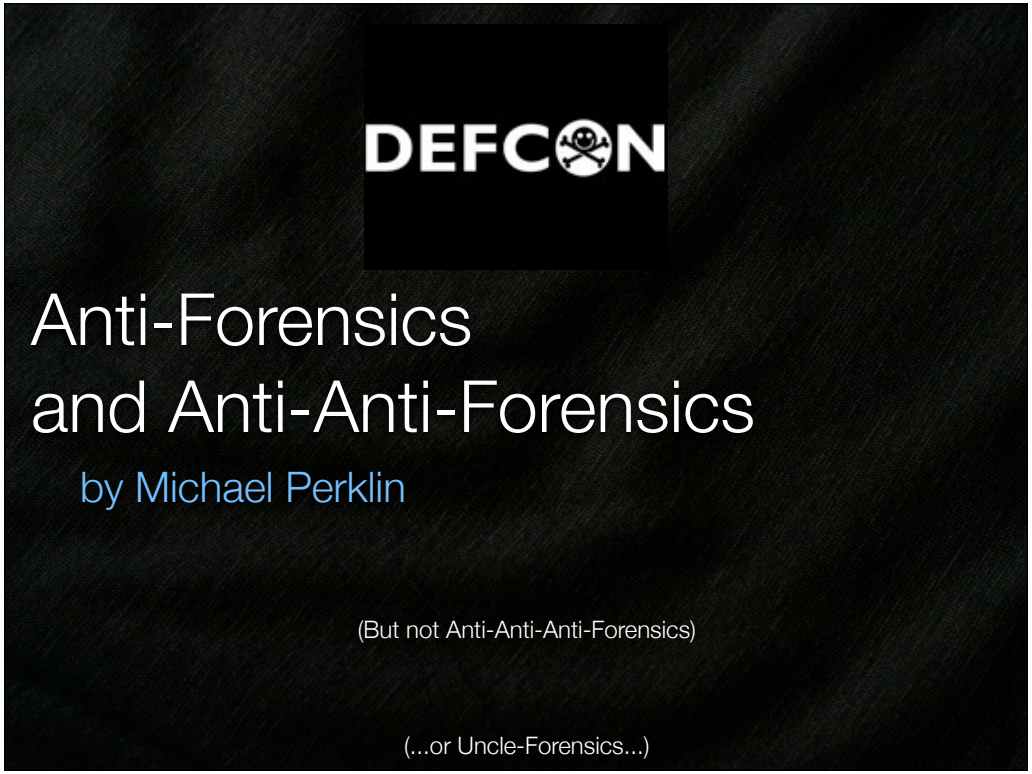
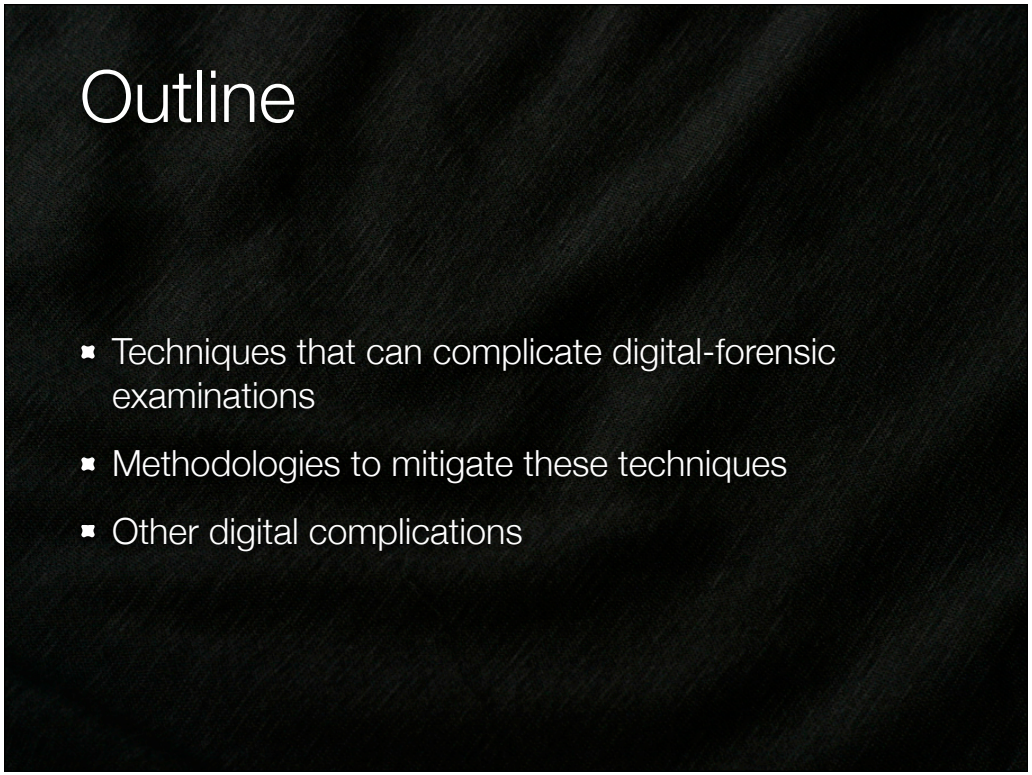

DEFCON
Anti-Forensics
and Anti-Anti-Forensics
by Michael Perklin
(But not Anti-Anti-Anti-Forensics)
(...or Uncle-Forensics...)

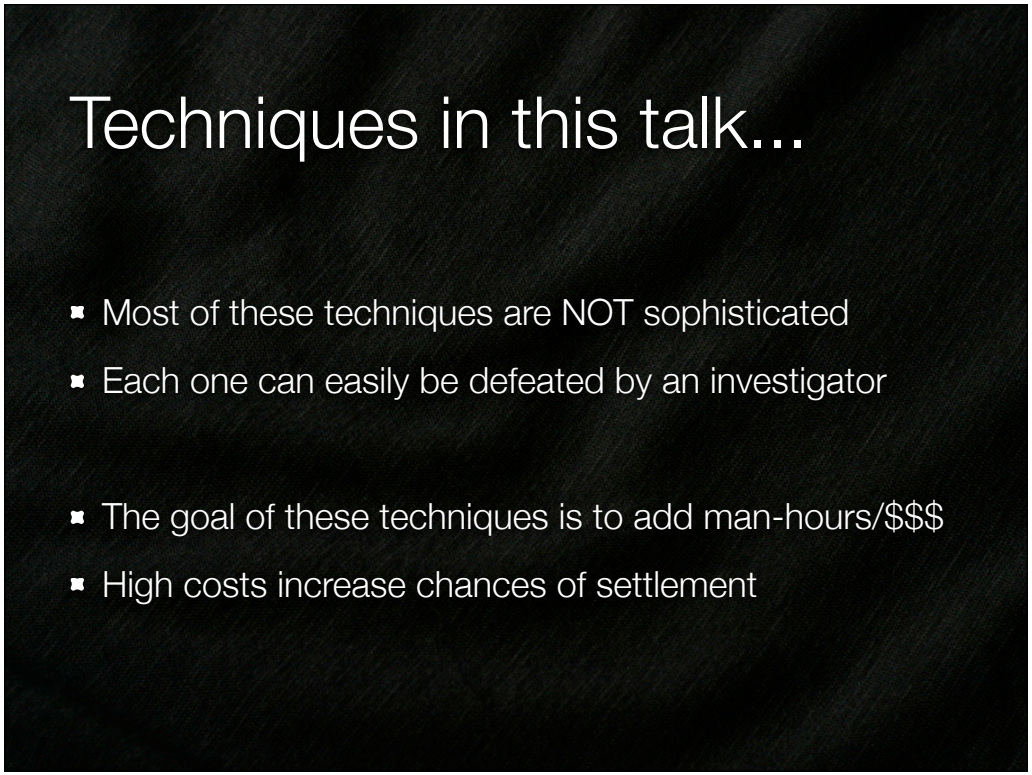


This talk will deal with a variety of complications that can arise in a digital investigation

Michael Perklin

- Digital Forensic Examiner
- Corporate Investigator
- Computer Programmer
- eDiscovery Consultant

- Basically - A computer geek + legal support hybrid



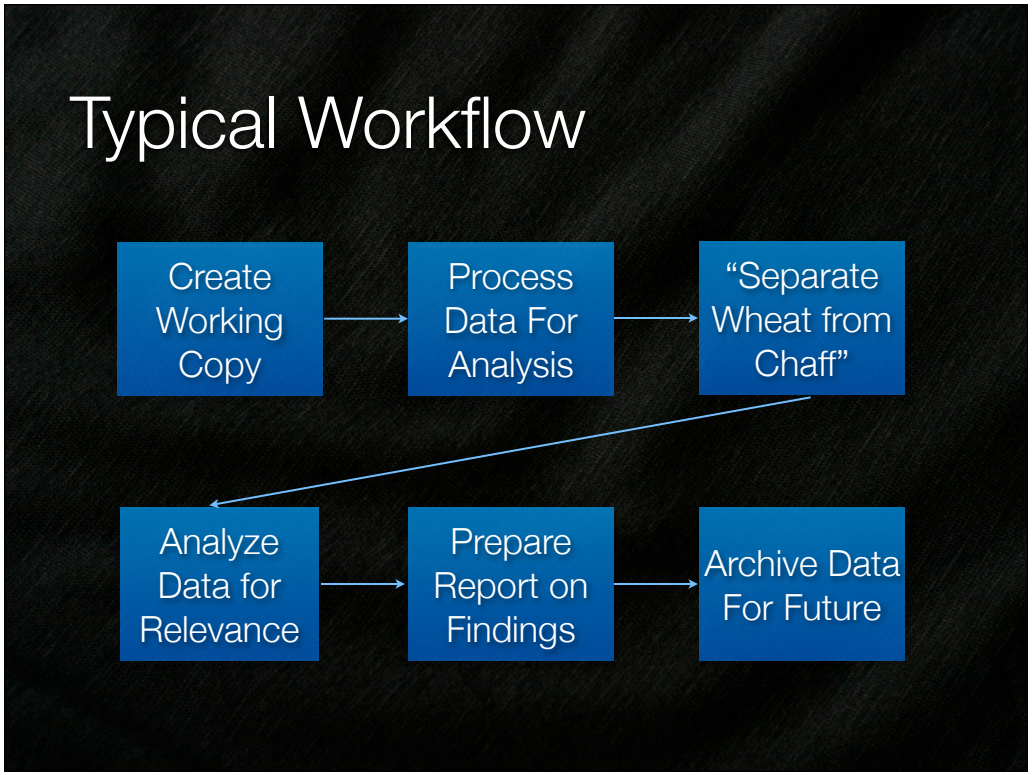
Wiping a hard drive, or using steganography will not be discussed because they've been around for decades

Typical Methodologies:

- Copy First, Ask Questions Later
  - Typically Law Enforcement
- Assess relevance first, copy relevant
  - All types of investigators
- Remote analysis of live system, copy targeted evidence only.
  - Enterprise, Private if they have help

Methodology #1 is typically used by police
Methodology #2 is used by all types of digital investigators
Methodology #3 is typically used by Enterprise companies on their employees

“Assess Relevance” method typically searches an HDD for one of a few specific keywords. If found, the HDD is imaged for further analysis.



