EECS 388

Introduction to Computer Security

Lecture 22:

Digital Forensics

November 14, 2024
Prof. Halderman

Forensics

Digital forensics is the process of preserving, identifying, extracting, documenting, and interpreting data in order to investigate past actions or obtain legal evidence. Used to **investigate crimes**, **recover from attacks**.

Four stages of forensic analysis:

1. Identification

Identify specific objects that store important data for the case analysis.

2. Collection

Preserve evidence, establish chain of custody, ensure data stays intact and unaltered.

3. Analysis

Examine the information stored on digital evidence and conduct an analysis of the incident.

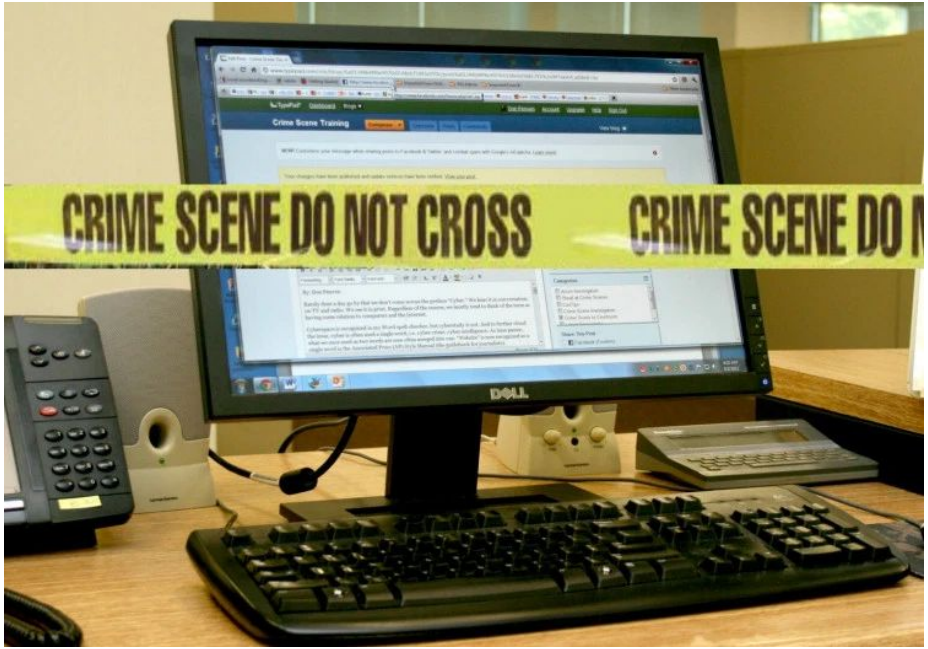
4. Reporting

Interpret findings, prepare and deliver an expert report and/or testimony.

Identification

You are the investigator, which objects do you think will be useful for investigations?

1. Computer (case and power supply)
2. Just the hard drive (without computer)
3. Monitor
4. Keyboard and mouse
5. Media (CD, DVD, USB drives, etc.)
6. Printer



Answer: All of the above!

Digital forensics does not replace traditional forensic analysis.

Collection

When collecting evidence, must take care not to *change* the evidence.

- Information on digital media is easily changed. Once altered, impossible to prove the original state.
- Computer or media is the "crime scene." Once evidence is contaminated, it can't be decontaminated.
- Examining a live system changes state of the evidence.
- Instead, work with a **forensic image** (carefully created copy) of the data.

Principles for collecting evidence:

Maintain a **chain of custody:**

- Physically secure items of evidence.
- Track possession step-by-step.
- Keep documentation (e.g., hash of image) to allow you to trace evidence back to the source.

