

# Informatica forense

---

*Disk forensics*

## Data carving

*Michele Ferrazzano*

24/04/2012

## Premessa

---

- I file, anche se cancellati, non vengono effettivamente rimossi dall'hard disk
- Gli hard disk, anche se formattati, non cancellano effettivamente il vecchio contenuto
- I file cancellati che risiedono sull'hard disk, finché lo spazio che occupavano non viene rimpiazzato, possono essere recuperati

## File system

- Un file system è una struttura atta ad organizzare in maniera logica i dati memorizzati in una memoria di massa
- I dischi sono organizzati per settori (tipicamente 512 byte o multipli); il file system tiene traccia di quali settori sono allocati e quali non allocati

## File system

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Settori da 512 byte															
0	1			2		3		4		5		6		7	
Settori da 1024 byte															
0				1				2				3			
Settori da 2048 byte															
0								1							
Settori da 4096 byte															



## File

---

- Il termine file è utilizzato per indicare un blocco di dati binari che costituiscono un aggregato omogeneo
- I file possono essere creati, modificati, spostati, copiati, cancellati
- Tipicamente il tipo di file è identificato dall'estensione; in realtà utile solo per associare il programma con il quale aprire il file

## File

---

- Il magic number è una costante che consente di identificare un file
  - Es: PDF comincia con **%PDF** e termina con **%EOF**
  - ES: JPG comincia con **0xFFD8** e termina con **0xFFD9**
- Comando `file` di linux

## Data carving - Definizione

---

- Il data carving è il processo di estrazione di un insieme di dati da un insieme molto ampio di dati
- La tecnica del data carving è utilizzata solitamente durante le indagini di analisi forense per analizzare lo spazio non allocato
- I file estratti dallo spazio non allocato sono catalogati sulla base dell'header e del footer
- La struttura del file system è ignorata durante questo procedimento

## Data carving

---

- Data carving base
  - L'header e footer del file non sono sovrascritti
  - Il file non è frammentato
  - Il file non è compresso
  - Il file estratto è l'insieme di bit contenuti tra header e footer
- Data carving avanzato
  - I frammenti non sono sequenziali
  - I frammenti non sono ordinati
  - Mancano dei frammenti

## Data carving - Considerazioni

---

- Consente di recuperare file o frammenti di file cancellati, anche quando il file system è danneggiato; è altresì utile per recuperare dati da memoria prive di una struttura dati definita come i dump della memoria RAM
- Può essere svolto su un dispositivo hardware o su un'immagine
- In alcuni file può mancare il footer
- Un inconveniente del processo di data carving è che ci possono essere numerosi falsi positivi
  - È consigliabile controllare ogni file estratto per verificarne la consistenza

## Carving – (alcuni) software

---

- Foremost
- Scalpel
- Photorec
  - Lavorano ad un livello inferiore, sui blocchi di dati
  - Es: i tool possono trovare documenti in immagini dd, dump della RAM o nel file di swap

## Carving – foremost e scalpel

---

- **foremost** un il più famoso tool di carving su linux
  - <http://foremost.sourceforge.net>
- Sui sistemi debian-link sono installabili con i seguenti comandi
  - `sudo apt-get install foremost`
  - `sudo apt-get install scalpel`
- foremost e scalpel sono in grado di lavorare su file immagine (dd, ewf...) o direttamente sul dispositivo
- Gli headers ed i footers possono essere specificati in file di configurazione o attraverso parametri della riga di comando

## Carving – foremost

---

- Una volta lanciato, foremost crea nella directory di output un file *audit.txt* ed una serie di sottocartelle nominate col tipo di file ricercato (es: doc,jpg, tiff...)
  - All'interno delle suddette cartelle ci sono i file recuperati sulla base degli headers e dei footers
    - File attivi, cancellati e non-allocati
- foremost agisce sulla parte dati del dispositivo da analizzare e non considera il file system
  - Utile per recuperare i dati dai supporti formattati
  - I files trovati non avranno il nome, ma saranno nominati con l'indirizzo del settore sul quale sono allocati
  - Sia i nomi dei files sia i loro offset sono scritti nel file *audit.txt*

