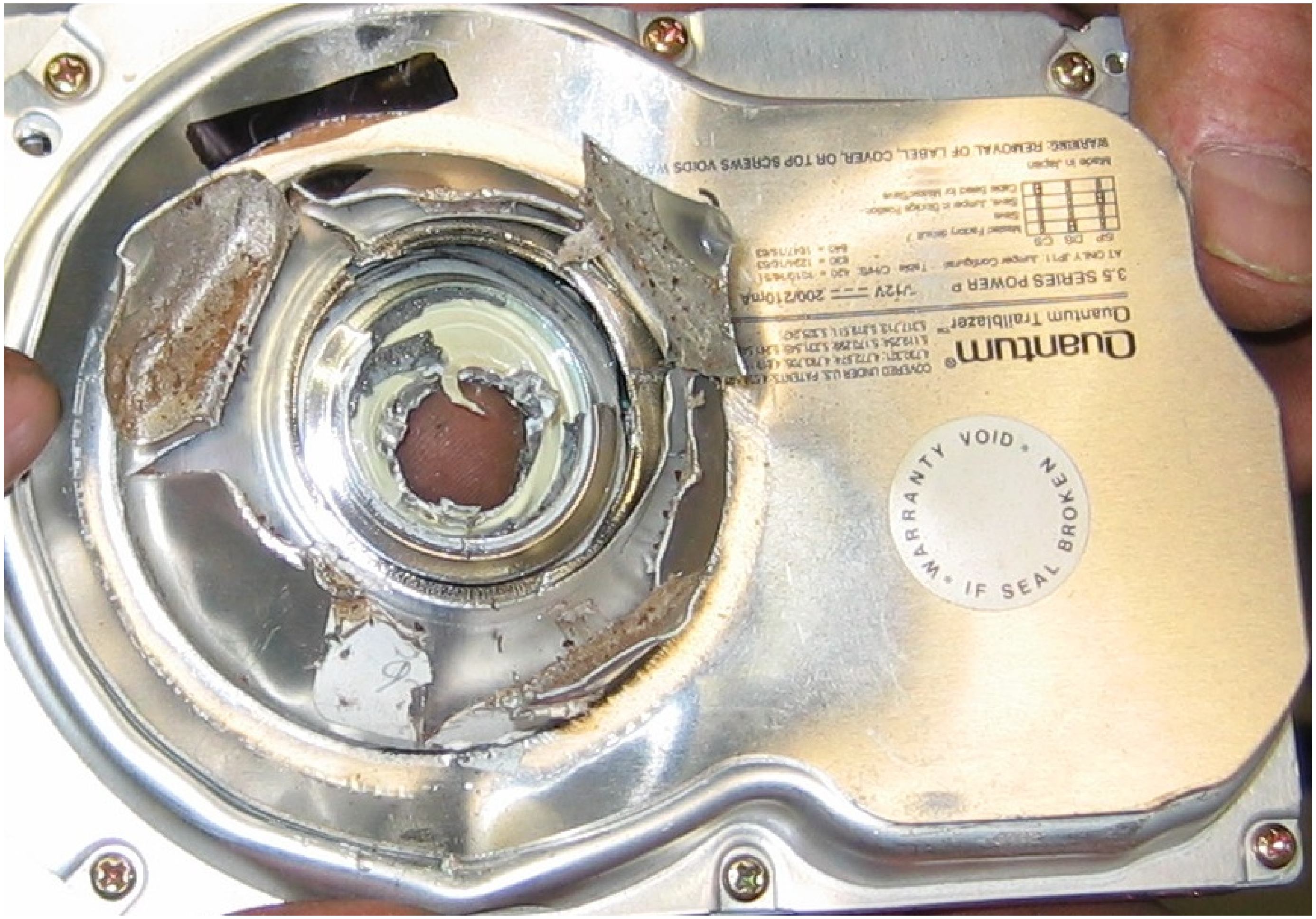
Ensure Proper Data Management with Discarded IT Assets


WARNING: REMOVAL OF LABEL, COVER, OR TOP SCREWS VOIDS WA
Made in Japan
3.5 SERIES POWER P
Quantum Trailblazer
Quantum
WARRANTY VOID IF SEAL BROKEN

Simson L. Garfinkel
Center for Research on Computation and Society
Harvard University

Thursday, November 3, 9:45am

Ten used computers were purchased in August 1998.

This is computer #1.

Computer #1 was a file server for a 15-attorney law firm.

Contents:

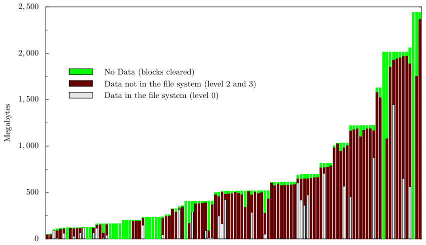
- client documents
- billing records
- correspondence

Computers #2 through #10 had:

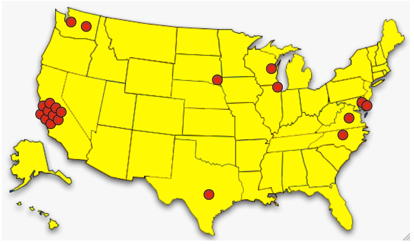
- Mental health records
- Home finances
- Draft of a novel

Was this a chance accident or common occurrence?

This talk explores the problem of data on discarded computers and presents real, practical solutions.



- Scale of the problem



- The Traceback Study

- Applicable Legislation

- Answers and Technology Choices

Data confidentiality is an important business goal.

Data can be:

- In flight
- Stored

Data spends most of its life in storage.
Therefore, stored data must be protected.

There are three techniques for assuring data confidentiality.

1. Physical security.

2. Logical access controls. (operating system)

3. Cryptography (disk & link)

When a disk is retired, techniques #1 and #2 don’t work.

1. ~~Physical security~~

2. ~~Logical access controls (operating system)~~

3. Cryptography (disk & link)

RSA
DATA SECURITY, INC.

Most organizations don’t encrypt their stored data.

Hard drives pose special problem for computer security.

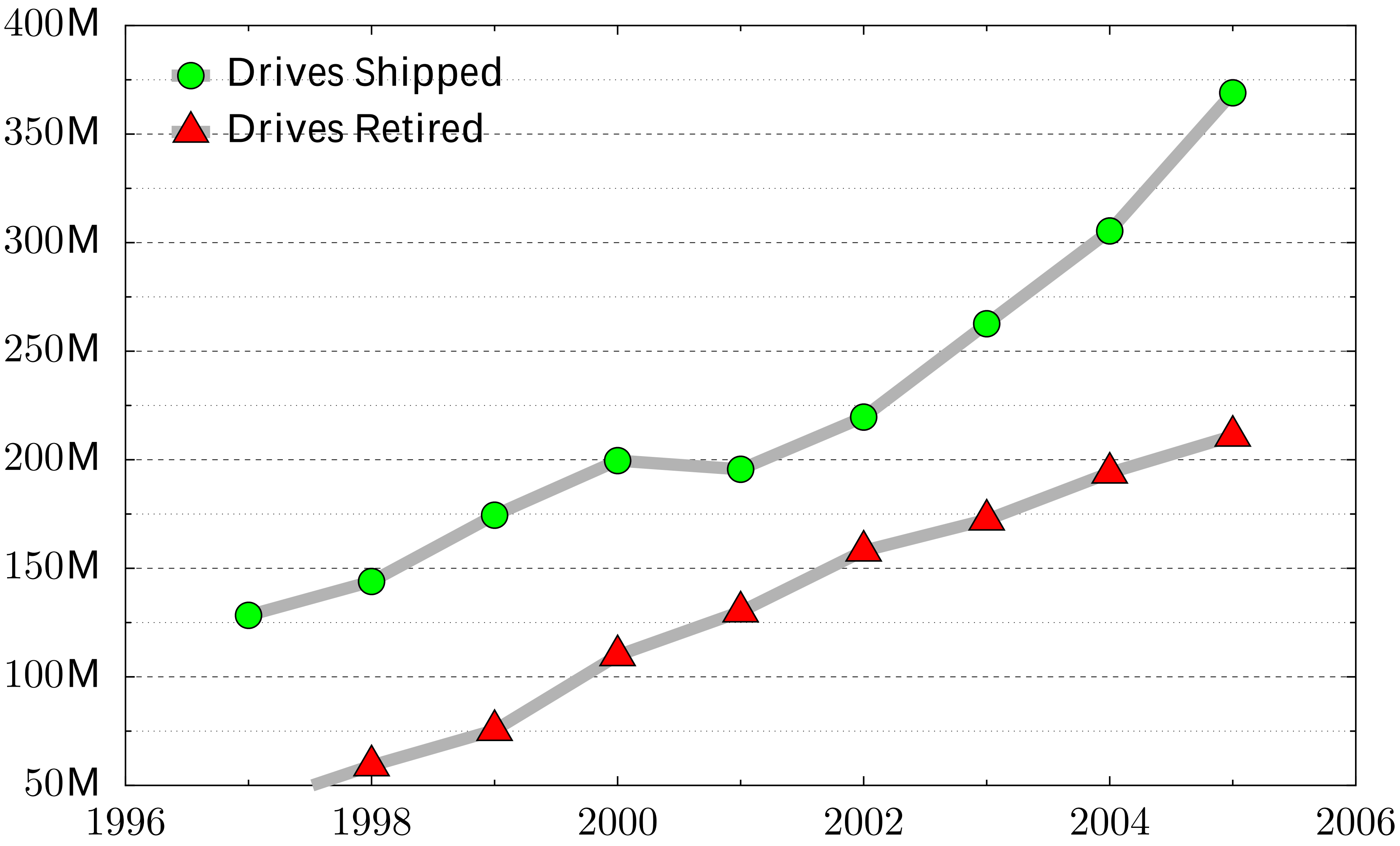
Do not forget data when power is removed.

Contain data that is not immediately visible.

Today’s computers can read hard drives that are 15 years old!

- Electrically compatible (IDE/ATA)
- Logically compatible (FAT16/32 file systems)
- Very different from tape systems

The problem: 210 million drives will be retired this year.



This is a staggering amount of data!

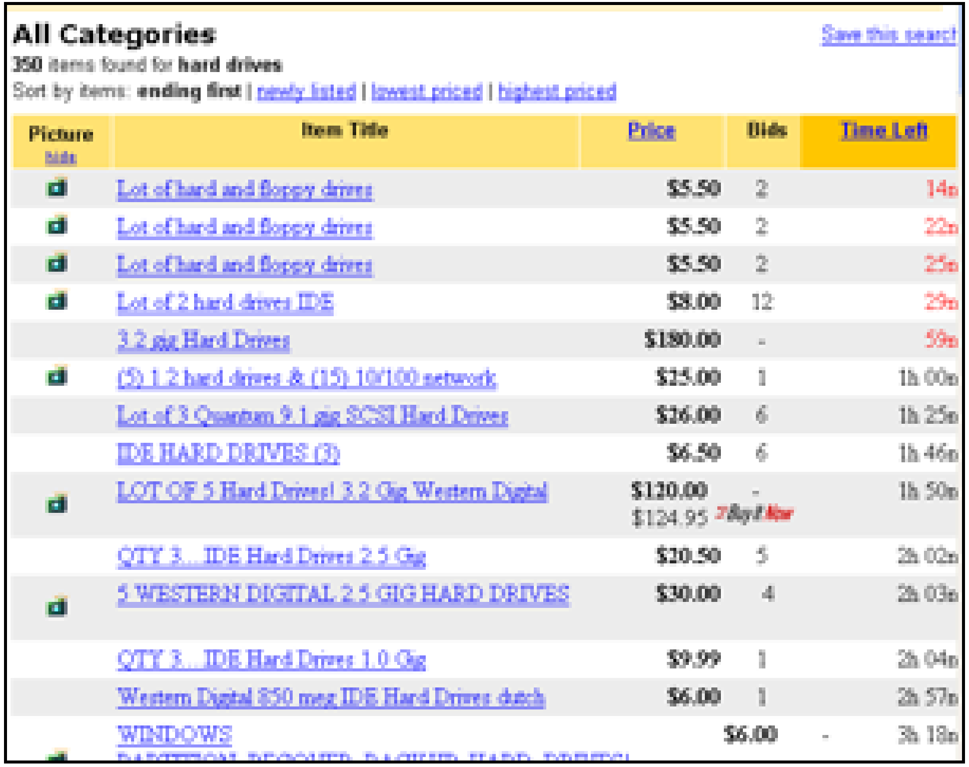
“Retire?”

Deckard (Harrison Ford) retiring a replicant.
“Blade Runner” (1982)

Many “retired” hard drives are actually resold or repurposed.