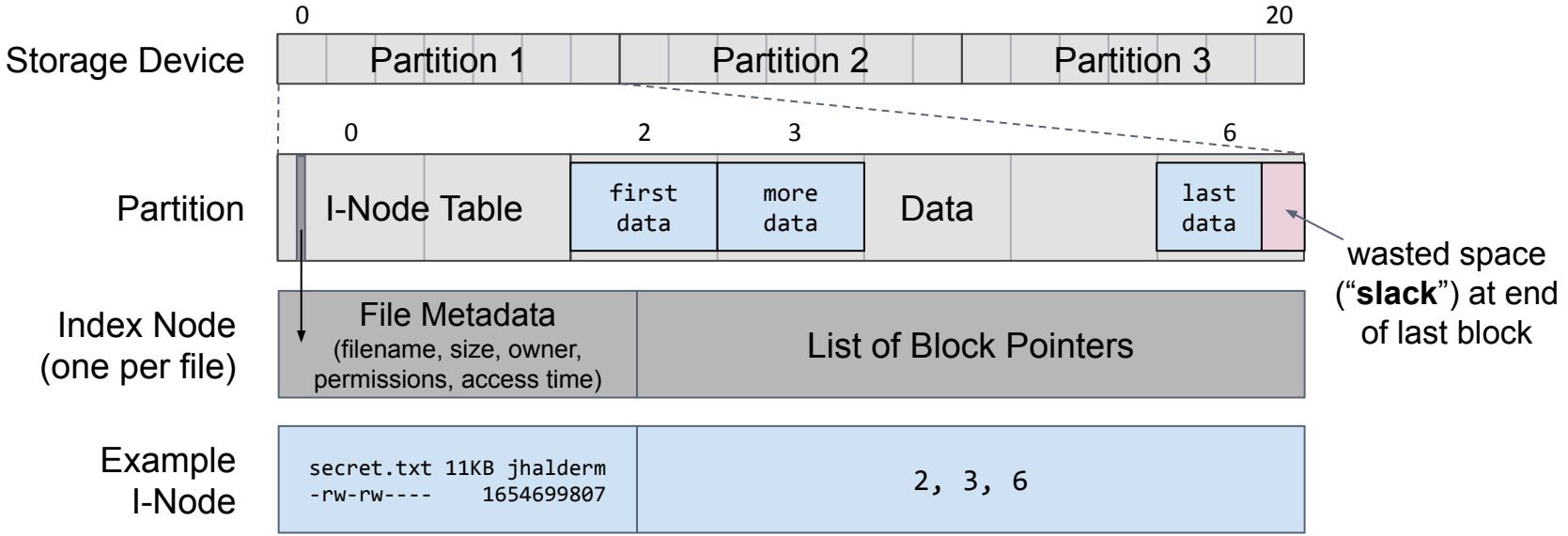
Prioritize collection by **volatility:**

- Some data is more **volatile**.
- **RAM > disk > external media**
- General idea: Capture more volatile evidence first. **[Why?]**

How Data is Stored on a Disk

Low-level storage devices present as arrays of fixed-sized **blocks** (typically 4 KB).

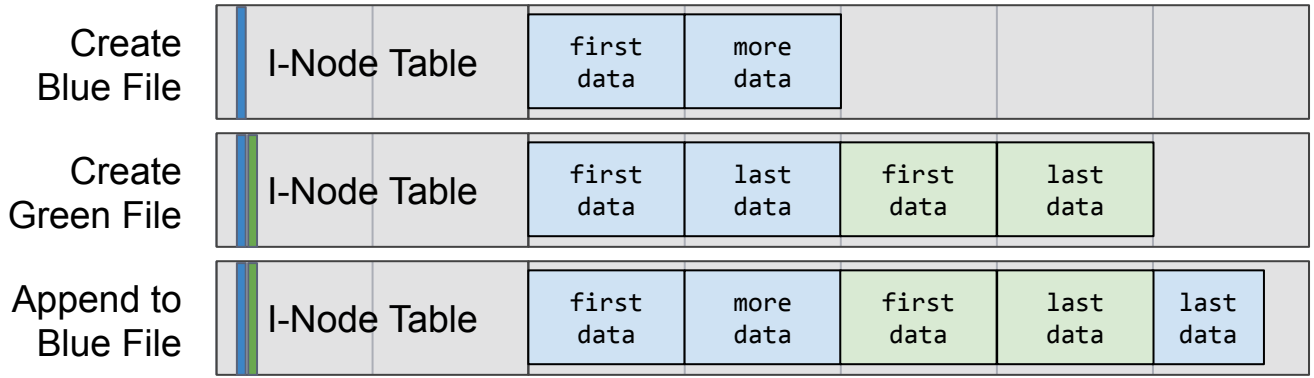
A **filesystem** organizes these blocks to provide abstractions like files and directories.



