The Caine 2.0 forensic distribution

Raising Caine

Caine is a Linux distribution based on Ubuntu 10.04 for forensic scientists and security-conscious administrators. Poised to do battle against IT ne'er-do-wells, Caine has a comprehensive selection of software, a user-friendly GUI, and responsive support.

By Hans-Peter Merkel and Markus Feilner

In 2007, IT systems were compromised with the Russian hacking framework MPack [1], infecting masses of servers, mostly at data centers in Italy. A short time later, 10 Italian open source developers fought back and launched the forensics CD Caine (Computer-Aided Investigative Environment). The Live CD is designed to capture images and analyze compromised systems on site. The most recent version is 2.0 [2].

Internet Crime Scene

At any crime scene, IT forensics experts or administrators always have to back up the data first. Typically this means creating forensically valid images that are suitable as evidence in court.

To support this, Caine 2.0 has the three most important formats, RAW, EWF [3], and AFF [4] on-board. Both GUI and command-line options are available, and you can decide which way to go when you boot the CD (Figure 1). If you have ever worked on a server farm, you will appreciate the command line option. Server farms have thousands of machines, but only a few with peripherals such as mice, keyboards or displays. The command line will be more useful, typically across the network.

After booting in text mode, an Ubuntu-style `sudo su` on Caine is all it takes to access the hard disk in the compromised server and back up the disk. The root password for the distribution is `caine`.

RAID controllers, commonly used by ISPs, don't typically pose a problem; the drivers they contain assure that the RAID system is properly detected. Image creation will form a single image and saves you the task of painstakingly setting up a mirror in the lab at a later stage. The fairly recent kernel 2.6.32 will detect and support modern controllers without any trouble, and the distribution has all the tools you would associate with creating images, such as `dd`, `dc3dd`, `dcfld`, `aimage`, or `ewfacquire`, on-board.

If you prefer a GUI-based approach, you will want to take the first option on the menu and work in X.org with Guymager [5]. Guy Voncken, one of two well-known forensics geeks from Luxembourg, created the reference program for capturing images (Figure 2).

Block Devices

Caine automatically ensures that the system mounts all media read-only while you are working in the GUI, thus preventing any kind of write access. But at the command line, you will need to take care of this yourself. In any case, one attached block device will need write ac-

LISTING 1: ntfsundelete

```
01 ntfsundelete /dev/sda1 -p 100 | awk &apos;{print $1}&apos; | egrep "^[[:digit:]]" | while read inode; do ntfsundelete
                        -u -i${inode} -d /tmp/recovered/ /dev/sda1 ; done
```

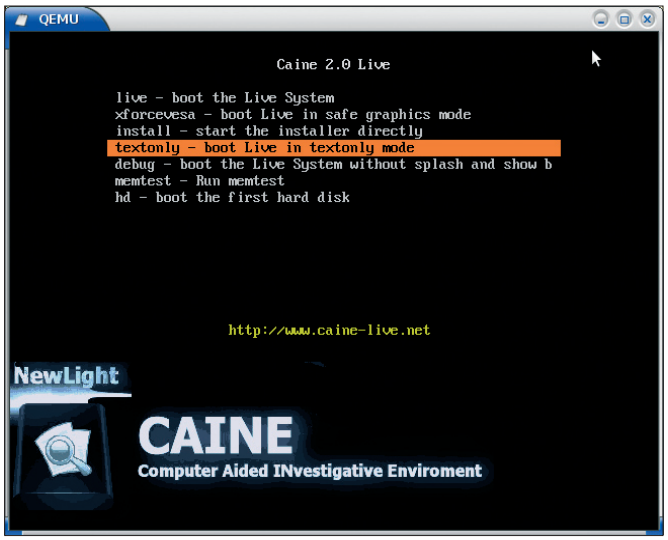

Figure 1: Choose between GUI and text-based modes at boot time.

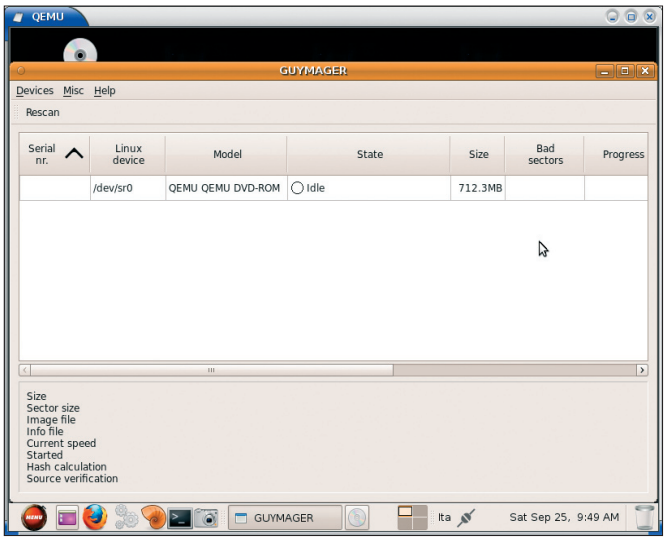

Figure 2: Guymager captures and analyzes system images on X.org.

cess: the forensic expert's external disk, where the image will be written. Investigators tend to prefer NTFS filesystems to support analysis with proprietary software on Windows later on. The NTFS-3g driver can help with this, but capturing the image to an NTFS filesystem has its drawbacks: The write speed is far slower than that of Ext filesystems, which is typically a problem with large volumes of data. Caine 2.0 can handle both.