Retired drives are:

- Re-used within organizations
- Given to charities
- Sold at auction



All Categories

Save this search

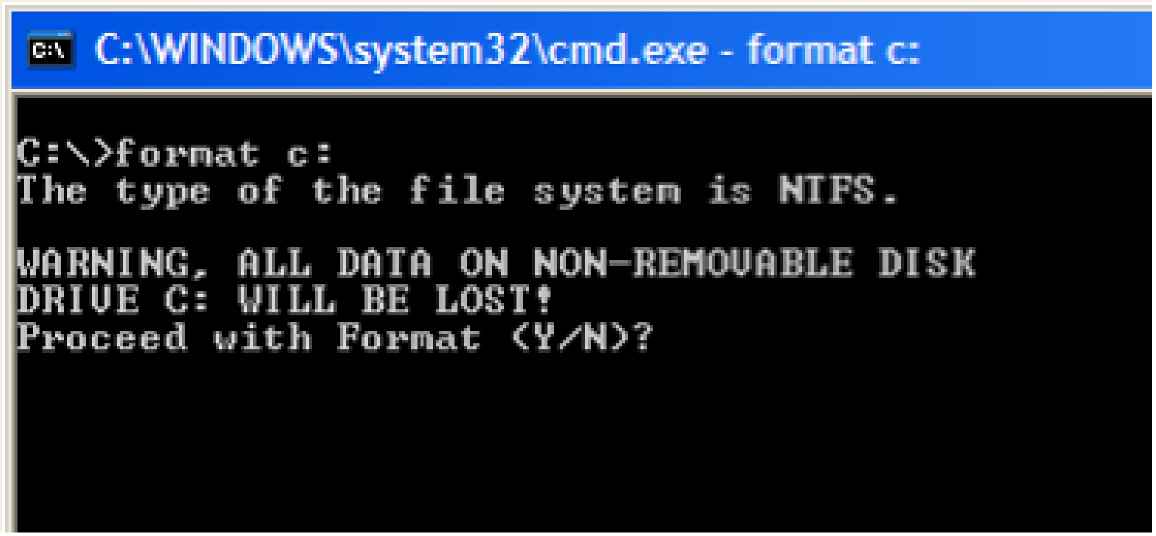
350 items found for **hard drives**

Sort by items: **ending first** | newly listed | lowest priced | highest priced

Picture hide	Item Title	Price	Bids	Time Left
📷	Lot of hard and floppy drives	$5.50	2	14m
📷	Lot of hard and floppy drives	$5.50	2	22m
📷	Lot of hard and floppy drives	$5.50	2	25m
📷	Lot of 2 hard drives IDE	$8.00	12	29m
	3.2 gig Hard Drives	$180.00	-	59m
📷	(5) 1.2 hard drives & (15) 10/100 network	$25.00	1	1h 00m
	Lot of 3 Quantum 9.1 gig SCSI Hard Drives	$26.00	6	1h 25m
	IDE HARD DRIVES (3)	$6.50	6	1h 46m
📷	LOT OF 5 Hard Drives! 3.2 Gig Western Digital	$120.00 $124.95 *Buy It Now*	-	1h 50m
	QTY 3....IDE Hard Drives 2.5 Gig	$20.50	5	2h 02m
📷	5 WESTERN DIGITAL 2.5 GIG HARD DRIVES	$30.00	4	2h 03m
	QTY 3....IDE Hard Drives 1.0 Gig	$9.99	1	2h 04m
	Western Digital 850 meg IDE Hard Drives dutch	$6.00	1	2h 57m
📷	WINDOWS PARTITION RECOVER BACKUP HARD DRIVES!	$6.00	-	3h 18m

About 1000 used drives/day sold on eBay.

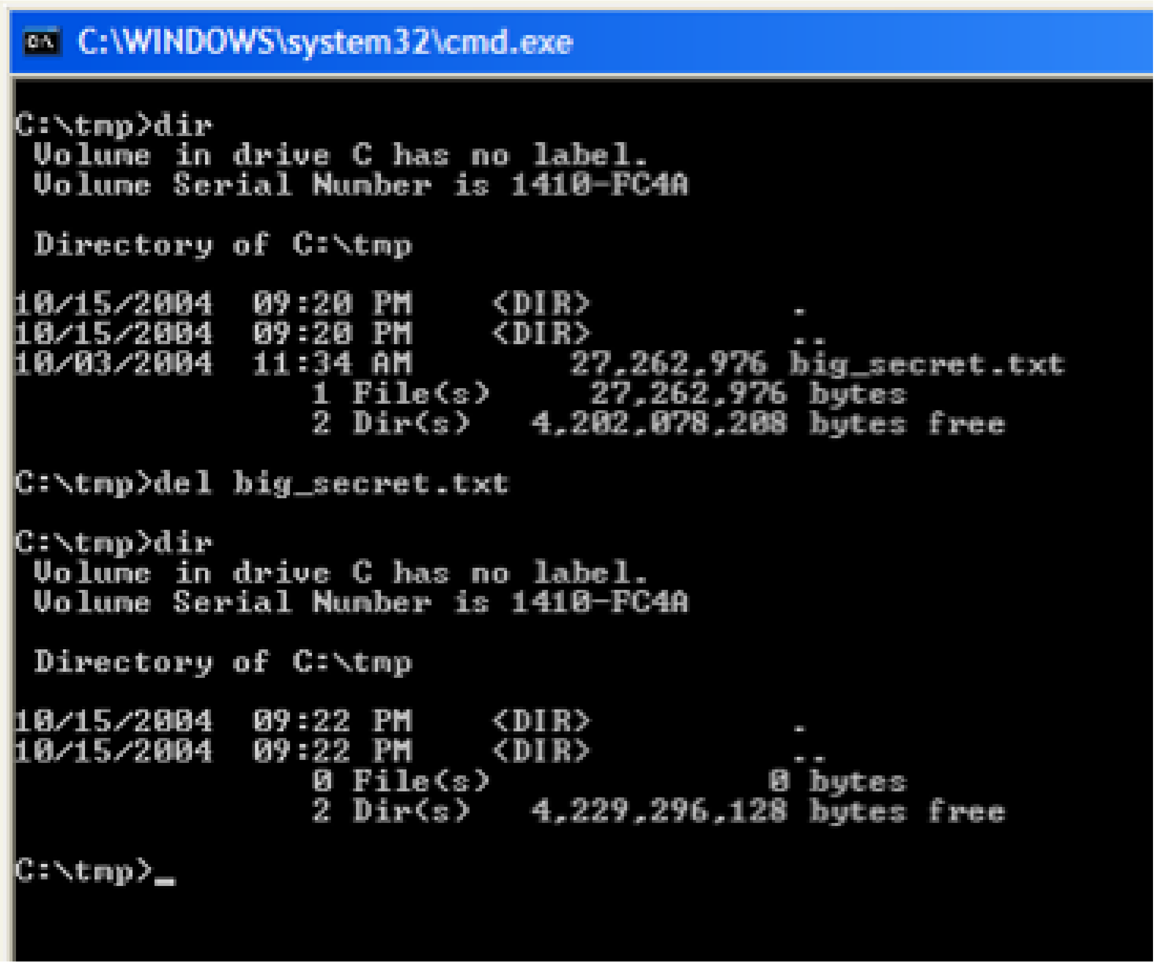
Bad news:
FORMAT C: doesn’t erase the hard drive.



```
C:\>format c:
The type of the file system is NTFS.

WARNING, ALL DATA ON NON-REMOVABLE DISK
DRIVE C: WILL BE LOST!
Proceed with Format (Y/N)?
```

FORMAT just writes a new root directory.

More bad news:
DEL doesn't delete files.



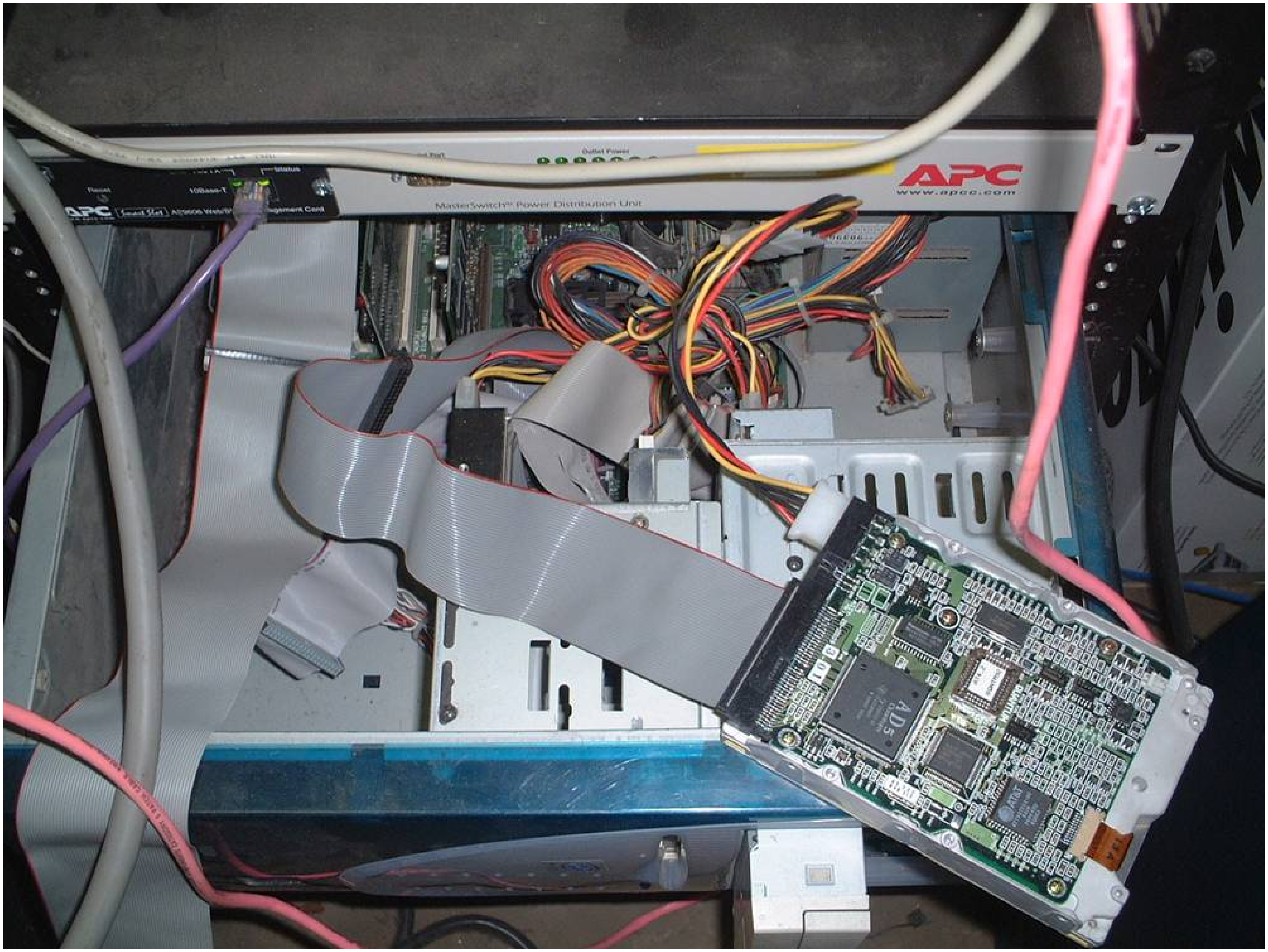
```
C:\tmp>dir
 Volume in drive C has no label.
 Volume Serial Number is 1410-FC4A

 Directory of C:\tmp

10/15/2004  09:20 PM    <DIR>          .
10/15/2004  09:20 PM    <DIR>          ..
10/03/2004  11:34 AM        27,262,976 big_secret.txt
               1 File(s)     27,262,976 bytes
               2 Dir(s)   4,202,078,208 bytes free

C:\tmp>del big_secret.txt

C:\tmp>dir
 Volume in drive C has no label.
 Volume Serial Number is 1410-FC4A

 Directory of C:\tmp

10/15/2004  09:22 PM    <DIR>          .
10/15/2004  09:22 PM    <DIR>          ..
               0 File(s)              0 bytes
               2 Dir(s)   4,229,296,128 bytes free

C:\tmp>_
```

DEL simply removes the file's name from the directory.

Between January 1999 and April 2002, I acquired 236 hard drives on the secondary market.