(Greatly simplified. Details vary by OS and kind of filesystem.)

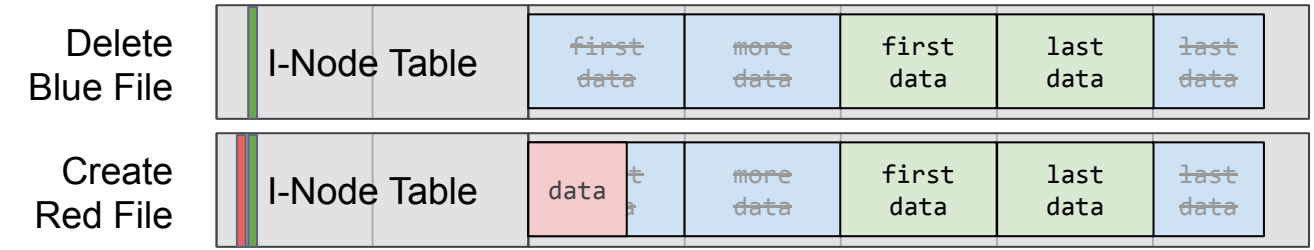
Forensic Clues in Low-Level Data

Low-level filesystem layout often contain important forensic clues.



Fragmentation shows some other data was there first.

Deleting file removes only the i-node! Residual data remains until overwritten by new files.



All blocks of deleted file still fully intact.

Part of first block still recoverable in slack space.

Collection: Forensic Images

A **forensic image** of a storage device is a file that contains a bit-for-bit copy of every block.

A file copy does not recover all data areas for examination, but a forensic image preserves all partitions and residual data (slack space, deleted files).

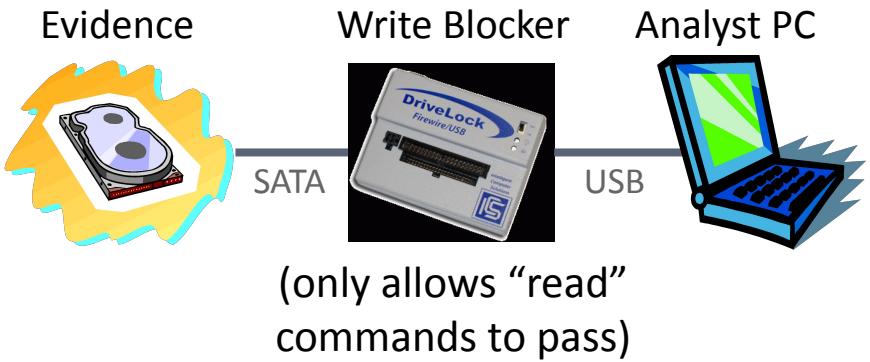
Preserves original evidence and allows recreation from duplicate if necessary.

Easily created in Linux:

```
$ dd if=/dev/sdc of=image.bin
```

More advanced: **`ewf-tools`** package

Don't use tools that write to the disk! Create image using a **write-blocker** device, ensures media is not modified by your PC.



(only allows "read" commands to pass)

Analyst should record original hash to later confirm image wasn't changed.

Collection: Imaging RAM

Live-memory forensics also considers the contents of RAM.