In state or federal investigations, the confiscated PCs will now be locked away in evidence vaults while the images are transferred to analysis stations. These powerful multicore machines with lots of RAM normally have the forensics software pre-installed, and access to the hard disks is much faster.

Caine can power a forensics workstation, too; just click the desktop icon to install on disk. After completing the installation, you can continue working with standard forensics tools. The front line of attack here includes The Sleuth Kit [6], with GUI support from Autopsy [7]. File carvers such as Photorec, Foremost, or Scalpel are also represented.

Nearly Complete

The increased use of virtual systems makes programs that handle EWF or AFF images indispensible. Although the Live CD has `mount-ewf`, the more powerful Xmount [8] by Luxembourg forensics geek, Daniel Gillen, is not in place. After contacting the Italian developers, we were assured the next version (2.5) will integrate Xmount, although the release date is not yet known. Guymager will also see an update to version 0.5.7. Admins can then mount forensic images directly in EWF or AFF format and evaluate the results at file level.

Windows operating systems might also need some help in the form of OpenGates [9], which removes the well-known initial bluescreen. A new sibling is Gillen's OpenJobs, an ISO image that helps jump-start virtualized Macs [9].

The tools on the Caine CD also help with issues typically faced by administrators. For example, the Ntfsundelete tool restores deleted files. In Listing 1 is a cryptic, but efficient, approach that saved a colleague the pain of losing around 2,000 files after prematurely unplugging a USB disk from a computer.

The command reconstructs all the files in partition `/dev/sda1` with a recovery potential of 100 percent. The reanimated files are found in `/tmp/recovered` after completing the process.

Caine can also help you find files you think you have lost. The Sleuth Kit includes the `fls` command:

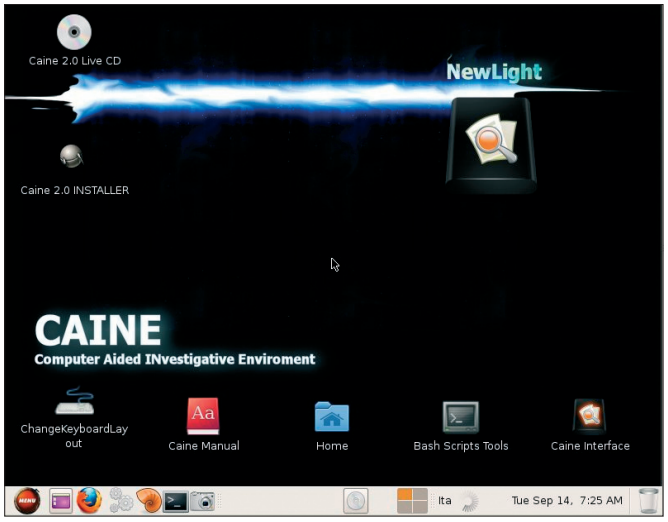

Figure 3: Hard disk installation, Bash Script Tools, and Caine Interface.

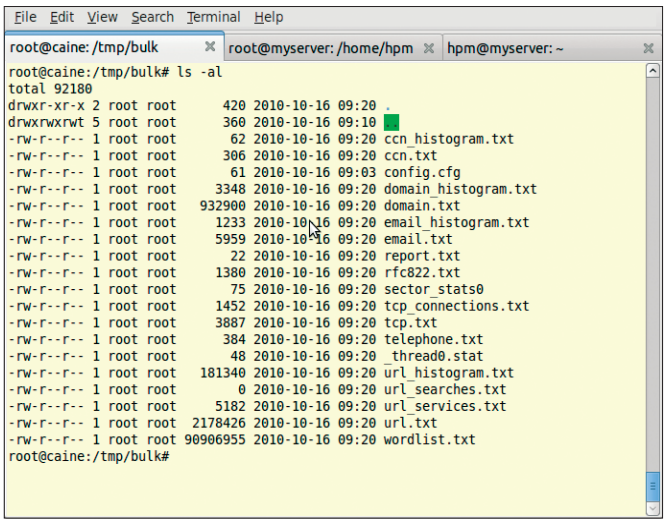

Figure 4: Captured from Windows hard disks by Bulk Extractor..

```
fls -r /dev/sda1 > /tmp/lost.log
```

For deleted files, use:

```
fls -rd /dev/sda1 > /tmp/deleted.log
```

Version 3.2.0, released October 2010, has new tools for automatic processing.