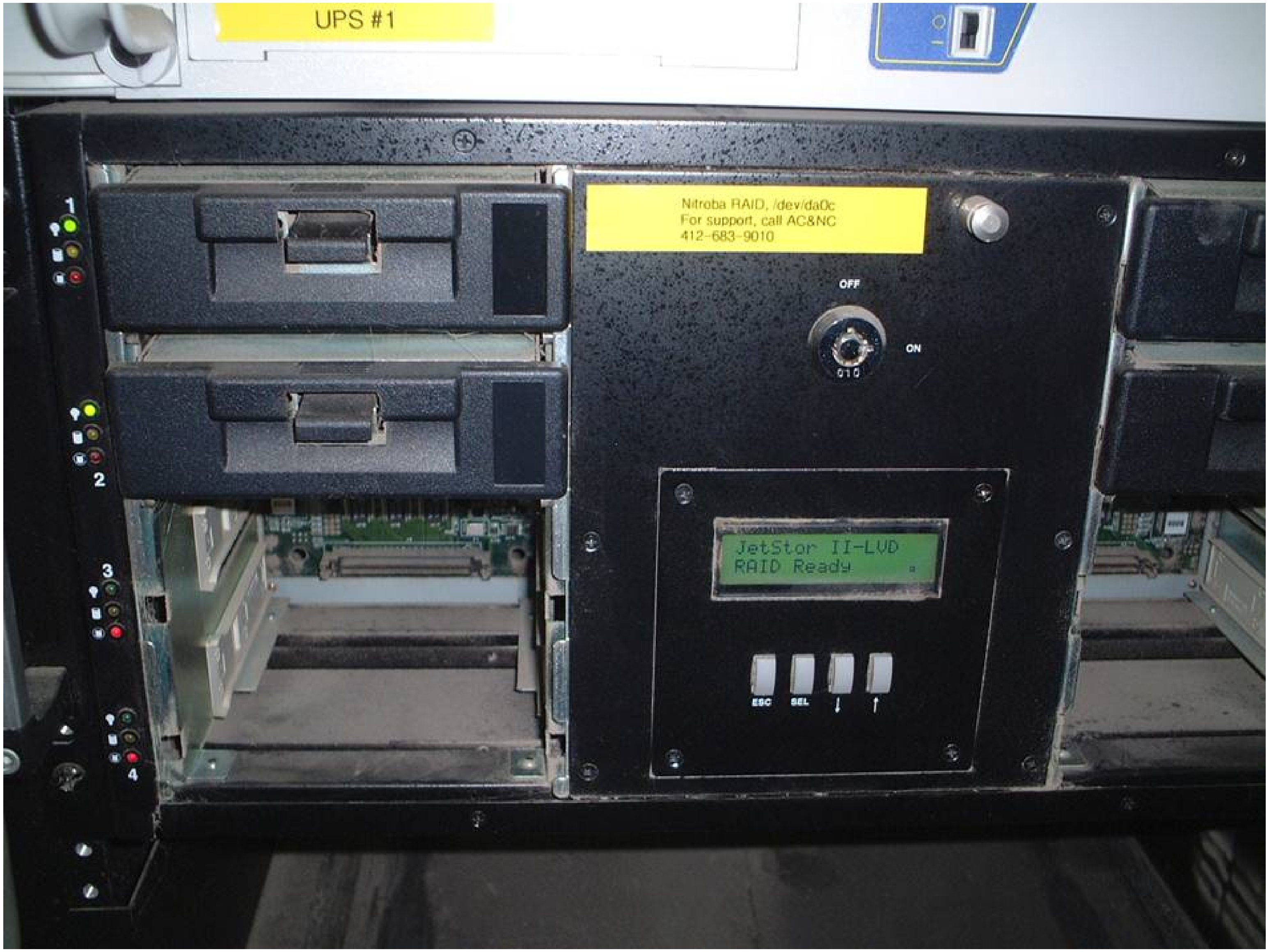
Drives arrived by UPS

Data on drives “imaged” using FreeBSD



```
dd if=/dev/ad0 of=file.img bs=65536 conv=noerror,sync
```

Images stored on a RAID



Example: Disk #70: IBM-DALA-3540/81B70E32

Purchased for $5 from a Mass retail store on eBay

Copied the data off: 541MB

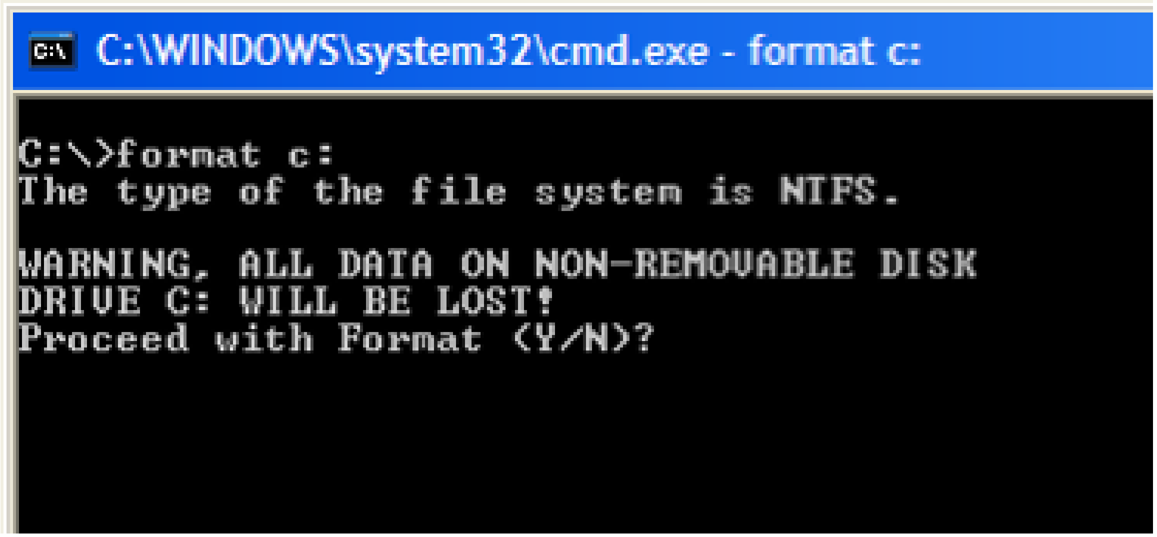
Initial analysis:

Total disk sectors:	1,057,392
Sectors with data:	989,514
Sectors all zero:	67,878
Total files:	3

The files:

```
drwxrwxrwx  0 root          0 Dec 31  1979 ./
-r-xr-xr-x  0 root     222390 May 11  1998 IO.SYS
-r-xr-xr-x  0 root          9 May 11  1998 MSDOS.SYS
-rwxrwxrwx  0 root      93880 May 11  1998 COMMAND.COM
```

Clearly, this disk had been FORMATed...



```
C:\>format c:
The type of the file system is NTFS.

WARNING, ALL DATA ON NON-REMOVABLE DISK
DRIVE C: WILL BE LOST!
Proceed with Format (Y/N)?
```

Windows FORMAT doesn’t erase the disk...
FORMAT just writes a new root directory.

UNIX “strings” reveals the disk’s previous contents...

```
Insert diskette for drive
 and press any key when ready
Your program caused a divide overflow error.
If the problem persists, contact your program vendor.
Windows has disabled direct disk access to protect your lo
To override this protection, see the LOCK /? command for m
The system has been halted.  Press Ctrl+Alt+Del to restart
You started your computer with a version of MS-DOS incompa
version of Windows. Insert a Startup diskette matching thi

OEMString = "NCR 14 inch Analog Color Display Enchanced SV
        Graphics Mode: 640 x 480 at 72Hz vertical refresh.
        XResolution             = 640
        YResolution             = 480
        VerticalRefresh         = 72
```

70.img con’t...

```
ling the Trial Edition
----------------------------
IBM AntiVirus Trial Edition is a full-function but time-li
evaluation version of the IBM AntiVirus Desktop Edition pr
may have received the Trial Edition on a promotional CD-RO
single-file installation program over a network.  The Tria
is available in seven national languages, and each languag
provided on a separate CC-ROM or as a separa
EAS.STCm
EET.STC
ELR.STCq
ELS.STC
```

70.img con’t...

```
MAB-DEDUCTIBLE
MAB-MOOP
MAB-MOOP-DED
METHIMAZOLE
INSULIN (HUMAN)
COUMARIN ANTICOAGULANTS
CARBAMATE DERIVATIVES
AMANTADINE
MANNITOL
MAPROTILINE
CARBAMAZEPINE
CHLORPHENESIN CARBAMATE
ETHINAMATE
FORMALDEHYDE
MAFENIDE ACETATE
```

[Garfinkel & Shelat 03] established the scale of the problem.

We found:

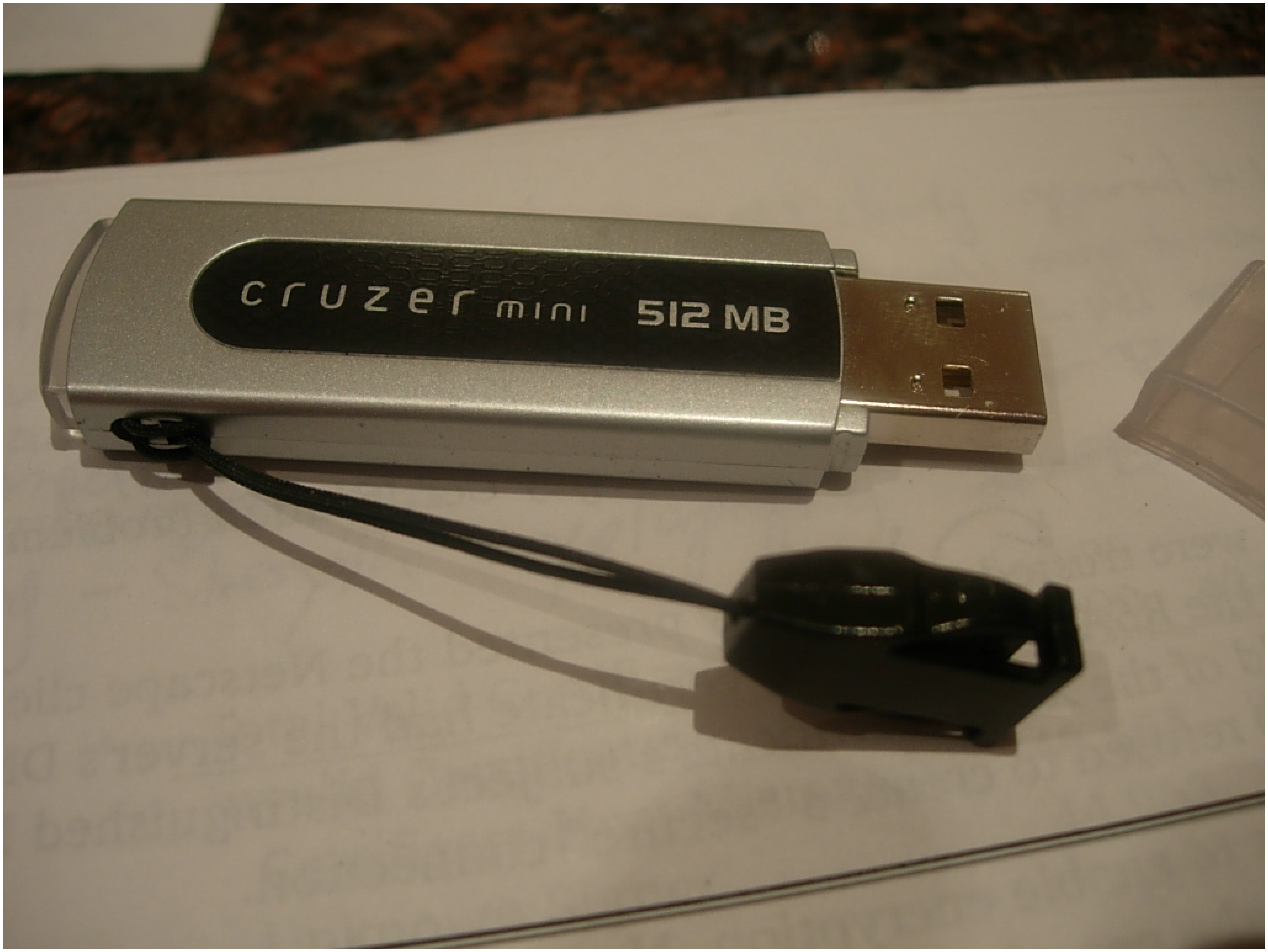
- Thousands of credit card numbers (many disks)
- Financial records
- Medical information
- Trade secrets
- Highly personal information

We did not determine why the data had been left behind.

There are roughly a dozen documented cases of people purchasing old PCs and finding sensitive data.

- A woman in Pahrump, NV bought a used PC with pharmacy records [Markoff 97]
- Pennsylvania sold PCs with “thousands of files” on state employees [Villano 02]
- Paul McCartney’s bank records sold by his bank [Leyden 04]
- O&O Software GmbH – 200 drives.[O&O 05]

Information is even left behind on USB drives...



A “new” drive contained images from a previous owner.