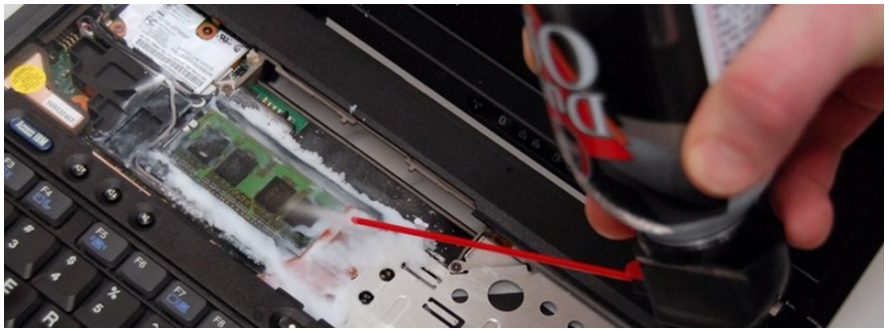
> Can be essential for decrypting data on disk, recovering passwords, or spotting in-memory malware.

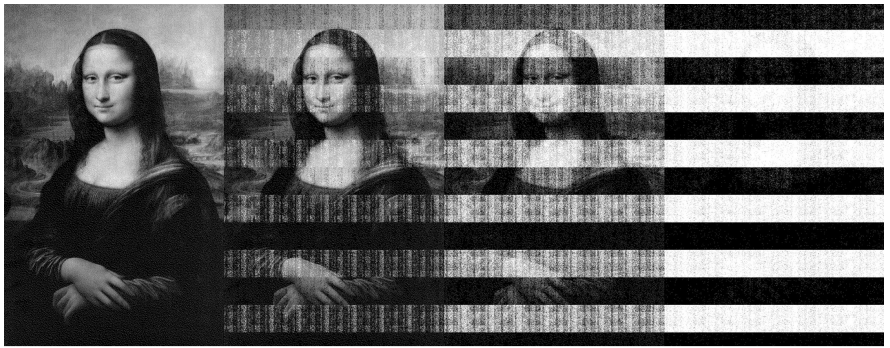
Specialized devices can image RAM by exploiting vulnerabilities in Thunderbolt.

Virtual machines can be snapshotted in order to image RAM and disk simultaneously.

Cold-boot attack: Systems can be reset and booted with special-purpose software designed to image RAM. (Typically, RAM not erased except when the normal OS loads.)

If unable to boot special OS, freeze memory chips and move them to different machine!





Collection: Mobile Devices

Mobile devices present special forensics challenges, due to radio connectivity and advanced security features.

Defeat remote wiping by placing device in a Faraday bag to shield RF signals.

Arms race: Mobile device makers implementing strong encryption, hardware-backed security.

Forensics companies make specialized devices to exploit vulnerabilities and recover data. (e.g., Cellebrite, GrayKey).

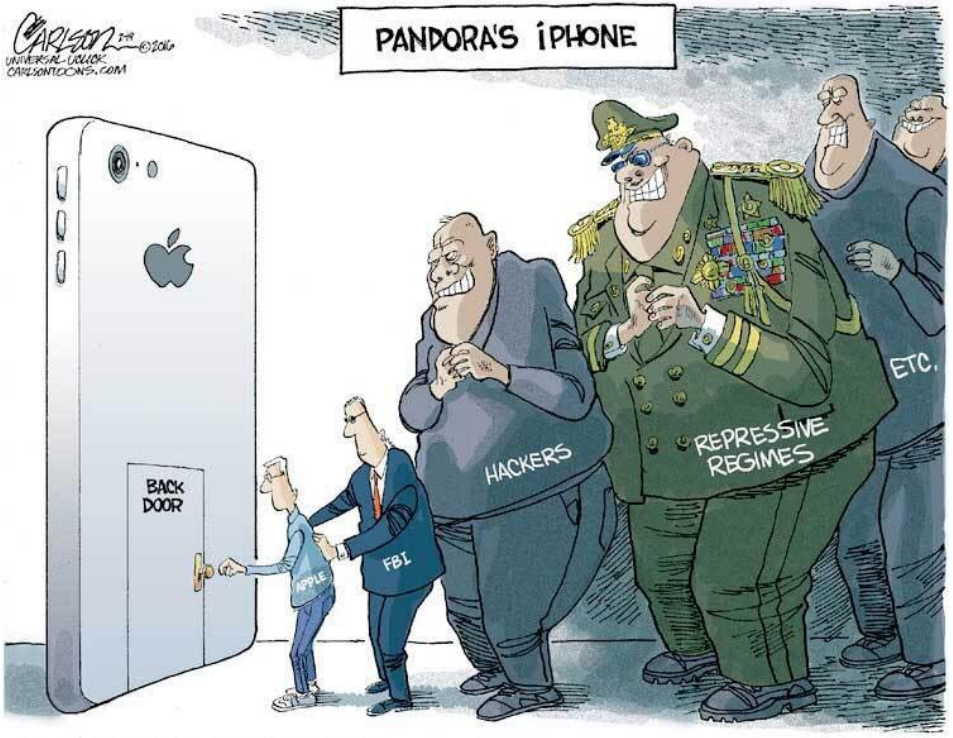
Latest device models/firmware may be unrecoverable, but probably not for long…