## Carving – foremost

---

- foremost consente di effettuare la ricerca di una singola stringa di testo
  - nel file */etc/foremost.conf* aggiungere una riga del tipo  
`testo n 1000 stringa`
  - inserendo al posto di *stringa* i caratteri da ricercare

## Carving – foremost

---

- La modalità quick (-q) obbliga il programma a ricercare l'header solo all'inizio di ogni settore
- Se non è impostata una directory di output (-o), foremost salva i file in una cartella predefinita chiamata *foremost-output*, contenuta nella directory corrente
- Il flag -s consente di partire da un determinato offset per dividere la scansione in più parti o perfezionare la ricerca di determinati file

## Carving – foremost (tipi di file)

- |  |  |
|--|--|
| <ul style="list-style-type: none"> <li>▪ jpg</li> <li>▪ gif</li> <li>▪ png</li> <li>▪ bmp</li> <li>▪ avi</li> <li>▪ exe</li> <li>▪ mpg               <ul style="list-style-type: none"> <li>▪ begin with 0x000001BA</li> </ul> </li> <li>▪ wav</li> <li>▪ riff</li> <li>▪ wmv</li> <li>▪ mov</li> <li>▪ pdf</li> </ul> | <ul style="list-style-type: none"> <li>▪ ole               <ul style="list-style-type: none"> <li>▪ This will grab any file using the OLE file structure. This includes PowerPoint, Word, Excel, Access, and StarWriter</li> </ul> </li> <li>▪ doc               <ul style="list-style-type: none"> <li>▪ Note it is more efficient to run OLE as you get more bang for your buck. If you wish to ignore all other ole files then use this</li> </ul> </li> <li>▪ zip               <ul style="list-style-type: none"> <li>▪ Note it will extract .jar files as well because they use a similar format. Open Office docs are just zip'd XML files so they are extracted as well. These include SXW, SXC, SXI, and SX? for undetermined OpenOffice files.</li> </ul> </li> <li>▪ rar</li> <li>▪ htm</li> <li>▪ cpp</li> <li>▪ all               <ul style="list-style-type: none"> <li>▪ Run all pre-defined extraction methods. [Default if no -t is specified]</li> </ul> </li> </ul> |
|--|--|

## Carving – foremost – Esempio d'uso

- Supponiamo si vogliano recuperare tutti i file cancellati di tutte le estensioni da **sda1**
- Supponiamo si debbano riversare i dati su **sdb1**
  - Montare la cartella destinazione (**/data-recovery**) in scrittura
  - `sudo mount /dev/sdb1 /data-recovery -o rw`
- Si esegue foremost con il seguente comando
  - `sudo foremost -i /dev/sda1 -o /data-recovery`



## Carving – foremost – Esempio d'uso

- Alcuni esempi in foremost

- Ricerca tutti i tipi di file

```
sudo foremost -t all -i image.dd
```

- Ricerca tutti i tipi di file saltando i primi 100000 blocchi

```
sudo foremost -s 100000 -i /dev/sda1 -o /data-recovery
```

- Ricerca tutti i tipi di file di tipo gif e jpg

```
sudo foremost -t jpg,gif -i /dev/sda1 -o /data-recovery
```

- Ricerca di immagini jpg saltando i primi 100 blocchi

```
sudo foremost -s 100 -t jpg -i image.dd
```

- Ricerca documenti di office e file jpg

```
sudo foremost -t ole,jpeg -i image.dd
```

## Carving – foremost – Esempio d'uso

- Esempio di file audit.txt

```
Foremost version 1.5.7 by Jesse Kornblum, Kris Kendall, and Nick Mikus
FileForemost started at Fri Apr 20 10:15:45 2011
```

```
Invocation: foremost /dev/sda1
```

```
Output directory: /data-recovery
```

```
Configuration file: /usr/local/etc/foremost.conf
```

```
-----
File: /dev/sda1
```

```
Start: Fri Apr 20 10:15:45 2011
```

```
Length: 962 MB (1009254400 bytes)
```

```
Num Name (bs=512)
```

```
Size File Offset Comment0:
```

```
0: 00000621.jpg 11 KB 312832
```

```
1: 00000743.jpg 8 KB 329216
```

```
2: 00000774.jpg 3 KB 345600
```

```
3: 00000877.jpg 9 KB 361984
```

```
4: 00000983.jpg 8 KB 378768
-----
```