Create Working Copy
- Image the HDD
- Copy files remotely for analysis

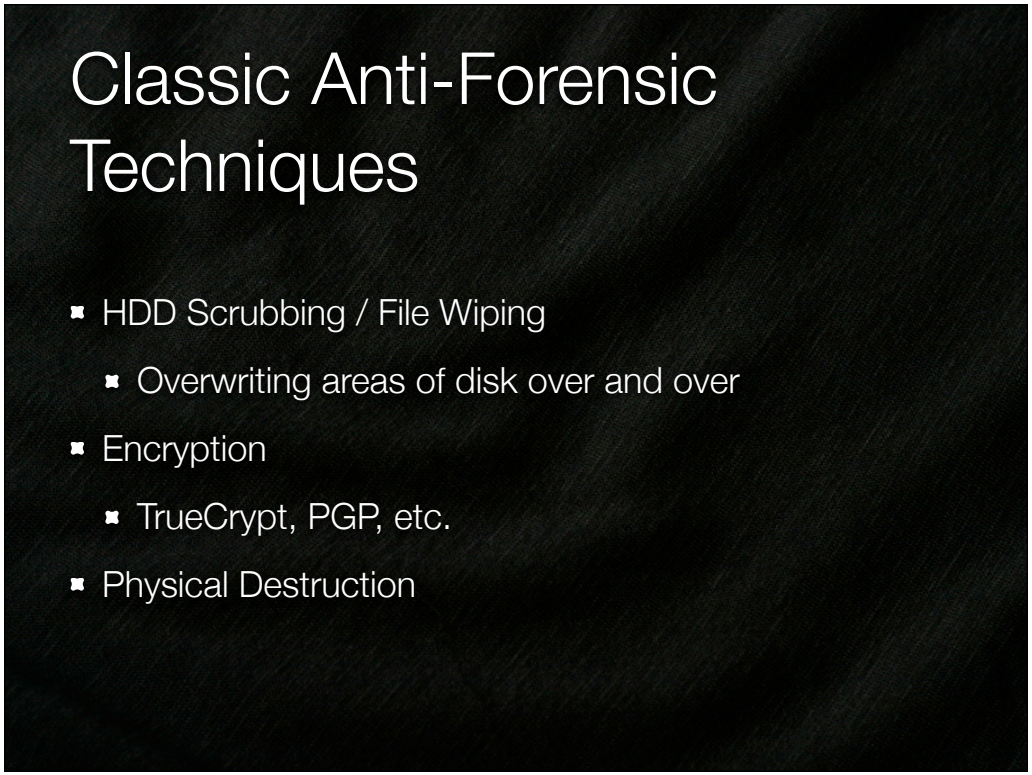
Process Data
- Hash files
- Analyze Signatures

Separate Wheat
- De-NIST or De-NSRL
- Known File Filter (KFF)
- Keyword Searches

Analyze For Relevance
- Good hits or false positives?
- Look at photos, read documents, analyze spreadsheets
- Export files for native analysis
- Bookmark, Flag, or otherwise list useful things

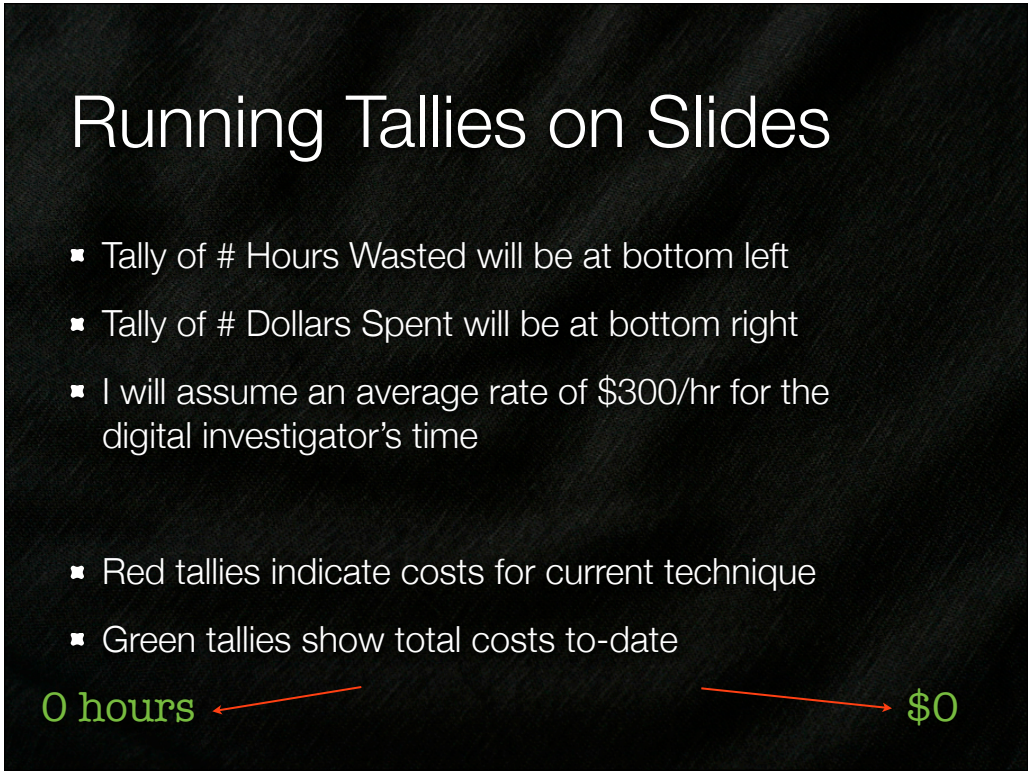
Prepare Report
- Include thumbnails, snapshots, or snippets
- Write-up procedures (Copy/Paste from similar case to speed up workload)
- Attach appendices, lists, etc

Archive Data
- Store images on central NAS
- Shelve HDDs for future use

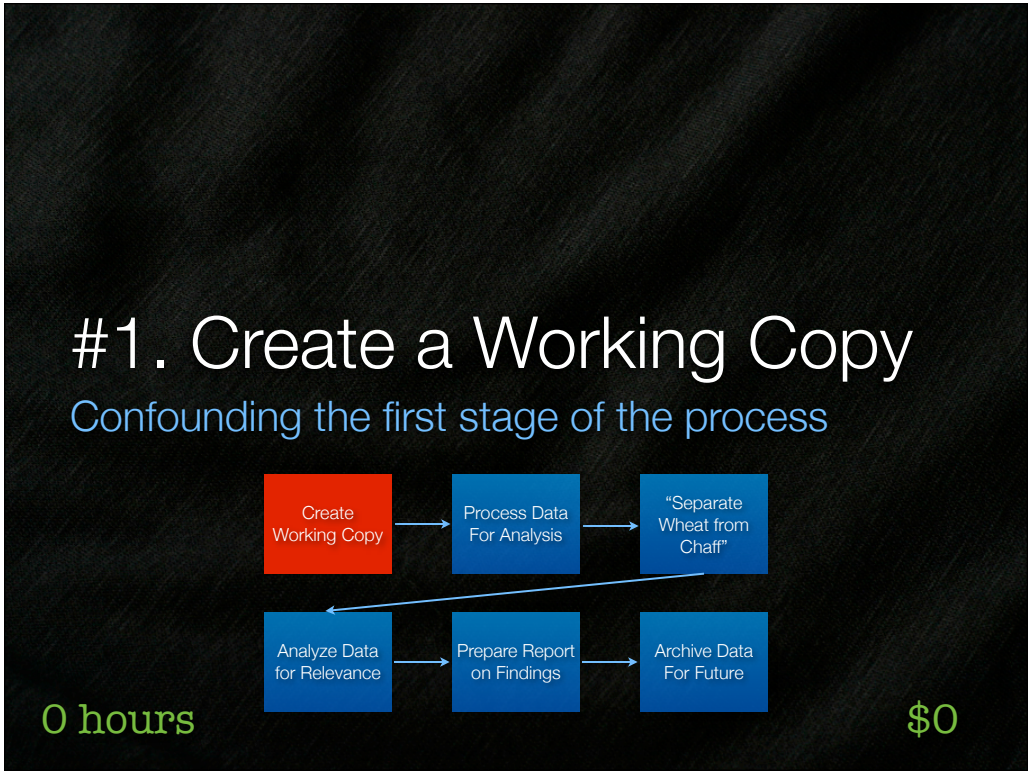


These 3 methods are fairly common amongst people like us
In reality, these are used rarely.

Each method implies guilt, and can be dealt with without tech.



$300/hr rate is fairly average for junior-intermediate investigators



Copy each device for later analysis
...or copy the file from the remote live machine

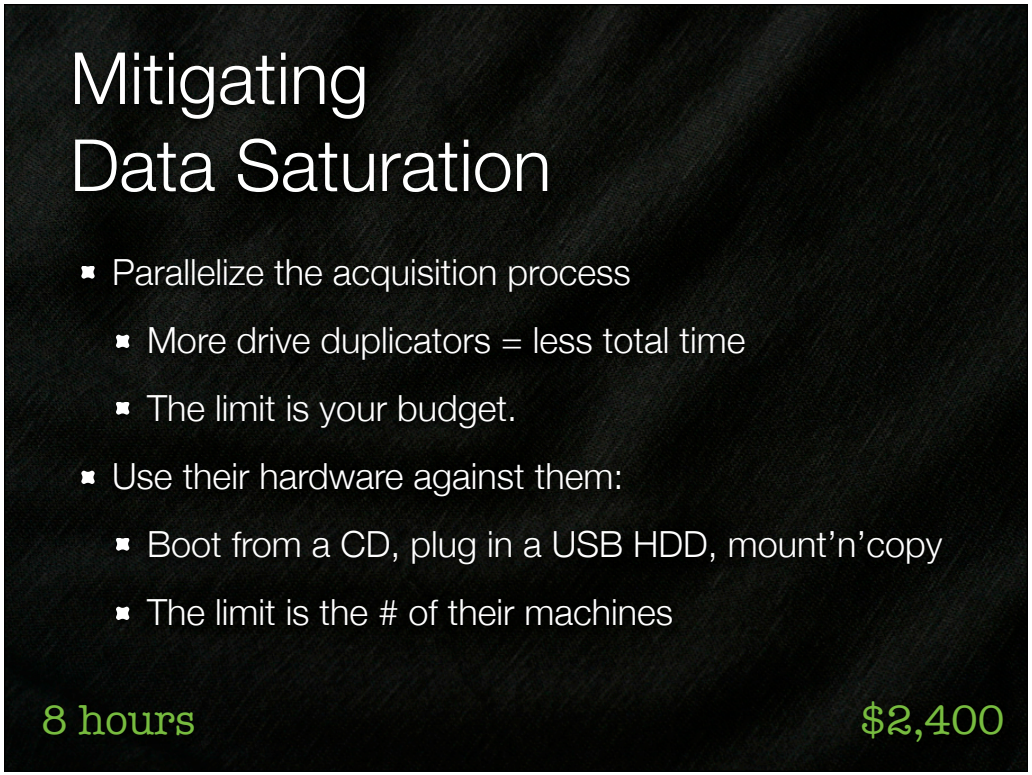
AF Technique #1
Data Saturation

- Let's start simple:
  - Own a LOT of media
  - Stop throwing out devices
  - Use each device/container regularly if possible
- Investigators will need to go through everything

8 hours

$2,400

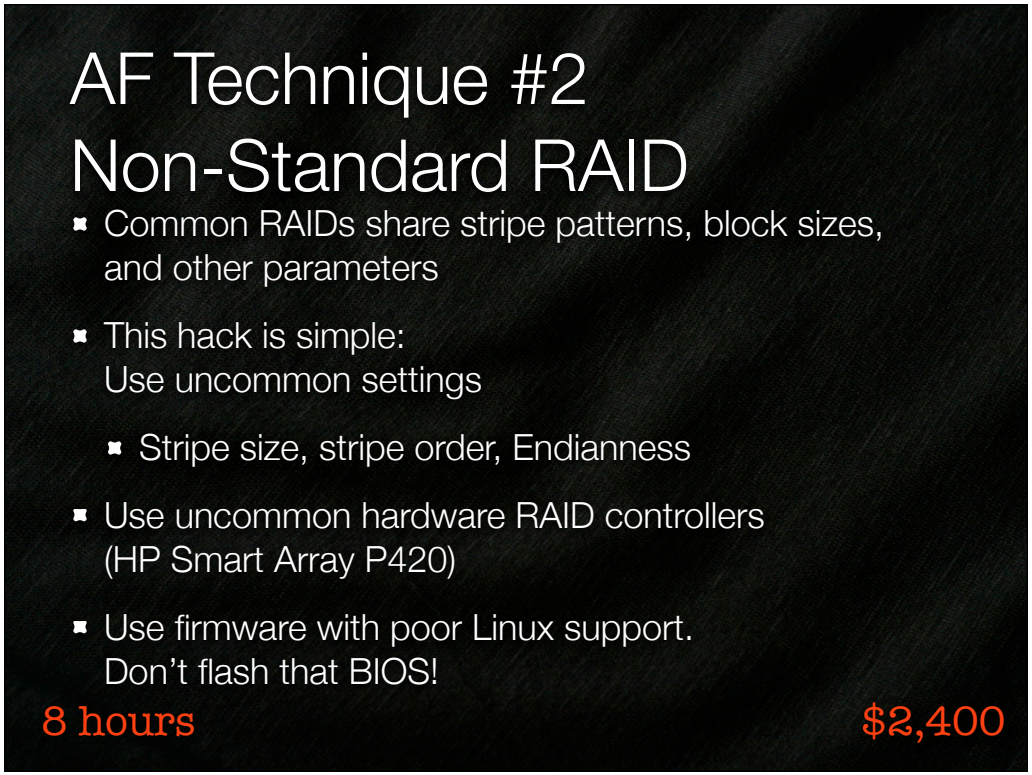
Cell Phones
Laptops
Old HDDs
USB Keys
Burned CD/DVDs



Incidentally, the # of their machines is typically equal to the number of machines you need to copy!!