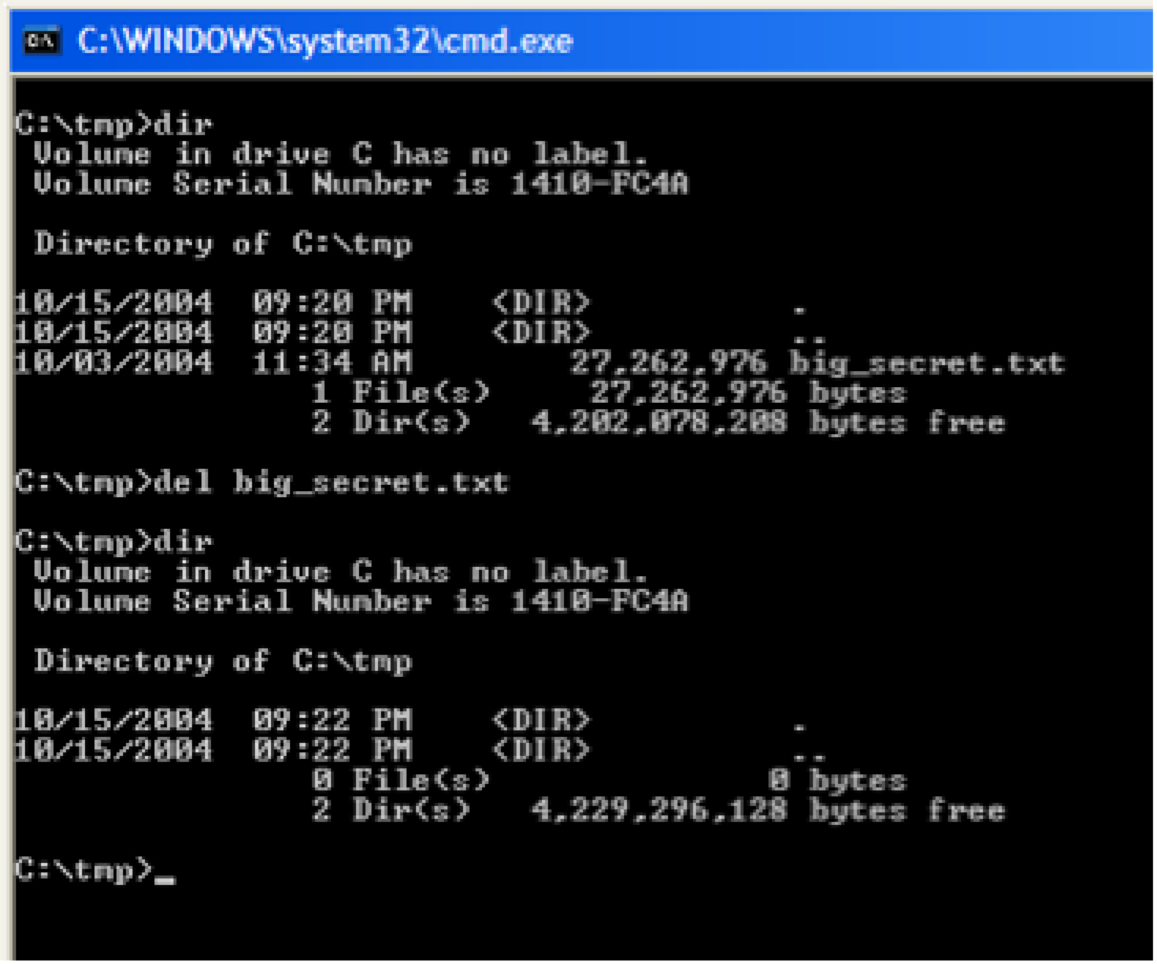
Why don’t we hear more stories?

Hypothesis #1: Disclosure of “data passed” is exceedingly rare because most systems are properly cleared.

Hypothesis #2: Disclosures are so common that they are not newsworthy.

Hypothesis #3: Systems aren’t properly cleared, but few people notice the data.

How could people not notice the data?



```
C:\tmp>dir
 Volume in drive C has no label.
 Volume Serial Number is 1410-FC4A

 Directory of C:\tmp

10/15/2004  09:20 PM    <DIR>          .
10/15/2004  09:20 PM    <DIR>          ..
10/03/2004  11:34 AM        27,262,976 big_secret.txt
               1 File(s)     27,262,976 bytes
               2 Dir(s)   4,202,078,208 bytes free

C:\tmp>del big_secret.txt

C:\tmp>dir
 Volume in drive C has no label.
 Volume Serial Number is 1410-FC4A

 Directory of C:\tmp

10/15/2004  09:22 PM    <DIR>          .
10/15/2004  09:22 PM    <DIR>          ..
               0 File(s)              0 bytes
               2 Dir(s)   4,229,296,128 bytes free

C:\tmp>_
```

DEL removes the file’s name; doesn’t delete the data.

Weird Stuff, Sunnyvale California,January 1999

10GB drive: $19 “tested”

500 MB drive: $3 “as is”

Q: “How do you sanitize them?”

A: “We FDISK them!”

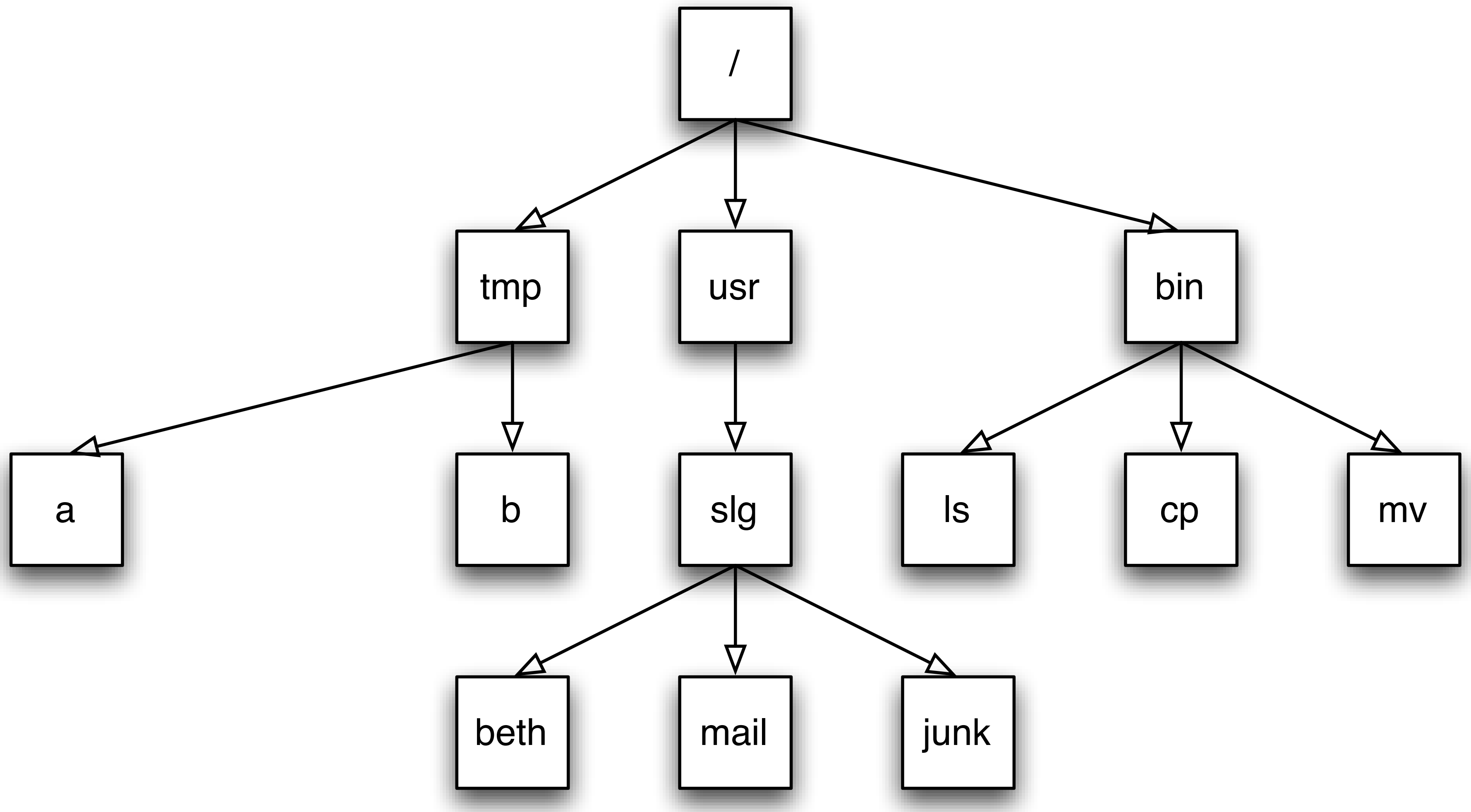
FDISK does not sanitize disks

10 GB drive: 20,044,160 sectors

Command	Sectors Written	%
FORMAT	21,541	0.11%
FDISK	2,563	0.01%

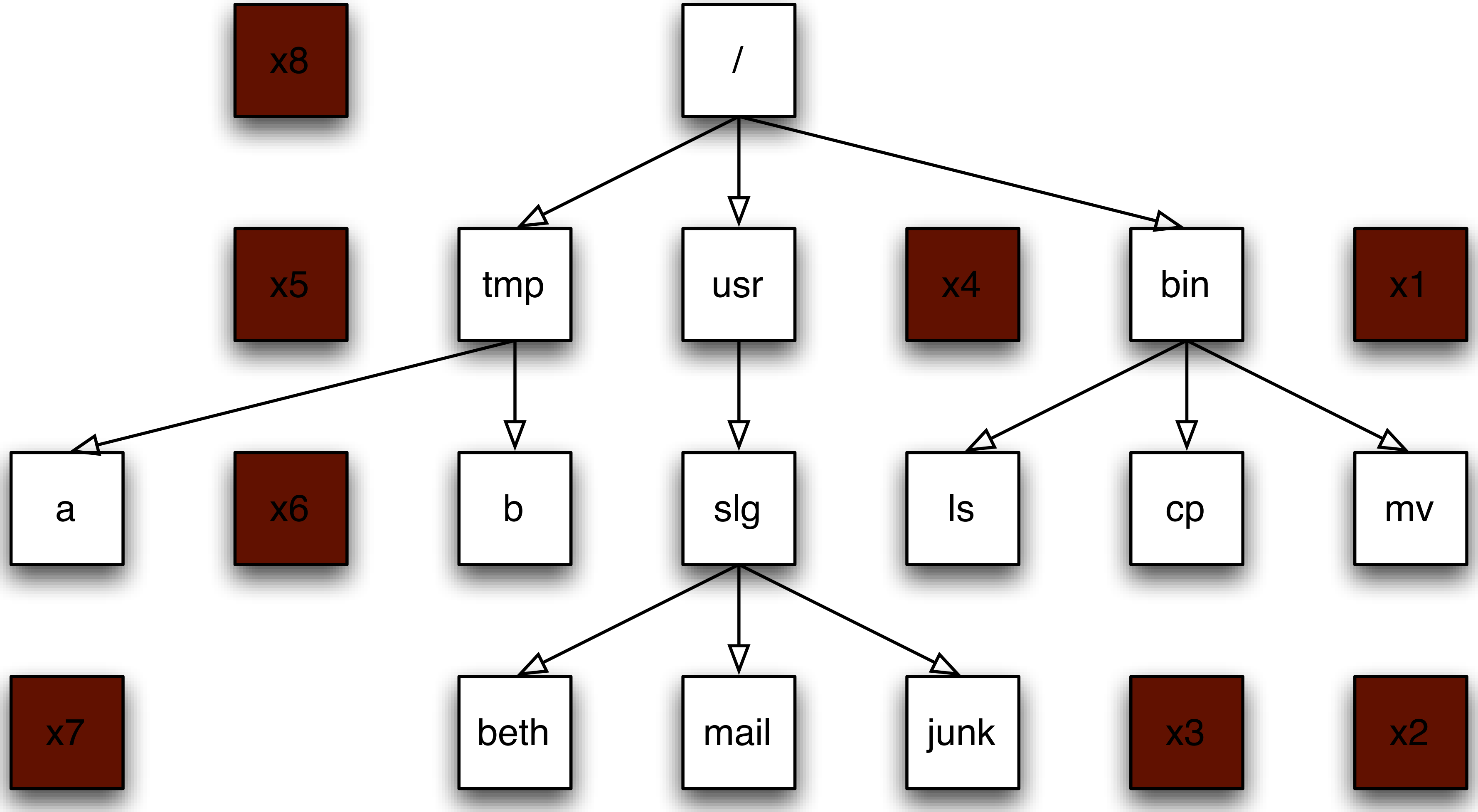
FORMAT erases the FAT,
complicating the recovery of fragmented files.

Data on a hard drive is arranged in sectors.



