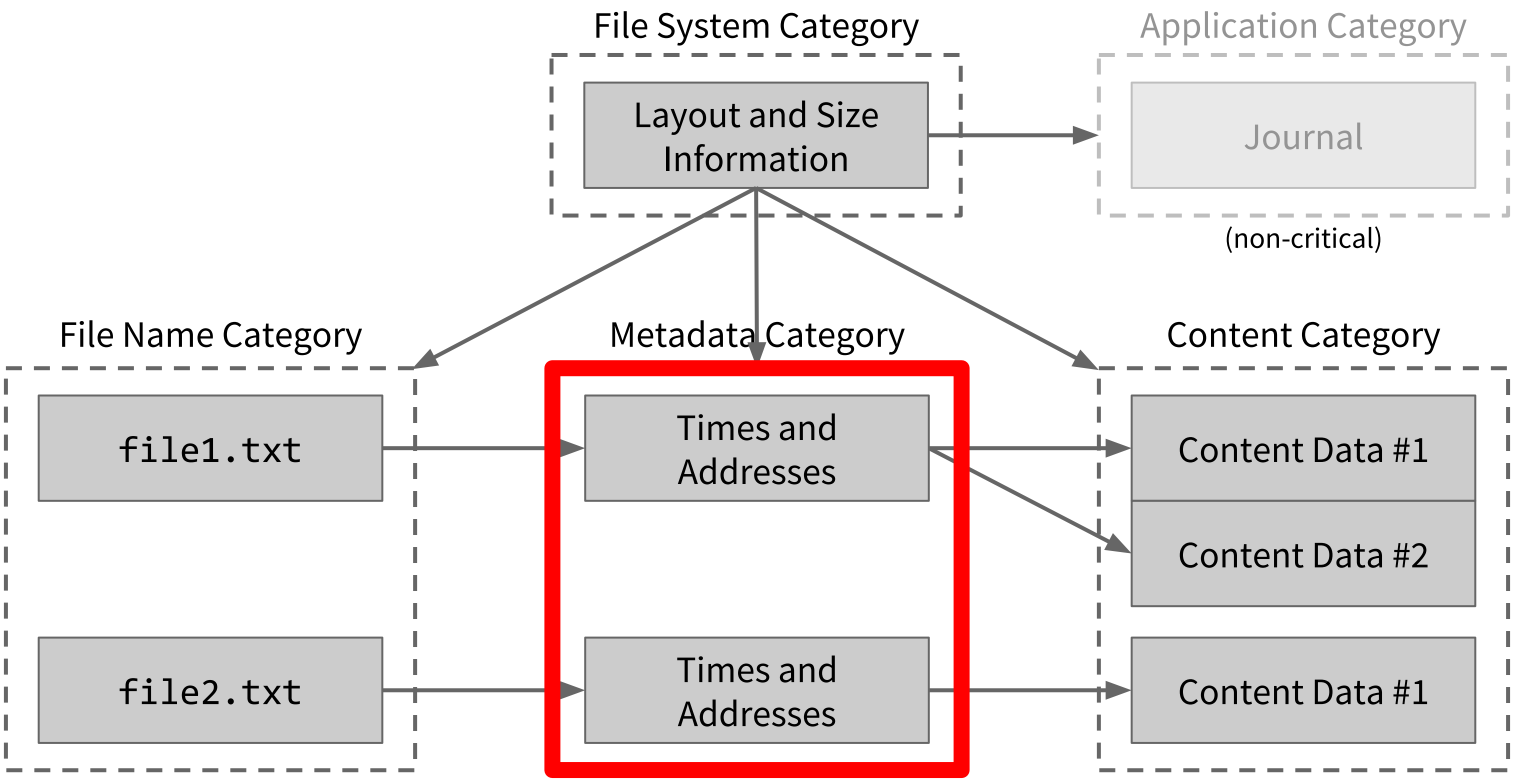
offset	inode	rec_len	name_len	file_type	name							
0	21	12	1	2	.	\0	\0	\0				
12	22	12	2	2	.	.	\0	\0				
24	53	16	5	2	h	o	m	e	1	\0	\0	\0
40	67	28	3	2	u	s	r	\0				
52	0	16	7	1	o	l	d	f	i	l	e	\0
68	34	4028	4	2	s	b	i	n				

The last record needs to point to the end of the block, so it will have a length much larger than normal.

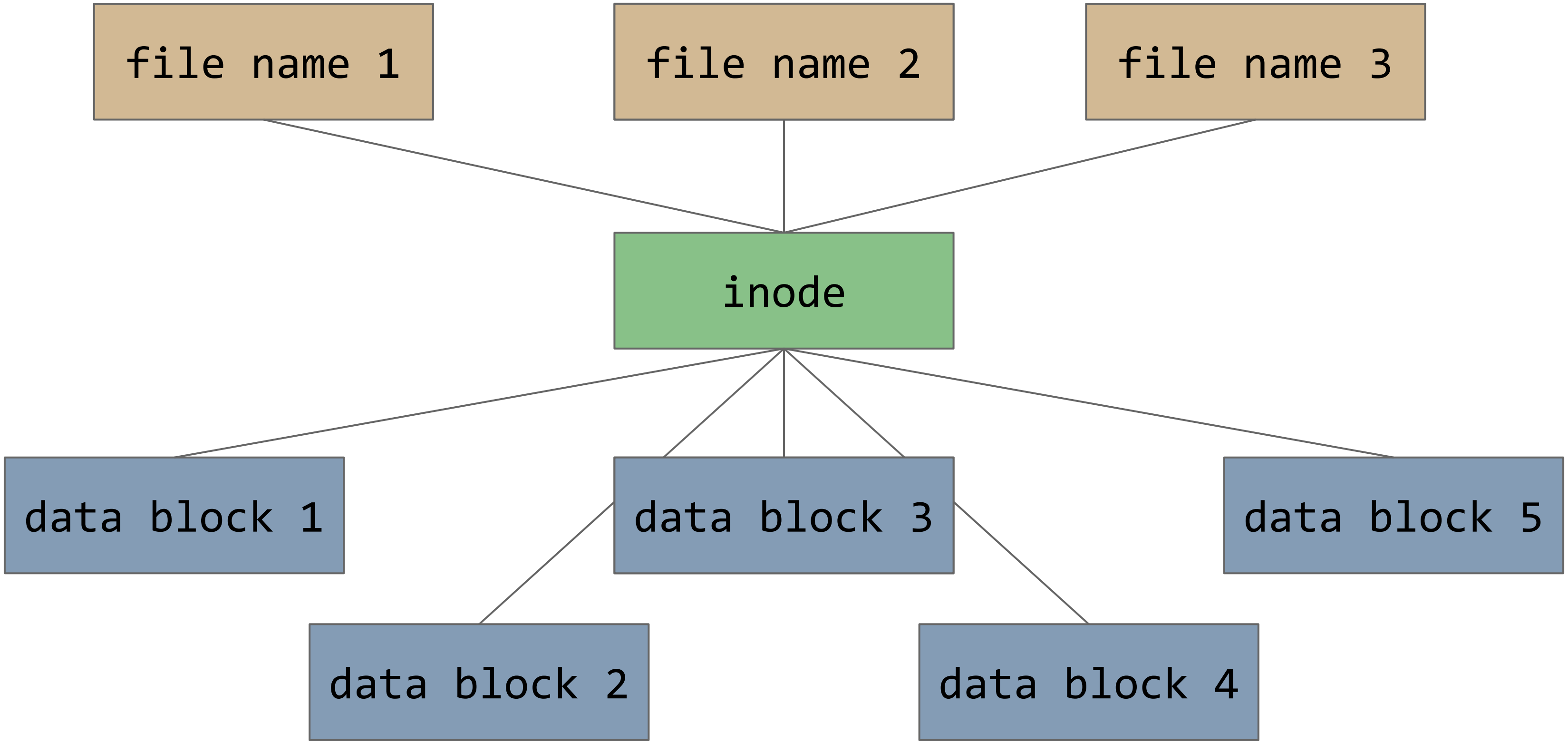
Deleted: There is no inode 0.