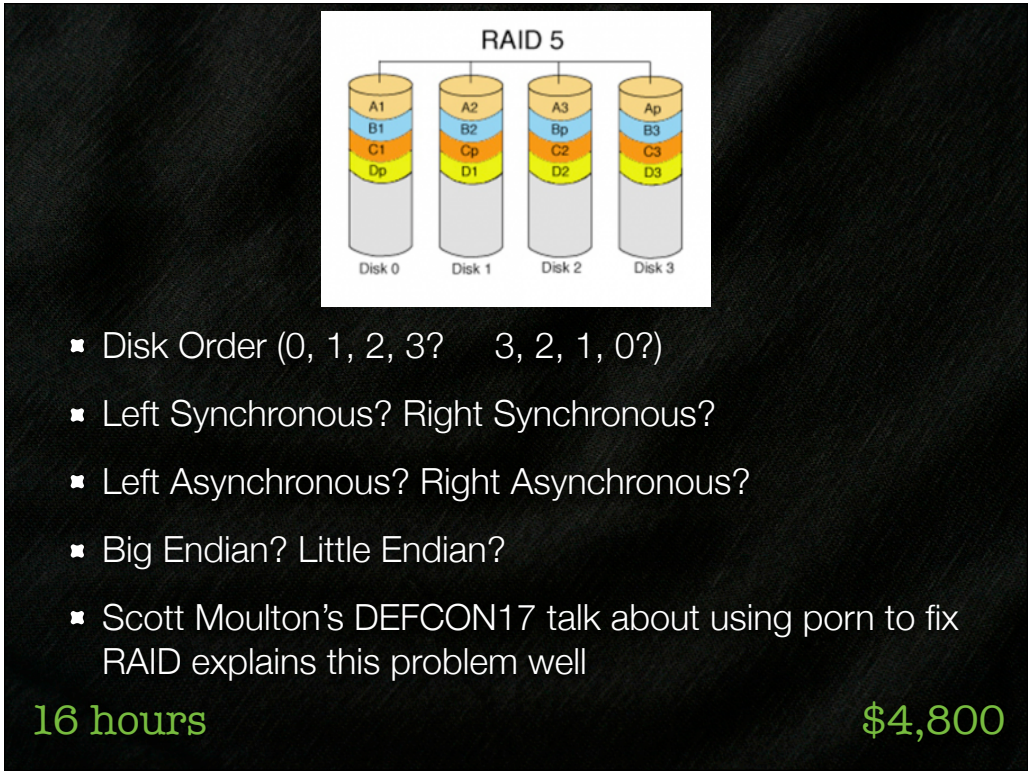
9 Machines imaging in parallel to an external USB drive.

Total time = time to image 1 drive.

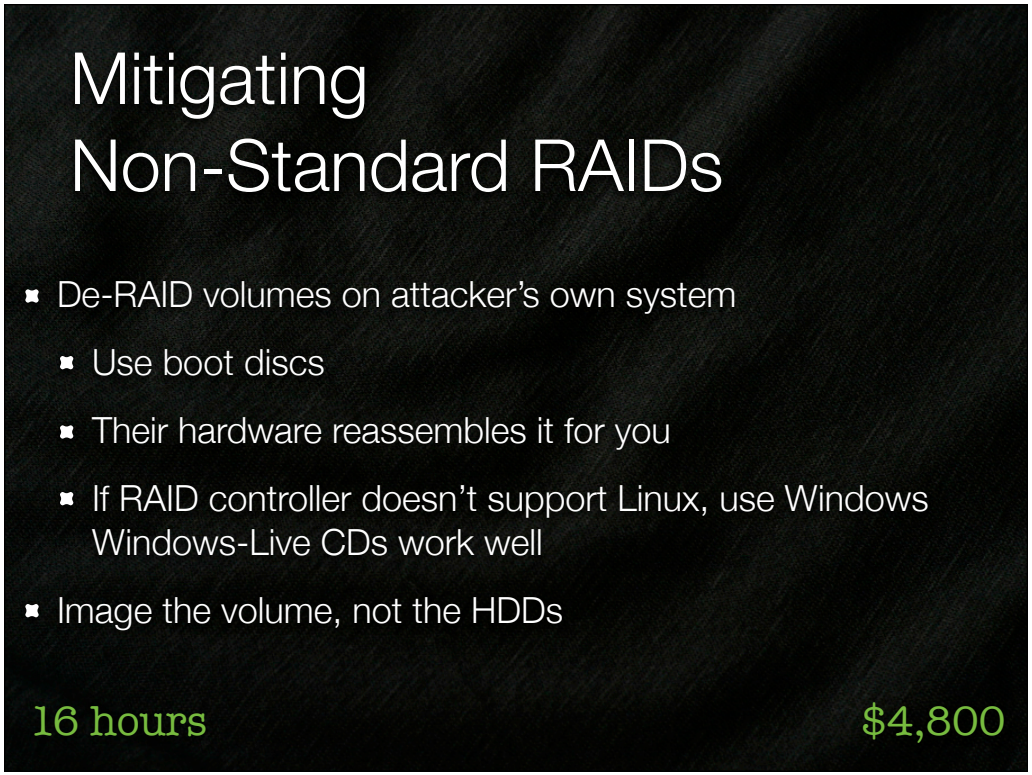


Non-standard RAID controllers sometimes allow you to choose arbitrary blocksizes
and other parameters that would otherwise be taken care of automatically.

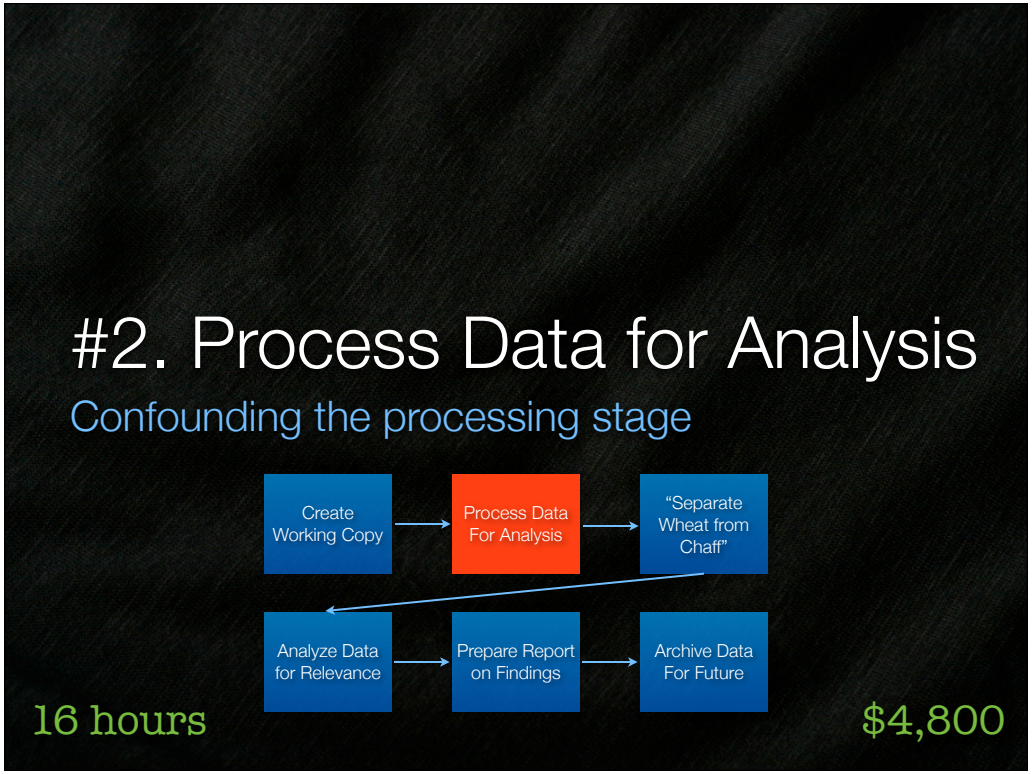
Less damaging for Public sector, can be very expensive for Private sector



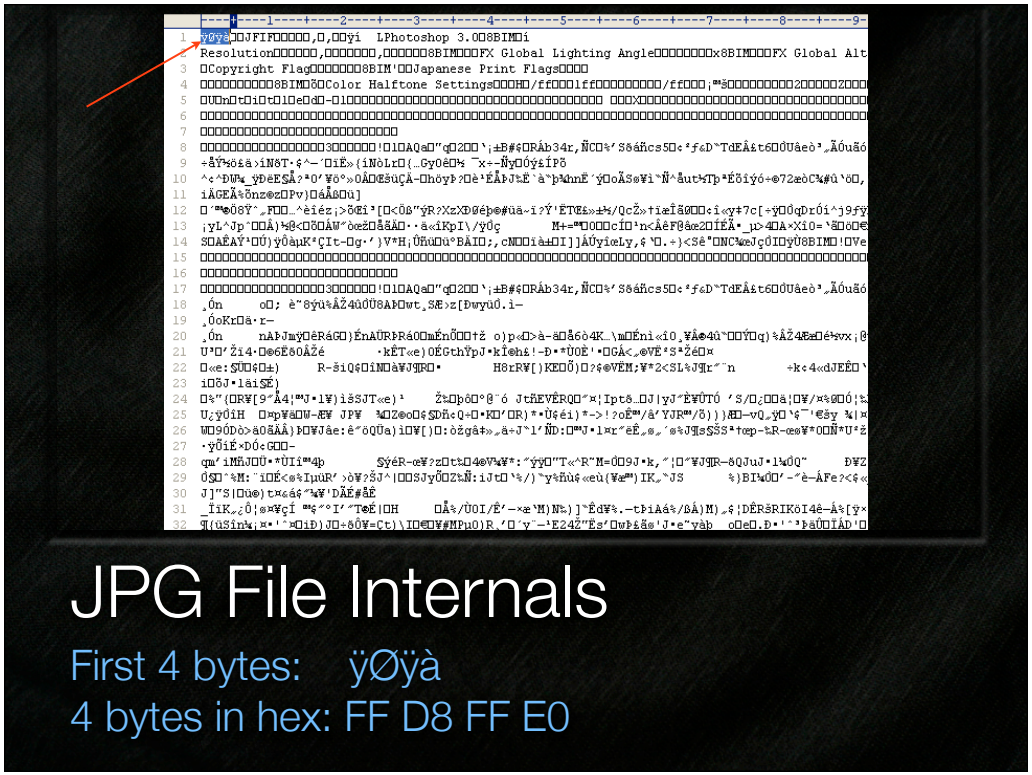
There are so many parameters used by RAID controllers that it can be quite time consuming to try all combinations in order to figure out the exact settings used by the original device



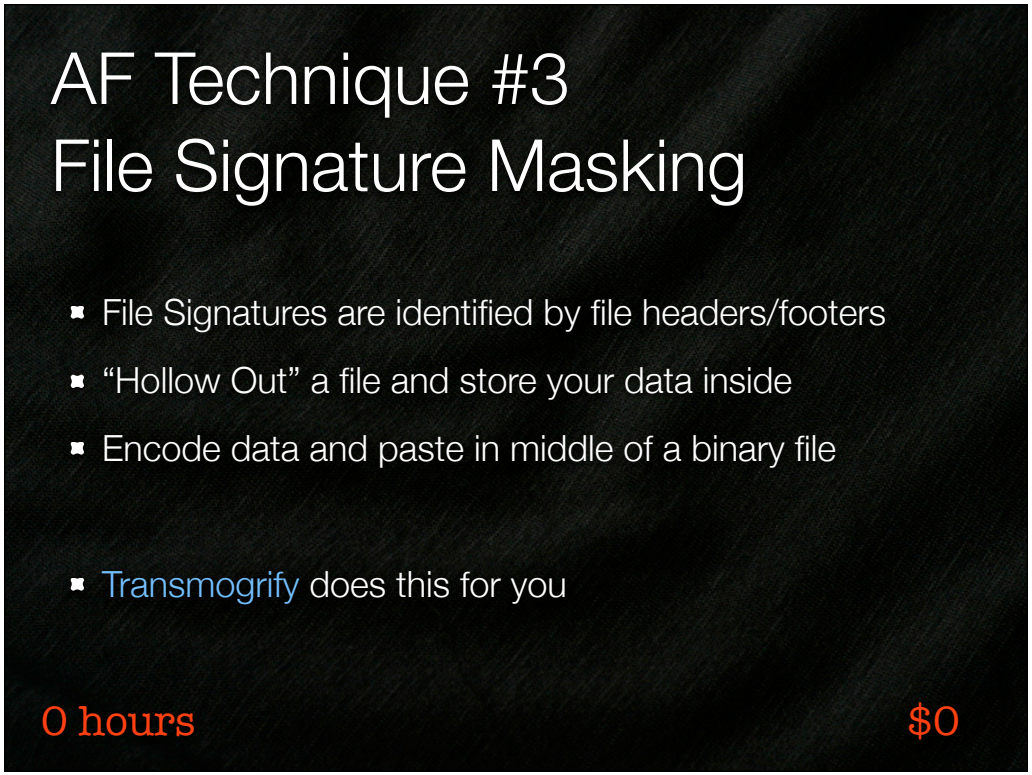
By recombining the RAID array on the attacker's system, their hardware does all the heavy lifting for you.
All you need to worry about is copying the data to your drive.