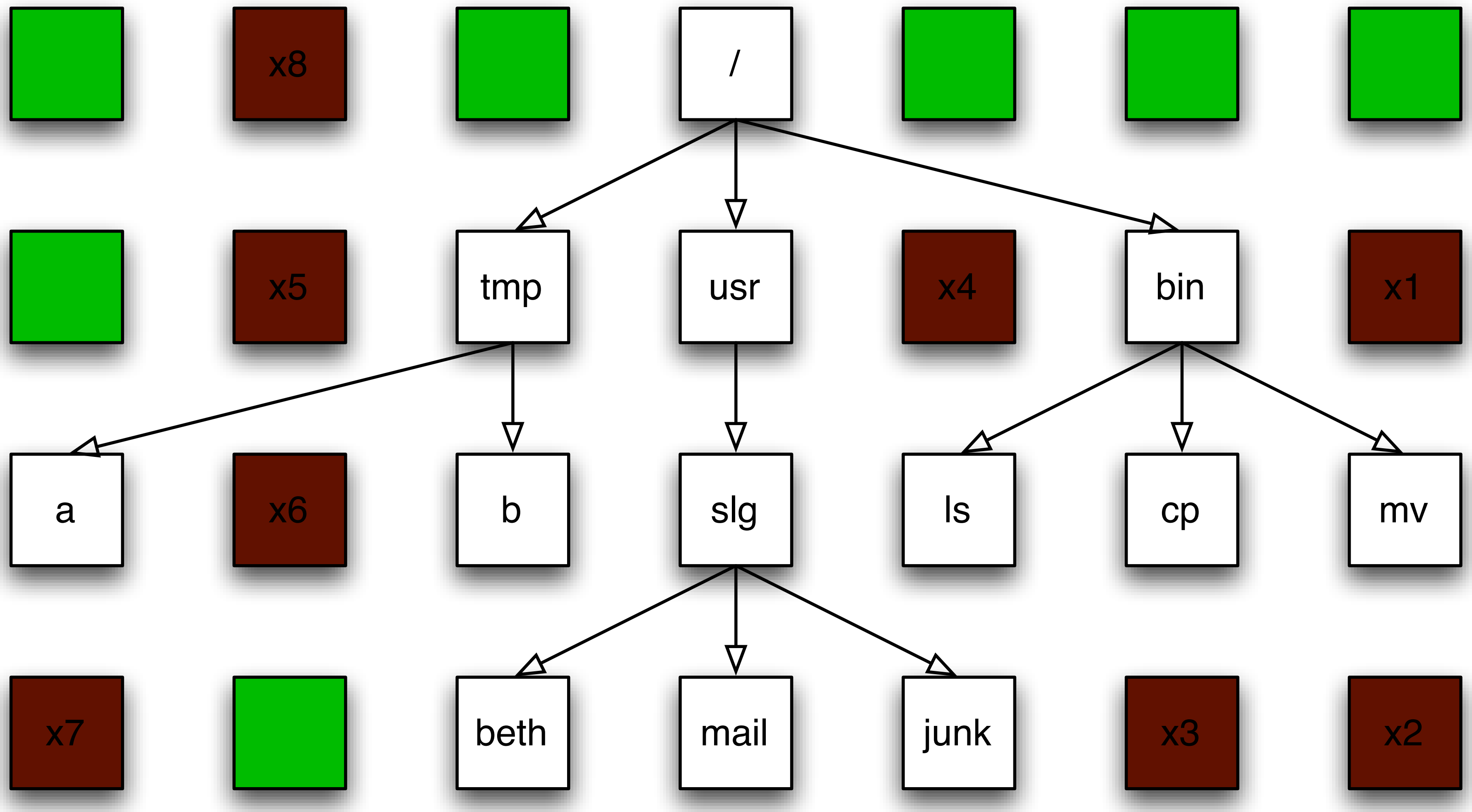
The white sectors indicate directories and files that are visible to the user.

Data on a hard drive is arranged in sectors.



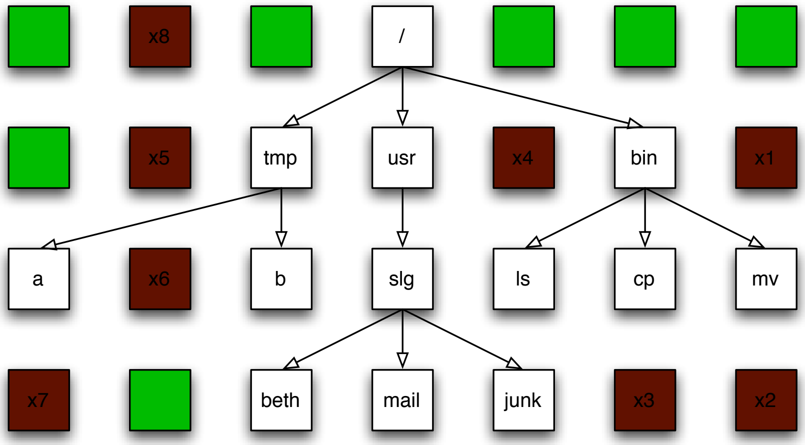
The brown sectors indicate files that were deleted.

Data on a hard drive is arranged in sectors.



The green sectors indicate sectors that were never used (or that were wiped clean).

Stack the disk sectors:

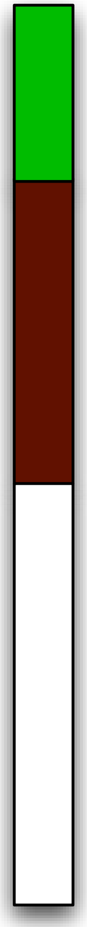


x8 /

x5 tmp usr x4 bin x1

a x6 b slg ls cp mv

x7 beth mail junk x3 x2

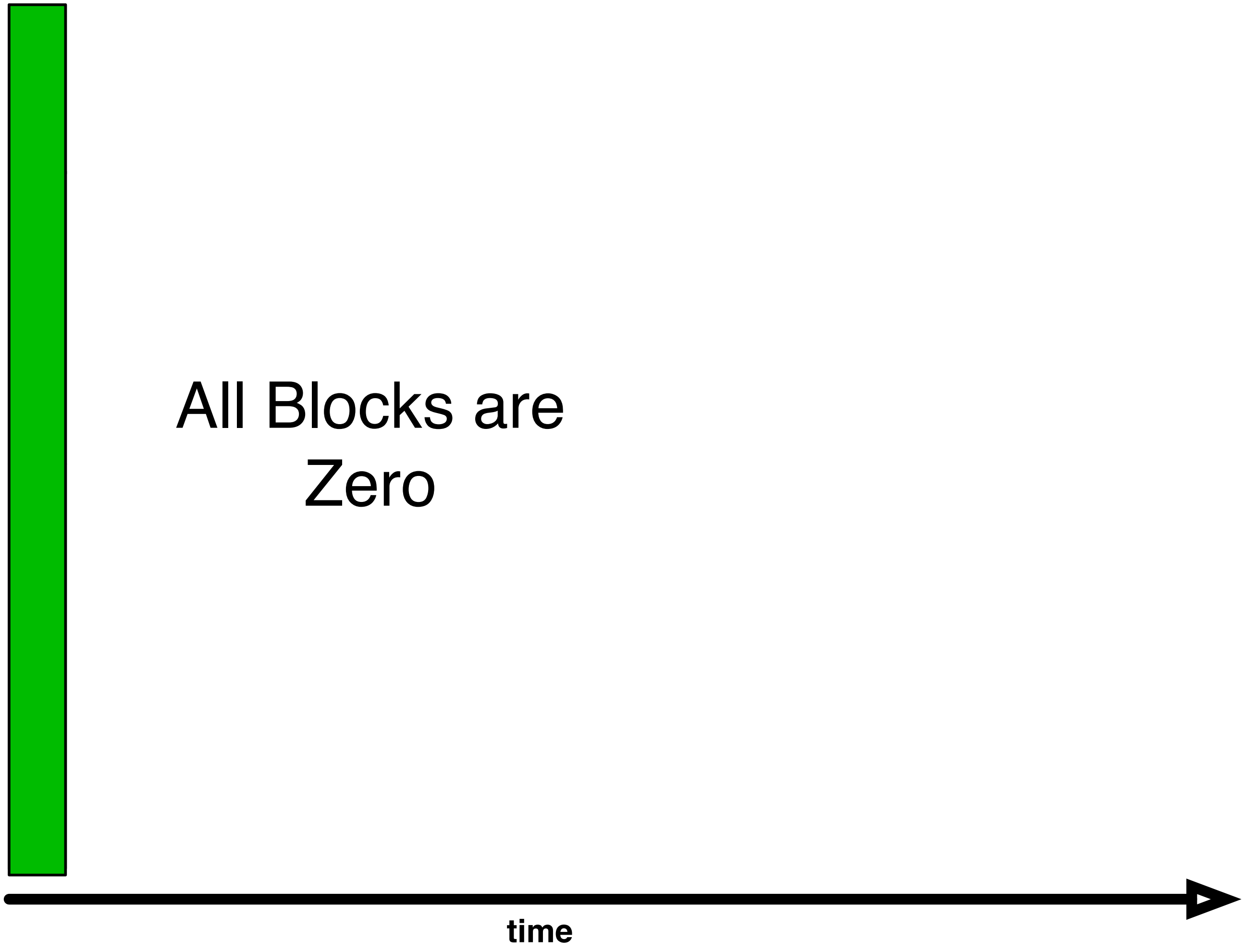


Zero Blocks

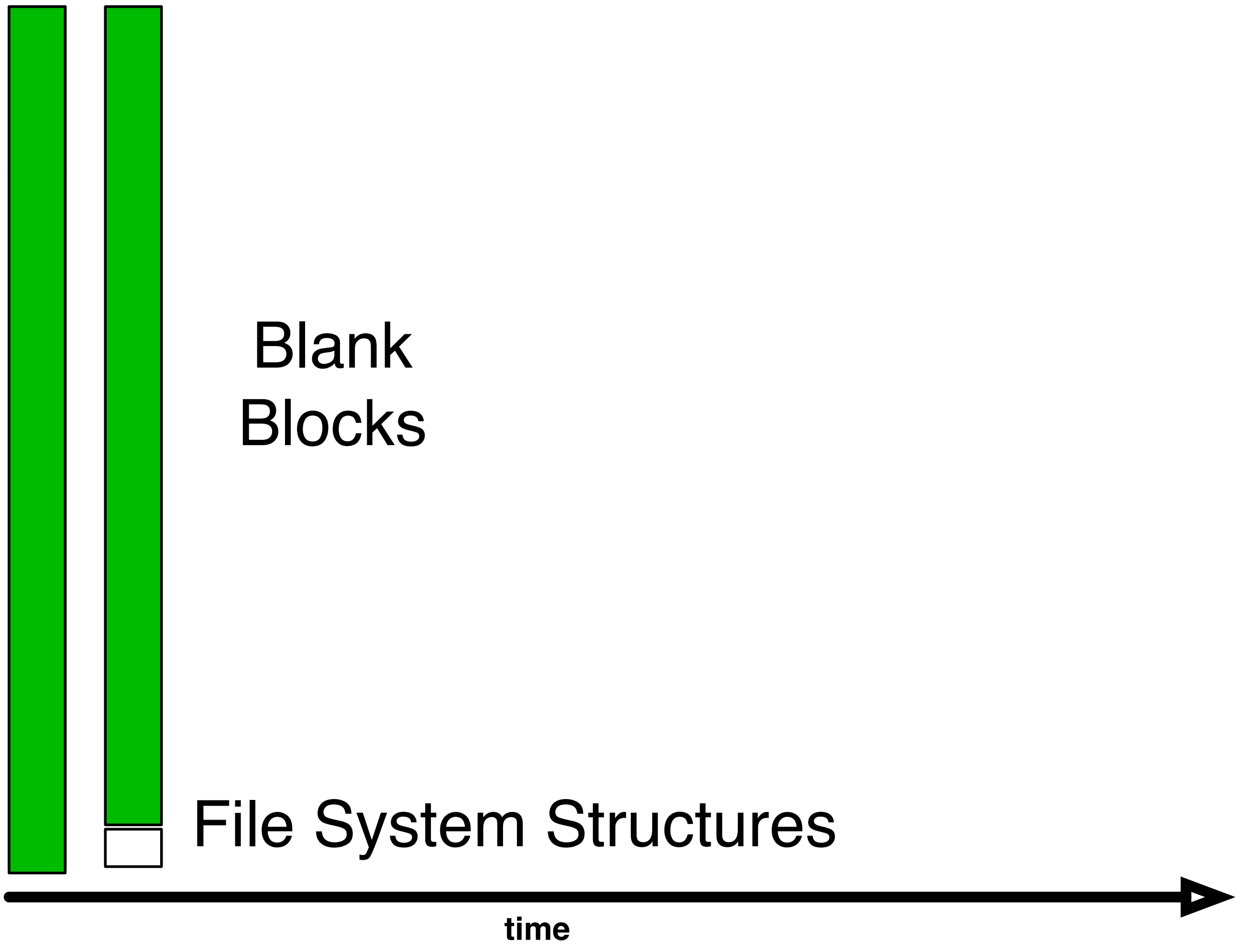
Deleted Files

Files

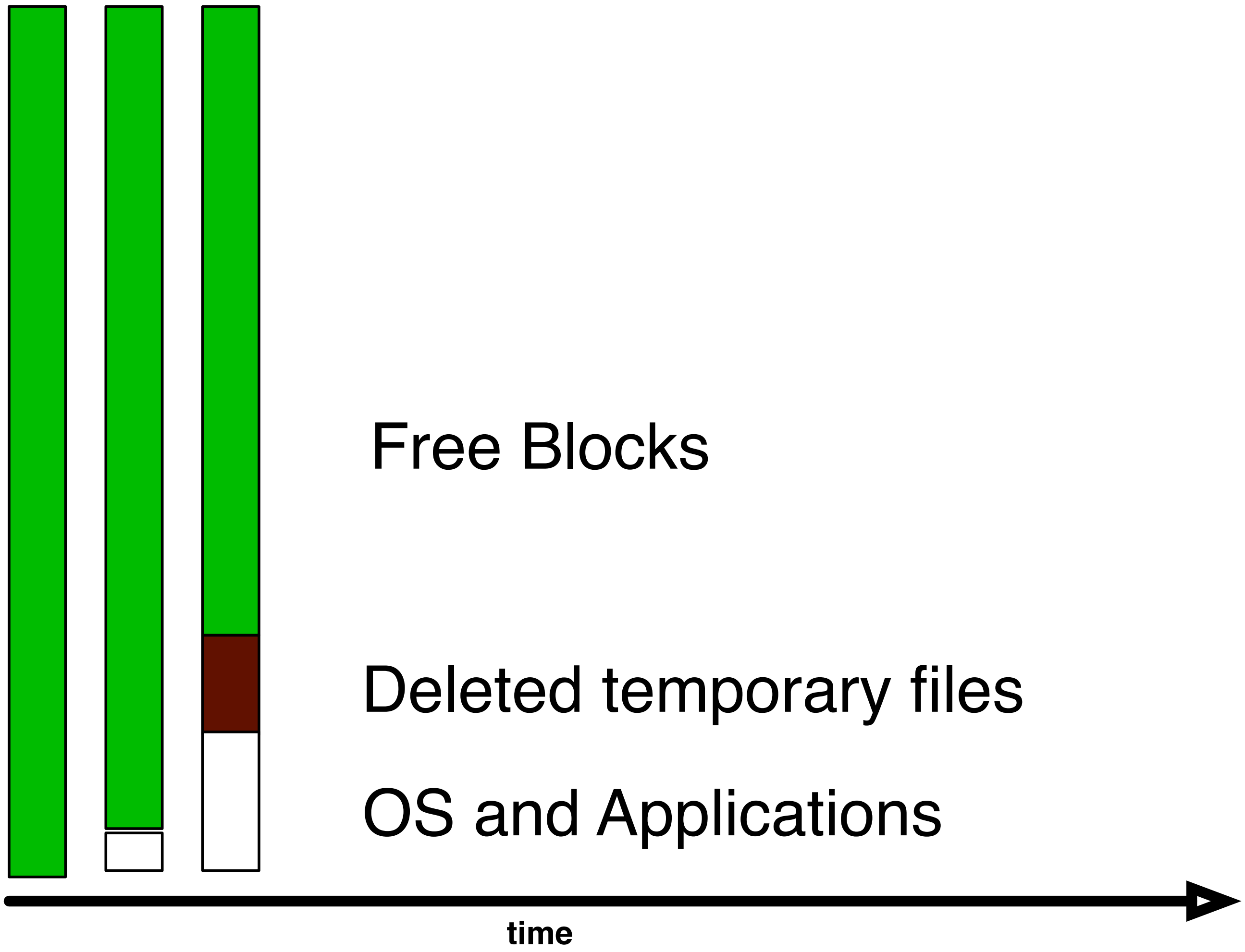
NO DATA: The disk is factory fresh.



FORMATTED: The disk has an empty file system



AFTER OS INSTALL: Temp. files have been deleted

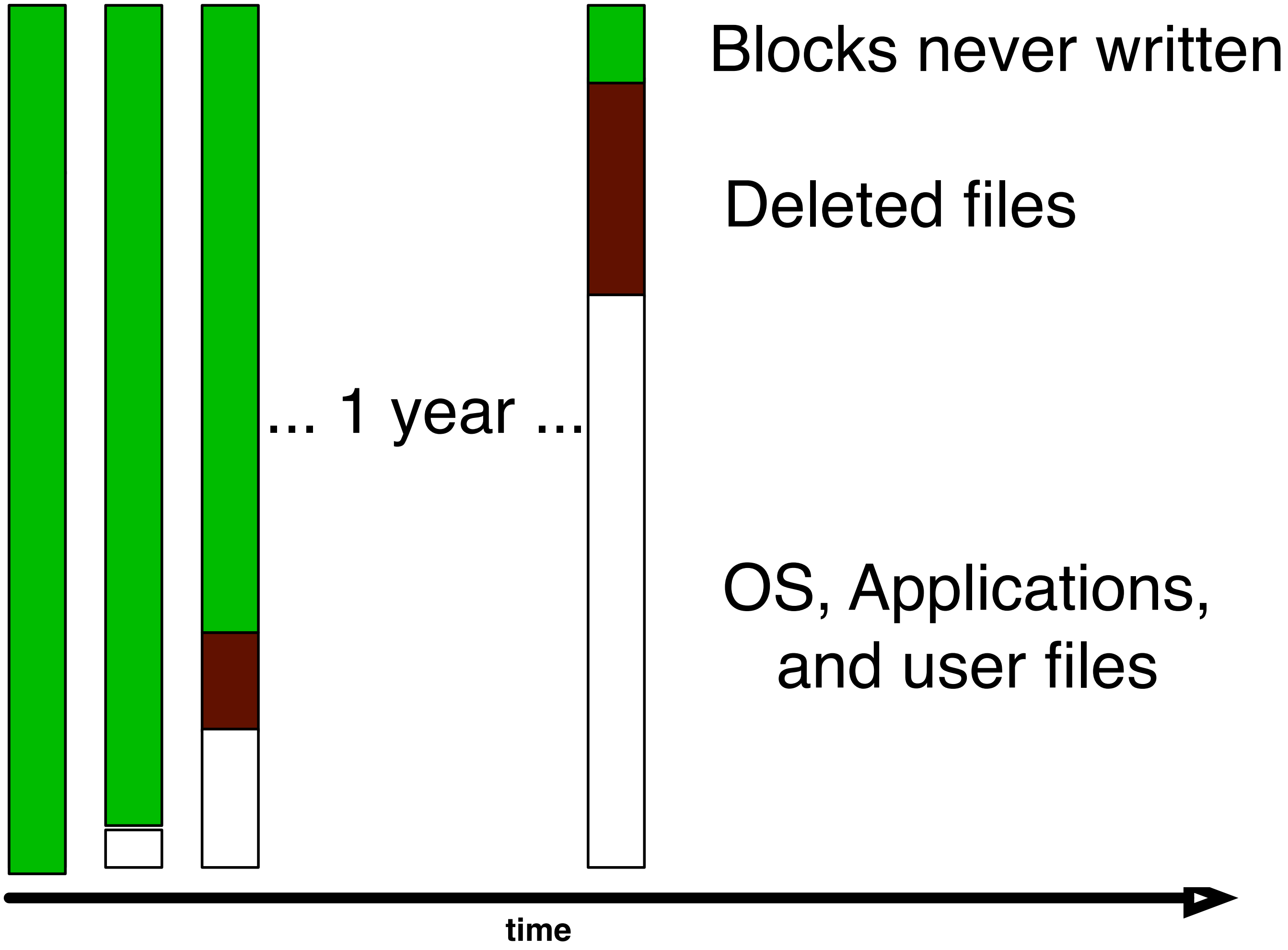


Free Blocks

Deleted temporary files

OS and Applications

AFTER A YEAR OF SERVICE



Blocks never written

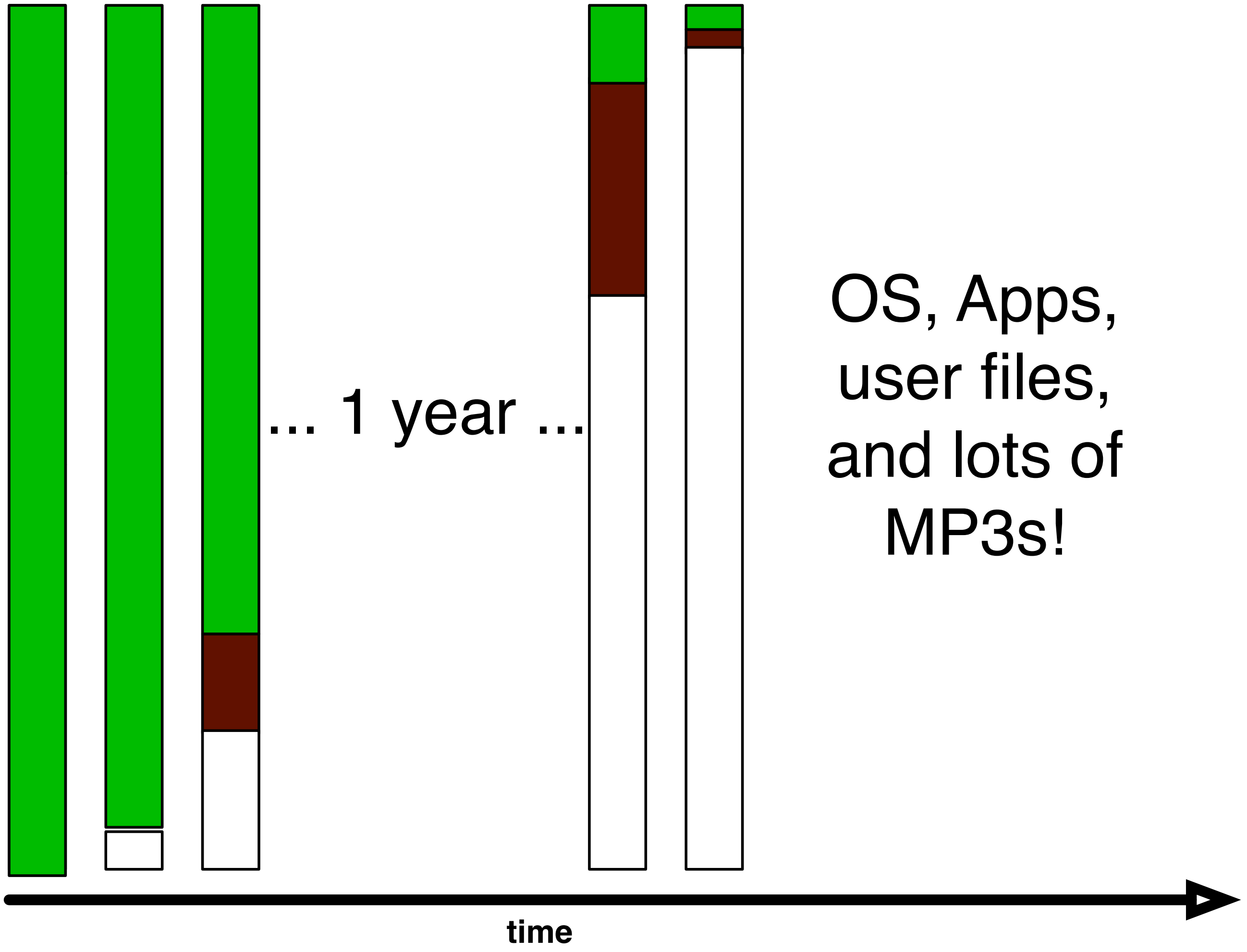
Deleted files

OS, Applications,
and user files

... 1 year ...

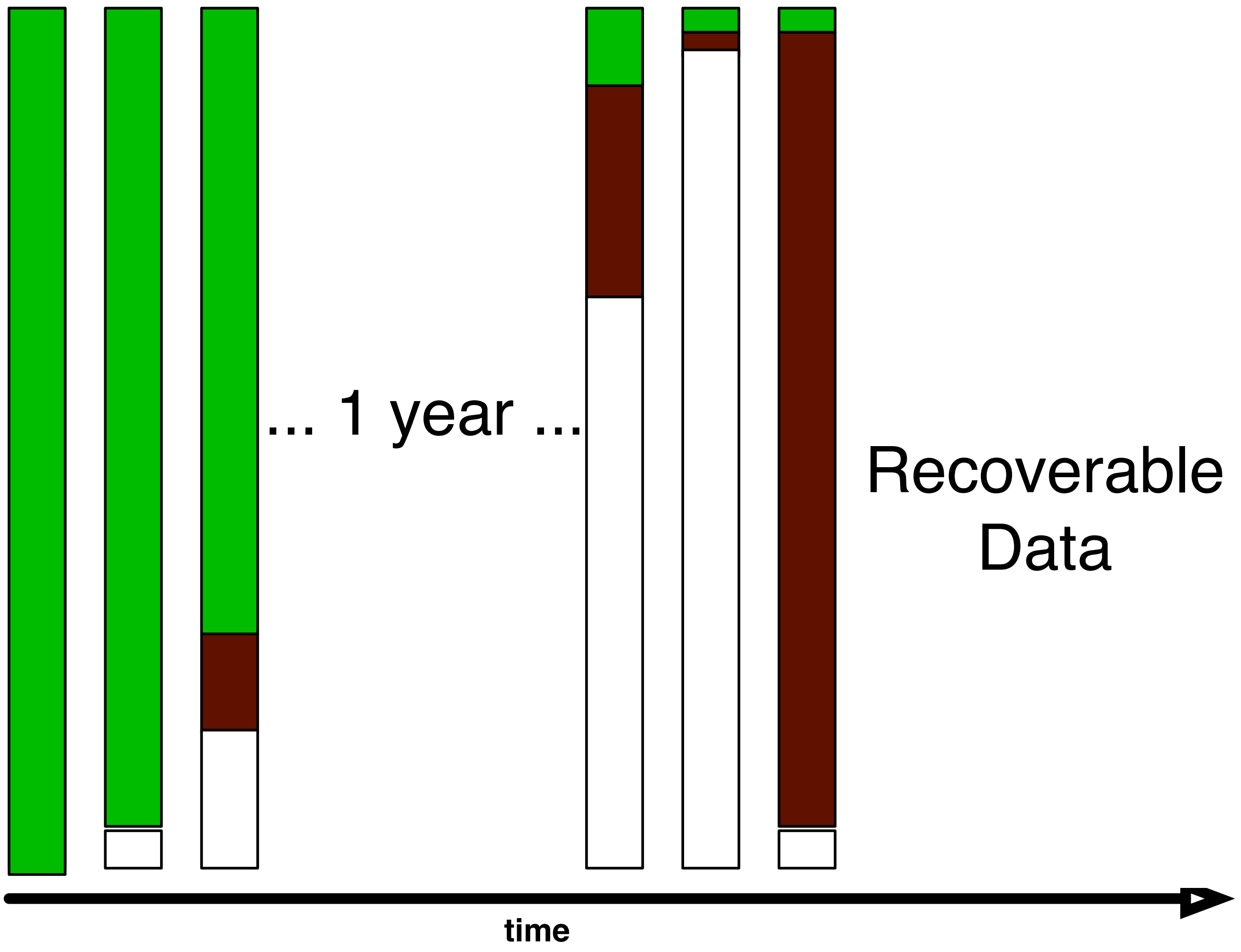
time

DISK NEARLY FULL!