Always 8 bytes

Always a multiple of 4 bytes

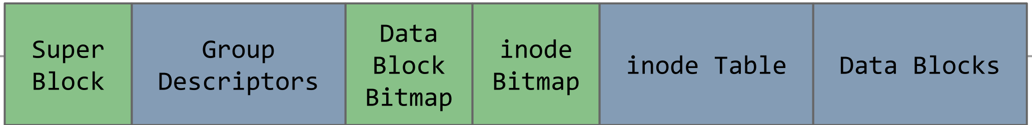

File System Category
Layout and Size Information
Application Category
Journal
(non-critical)
File Name Category
file1.txt
file2.txt
Metadata Category
Times and Addresses
Times and Addresses
Content Category
Content Data #1
Content Data #2
Content Data #1

inodes


file name 1
file name 2
file name 3
inode
data block 1
data block 3
data block 5
data block 2
data block 4

inode Fields (Selected) (1)

Offset	Bits	Name	Description
0x0	16	i_mode	Mode (9 bits). Sticky bit, setgid, setuid (3 bits). File type (4 bits).
0x2	16	i_uid	Owner's user identifier (UID).
0x18	16	i_gid	Group identifier (GID).
0x8	32	i_atime	Last access time, in seconds since the epoch.
0xC	32	i_ctime	Last inode change time, in seconds since the epoch.
0x10	32	i_mtime	Last data modification time, in seconds since the epoch.
0x14	32	i_dtime	Deletion Time, in seconds since the epoch.
0x1A	16	i_links_count	Hard link count. With the DIR_NLINK feature enabled, ext4 supports more than 64,998 subdirectories by setting this field to 1 to indicate that the number of hard links is not known.
0x28	60	i_block	Extent tree.



See also https://ext4.wiki.kernel.org/index.php/Ext4_Disk_Layout#Inode_Table

inode Fields (Selected) (2)

Offset	Bits	Name	Description
0x4	32	i_size_lo	Lower 32-bits of size in bytes.
0x6C	32	i_size_high	Upper 32-bits of file/directory size.
0x1C	32	i_blocks_lo	Lower 32-bits of "block" count.
0x74	16	i_blocks_hi	Upper 16-bits of the block count.
0x84	32	i_ctime_extra	Extra change time bits. This provides sub-second precision.
0x88	32	i_mtime_extra	Extra modification time bits. This provides sub-second precision.
0x8C	32	i_atime_extra	Extra access time bits. This provides sub-second precision.
0x90	32	i_crtime	File creation time, in seconds since the epoch. (Creation time of inode.)
0x94	32	i_crtime_extra	Extra file creation time bits. This provides sub-second precision.

Note: Every field with an offset >=0x80 is an ***extended field***, meaning it was introduced in ext4 and is not backwards compatible with ext2/3.

See also https://ext4.wiki.kernel.org/index.php/Ext4_Disk_Layout#Inode_Table

Super Block	Group Descriptors	Data Block Bitmap	inode Bitmap	inode Table	Data Blocks

Mode

- ext4 stores file permissions for the **user** (the owner of the file), the **group** the file is a part of, and all **others** (world).
- 3 bits for each ↑ represent the ***read***, ***write***, and ***execute*** permissions: 1 means they can, 0 means they can't.

Example Mode: `0754`

`0`: Means number is displayed in octal

`111`
`1`: Owner can read
`1`: Owner can write
`1`: Owner can execute

`101`
`1`: Group can read
`0`: Group cannot write
`1`: Group can execute

`100`
`1`: World can read
`0`: World cannot write
`0`: World cannot execute

File Types

0. Unknown
1. Regular file
2. Directory
3. Character device
4. Block device
5. Named pipe
6. Socket
7. Symbolic link

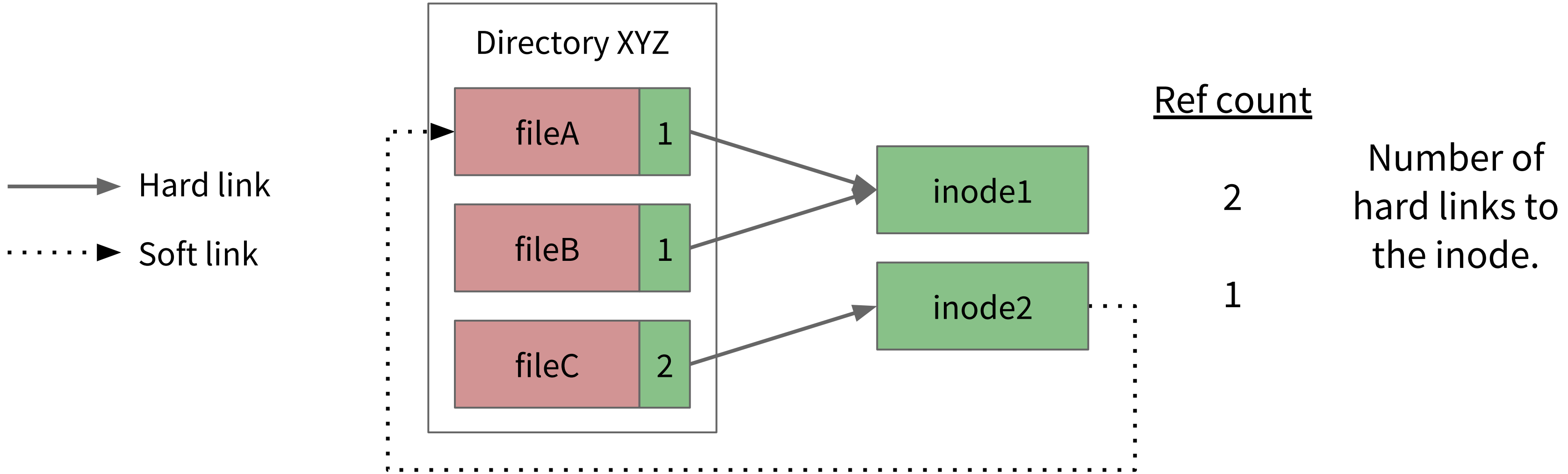
The only 2 types that allocate data blocks in the file system (except symbolic links, sometimes).

Require all read/write operations to work on an entire block at a time.

Contents of the file are the path to the file pointed to. Path is stored in inode if <60 characters, uses a data block otherwise.

Hard and Soft Links

- Hard link: A **filename** that points to an **inode**.
  - *Everything* has a hard link to it.
- Soft link: An **inode** that points to a **filename**.
  - Optional.



Time Attributes

- Allow an investigator to develop a timeline of the incident
- M-A-C
  - **m**time: Modified time
    - Changed by modifying a file's content.
  - **a**time: Accessed time
    - Changed by reading a file or running a program.
  - **c**time : changed time
    - Keeps track of when the meta-information about the file was changed (e.g., owner, group, file permission, or access privilege settings).
    - Can be used as approximate ***dtime*** (deleted time).

This slide is from
Topic 1: Forensics Intro

ext4: Extra Time Attributes

- ext4 introduces two additional time attributes:
  - **d**time: deletion time
  - **cr**time: creation time
- ext4 extends the time values from 32 bits to 64.
  - Overcomes the 2038 problem (puts it off until 2446).
    - 32 bits is a signed int to allow referencing dates *before* January 1, 1970 by using negative numbers.
  - Does not apply to dtime (remains 32 bits).