This stage involves:
Hashing
Full-Text Indexing
FileType identification
etc.



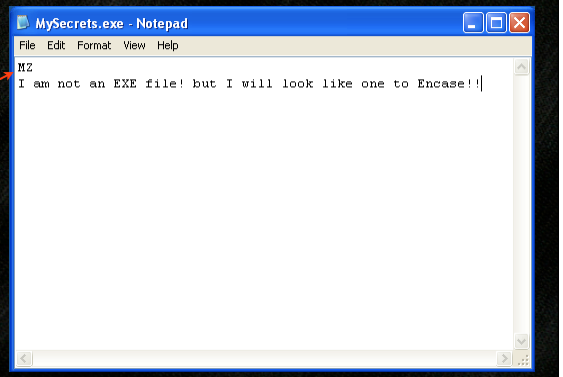
ZIP Files: PK
EXE Files: MZ
PDF Files: PDF



File signatures are identified by the first few bytes
This makes it easy to fake a file match

File Signatures (cont.)

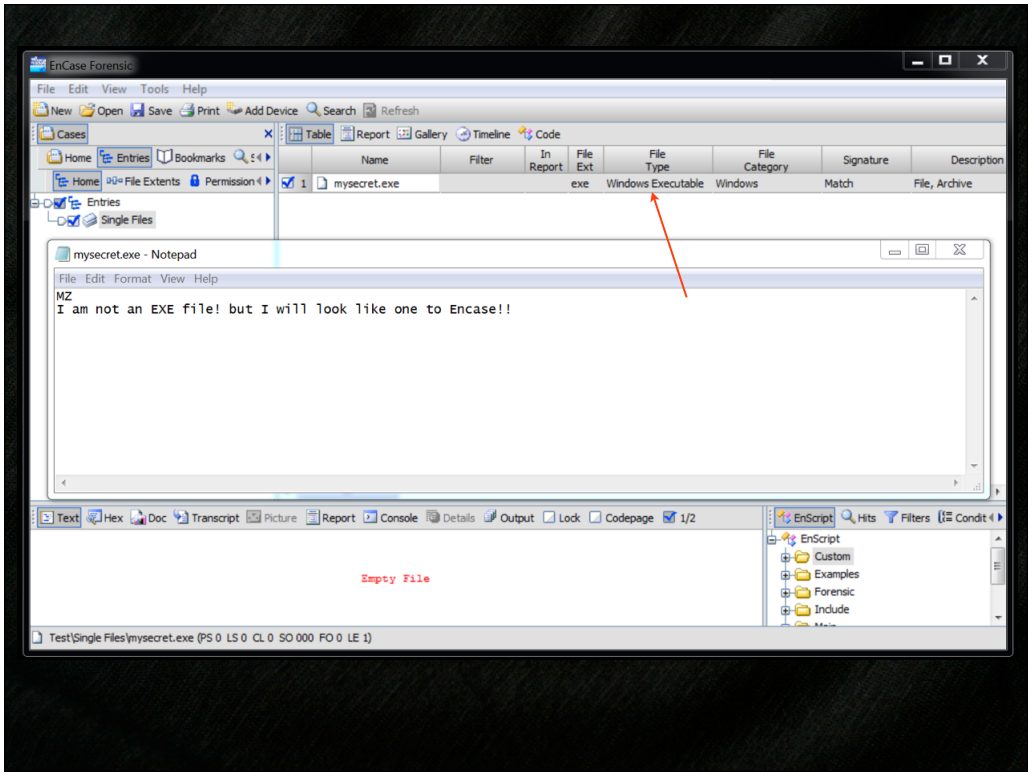
- EXE files begin with bytes MZ
- It's dead easy to make files match their extensions



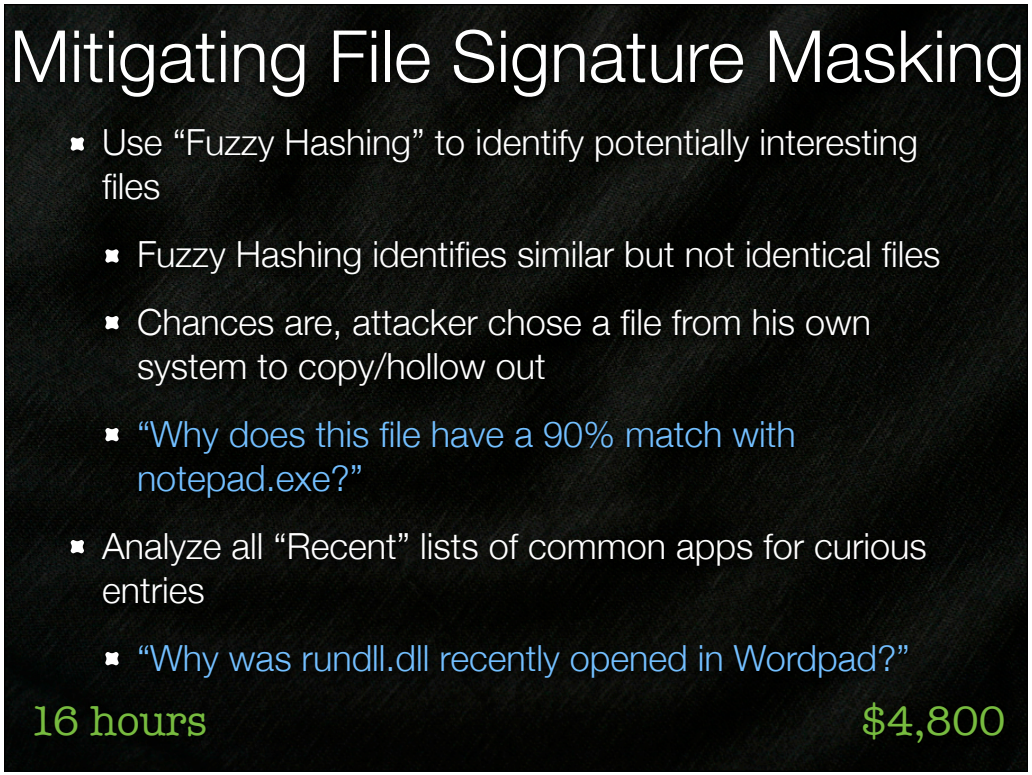
16 hours

$4,800

This TXT file shows how easy it is to make a file match its extension despite
having contents that are vastly different

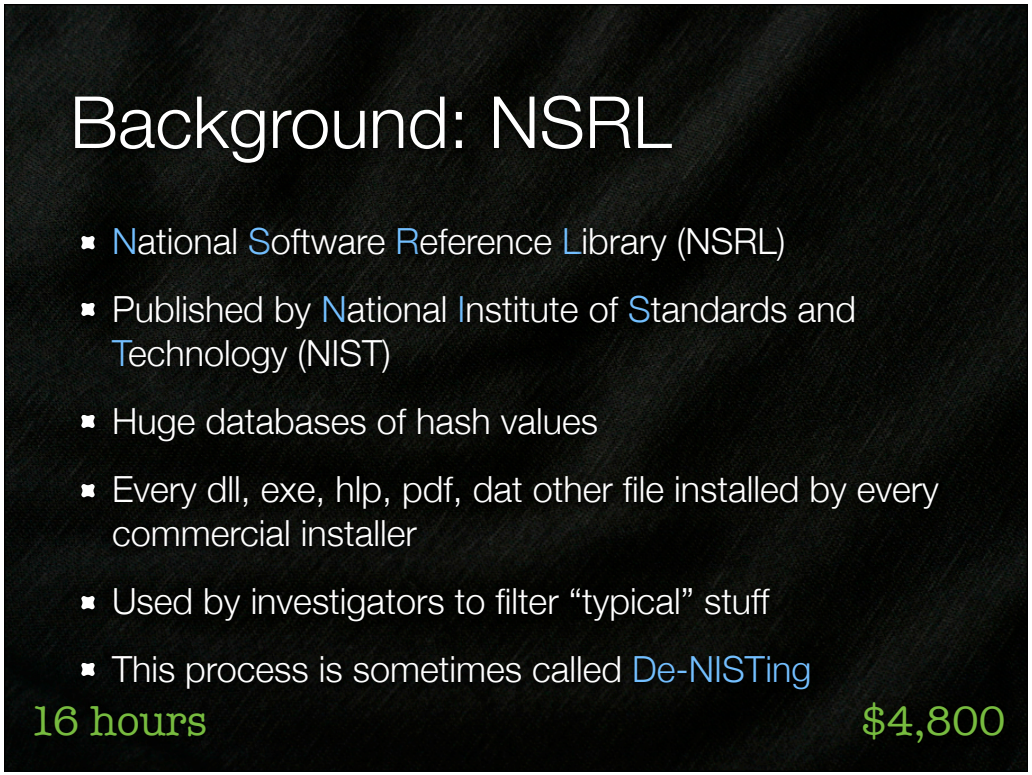


Even though this txt file was created in notepad, it is recognized as a ‘Windows Executable’
because the first two characters are MZ.