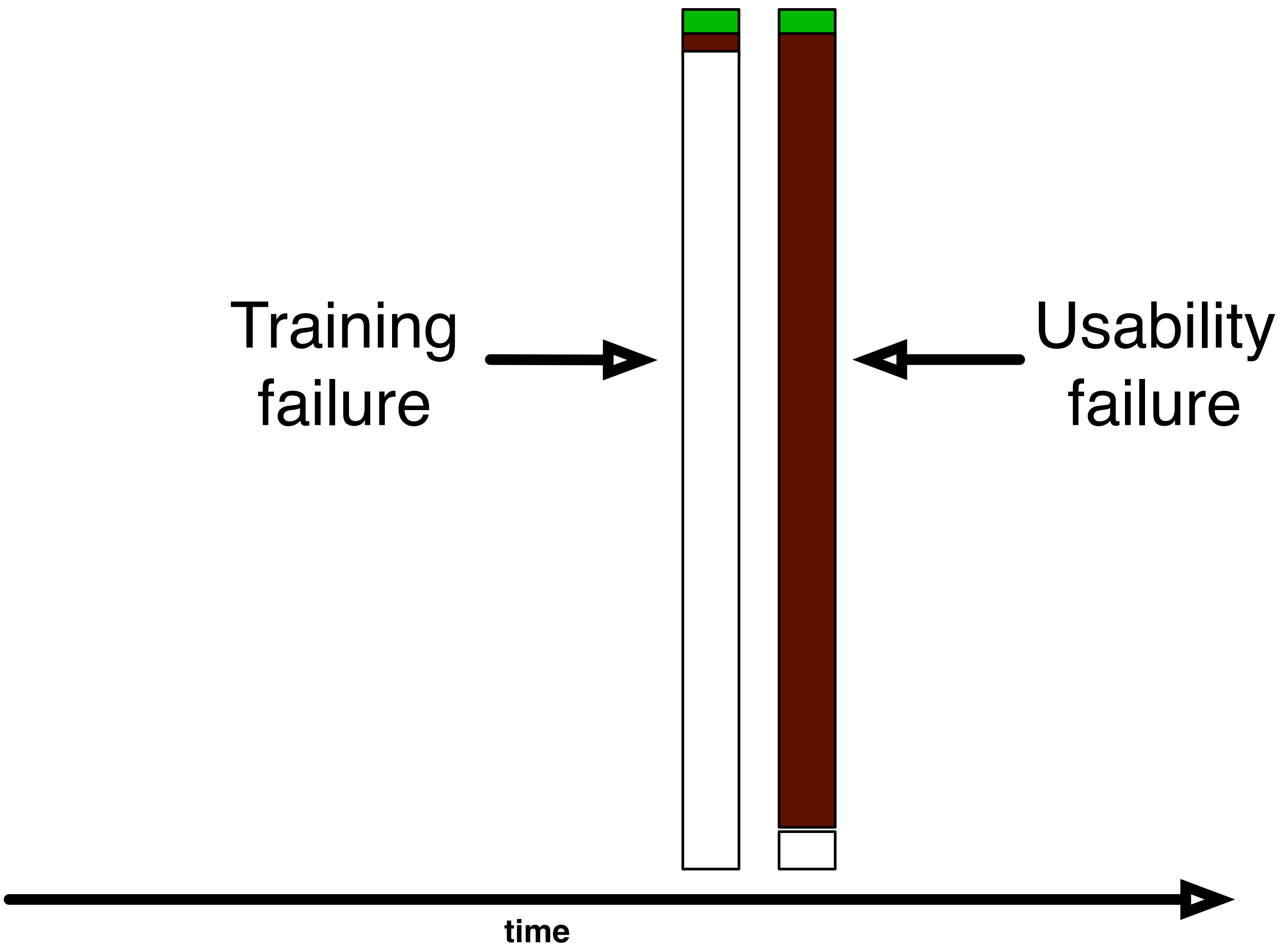


OS, Apps, user files, and lots of MP3s!

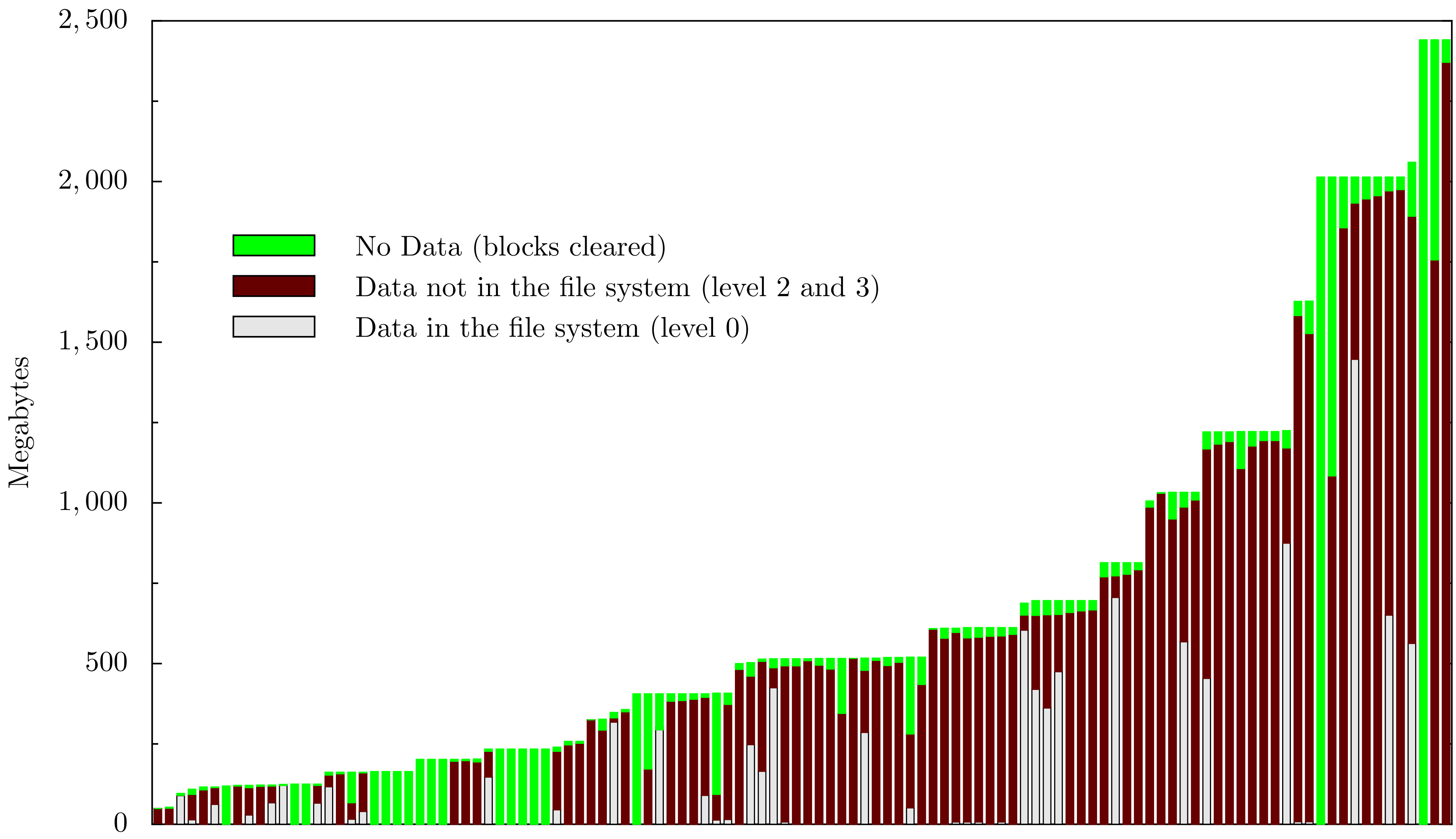
FORMAT C:\ (to sell the computer.)



We can use forensics to reconstruct motivations:



The drives are dominated by failed sanitization attempts...



..but training failures are also important.

Overall numbers

Drives Acquired:	236
Drives DOA:	60
Drives Images:	176
Drives Zeroed:	11
Drives “Clean Formatted:”	22
Total files:	168,459
Total data:	125G

Only 33 out of 176 working drives were properly cleared!

- 1 from Driveguys — but 2 others had lots of data.
- 18 from pcjunkyard — but 7 others had data.
- 1 from a VA reseller — 1 DOA; 3 dirty formats.
- 1 from an unknown source — 1 DOA, 1 dirty format.
- 1 from Mr. M. who sold his 2GB drive on eBay.

But what *really* happened?

?

I needed to contact the original drive owners.

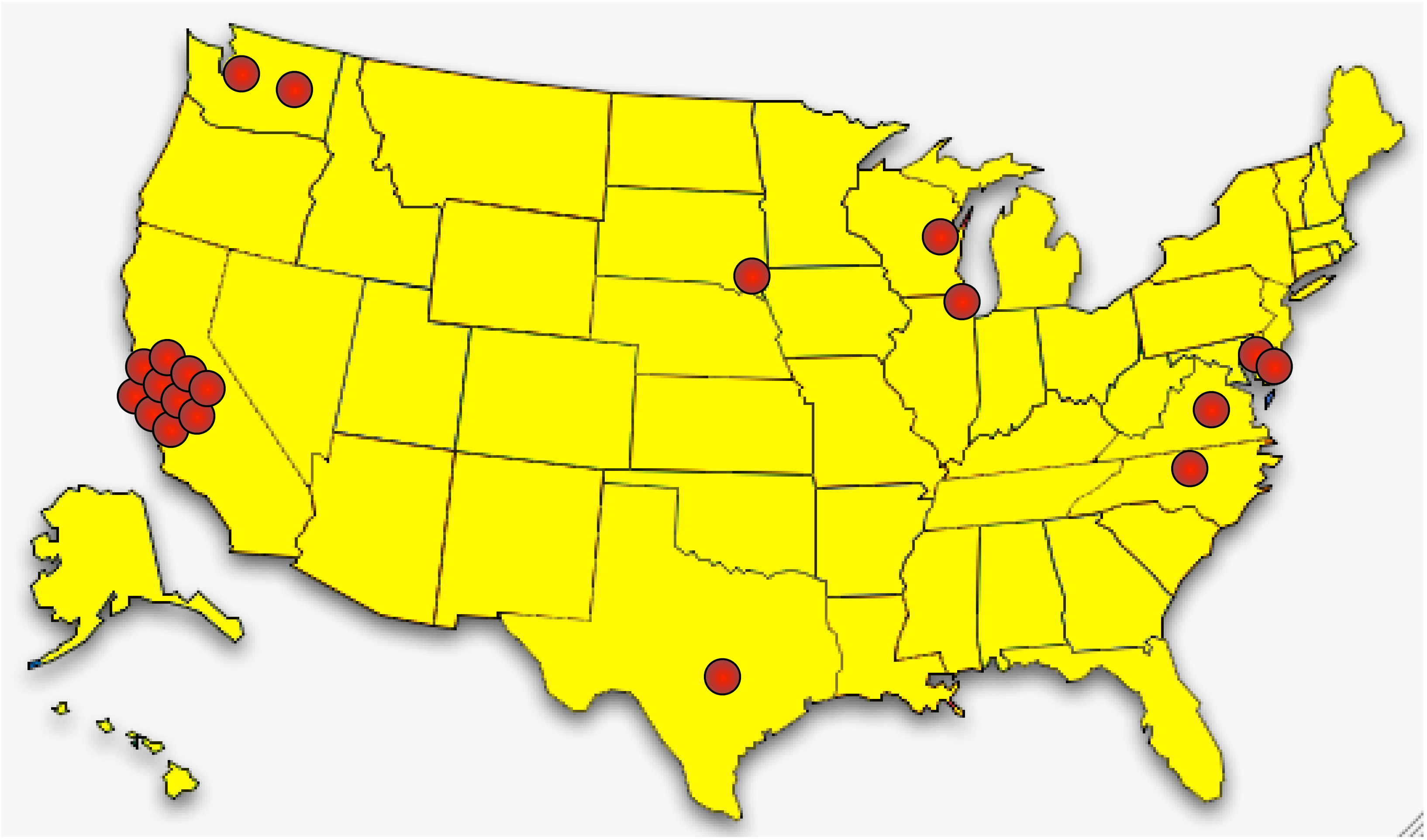
The *Remembrance of Data Passed Traceback Study.* [Garfinkel 05]