## Carving – foremost – Esempio d'uso

---

- Il nome rappresenta il settore in cui si trova il file
- Nell'esempio, il primo file si chiama 611
  - 611 è il settore in cui si trova il file
  - 512 è la dimensione in byte del settore
  - L'offset è  $512 \times 611 = 312832$ 
    - Rappresenta l'indirizzo assoluto in byte sul disco

## Carving – foremost – Esempio d'uso

---

- Controverifica su disco

```
root@deft:~/$ xxd -s 312832 -l 512 /dev/sda1
004c600: ffd8 ffe0 0010 4a46 4946 0001 0100 0001 .....JFIF.....
004c610: 0001 0000 ffdb 0043 0005 0304 0404 0305 .....C.....
004c620: 0404 0405 0505 0607 0c08 0707 0707 0f0b .....
004c630: 0b09 0c11 0f12 1211 0f11 1113 161c 1713 .....
004c640: 141a 1511 1118 2118 1a1d 1d1f 1f1f 1317 .....!.....
004c650: 2224 221e 241c 1e1f 1eff db00 4301 0505 "$".$.C...
004c660: 0507 0607 0e08 080e 1e14 1114 1e1e 1e1e .....
```

## Carving – scalpel

---

- scalpel è stato riscritto a partire da foremost per migliorare performance e ridurre uso di memoria
- Nel caso di scalpel, i file da identificare devono essere specificati nel file di configurazione `/etc/scalpel/scalpel.conf`
- Di default, i file sono commentati; per specificare i file da sottoporre a carving occorre rimuovere il simbolo di commento #

## Carving – scalpel – Esempio d'uso

---

- Supponiamo si vogliano recuperare tutti i file cancellati di tutte le estensioni da **sda1**
- Supponiamo si debbano riversare i dati su **sdb1**
  - Montare la cartella destinazione (**/data-recovery**) in scrittura
  - `sudo mount /dev/sdb1 /data-recovery -o rw`
- Si esegue foremost con il seguente comando
  - `sudo scalpel /dev/sda1 -o /data-recovery`

## Carving – foremost vs. scalpel

---

- Paper su carving con alcuni test di confronto
  - [http://www.sans.org/reading\\_room/whitepapers/forensics/data-carving-concepts\\_32969](http://www.sans.org/reading_room/whitepapers/forensics/data-carving-concepts_32969)
- foremost fornisce più risultati di scalpel testato sullo stesso device

## Carving – Photorec

---

- PhotoRec è un software di data recovery per tutti i tipi di file (non solo foto).
- PhotoRec può lavorare a stretto contatto con TestDisk
- PhotoRec è open source (GPLV v2+) ed è eseguibile su una molteplicità di sistemi
  - DOS/Win9x, Windows NT 4/2000/XP/2003/Vista/2008/7, Linux, FreeBSD, NetBSD, OpenBSD, Sun Solaris, Mac OS X