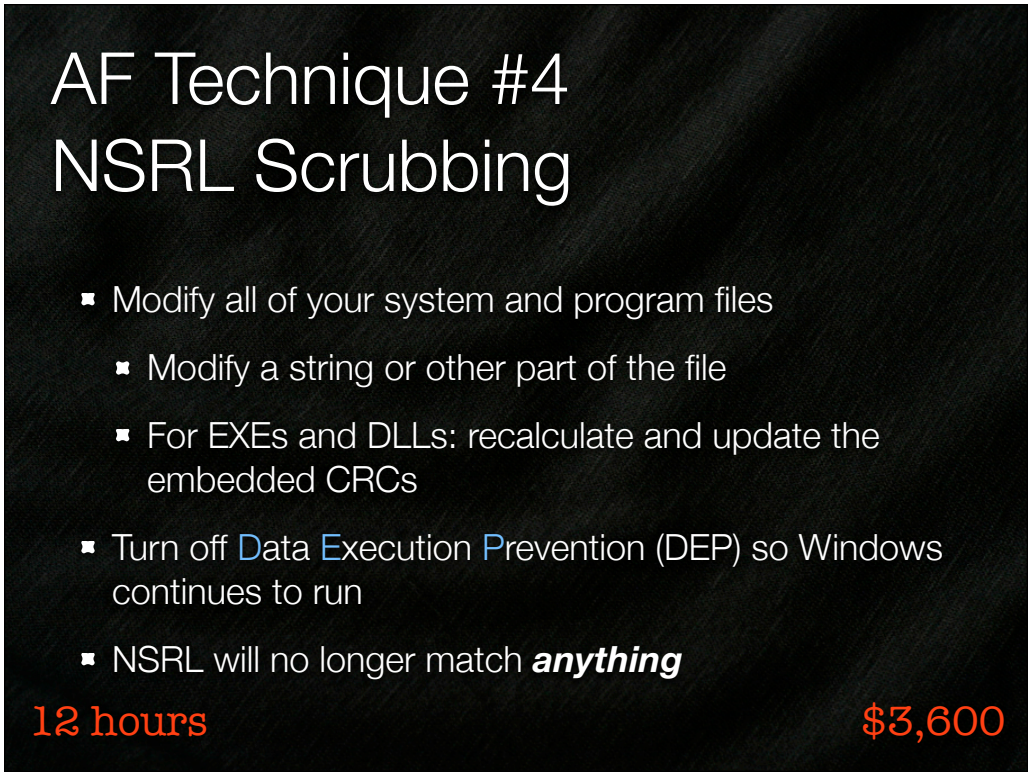
Fuzzy Hashing and Recent File analysis can mitigate false file signatures fairly easily by asking simple questions

Data Deduplication
Date Filtering
NSRL

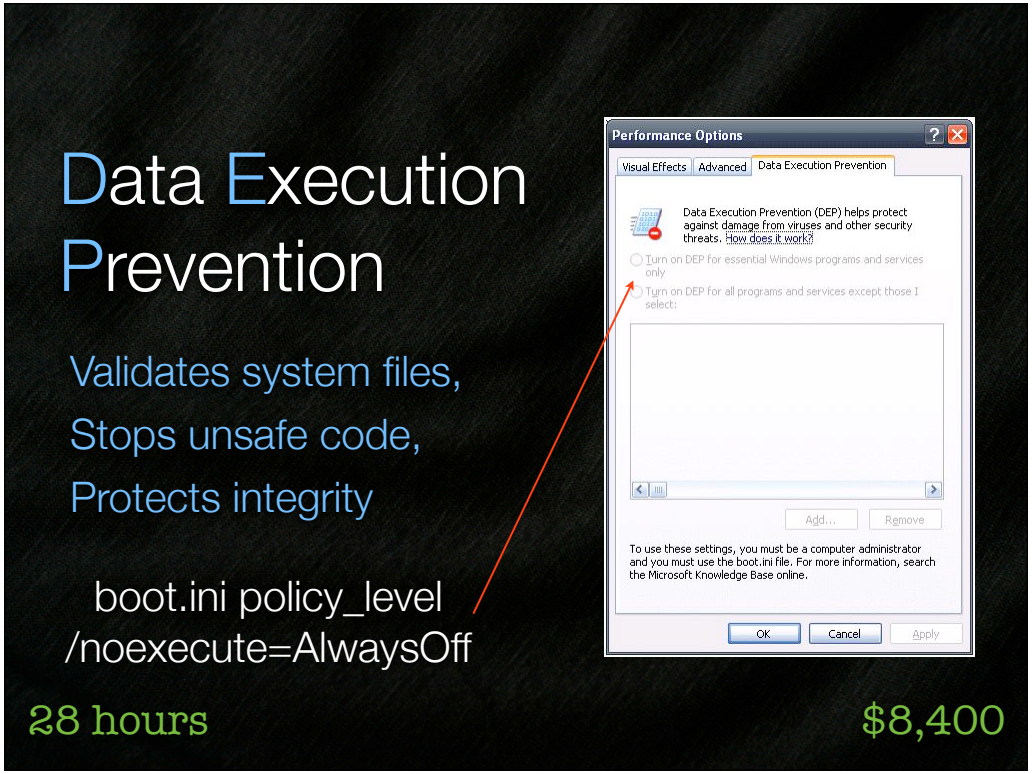


De-NISTing a drive may bring 120GB of data down to about 700MB
Only user-created content will remain

Hundreds of gigabytes can be reduced to a few hundred megabytes

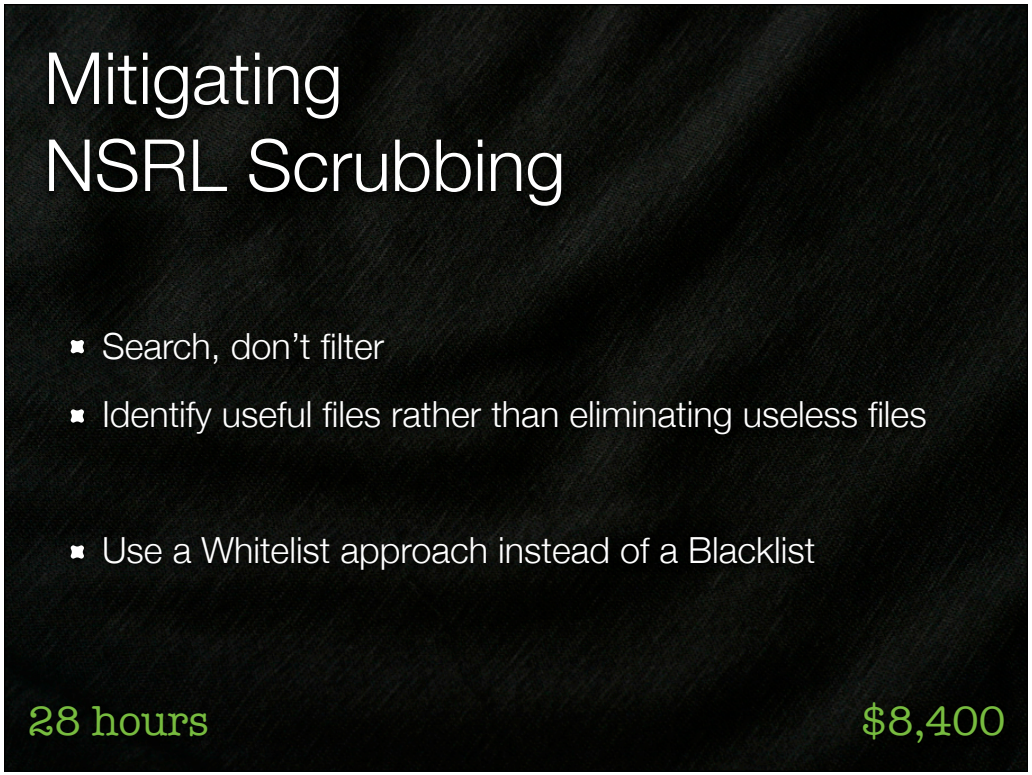


Most files won't need a lot of work: simply change a character and you're good.
Executable files (DLLs, EXEs) have embedded Cyclical Redundancy Checks (CRCs) that make sure they are still good
You will need to recalculate the CRCs for these files in order to change them in a way that will keep them running



DEP will stop Windows from running if it sees parts of Windows being modified.

So turn it off! You can then run your modified version of Windows without restriction.



Whitelist approach looks for things that match
Blacklist approach suppresses things that don't

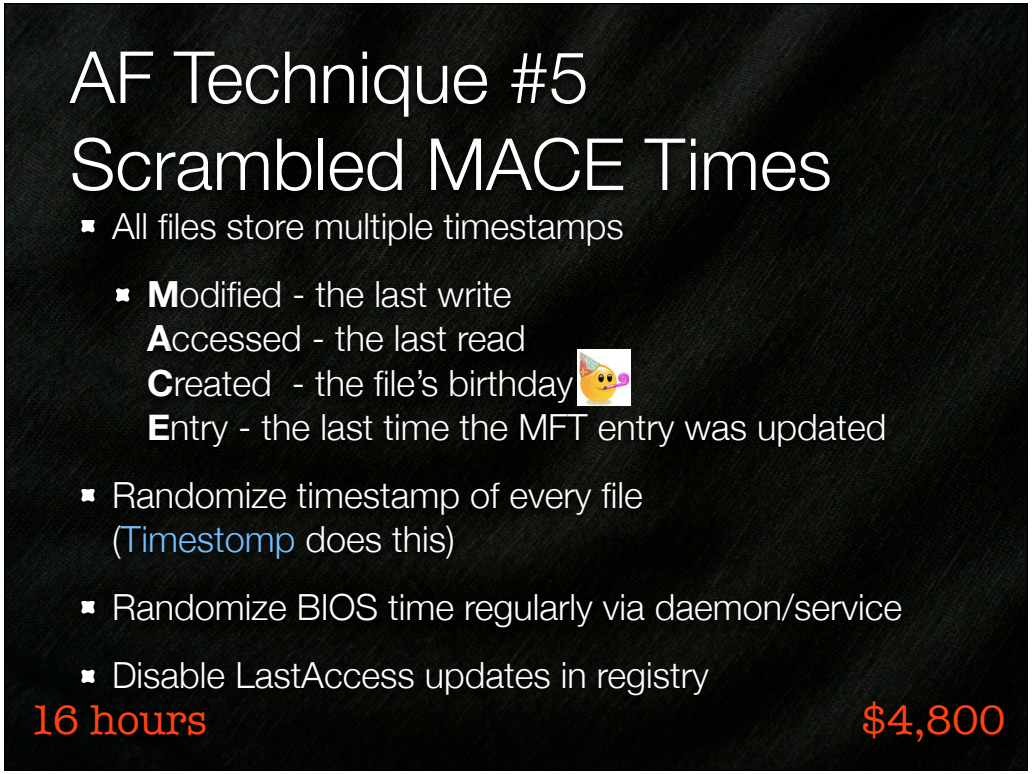
Use a whitelist approach

Background: Histograms

- Investigators use histograms to identify which dates have higher-than-average activity
- e.g. VPN Logins, Firewall alerts, even FileCreated times