Encryption and Public Policy

The New York Times

Barr Revives Encryption Debate, Calling on Tech Firms to Allow for Law Enforcement

The attorney general, reopening the conversation on security vs. privacy, said that encryption and other measures effectively turned devices into "law-free zones."



Common Collection Mistakes

What is the first step to collect evidence, when you find:

- A computer/device **turned on**?
- A computer/device **turned off**?

Any device at a crime scene is potentially **fully adversarial!**

Analysis and Evaluation

Forensic analysis is “detective work.”

- Define goals for analysis (what you are looking for)
- Know where evidence can be found
- Follow clues and reconstruct the incident piece by piece
- Understand techniques used to conceal or obscure data (e.g., encryption, steganography).
- Carefully document steps taken and results obtained

Many places to look for evidence:

- Apps (email, chat, photos, …)
- Browsing history and cache
- Existing files (including “hidden” files)
- Deleted files / slack space
- System logs
- Configuration files / Registry
- Temp files
- Backups
- Hibernation and swap files
- Alternate or hidden partitions
- Accessible cloud accounts, if permitted

You’ll conduct a forensic analysis in Project 5

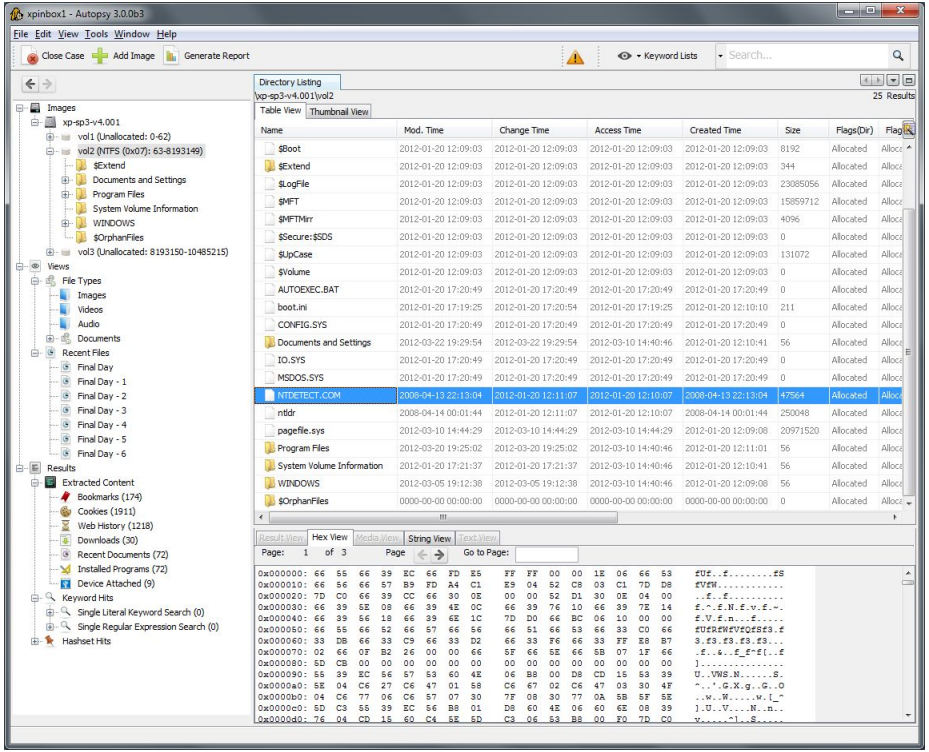
Forensic Analysis Software

Digital forensics tools help analysts manage cases and process large volumes of collected data

- Designed to work with forensic images
- Support quickly searching for relevant evidence, ignoring irrelevant data
- Toolbox of techniques to discover hidden data and recover deleted files

Examples: Encase, **Autopsy** (open source)

Others important kinds of tools:
password crackers, virtual machines



Autopsy *(tutorial during this week's lab)*

Techniques for Hiding Data

Encryption