FUSE, ZFS, NFTS, sshfs

FUSE [10] is one of the most important tools in a forensic investigators toolkit, even if you don't work with it directly – which brings us back to NTFS-3g, which also relies on FUSE. The effect that Oracle's ZFS filesystem will have on the operating system world is uncertain. Kernel-based solutions are currently only available for FreeBSD, not Linux, because of the GPL.

The Caine CD already integrates a FUSE driver for the ZFS filesystem in userspace, which means you can analyze Solaris and FreeBSD machines with ZFS filesystems. The performance hit associated with this userspace-based solution can typically be ignored in forensics applications.

Unfortunately, another important FUSE driver, `sshfs`, that often plays an important role in creating logical backups is also missing; you can mount complete root servers on server farms with commands like

```
sshfs -p port IP address:/ /mnt
```

and explore up front.

Sshfs is a very practical driver that is always a useful option if time is of the essence. According to the Caine developers, sshfs will be on-board in the next version.

GUI

Even without sshfs, Caine has much to offer; especially with its GUI (Figure 3). The user-friendly Caine graphical interface makes life easy for newcomers; Caine's GUI tool resides as an icon directly on the desktop and supports detailed reporting.

Returning to the command line, you will find Bash Script Tools, a collection of Bash and Perl scripts for advanced forensics (in the `/usr/share/caine/pacchetti/scripts` directory.) Under the hood are even more treasures, such as the Bulk Extractor from Afflib, which helps official sources in Afghanistan filter suspicious traces from Windows systems (see the box titled "Caine in Afghanistan").

Conclusions

Caine 2.0 is a compact and useful forensic distribution. The support offered by the Caine developer team is also very promising. When our editorial office contacted the Caine developers about the possibility of adding missing programs, we received replies within 24 hours in all cases. The ability to install the Live CD or the USB stick version permanently on a hard disk, in particular, makes it easier for newcomers to gain entry quickly to the fascinating field of computer forensics. ■■■

CAINE IN AFGHANISTAN

Linux Magazine author Hans-Peter Merkel tested Caine's capabilities during open source training of an Afghani ministry in Kabul. Participants with minimal knowledge of Linux were asked to capture images and familiarize themselves with some initial forensic tasks.

The test objects were lean Acer Aspire netbooks with double keyboard layouts (Persian and US). The devices were provided by the Skateistan project [11]; Linux4Afrika [12] later integrated them with a terminal server-based solution for Afghani street kids. The Live CD initially failed because of the lack of an optical drive, but USB sticks got the session up and running.

Email address and IP addresses

The participants then used Ewfacquire to create EWF images of pre-installed (Persian) Windows XP on external hard disks. The participants' desire to extract all email and IP addresses on the disk at a single pass turned out to be a surprisingly difficult nut to crack. After an introduction to unallocated space, the RAM or file stack [13] of a disk, it quickly became clear that checking the file content is not enough, and even the use of regular expressions to discover email addresses turned out to be counterproductive.

Bulk Extractor

The Afflib project provides the Bulk Extractor for this task, and the tool is included with the Caine distribution. The command

```
bulk_extractor /dev/sda1 -o /tmp/bulk
```

creates a `bulk` directory in which all the email addresses, IP addresses, URLs, and even CCNs (potential credit card numbers) are listed in neat groups. Figure 4 shows the tool running on a newly installed Windows 7. The `email_histogramm.txt` file contains the following:

```
n=48 tj@.tjH.tj
n=18 yourname@example.com
n=16 SzX@Szh.Sz
n=11 Sz@.SzH.SzX.Sz
n=10 jemand@example.com
n=8 DefaultUser@DefaultDomain.De
n=8 anonymous@discussions.microsoft.com
n=6 CPS-requests@verisign.com
n=6 someone@microsoft.com
n=5 username@domain.com
n=4 4M7@T.UK
n=3 itfinc@libertynet.org
n=3 jemand@microsoft.com
n=1 gates@microsoft.com
```

The file lists the names and the number of email address instances found on the system. Like Facebook's more controversial features, this lets you find out more about the user's social networking habits.

INFO

[1] MPack: *http://www.itrportal.com/absolutenm/templates/article-security.aspx?articleid=4219&zoneid=18*

[2] Caine: *http://www.caine-live.net*

[3] libewf: *http://sourceforge.net/projects/libewf*

[4] AFF forensics wiki: *http://www.forensicswiki.org/wiki/AFF*

[5] Guymager: *http://guymager.sourceforge.net*

[6] Sleuthkit: *http://sleuthkit.org*

[7] Autopsy: *http://www.sleuthkit.org/autopsy*

[8] Xmount forensics wiki: *http://www.forensicswiki.org/wiki/Xmount*

[9] OpenGates and OpenJobs: *https://www.pinguin.lu/index.php*

[10] FUSE: *http://fuse.sourceforge.net*

[11] Skateistan: *http://www.skateistan.org*

[12] Linux4Afrika: *http://www.linux4afrika.de/vision.html?L=0*

[13] "Investigating Windows Systems" by Hans-Peter Merkel, *Linux Magazine*, August 2008, *http://www.linux-magazine.com/Issues/2008/93/WINDOW-KIT*