28 hours

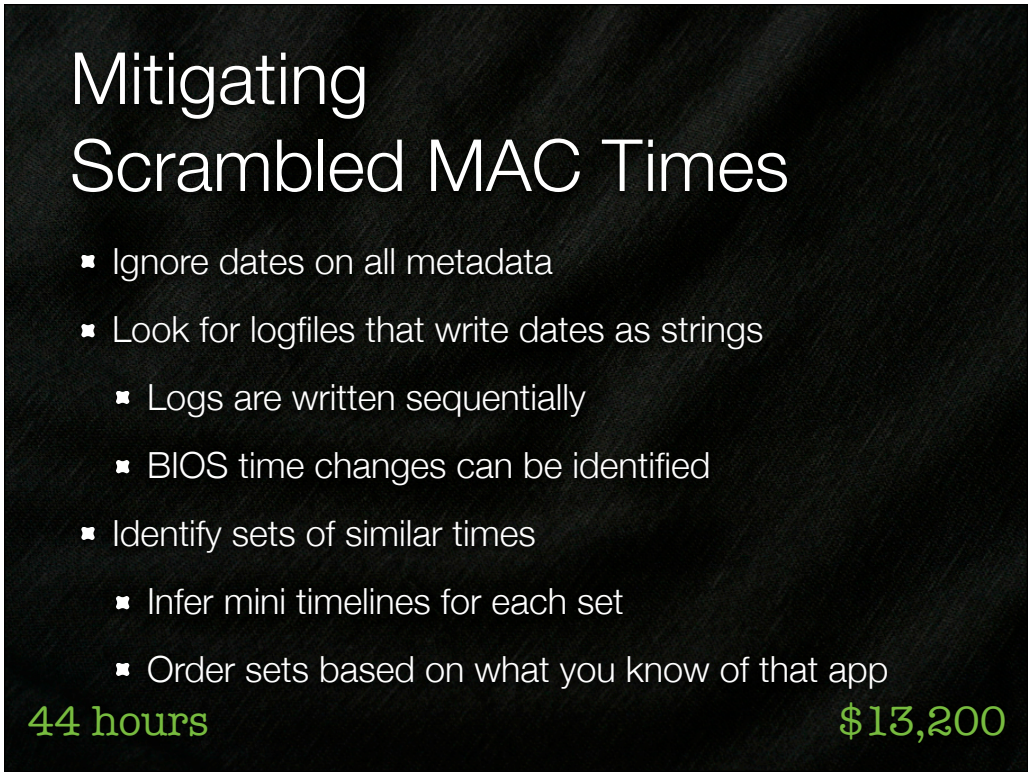
$8,400



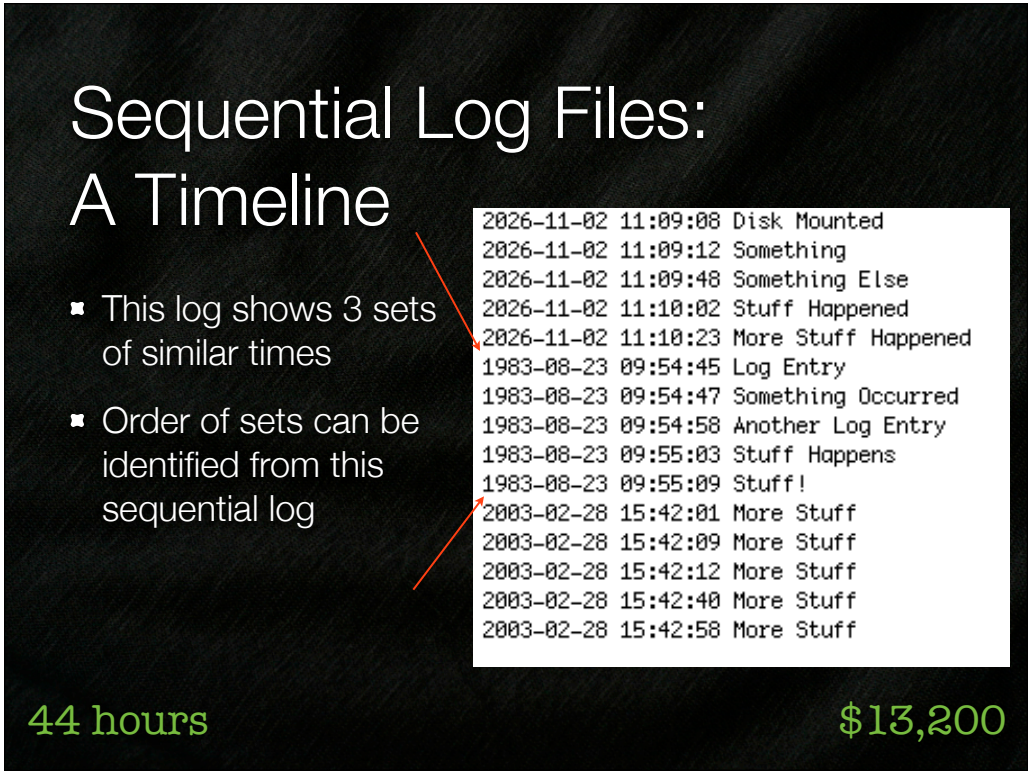
Most of an investigator's "Timeline Assembly" revolve around MACE times
MACE times can be modified easily

A malicious person can modify EVERY MACE time across an entire system

Two ways to suppress “Last Accessed Time” updates



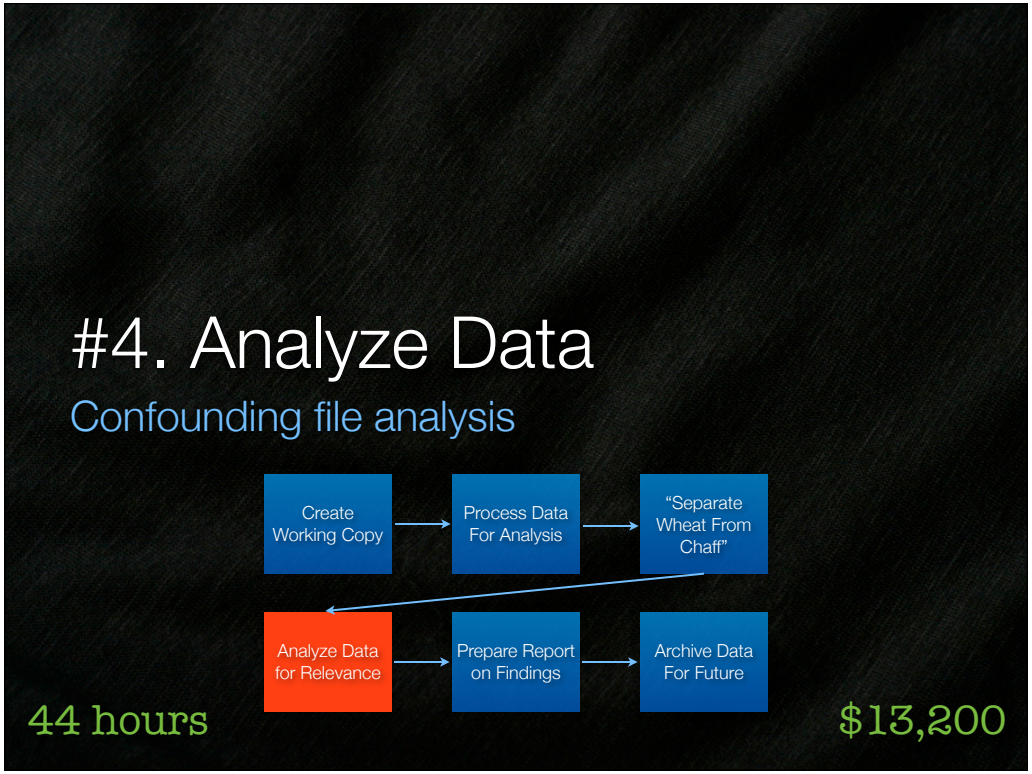
Sequential logfiles can help identify timelines



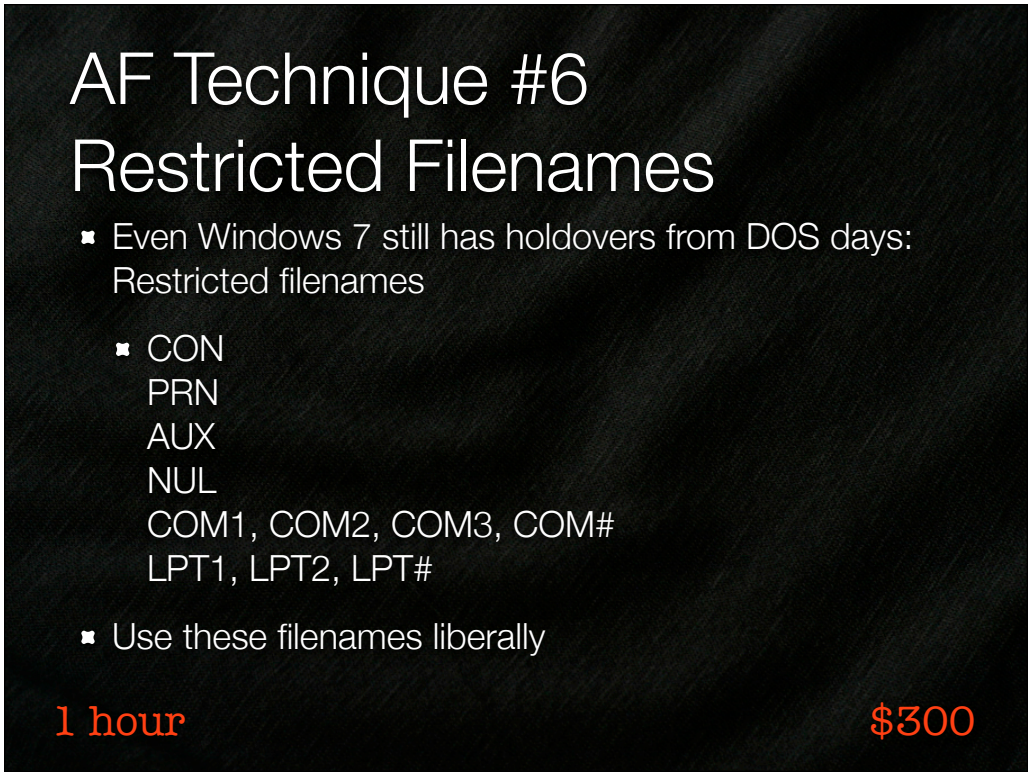
This logfile shows 3 sets of similar times
It also shows the ordering of each set

The BIOS time was changed twice

Smart investigators never say “This occurred at this time”
They say ‘Logs show it occurred at this time’
and “This time is consistent with these other logs which reference this action”

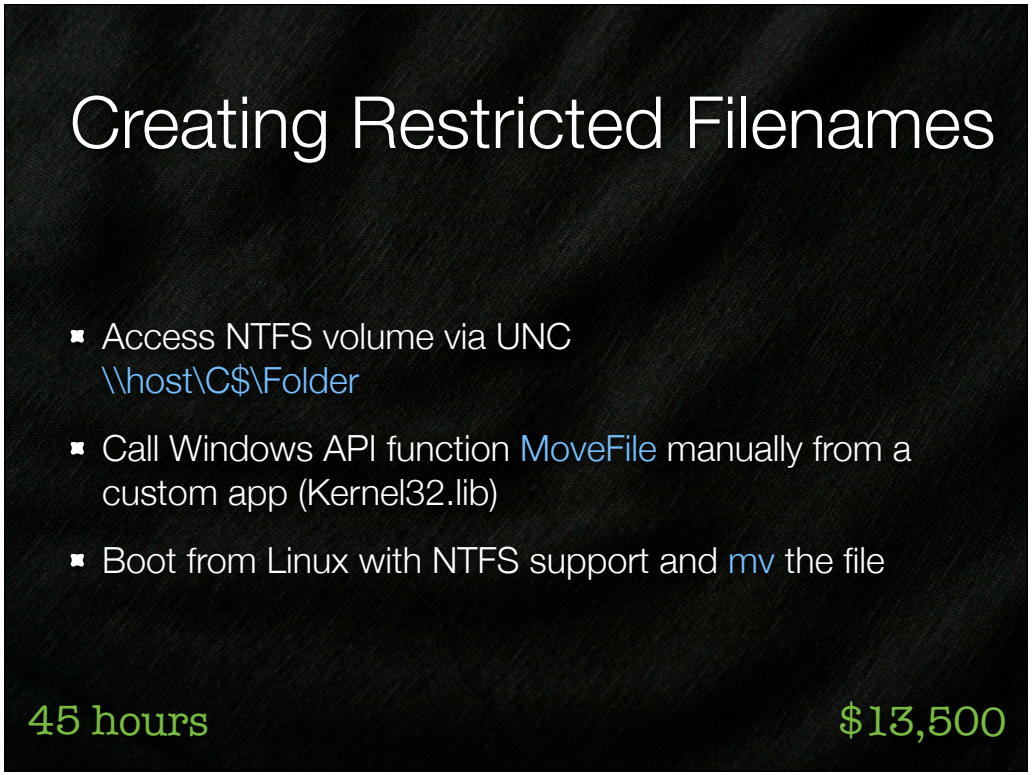


When the suite you're using doesn't show you everything you want to see, you typically take the file out of the image to your workstation
You can then use your own app to analyze the file



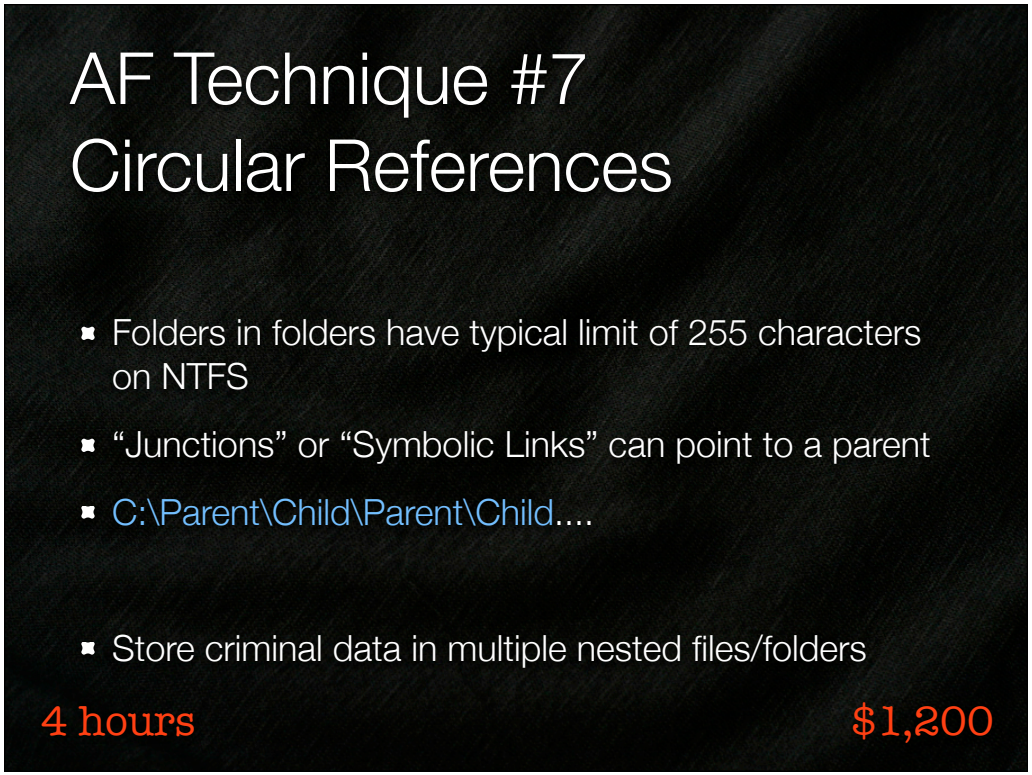
Windows 7 still has parts of DOS in it

This won't take up too much time but will still frustrate the investigator. He'll likely figure out what's wrong in less than an hour, but will bill a full hour of work for it.