64-bit Time Values in ext4

Extra time field: 32 bits

Original time field: 32 bits

00010100101001010010100101001001 10010100101001001100101001010010

Number of seconds since the epoch (Jan 1, 1970 UTC)

New whole-second value:

February 16, 2185 00:22:42 6788794962 == 0110010100101001001100101001010010

Nanoseconds means 9 decimal places

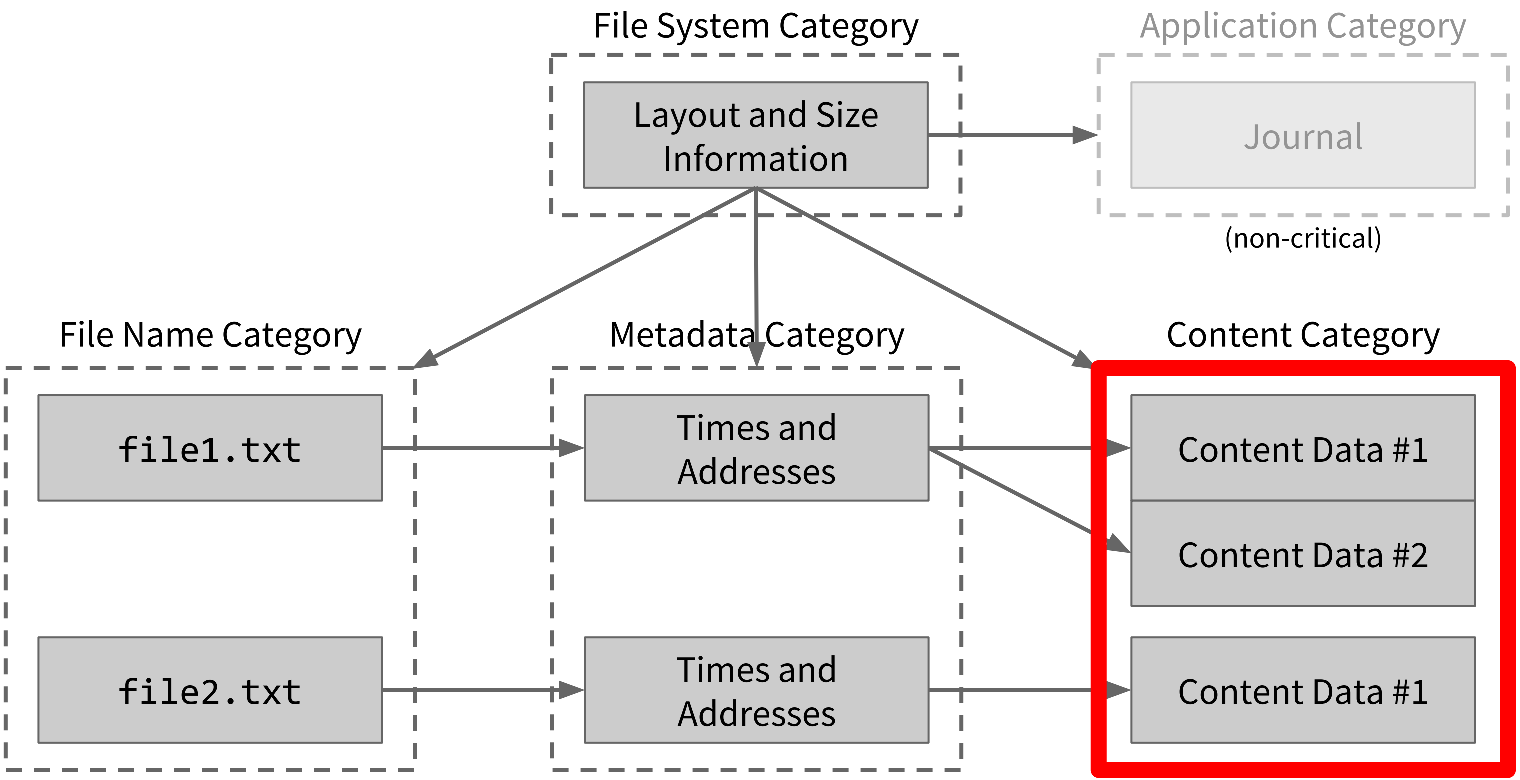
Nanosecond value:

000101001010010100101001010010 == 86592082 0.086592082

Don't forget you have to convert the bytes from Little Endian first!

Final date value:

February 16, 2185 00:22:42.086592082


File System Category
Layout and Size Information
Application Category
Journal
(non-critical)
File Name Category
file1.txt
file2.txt
Metadata Category
Times and Addresses
Times and Addresses
Content Category
Content Data #1
Content Data #2
Content Data #1

Block Bitmap / inode Bitmap

- 0 == available.
- 1 == in use.
- One bit per block/inode.
  - Denotes ***allocation status***.
- Number of **data blocks in a group** is always equal to the number of **bits in a block**.
- Far fewer inodes than blocks per group.
  - User-configurable.
  - Makes sense since most files will occupy more than one block, only need one (initial) inode per file.

Super Block	Group Descriptors	Data Block Bitmap	inode Bitmap	inode Table	Data Blocks

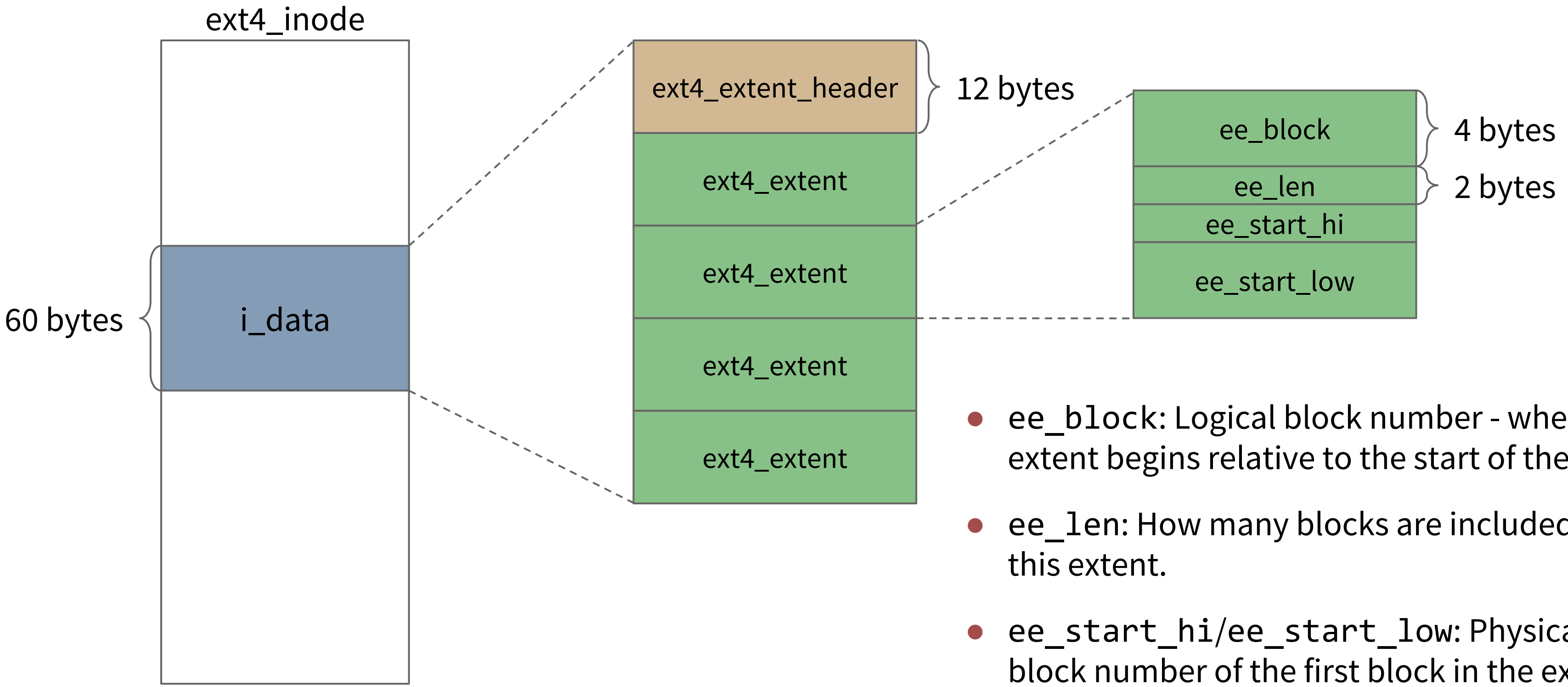
Extents

- The unit of allocation in ext4.
  - Described by its **starting** and **length** in blocks.
  - One file fragment only uses one extent.