Encodes data so it can't be read without a key.

Obfuscation

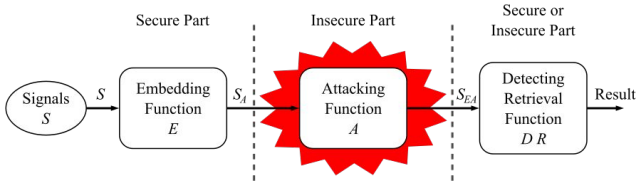
Transforms code or binaries to conceal their purpose or mode of operation, while preserving their function.

```
if(t){var A=t.indexOf(":"),O=-1!=A,ha=O?Q(t.substr(0,A)):"click";t=O?Q(t.
substr(A+1)):t;p[ha]=t}}u[v]=p}r.__jsaction=p}else p=ia,r.__jsaction=p}r=p;
"maybe_click"==m&&r.click?(n=m,m="click"):"clickkey"==m?m="click":"click"!=
m||r.click||(m="clickonly");n={m:n?n:m,action:r[m]||"",event:null,s:!1};f=R
(n.m,n.event||a,e,n.action||"",k,f.timeStamp);if(n.s||n.action)break}f&&"to
uchend"==f.eventType&&(f.event._preventMouseEvents=ba);if(n&&n.action){if(e
```

Commonly used by malware, but also by some legitimate software to prevent reverse engineering.

Watermarking

Add a durable mark to a file that resists removal attempts. Uses include tracking documents, combating "piracy".



Steganography

Encodes hidden data inside other data so people don't even suspect it's there.

Example: Hide message in least-significant bits of a bitmap image or sound file.


Original JPEG image (Prof. Honeyman's kids)


With message hidden in LSB of JPEG coefficients

Tools like stegdetect can analyze images with statistical tests to detect some forms of steganography.

Decoding reveals a second image

Reporting

The output of a forensic investigation is often a forensic report documenting the results of the analysis.

A well written report will:

- Accurately describe the details of an incident.
- Document the process the analyst followed.
- State the evidence that was uncovered.
- Draw conclusions founded on the evidence.
- State the level of confidence for each conclusion.
- Offer recommendations for further action when appropriate.
- Be understandable to decision makers.
- Be able to withstand legal scrutiny, when applicable.