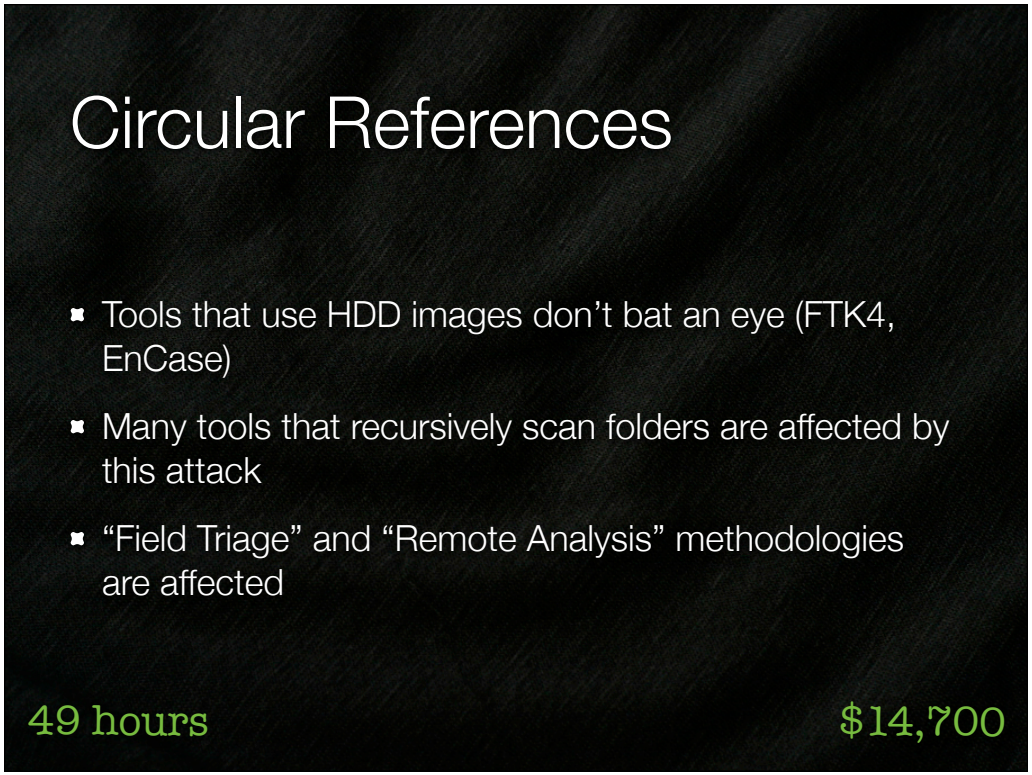


You can't just create a file with a restricted name. You need to trick Windows into doing it

Your analysis machine should go by your rules. You make up the filenames.



When a circular reference is followed, it could cause programs to enter an infinite loop.
Other programs may detect that the path they’re trying to access is > 255 characters and throw an exception



Reminder: The 3 Methodologies are:
* Image Everything, Analyze Later
* Field Triage to decide what to image
* Remote Analysis, target only evidence you need

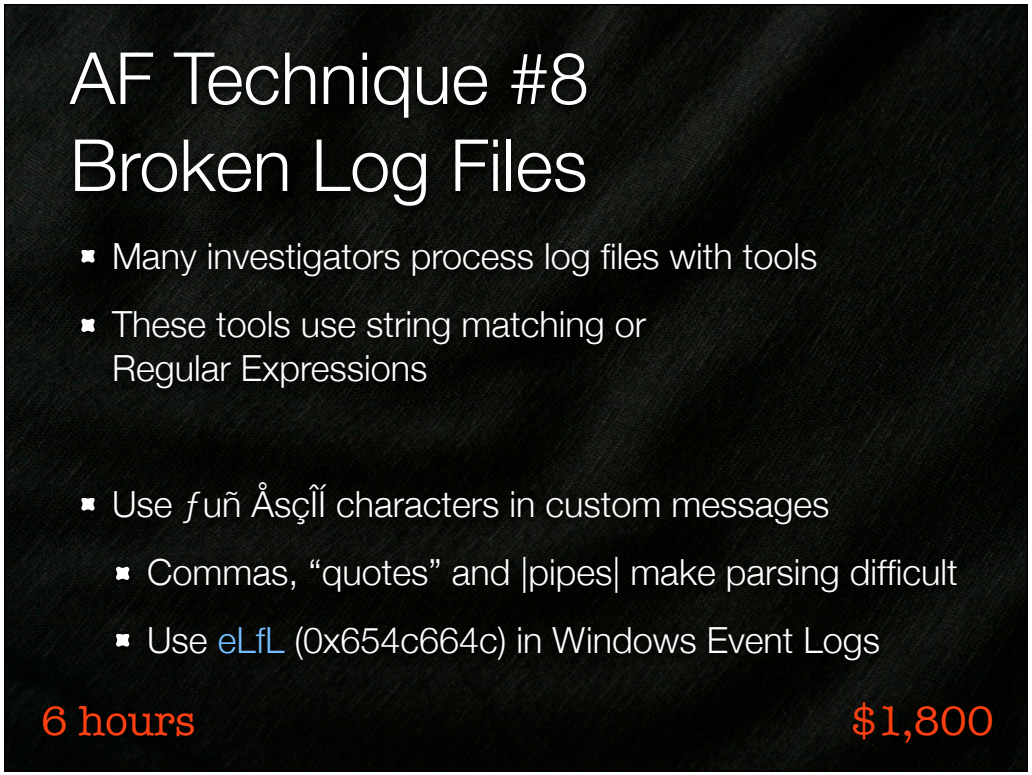
Mitigating Circular References

- Always work from an image
- Be mindful of this attack when dealing with an attacker's live system

- Just knowing about it will help you recognize it

49 hours

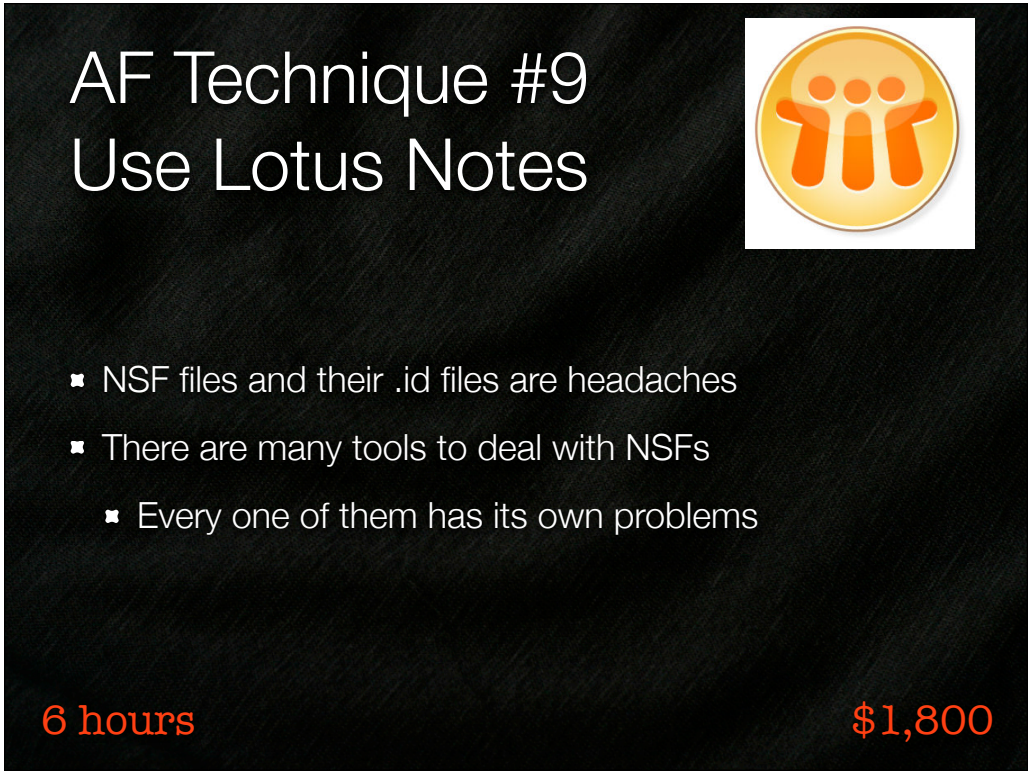
$14,700



eLfL is the 4byte header for Windows Event Logs
It marks the start of an eventlog record

Throwing these characters in the middle of a record will confuse some parsers into thinking a new entry has begun

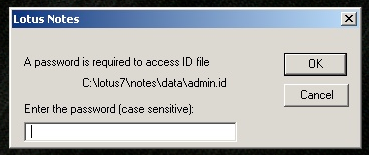
Zeroing in on the specific records you need is a lot better than parsing the whole log



Lotus Notes uses NSF files to hold emails, similar to PST files.
ID files include a user ID and an encryption key that can be unlocked with the user's password
2hrs per custodian

Lotus Notes

- Most apps use IBM's own Lotus Notes dlls/API to work with NSF files
- When opening each encrypted NSF, the API raises the password dialog:

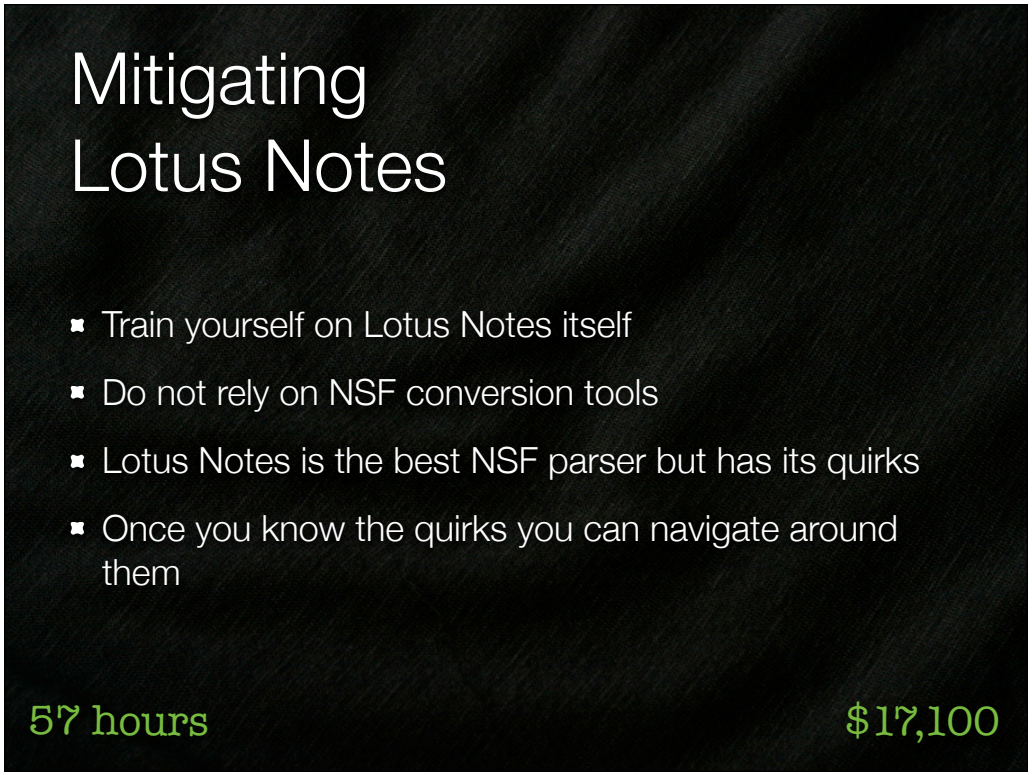


- Examiners/eDiscovery operators must select the user's ID file and type the associated password for every NSF file being processed

57 hours $17,100

The password dialog is raised in an interactive context, even when automated
The moment the API is used to open an NSF file, this dialog is presented to the user

This means you can't easily script NSF processing



Load up each NSF manually and deal with it using your own keyboard and mouse
Print the notable emails to PDF to be included in your report/affidavit

Reporting didn't seem to have many hacks at first until I started thinking about it....