- Previous "block mapping" scheme (<=ext3) stored each block address used by the file.

Super Block	Group Descriptors	Data Block Bitmap	inode Bitmap	inode Table	Data Blocks

Extent Structure



- `ee_block`: Logical block number - where this extent begins relative to the start of the file.
- `ee_len`: How many blocks are included in this extent.
- `ee_start_hi/ee_start_low`: Physical block number of the first block in the extent- i.e., where the extent actually begins on disk.

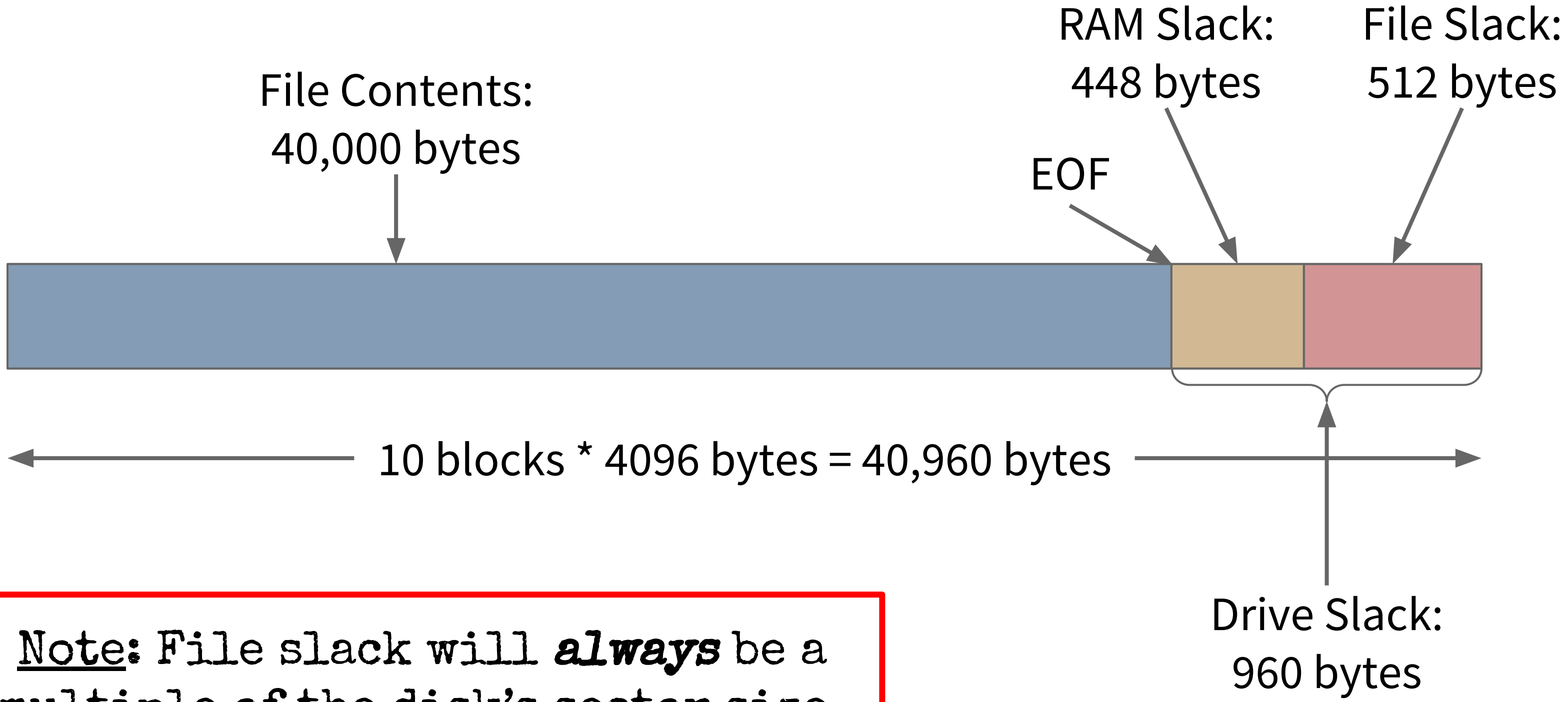
Drive Slack

- Drive Slack: The area on a disk that is ***allocated*** to a file, but doesn't store any of the file's data.
- Example:
  - File system with 4K blocks on a disk with 512 byte sectors.
  - File that is 40,000 bytes long occupies 10 blocks.
  - 10 blocks * 4096 bytes = 40,960 bytes allocated for the file.
  - The excess space of 960 bytes is called **drive slack**.
- Drive slack is divided into two parts: File slack and RAM slack.

File and RAM Slack

- Block devices: Require all read/write operations to work on an **entire block** at a time.
  - Cannot read/write a character at a time the way ***character devices*** do.
- Legacy operating systems used to read an entire block of data from RAM when writing to disk, *whether or not the entire block was part of the file being written!*
  - This is **RAM slack**. The size of the RAM slack is determined by how much of the disk's sector is leftover after writing the file.
  - The part of drive slack that isn't RAM slack is **file slack**.
- RAM slack Could be anything stored in memory: logon IDs, passwords, file fragments, ... anything!

Slack: Illustrated



<u>Note</u>: File slack will ***always*** be a multiple of the disk's sector size.

Review:

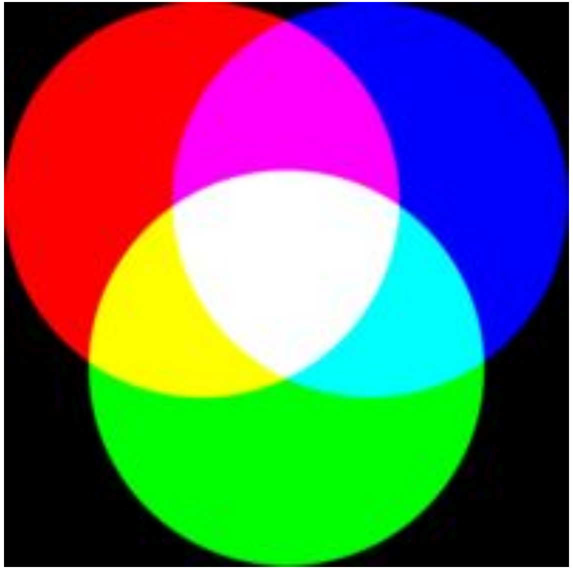
Topic 5: Image Forensics

Bit Depth

- Number of bits per pixel:
  - 1 bit – black and white
  - 4 bits – 16 colors (2^4)
  - 8 bits – 256 colors (2^8)
  - 16 bits – 65,536 colors (2^{16})
  - 24 bits – 16,777,216 colors (2^{24})
- Bit depth controls image file size:
  - Higher the bit depth = larger file

RGB Color Model

- Red – Green – Blue
- Additive model combines varying amounts of these 3 colors:



RGB Value Storage

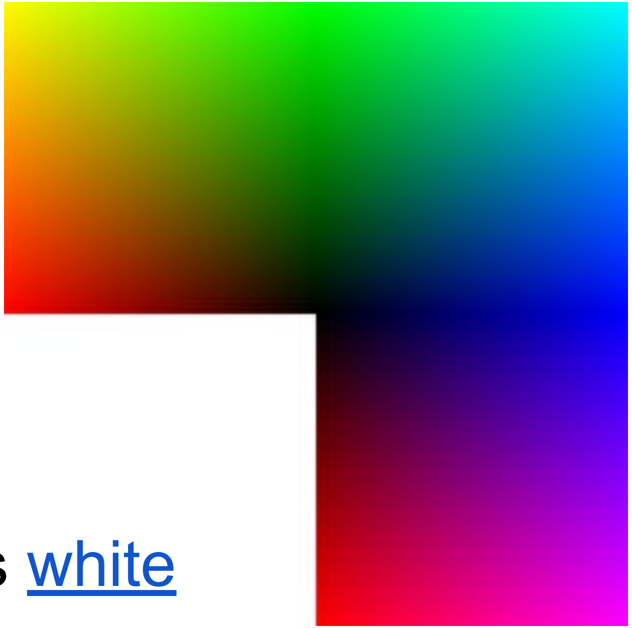
- Individual pixels represented in memory as a
  - Red value
  - Green value
  - Blue value
- Values represent **intensity**:
  - If red is more intense, the color perceived is towards the red.
- 24-bit pixel value means:
  - 8 bits for each RGB value
    - Values expressed as 0 – 255
  - 256 possible values for each primary color

Image Basics

(0, 255, 0) is green

(255, 255, 0) is yellow

(0, 255, 255) is cyan



(255, 0, 0) is red

(0, 0, 255) is blue

(0, 0, 0) is black
(255, 255, 255) is white

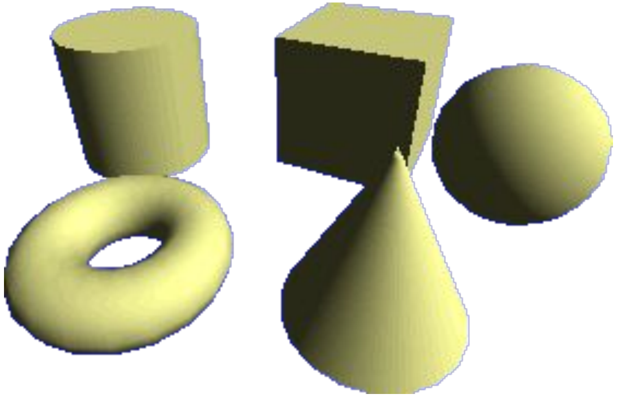
(255, 0, 255) is magenta

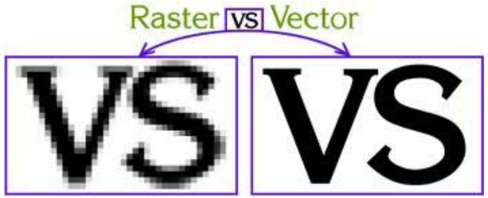
Recognizing a Graphics File

- Contains digital photographs, line art, three-dimensional images, and scanned replicas of printed pictures.
  - Bitmap images: collection of dots
  - Vector graphics: based on mathematical instructions
  - Metafile graphics: combination of bitmap and vector

Vector Graphics

- Characteristics:
  - Lines and geometric primitives instead of dots.
  - Store only the calculations for drawing lines and shapes.
  - For example: CorelDraw, Adobe Illustrator, Inkscape.





Examining the Raw File Format

- Raw file format:
  - Referred to as a digital negative.
  - Typically found on many higher-end digital cameras.
- Sensors in the digital camera simply record pixels on the camera's memory card.
- Raw format maintains the **best picture quality**.
- The biggest disadvantage is that it's **proprietary**:
  - Not all image viewers can display these formats.
- The process of converting raw picture data to another format is referred to as ***demosaicing***.

Examining EXIF Format

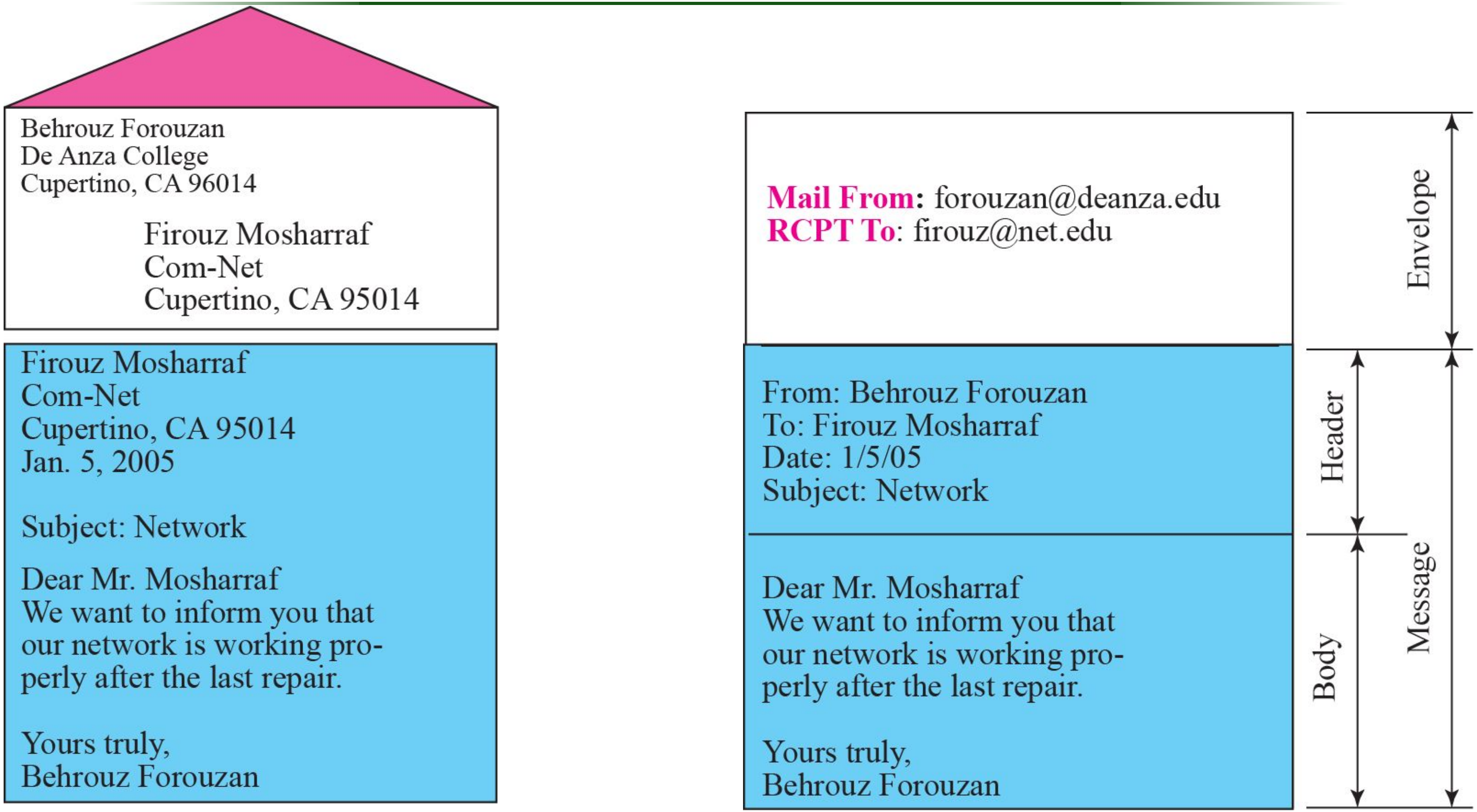
- Exchangeable Image File (EXIF) format:
  - Developed by JEIDA as a standard for storing metadata in JPEG and TIFF files.
  - Stores **metadata** at the beginning of the file:
    - Investigators can learn more about the type of digital camera and the environment in which pictures were taken.