Example: Forensic Report

Analysis of the Antrim County, Michigan November 2020 Election Incident

J. Alex Halderman

Description	SHA-256 Hash of Forensic Image
Computer hard drive:	
County EMS	1d0d7248a0d1db99051a164766a08c895f67f358a58046102e06c20ad4785d81
Ballot scanner memory cards:	
Banks	784ccc460346ba85554c4798f9a1711cd73c860eaea58fa458ac241b049d2510
Central Lake	5bd0798b4a21edd390bee784519764fccc4369bdda6dbe1cafbdf28c11a098bd
Chestonia	48a55e328dcf1816b42a0163a334bb4cf35fa964f1c5460ebc7a4b3fe1a2a474
Custer	1cc9a044a69567a7a38f45892b91c32a4acf631699ad1b4d9d9fbfe72e28e433
Echo	371eb895e922cd2d36cf1859c1d84df01e6fd9176132ec86c889930a20c1a8f8
Elk Rapids 1	ad69dfcadf17b5bb3a744417daae2251aca0f19ccf34af7b5e732a688a4f68f5
Elk Rapids AV	3d4ffc1d8f3ef2b336e5934f0ab98048ad0e24c96efd027f539c68239f6cdbac
Forest Home	a93c1021367b93ebc89957b5d0c5df6828c885877b2640bf4495441bbe7df474
Helena	529bc91c0d012ef4df947898d8fbde3e0d1c5f430e374080c2e52fb29d02e565
Jordan	c2fc4e0e50ca56d55cfa9b2111b120f319be602e357dd09fa12c14203297f3d7
Kearney	2cd3fccb9640738fb062da32d3bb1ea1e4bbafc2e97ca70d87863d12cc8fd38c
Mancelona 1	35f7d069f5556ea9aed3727a0433819b6940cb3474579cf969d2cc208324fbae
Mancelona 2	8931572f6aaeff7c7f80000ca6958da172660b9e3e8d40c073c48008593aa572
Milton 1	386390a3edfa366bb12c8263825d96b51d63a9a4711685edbd7c63fe38e2ba4a
Milton AV	255e1e27daadadbfc7ecc64d1a2c9a8e4f9cf7c65c21e9f18a2ff615dc041d09
Star	9aed3328e89ac4a98ac8ff8877a99b400f092f04a71b738b3e56579384a60379
Torch Lake	60bf46c9fb769fa6a2d238bb1d34387c0854594c9b83c72723ff306dd6abd775
Warner	8132a7e3bf7ac839152d7c6f6f68e9d1316911039241a49f23489a1c2d8e1801

Table 2: **Data Sources.** These are the hashes of the forensic images I examined. The data was collected from the EMS hard drive and 18 scanner memory cards.

	Final Results				Reproduced Error		Δ	
	Biden	Trump	Jorgen.		Biden	Trump	B	T
Banks	349	756	11	(a)	349	756	0	0
Central Lake	549	906	16	(a)	549	906	0	2
Chestonia	93	197	3	←	197	3	0	0
Custer	240	521	11	←	521	11	2	0
Echo	198	392	8	←	392	8	0	0
Elk Rapids 1	784	611	5	(b)	784	611	0	0
Elk Rapids AV	202	414	12	←	414	12	0	2
Forest Home	610	753	19	←	753	19	2	0
Helena	306	431	4	←	431	4	1	0
Jordan	183	371	13	←	371	13	1	0
Kearney	471	743	16	←	743	16	1	0
Mancelona 1	276	835	20	(c)	276	835	0	0
Mancelona 2	247	646	13	(c)	247	646	0	0
Milton 1	143	478	12	(b)	143	478	0	0
Milton AV	626	543	6	←	543	6	0	0
Star	161	462	10	←	462	10	0	0
Torch Lake	462	526	7	←	526	7	1	1
Warner	60	163	3	(b)	60	163	0	0
Total	**5960**	**9748**	**189**		**7761**	**4504**	**8**	**5**

Precinct notes: (a) IDs not shifted; (b) Entered manually; (c) Used updated card.

Table 5: **Approximating the Erroneous Presidential Results.** A simple rule closely reproduces the erroneous initial presidential results. Working backwards from the final results (*left*), shift Trump's votes into Biden's column and Jorgensen's votes into Trump's (*right*), except for in precincts that were unaffected by the election definition mismatch for reasons noted. This yields totals that differ from the initial reported results by only 13 votes, or 0.1% (Δ).

Full report available here: https://www.michigan.gov/documents/sos/Antrim_720623_7.pdf

Anti-Forensic Techniques

Anti-forensic techniques try to frustrate forensic investigators and their techniques.