AF Technique #10
HASH Collisions

- MD5 and SHA1 hashes are used to locate files
- Add dummy data to your criminal files so its MD5 hash matches known good files
- Searching by hash will yield unexpected results

- badfile.doc e4d909c290d0fb1ca068ffaddf22cbd0
  goodfile.doc e4d909c290d0fb1ca068ffaddf22cbd0

2 hours $600

What if you match your bad stuff with rundll.dll?
NSRL will suppress it!

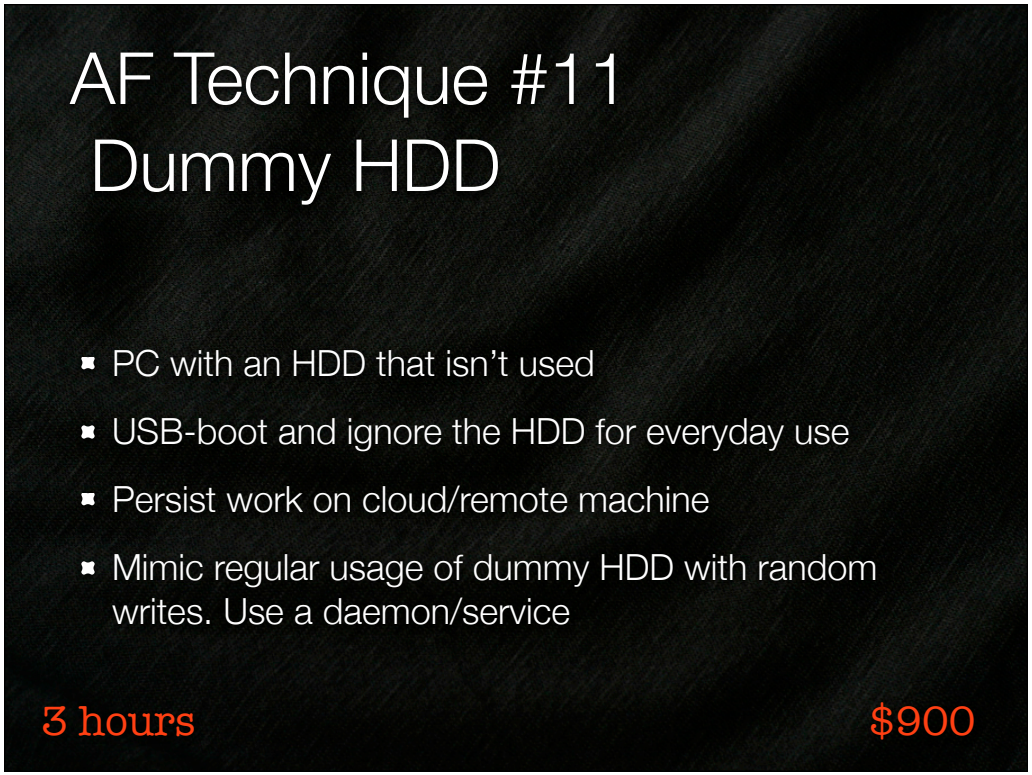
Hash Collisions

- Of course, this would only be useful in a select few cases:
  - i.e. you stole company data and stored on a volume they could seize/search
- Try explaining why GoodFile.doc and BadFile.doc have identical hashes to judges/justices/arbiters/non-techies
  - could provide just-the-right-amount of ‘reasonable doubt’

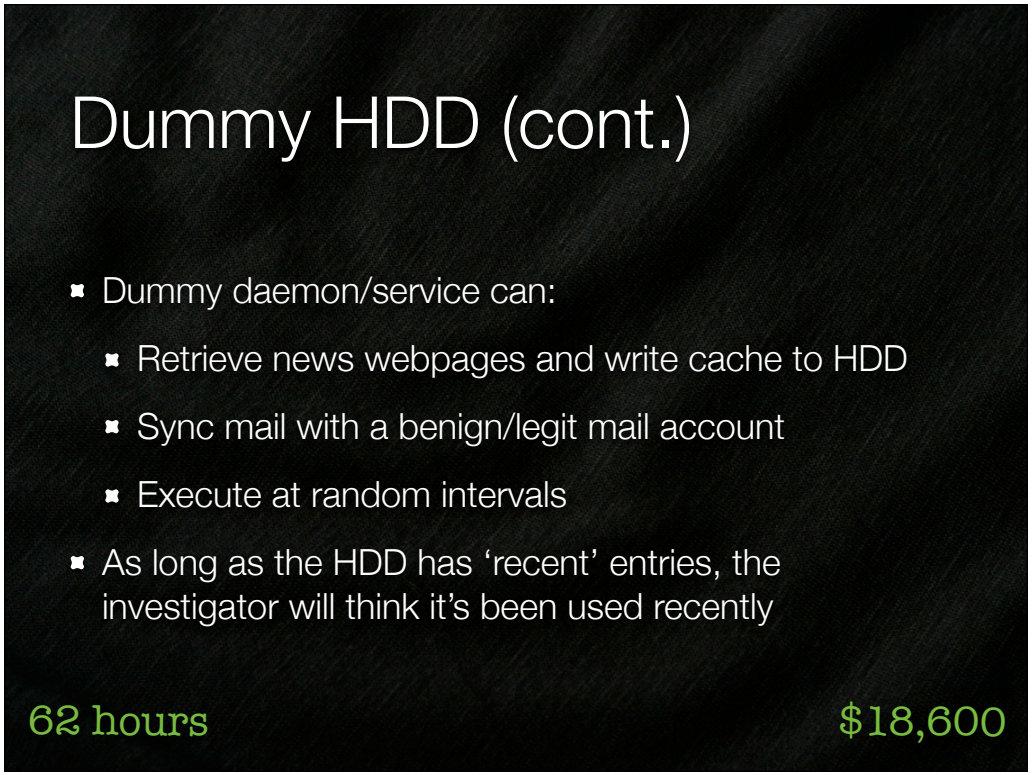
59 hours $17,700

Most of the research into MD5 collisions is a result of Marc's 2008 paper

Always doublecheck your findings!
Never rely on hash matches to guarantee you've found the file you're looking for

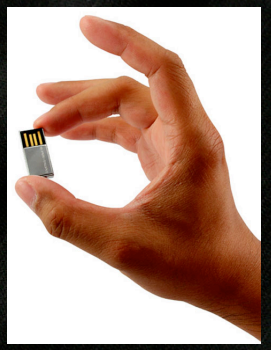


Using your computer without a hard drive is very easy nowadays thanks to large removable media



Creating a "dummy service" can simulate recent usage of a computer

Mitigating Dummy HDDs



- Always check for USB drives in USB slots AND on motherboards.
  They can be SMALL these days...
- Pagefile on USB drive may point to network locations (if the OS was paging at all...)
- If possible, monitor network traffic before seizure to detect remote drive location

62 hours $18,600

Look on the motherboard itself and identify every USB header. Follow all USB cables to the external ports on the case.

Don’t be fooled!

In the event a case is challenged in a year or two, firms need to archive data for the future

Technique #1
Data Saturation

- Same as Technique #1
  - The more data you have, the more they need to keep
- We've come full circle

1 hours

$20/mo per HDD

3 HDDs per month = $60
3 HDDs per year = $720

Budget Overrun

- We've taken up roughly 63 hours of an investigator's time
  - That's more than 8 workdays, without overtime
- This extra time was spent trying to image drives, export files, read email, and perform other menial tasks
  - The investigator still needs to do his regular work!
- Increased likelihood that opposing council will settle

63 hours

$18,900 + $~720/yr

Questions

- Have you encountered frustration in your examinations?
- How did you deal with it?

- I'd love to hear about it in the speaker Q&A room!

Thanks!

- Thanks DEFCON for letting me speak
- Thanks:
  - Forensic Friends (Josh, Joel, Nick)
  - Family
  - Coworkers
  - You!

Slide Availability

- The slides on your CD are outdated
- You can grab this latest version of these slides from: http://www.perklin.ca/~defcon20/perklin_antiforensics.pdf

References

- Berinato, Scott. June 8, 2007. The Rise of Anti-Forensics.
  Last accessed on June 12, 2012 from <http://www.csoonline.com/article/221208/the-rise-of-anti-forensics>
- Max. July 3, 2011. Disk Wiping with dcfldd.
  Last accessed on June 12, 2012 from <http://www.anti-forensics.com/disk-wiping-with-dcfldd>
- The grugq. Unknown Date. Modern Anti-Forensics.
  Last accessed on June 12, 2012 from <http://sebug.net/paper/Meeting-Documents/syscanhk/Modern%20Anti%20Forensics.pdf>
- Henry, Paul. November 15, 2007. Anti-Forensics.
  Last accessed on June 12, 2012 from <http://www.techsec.com/pdf/Tuesday/Tuesday%20Keynote%20-%20Anti-Forensics%20-%20Henry.pdf>
- Garfinkel, Simson. 2007. Anti-Forensics: Techniques, Detection, and Countermeasures.
  Last accessed on June 12, 2012 from <http://simson.net/ref/2007/slides-ICIW.pdf>
- Kessler, Gary. 2007. Anti-Forensics and the Digital Investigator.
  Last accessed on June 12, 2012 from <http://www.garykessler.net/library/2007_ADFC_anti-forensics.pdf>
- Hilley, S. 2007. Anti-Forensics with a small army of exploits.
  Last accessed on June 12, 2012 from <http://cryptome.org/0003/anti-forensics.pdf>

References (cont.)

- Dardick, G., La Roche, C., Flanigan, M. 2007. BLOGS: Anti-Forensics and Counter Anti-Forensics.
  Last accessed on June 12, 2012 from < http://igneous.scis.ecu.edu.au/proceedings/2007/forensics/21-Dardick%20et.al%20BLOGS%20ANTI-FORENSICS%20and%20COUNTER%20ANTI-FORENSICS.pdf>

- Hartley, Matthew. August, 2007. Current and Future Threats to Digital Forensics.
  Last accessed on June 12, 2012 from < https://dev.issa.org/Library/Journals/2007/August/Hartley-Current%20and%20Future%20Threats%20to%20Digital%20Forensics.pdf>

- Perklin, Michael. April 26, 2012. Anti-forensics: Techniques that make our lives difficult, and what we can do to mitigate them.
  Presented at HTCIA Ontario Chapter, Brampton, ON. Canada.

- Peron, C., Legary, M. . Digital Anti-Forensics: Emerging trends in data transformation techniques.
  Last accessed on June 12, 2012 from <http://www.seccuris.com/documents/whitepapers/Seccuris-Antiforensics.pdf>

- Stevens, M. June, 2007. On Collisions for MD5.
  Last accessed on June 12, 2012 from <http://www.win.tue.nl/hashclash/On%20Collisions%20for%20MD5%20-%20M.M.J.%20Stevens.pdf>

- Foster, J., Liu, V. July 2005. Catch Me If You Can: Exploiting Encase, Microsoft, Computer Associates, and the rest of the bunch…
  Last Accessed on June 12, 2012 from <http://www.blackhat.com/presentations/bh-usa-05/bh-us-05-foster-liu-update.pdf>

- Moulton, S. July 2009. RAID Recovery: Recover your PORN by sight and sound.
  Last Accessed on June 12, 2012 from <http://www.defcon.org/images/defcon-17/dc-17-presentations/defcon-17-scott_moulton-raid_recovery.pdf>