## Carving – Photorec

---

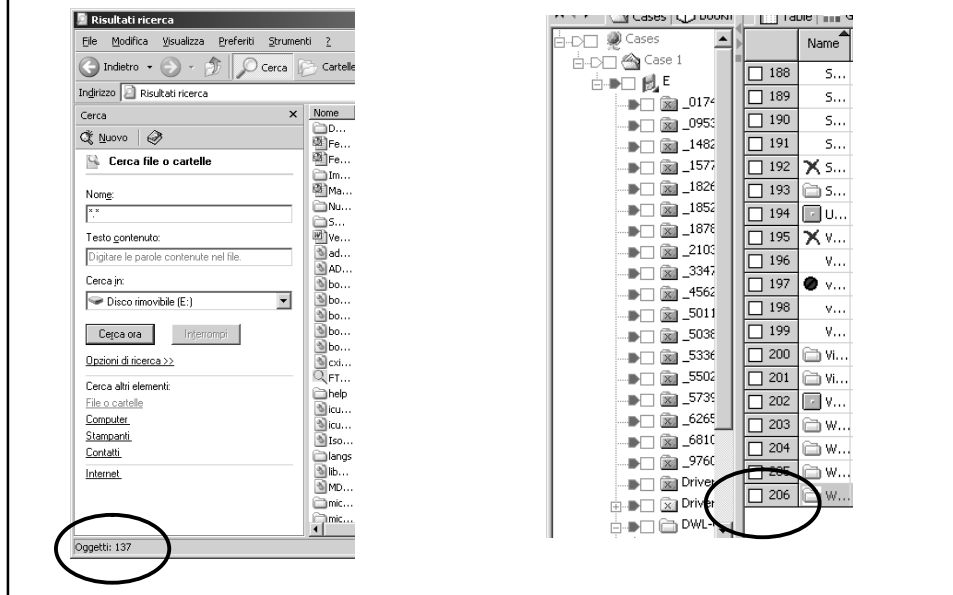
- Riconosce circa 400 formati di file noti, tra i quali
  - ZIP
  - Office
  - PDF
  - HTML
  - JPEG e altri formati di immagini
  - ...

## Carving – Photorec

---

- Esperienze positive con le seguenti fotocamere digitali
  - Canon EOS300D, 10D
  - Casio Exilim EX-Z 750
  - HP PhotoSmart 620, 850, 935
  - Nikon CoolPix 775, 950, 5700
  - Olympus C350N, C860L, Mju 400 Digital, Stylus 300
  - Sony Alpha DSLR, DSC-P9
  - Pentax K20D
  - Praktica DCZ-3.4

## Carving – Photorec – Esempio d'uso



## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP. Data Recovery Utility. April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

PhotoRec is free software, and
comes with ABSOLUTELY NO WARRANTY.

Select a media (use Arrow keys, then press Enter):
Disk /dev/sda - 250 GB / 232 GiB (RO) - SAMSUNG HD251HJ
>Disk /dev/sdb - 4110 MB / 3920 MiB (RO)

>[Proceed] [Quit]

Note:
Disk capacity must be correctly detected for a successful recovery.
If a disk listed above has incorrect size, check HD jumper settings, BIOS
detection, and install the latest OS patches and disk drivers.

```

## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

Disk /dev/sdb - 4110 MB / 3920 MiB <RO>

Partition      Start      End      Size in sectors
Unknown        0  0  1    499 186  ?    8028160 [Whole disk]
> P FAT32      0  0  1    499 186  ?    8028160 [NO NAME]

>[ Search ] [ Options ] [ File Opt ] [ Quit ]
Start file recovery_

```

## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

Paranoid : Yes <Brute force disabled>
Keep corrupted files : No
Expert mode : No
Low memory: No
>Quit

Return to main menu_

```

## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

PhotoRec will try to locate the following files

>[X] custom Own custom signatures
[X] lcd Russian Finance IC:Enterprise 8
[X] 7z 7zip archive file
[X] DB
[X] a Unix Archive/Debian package
[X] abcdp Apple binary property list
[X] accdb Access Data Base
[X] ace ACE archive
[X] ab MAC Address Book
[X] ahn Ahnenblatt
[X] aif Audio Interchange File Format
[X] all Cubase Song file: .all
[X] als Ableton Live Sets
[X] amd AlphaCAM (amd/amt/atd/att)
[X] amr Adaptive Multi-Rate
Next
Press s to disable all file families, b to save the settings
>[ Quit ] Return to main menu
  