1. Find data on hard drive
2. Determine the owner
3. Get contact information for organization
4. Find the right person *inside* the organization
5. Set up interviews
6. Follow guidelines for human subjects work

```
06/19/1999 /:dir216/Four H Resume.doc
03/31/1999 /:dir216/U.M. Markets & Society.doc
08/27/1999 /:dir270/Resume-Deb.doc
03/31/1999 /:dir270/Deb-Marymount Letter.doc
03/31/1999 /:dir270/Links App. Ltr..doc
08/27/1999 /:dir270/Resume=Marymount U..doc
03/31/1999 /:dir270/NCR App. Ltr..doc
03/31/1999 /:dir270/Admissions counselor, NCR.doc
08/27/1999 /:dir270/Resume, Deb.doc
03/31/1999 /:dir270/UMUC App. Ltr..doc
03/31/1999 /:dir270/Ed. Coordinator Ltr..doc
03/31/1999 /:dir270/American College ...doc
04/01/1999 /:dir270/Am. U. Admin. Dir..doc
04/05/1999 /:dir270/IR Unknown Lab.doc
04/06/1999 /:dir270/Admit Slip for Modernism.doc
04/07/1999 /:dir270/Your Honor.doc
```

This was a lot harder than I thought it would be.

Ultimately, I contacted 20 organizations between April 2003 and April 2005.



The leading cause: betrayed trust.

Trust Failure: 5 cases

- ✔ Home computer; woman’s son took to “PC Recycle”
- ✔ Community college; no procedures in place
- ✔ Church in South Dakota; administrator “kind of crazy”
- ✔ Auto dealership; consultant sold drives he “upgraded”
- ✔ Home computer, financial records; same consultant

This specific failure wasn’t considered in [GS 03]; it was the most common failure.

Second leading cause: Poor training and supervision

Trust Failure: 5 cases

Lack of Training: 3 cases

- ✔ California electronic manufacturer
- ✔ Supermarket credit-card processing terminal
- ✔ ATM machine from a Chicago bank

Alignment between the interface and the underlying representation would overcome this problem.

Sometimes the data custodians just don’t care.

Trust Failure: 5 cases

Lack of Training: 3 cases

Lack of Concern: 2 cases

- ✔ Bankrupt Internet software developer
- ✔ Layoffs at a computer magazine

Regulation on resellers might have prevented these cases.

In seven cases, no cause could be determined.

Trust Failure: 5 cases
Lack of Training: 3 cases
Lack of Concern: 2 cases

Unknown Reason: 7 cases

- ✘ Bankrupt biotech startup
- ✘ Another major electronics manufacturer
- ✘ Primary school principal’s office
- ✘ Mail order pharmacy
- ✘ Major telecommunications provider
- ✘ Minnesota food company
- ✘ State Corporation Commission

Regulation might have helped here, too.

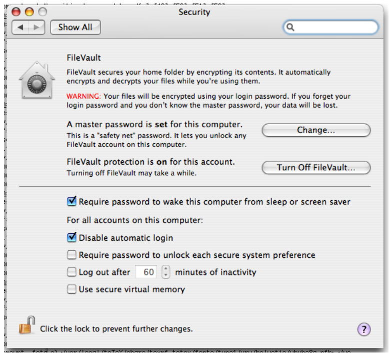
Legislative reactions to this research: “Fair and Accurate Credit Transactions Act of 2003” (US)

- Introduced in July 2003. Signed December 2003.
- Regulations adopted in 2004, effective June 2005.
- Amends the FCRA to standardize consumer reports.
- Requires destruction of paper or electronic “consumer records.”

The adopted rules specifically exempt resellers.

There are three options for solving this problem.

1. Encrypt stored data.



FileVault

FileVault secures your home folder by encrypting its contents. It automatically encrypts and decrypts your files while you're using them.

WARNING: Your files will be encrypted using your login password. If you forget your login password and you don't know the master password, your data will be lost.

A master password is **set** for this computer.
This is a "safety net" password. It lets you unlock any FileVault account on this computer.

FileVault protection is **on** for this account.
Turning off FileVault may take a while.

☑ Require password to wake this computer from sleep or screen saver

For all accounts on this computer:

☑ Disable automatic login

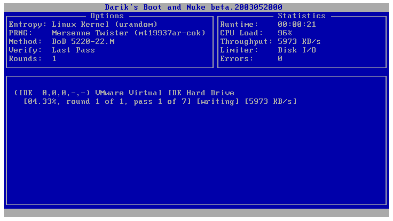
□ Require password to unlock each secure system preference

□ Log out after 60 minutes of inactivity

□ Use secure virtual memory

Click the lock to prevent further changes.

2. Clear drives before discarding them.



```
                 Darik's Boot and Nuke beta.2003052000
 ─────────── Options ──────────────     ────── Statistics ──────
 Entropy: Linux Kernel (urandom)        Runtime:    00:00:21
 PRNG:    Mersenne Twister (mt19937ar-cok)  CPU Load:   96%
 Method:  DoD 5220-22.M                 Throughput: 5973 KB/s
 Verify:  Last Pass                     Limiter:    Disk I/O
 Rounds:  1                             Errors:     0

  (IDE  0,0,0,-,-) VMware Virtual IDE Hard Drive
    [04.33%, round 1 of 1, pass 1 of 7] [writing] [5973 KB/s]
```

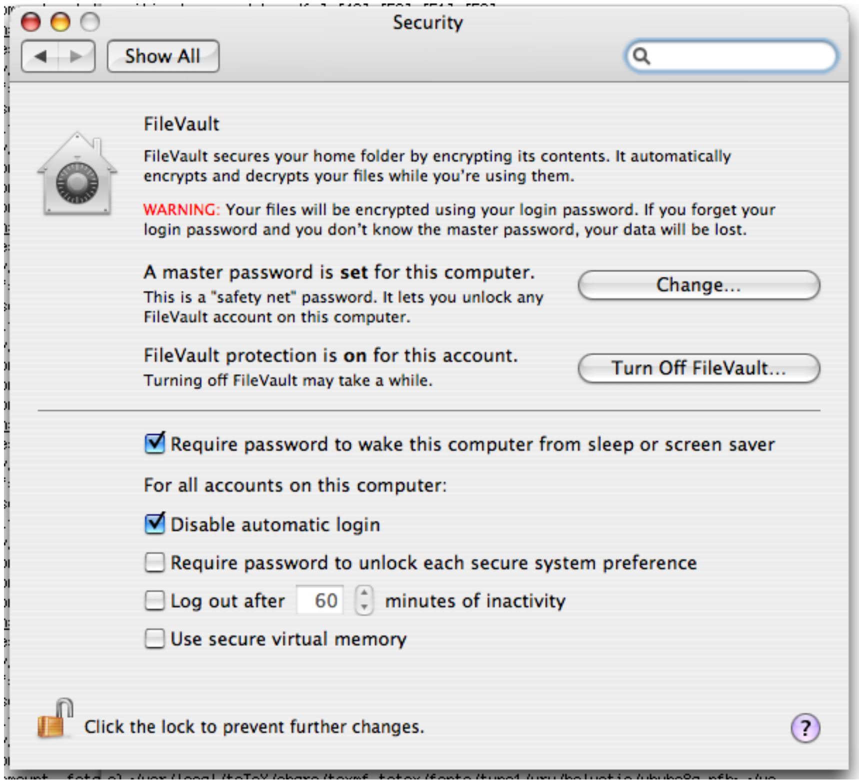
3. Physically destroy drives when discarding.

Of these options, encryption is the best alternative.

Encryption protects protects data against anyone who doesn’t have the key

Options:

- Record-level encryption
- Database encryption
- Encrypted file systems



FileVault

FileVault secures your home folder by encrypting its contents. It automatically encrypts and decrypts your files while you’re using them.

WARNING: Your files will be encrypted using your login password. If you forget your login password and you don’t know the master password, your data will be lost.

A master password is **set** for this computer.
This is a "safety net" password. It lets you unlock any FileVault account on this computer.

FileVault protection is **on** for this account.
Turning off FileVault may take a while.

☑ Require password to wake this computer from sleep or screen saver

For all accounts on this computer:

☑ Disable automatic login

□ Require password to unlock each secure system preference

□ Log out after 60 minutes of inactivity

□ Use secure virtual memory

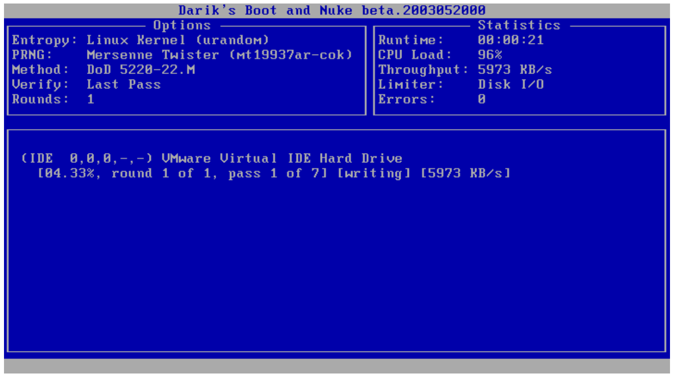
Click the lock to prevent further changes.

Apple’s File Vault is an easy-to-use encrypted file system.

Drives can be individually “cleared” before they are repurposed or retired.

Commercial Software

- Wipe Drive, White Canyon Software.
  `http://www.whitecanyon.com/`
- Blancco Data Cleaner
  `http://www.blancco.com/`



```
                 Darik's Boot and Nuke beta.2003052000
─────────────── Options ──────────────  ───────── Statistics ─────────
Entropy: Linux Kernel (urandom)          Runtime:    00:00:21
PRNG:    Mersenne Twister (mt19937ar-cok) CPU Load:   96%
Method:  DoD 5220-22.M                   Throughput: 5973 KB/s
Verify:  Last Pass                       Limiter:    Disk I/O
Rounds:  1                               Errors:     0

 (IDE  0,0,0,-,-) VMware Virtual IDE Hard Drive
   [04.33%, round 1 of 1, pass 1 of 7] [writing] [5973 KB/s]
```

DBAN

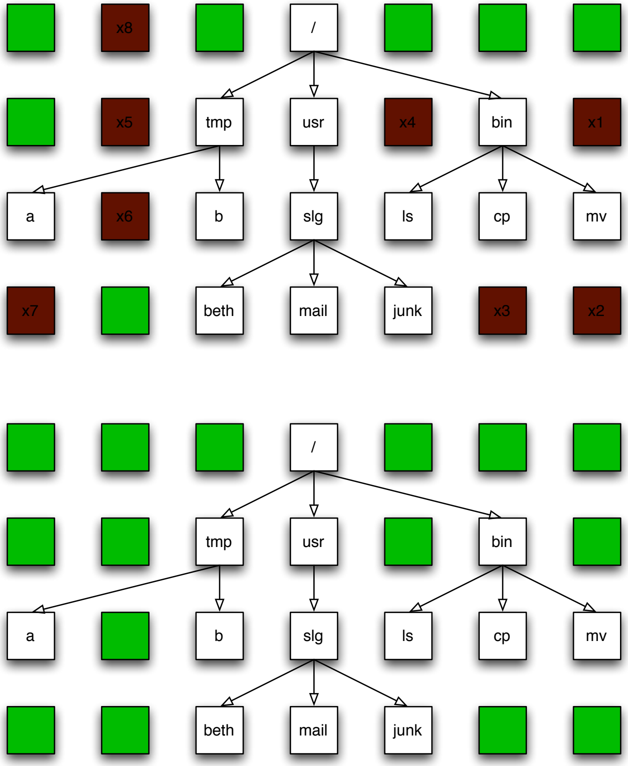
Free Software:

- DBAN
  `http://www.dban.sourceforget.net`

Wiping an entire drive is easy.
Selectively wiping is hard.

Features to look for in a wiping program:

- Cost: Free? Per copy? Per use?
- Bootable disk?
- Wipe whole drive?
- Wipe slack space?



Program that wipe slack space can be installed on every desktop.

Physical destruction offers the most security.

Destruction is the only technique approved for use with classified information.

Summary

Information left on hard drives:

- is a serious problem.
- is invisible
- violates Federal law
- is easily address

This is largely the fault of Microsoft, Apple, and Linux.

Destroying this information is your organization's *legal* and *moral* obligation.

Questions?