EXIF Information			
File name:	DSC_0260.JPG	File size:	922866 bytes
File date:	2006:04:22 22:06:16	Camera make:	NIKON CORPORATION
Camera model:	NIKON D70s	Date/Time:	2006:04:17 18:06:08
Resolution:	3000 x 2632	Flash used:	No
Focal length:	18.0mm (35mm equivalent: 27mm)	Exposure time:	0.0008 s (1/1250)
Aperture:	f/8.0	Whitebalance:	Manual
Metering Mode:	matrix	Exposure:	Manual
Exposure Mode:	ManualAuto bracketing		

Review:

Topic 6: Email Forensics

Format of Email



Corporate vs Public Email

- Tracing **corporate** emails is easier:
  - Standard names.
  - Assigned by local administrator.
- Contrast with **public** email:
  - Non-standard names.
  - Usually not informative.

Identifying Email Crimes/Violations

- "Crime" may depend on jurisdiction:
  - Spam:
    - Illegal in Washington state
    - Elsewhere?
- Email crime is becoming commonplace:
  - Narcotics trafficking
  - Sexual harassment
  - Child pornography
  - Fraud
  - Terrorism

Email Headers

- **From**: Who the message is from. This is the easiest to forge, and thus the least reliable.
- **Reply-To**: The address to which replies should be sent. Often absent from the message, and very easily forgeable.
- **Return-Path**: The email address for return mail. Same as Reply-To:
- **Message-ID**: A unique string assigned by the mail system when the message is first created. The format of a Message-ID: field is <uniquestring>@<sitename>
- **Received**: They form a list of all sites (MTA) through which the message traveled in order to reach you.

Examining Email Headers

- Gather supporting evidence and track suspect:
  - Return path.
  - Recipient's email address.
  - Type of sending email service.
  - IP address of sending server.
  - Name of the email server.
  - Unique message number.
  - Date and time email was sent.
  - Attachment files information.

Tracing an Email Message

- Preliminary Steps:
  - Examine each field in the email header, especially the recorded IP address of sender.
  - Content analysis on suspicious email(s):
    - Determine if crime/violation of policy has been committed.
  - Investigate attachments.
- Verification and validation
  - Email route - may include clues about sender’s origin, location, methods.
  - Analyze domain name’s point of contact.
  - Aggregate suspect’s contact information.
  - Acquire attributes against network logs.

Review:

Topic 7: Mobile Forensics

What is Mobile Forensics?

- A branch of digital forensics relating to recovery of digital evidence or data from a mobile device under forensically sound conditions.
- Involves recovering data specific to mobile platforms.
- Can refer to any device with internal memory and communication ability, like PDA or GPS devices.
- There are multiple methods / tools for data extraction, and no single method is best.

What data is obtainable?

- FROM SIM Cards:
- IMSI: International Mobile Subscriber Identity
- ICCID: Integrated Circuit Card Identification (SIM Serial No.)
- MSISDN: Mobile Station Integrated Services Digital Network (phone number)
- LND: Last Number Dialled (sometimes, not always, depends on the phone)
- SMS: Text Messages, Sent, Received, Deleted, Originating Number, Service Center (also depends on Phone)

What data is obtainable?

- Phonebook
- Call History and Details (To/From)
- Call Durations
- Text Messages with identifiers (sent-to, and originating) Sent, received, deleted messages
- Multimedia Text Messages with identifiers
- ***Photos and Video (also stored on external flash)***
- Sound Files (also stored on external flash)
- Network Information, GPS location
- Phone Info (CDMA Serial Number)
- ***Emails***, memos, calendars, documents, etc. from PDAs.
- ***Facebook Contacts, Skype, YouTube data, Username and Passwords***
- Location from GPS, Cell Towers and Wi-Fi networks

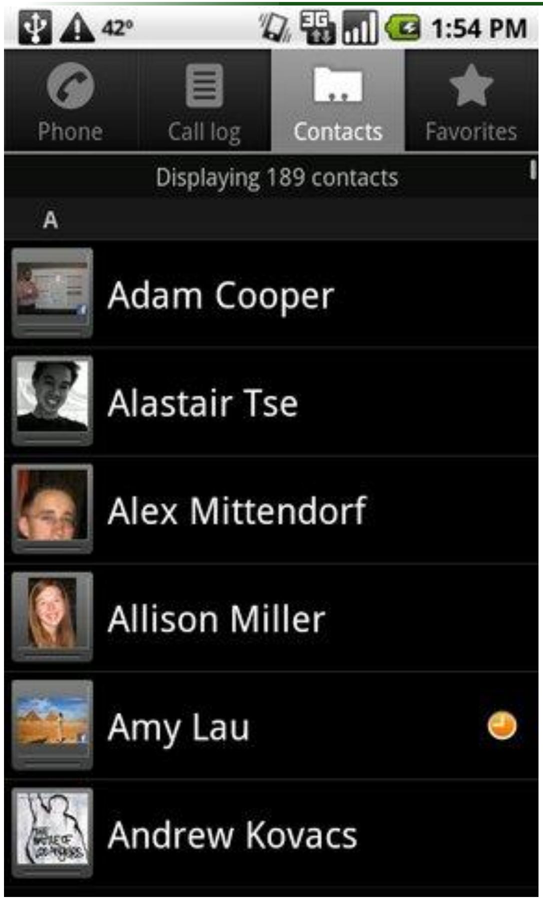
Mobile Forensics Process

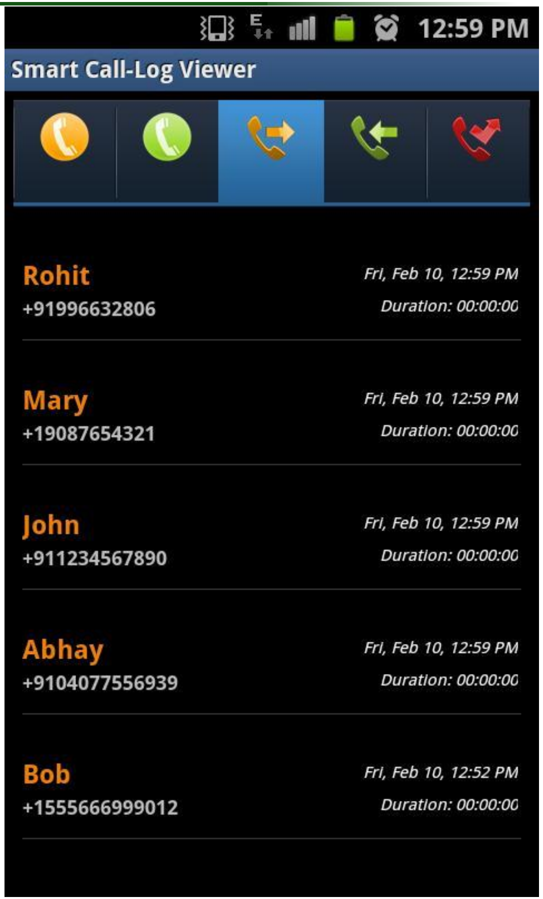
- Differences and Challenges
  - Lose – Lose – Lose situation:
    - Investigator does not alter device state after seizure to ensure data integrity.
      - Suspect uses remote wipe to erase evidence.
    - Investigator uses Faraday Bag to block communications
      - Battery is drained causing device to power down.
    - Investigator switches device to Airplane mode.
      - Memory is slightly changed.