```

## Carving – Photorec – Esempio d'uso

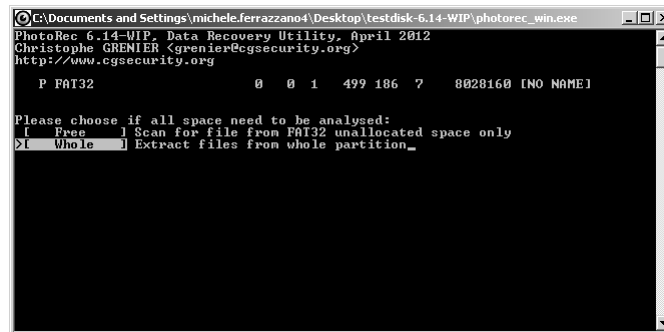
```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

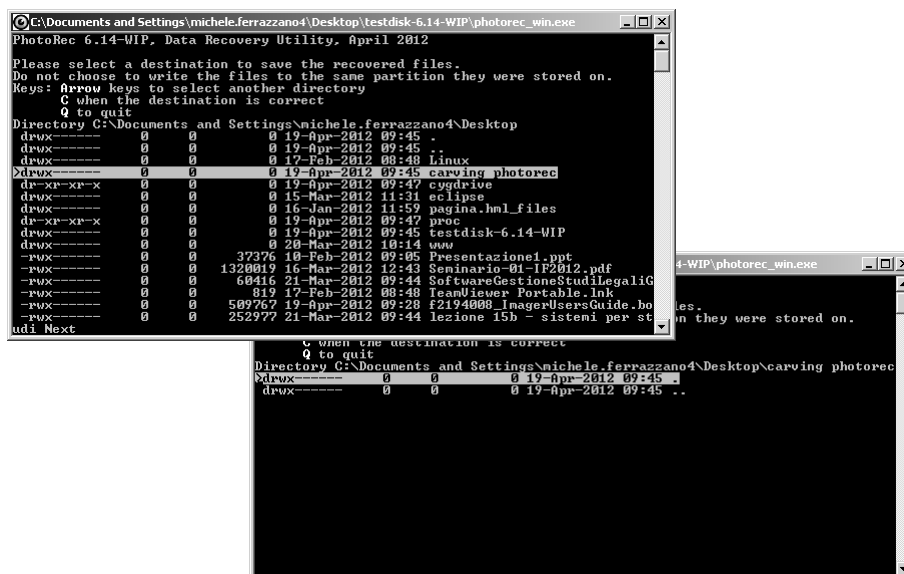
P FAT32 0 0 1 499 186 7 8028160 [NO NAME]
To recover lost files, PhotoRec need to know the filesystem type where the
file were stored:
[ ext2/ext3 ] ext2/ext3/ext4 filesystem
>[ Other ] FAT/NTFS/HFS+/ReiserFS/...
  
```



## Carving – Photorec – Esempio d'uso



## Carving – Photorec – Esempio d'uso



## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

Disk /dev/sdb - 4110 MB / 3920 MiB <RO>
Partition      Start      End      Size in sectors
P FAT32         0      1  499 186  7      8028160 [NO NAME]

Pass 1 - Reading sector 260160/8028160, 22 files found
Elapsed time 0h00m13s - Estimated time to completion 0h06m28
doc: 7 recovered
bmp: 5 recovered
zip: 4 recovered
txt: 3 recovered
mdb: 2 recovered
exe: 1 recovered

[ Stop ]

```

*recovering...*

## Carving – Photorec – Esempio d'uso

```

C:\Documents and Settings\michele.ferrazzano4\Desktop\testdisk-6.14-WIP\photorec_win.exe
PhotoRec 6.14-WIP, Data Recovery Utility, April 2012
Christophe GRENIER <grenier@cgsecurity.org>
http://www.cgsecurity.org

Disk /dev/sdb - 4110 MB / 3920 MiB <RO>
Partition      Start      End      Size in sectors
P FAT32         0      1  499 186  7      8028160 [NO NAME]

2639 files saved in /carving photorec/recup_dir directory.
Recovery completed.

You are welcome to donate to support further development and encouragement
http://www.cgsecurity.org/wiki/Donation

[ Quit ]

```

## Carving - tcpextract

---

- `tcpextract` è un tool open source che consente di estrarre file da un traffico di rete
  - <http://tcpextract.sourceforge.net>
  - `$ sudo apt-get install tcpextract`
  - Come foremost ma concentrato sul traffico di rete
    - Utilizza `libpcap`
  - 26 tipi di file noti
- Uso
  - `$ tcpextract -f sniffed -o ./output`

## Carving – Altri tool

---

- Chaosreader
  - <http://chaosreader.sourceforge.net/>
- msramdmp
  - <http://www.mcgrewwsecurity.com/tools/msramdmp/>