Difficult:

Securely deleting data, so that it cannot be restored with forensic methods. Can be very hard to remove all traces. [Why?]

Better:

Prevent certain data from being recorded in the first place. Data that was never stored cannot be recovered.

Examples:

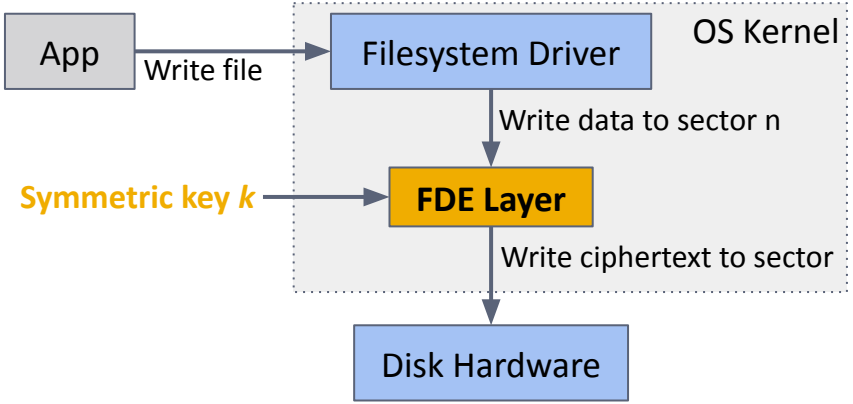
Tails Linux distro (disk read-only by default)

Browser incognito mode
(history and cache not stored to disk)

Protecting Yourself: Full Disk Encryption

Full-disk encryption (FDE) protects **data at rest** by encrypting every sector of a storage partition using a secret key.

Threat model: Device is stolen or lost, attacker has physical access but ***not*** password/PIN.



Provided by most OSes: e.g. Windows BitLocker, MacOS FileVault, similar iOS and Android features

Complication: *How to store the key?*

Approach #1: Encrypt ***k*** with user-selected password or PIN ***p***, applying a **key derivation function (KDF)**: $\mathbf{E}_{\mathrm{KDF}(\boldsymbol{p})}(\boldsymbol{k})$.

Drawback: If boot partition is encrypted, OS can't start until user enters password. (GUI? Servers?)

Approach #2: Encrypt key using a **trusted platform module (TPM)**, a tamper-resistant chip that measures boot-time code. TPM will only provide the key to OS if kernel is unmodified.

Kernel decrypts disk and boots OS, which applies user authentication and access control normally.

Drawback: Booting automatically loads key into RAM, where it's vulnerable to cold boot attacks.

Securely Erasing Media

Reliably erasing data is ***surprising hard***.

Example: **Flash**. Can't just overwrite the data:

- Writes can change a 1 to 0 anywhere, but can only change a 0 to 1 by erasing an entire region (~32 KB). Erases are slow (ms) / wear out cells.
- Flash drives have more space than OS sees. Wear- leveling algorithm assigns writes to blocks in way OS doesn't control.
- Worn blocks become unwritable. Controller remaps them to spares, but still readable with specialized hardware.
- Erased flash cells store residual charge from previous value, may be recovered in a lab.

Recommended erasure strategies:

1. **Encrypt data and discard key to erase.**

   Reduces the problem to erasing the key. Can store in special hardware (TPM, HSM, enclave).

   Works for other media, backups, cloud.

2. **Physically destroy the media.**

   Burning, grinding, degaussing (modern disks store ~2 gigabits per sq mm!)

   For flash, be sure to destroy the silicon die, not just the plastic carrier.

   Doesn't work in the cloud!

Coming Up

Reminders:

AppSec Project due today at 6 PM

Lab Assignment 5 available today, due Nov. 21

Forensics Project available today, due Dec. 5

Tuesday

Privacy

Guest Lecture by
Prof. Amrita Chowdhury

Thursday

Data center security

Remote direct memory access, kernel bypass, access control; memory introspection