Acquisition Techniques

- Manual Acquisition:
  - Manually interfacing with the device.
- File System Acquisition:
  - Can obtain some deleted data through synchronization.
- Physical Acquisition:
  - Bit-by-bit copy of the device’s flash memory / disk.

Manual Acquisition

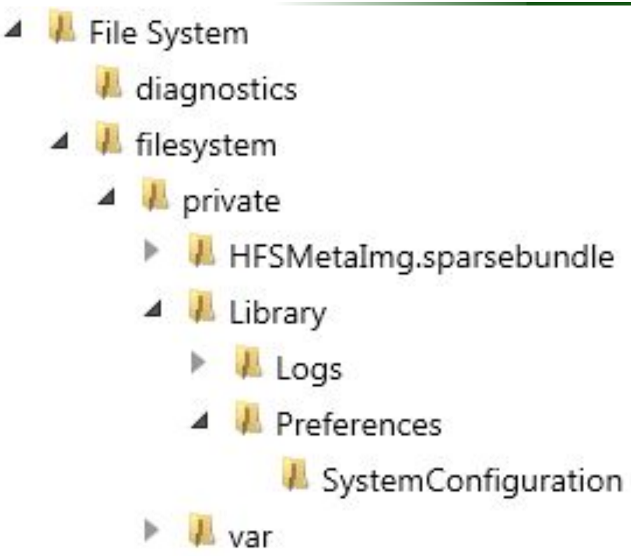


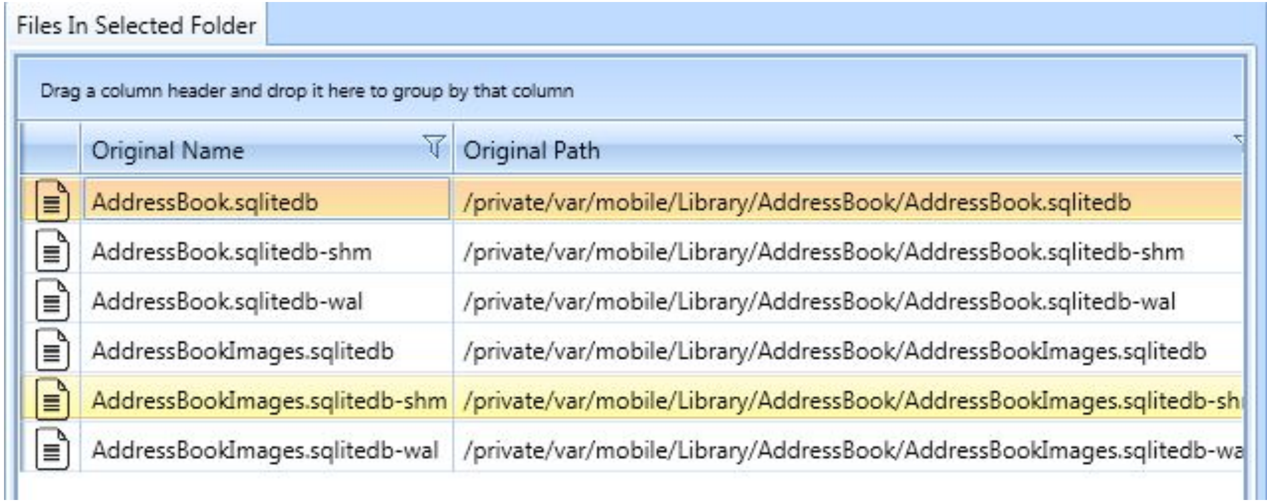


Manual Acquisition and Analysis

- Pros:
  - No prior setup / external tools required
  - Easily performed
- Cons:
  - Very slow at extracting large quantities of information.
  - Compromises data integrity
  - Can be halted if the device is locked.
  - Cannot recover hidden /deleted information.

File System Acquisition




Files In Selected Folder

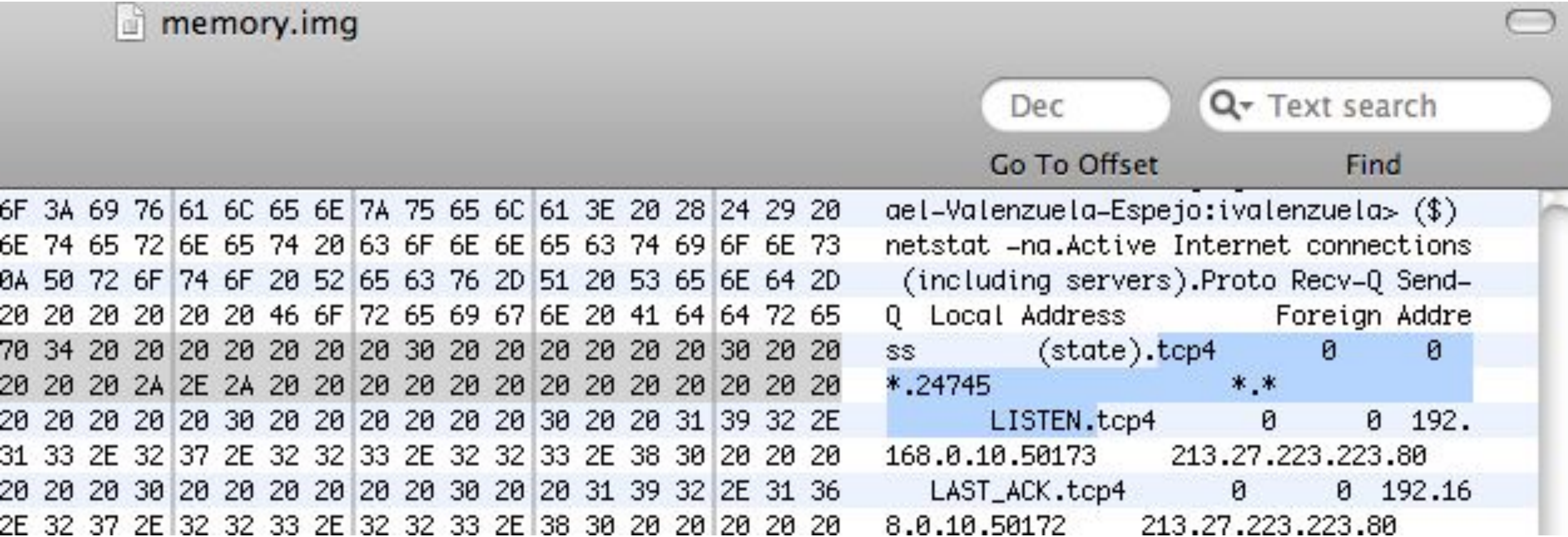
Drag a column header and drop it here to group by that column

Original Name	Original Path
AddressBook.sqlitedb	/private/var/mobile/Library/AddressBook/AddressBook.sqlitedb
AddressBook.sqlitedb-shm	/private/var/mobile/Library/AddressBook/AddressBook.sqlitedb-shm
AddressBook.sqlitedb-wal	/private/var/mobile/Library/AddressBook/AddressBook.sqlitedb-wal
AddressBookImages.sqlitedb	/private/var/mobile/Library/AddressBook/AddressBookImages.sqlitedb
AddressBookImages.sqlitedb-shm	/private/var/mobile/Library/AddressBook/AddressBookImages.sqlitedb-sh
AddressBookImages.sqlitedb-wal	/private/var/mobile/Library/AddressBook/AddressBookImages.sqlitedb-wa

File System Acquisition and Analysis

- Pros:
  - Quickly extracts large amounts of information for analysis.
  - Can recover some deleted information via database analysis – Some OS's mark data in databases as "deleted" w/o removing.
- Cons:
  - Use of this technique is limited as it requires the OS to keep track of deleted files.
  - Does not recover all deleted information.

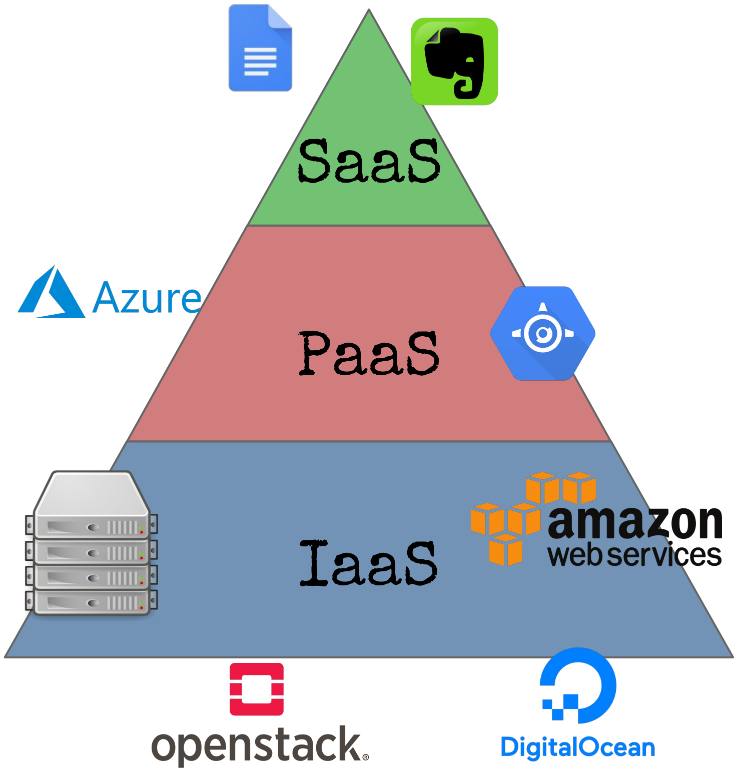
Physical Acquisition



Review:

Topic 8: Cloud and Web Forensics

Cloud Service Levels



- Software as a Service (Saas)
  - Applications are delivered via the Internet, such as Google Docs.
  - Target is the end user of an application.
- Platform as a Service (Paas)
  - OS installed on a cloud server, users can install their software and tools.
  - Target is the application developer.
- Infrastructure as a Service (IaaS)
  - Customer rents hardware, installs OS of choice. Highly configurable network options. Tremendous scaling ability.
  - Target is the system administrator.

Cloud Deployment Methods

- Public Cloud:
  - Cloud services are available to anyone.
- Private Cloud:
  - Limited-access, typically on-premises.
  - Uses a cloud architecture such as OpenStack.
- Community Cloud:
  - A way to bring people together for a specific purpose.
- Hybrid Cloud:
  - A public and private cloud that talk to each other.
  - Gives companies more control over data and services.

Cyber Crimes Using the Cloud

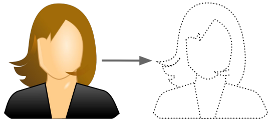
- Cloud assisted:
  - Using cloud VMs as bots or Command and control servers
  - Data breach (tool)
- Cloud targeted:
  - Cyber attack against a cloud
  - Policy violations in accessing a cloud
  - Data breach (victim)
- Cloud incidental:
  - Fraud
  - Data breach (storage)

A Framework for Web Environment Forensics

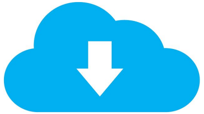
Unique Web Forensic Challenges

C0. Complying with the Rule of Completeness

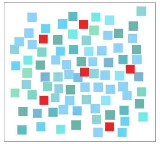
C1. Associating a suspect with online personas

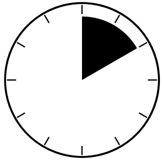


C2. Gaining access to the evidence stored online

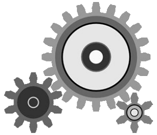


C3. Contextualizing evidence in terms of content (*thematic context*) and time (*temporal context*)





C4. Integrating tools to perform advanced analyses



Framework

F1. Evidence Discovery and Acquisition

- Connect suspect and persona (**C1**)
- Gain access to evidence from web services (**C2**)*

F2. Analysis Space Reduction

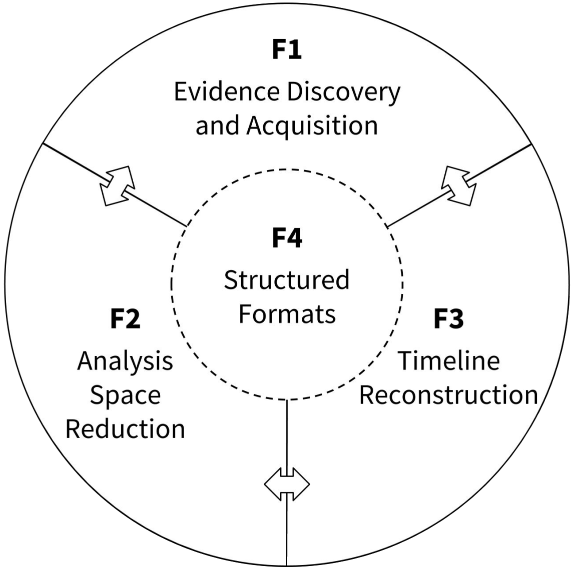
- Filter irrelevant artifacts (**C3** *Thematic* Context)*

F3. Timeline Reconstruction

- Reconstruct timeline (**C3** *Temporal* Context)*

F4. Structured Formats

- Bridges the other three components
- Facilitate tool interoperability (**C4**)



	F1	F2	F3	F4
C0: Rule of Completeness	●	●	●	○
C1: Associating Personas	●	○	○	○
C2: Evidence Access	●	○	○	○
C3: Relevant Context	○	●	●	○
C4: Tool Integration	○	○	○	●

* Also addresses **C0**: Rule of Completeness

Considerations for Forensic Investigations in the Cloud

Legal Challenges

- Service Level Agreements (SLAs):
  - Among other things, these state **who** is authorized to access data and **what the limitations are** in conducting acquisitions for an investigation.
- Jurisdiction issues:
  - Perpetrator, victim, and instrument of the crime can all be in **different locations** with **different laws** applying to each in **different ways**.
- Accessibility:
  - **Search Warrant**: Used only in **criminal** cases, requested by **law enforcement** with probable cause of a crime. Used to **seize hardware**.
  - **Subpoenas and Court Orders**: Used when **information** (or **data**) is needed, not the original equipment.

ARIZONA STATE UNIVERSITY

Technical Challenges (1)

- Cloud architectures vary:
  - No two providers are alike.
- Data collection and authentication:
  - Remote acquisitions are hard.
  - Virtual network switches == duplicate IPs, IP spaces.
  - Encrypted data (now common) requires cooperation of cloud provider to access the data.
- Analysis of cloud forensic data:
  - Verifying integrity, reconstructing timeline is even harder.

Technical Challenges (2)

- Anti-forensics:
  - Myriad ways for criminals to undermine evidence collection and analysis.
- Incident first responders:
  - Will they be cooperative, well-trained, and capable?
- Role management:
  - Who has what roles (owner, user, etc.)?
- Standards and training:
  - Never-ending struggle to keep up with current technologies and approaches.

Levels of Investigation

- Cloud Service Provider (CSP):
  - Requires detailed knowledge of the cloud's topology, policies, data storage methods, and devices available.
- Cloud customers:
  - Data may be stored on computers, mobile devices, in web browser cache, etc.
- Locally-stored cloud data:
  - Popular cloud storage services have sync clients that leave artifacts even when uninstalled.
  - May include info about files that were never synced.