CS4677 Computer Forensics: Lab 5 – deleted file recovery and introduction to Autopsy

The purpose of this lab is to familiarize you with the basics of deleted file recovery. You will be given two images of file system partitions. Each contains a deleted file. Your task is to recover the deleted file from each image. **NOTE: All of the software that you will use in this lab is available on the ftp servers in the lab at /pub/cs4677/software.**

Part I: Data Recovery with Debugfs, Scalpel and Foremost:

1. Boot up and login to your Fedora VMware image, then su to become root and change to root's home directory
    a. su root
    b. cd

2. Create two subdirectories, lab5 and forensics then change into the lab5 subdirectory
    a. mkdir lab5
    b. mkdir forensics
    c. cd lab5

3. Connect to the ftp server at 192.168.100.6 and download the images for the lab along with the file carving utilities listed below.
    a. /pub/cs4677/lab5/image1
    b. /pub/cs4677/lab5/image2
    c. /pub/cs4677/software/foremost-1.5.7.tar.gz
    d. /pub/cs4677/software/scalpel-2.0.tar.gz

4. Use the file utility to identify the type of each image file. Save the output or take notes for reference later (see Question 1).

    ```
    file <file name>
    ```

5. Use debugfs to view super block information for each image. Save the output or take notes for reference later (see Question 2).

    ```
    debugfs -R show_super_stats <image name>
    ```

6. Using notes from class, recover the deleted files from one of the images with debugfs. Your knowledge of the file system type from 4 above should assist you in determining which image file is most suited to debugfs assisted recovery. Once you have decided which image you want to use, launch debugfs as follows:

    ```
    debugfs <image name>
    ```

7. After you have recovered the files, save the sha1 hash of each. You will turn these in.**

8. Scalpel and Foremost are file carving utilities. You will find the source code for each on the ftp server in /pub/cs4677/software. Scalpel was a fork of Foremost version 0.69 so there are some similarities between the two tools. As part of your ongoing introduction to open source tools, you are to extract, compile, and use both tools to attempt to recover the same deleted files that you recovered in step 6. You must use both tools. Make a note of the similarities, differences, and effectiveness of each (see Question 5).

a. Use tar and make to extract and build both tools. I strongly recommend reading through the ***entire*** README file included with each tool (though you can skim the bit about using GPUs to enhance performance in scalpel). If you run into trouble, please follow the standard troubleshooting instructions (see the appendix to this lab if you need to refresh your memory about the standard instructions).

b. After you've installed the tools and read their README files, familiarize yourself with their man pages. **Pay particular attention to the use of configuration files (foremost.conf and scalpel.conf) to help direct both scalpel and foremost**. You can read the man page while in each tool's subdirectory as follows:

```
man ./foremost.8.gz
man man/scalpel.1
```

9. Use both scalpel and foremost to carve files from **the same disk image** from which you recovered the deleted files in step 6. **Note: this is not extremely easy—it may take some trial and error to get it working well.** It is important to remember that when using a disk carving utility, both allocated and deleted files will be recovered, so you will need to perform additional analysis in order to determine which recovered files may be the deleted files. You may need to review the configuration files for both foremost and scalpel in order to get them to find what you are interested in. Foremost contains some command line parameters to direct its searches, while scalpel relies entirely on its configuration file. In particular, scalpel will only recover file types that have uncommented signature lines in scalpel.conf. You may add your own lines to these configuration files if you have some idea how to recognize a particular file format. For example, you might add gzip file support in the following way:

```
gzip y     10000000   \x1f\x8b\x08\x08
```

which states that a gzip file begins with the indicated four byte sequence and that you wish to grab 10Mb of data beginning with any sector that starts with these four bytes. Some file formats also have recognizable tail bytes, but this is not the case for gzip files.

10. Use dd to hash the first 1024 bytes of **each** file you recover with scalpel and foremost:

```
dd if=file bs=1024 count=1 | sha1sum > fore_imX_recov1.sha1
dd if=file bs=1024 count=1 | sha1sum > scal_imX_recov1.sha1
```

You will turn these in.**

11. You must now attempt to recover the same set of deleted files that you recovered in step 6 from the ***other*** image (the one that you did not use in step 6). Use what you have learned in steps 6 and 8. **Keep notes on ALL difficulties that you encounter and be prepared to discuss all possible reasons for the difficulties (see Question 9).** Note that while the file content of the two images are the same, and the deleted files are the same within each of the images, the sequence in which the files were created and/or deleted was not the same for each image.

12. Again, use dd to hash the first 1024 bytes of **each** file recovered in step 10.

```
dd if=file bs=1024 count=1 | sha1sum > fore_imY_recov1.sha1
dd if=file bs=1024 count=1 | sha1sum > scal_imY_recov1.sha1
```

You will turn these in.**

Part II: Image Analysis with Autopsy

1. Type “exit” to stop being root and change back to student and cd to your home directory. Create another lab5 directory here. cd into it, then reconnect to the ftp server at 192.168.100.6 and download the required items for this portion of the lab.
    a. /pub/cs4677/lab5/image3.gz
    b. /pub/cs4677/software/autopsy-2.24.tar.gz
    c. /pub/cs4677/software/sleuthkit-3.2.1.tar.gz

2. Use gunzip to extract the lab
    a. gunzip image3.gz
    b. Use the tools we've discussed in class to get some information about this image file. Is it a disk image or a partition (hint: does it have a partition table)? What type of file system (or systems) is it using? Make notes of your answers—you will need them to set up the image for analysis in step 6d and e.
    c. run md5sum on the image and record the output for use in step 6e.

3. Extract and install sleuthkit, becoming root when needed and typing exit to become a regular user again afterward.

4. Autopsy requires that you designate a directory as an “evidence locker”. Again, make sure you are NOT root and create /home/student/evidence, then configure autopsy using ./configure. When asked about the NSRL, answer “n”. When prompted for the Evidence locker directory, enter /home/student/evidence

5. Run autopsy. There are two ways to do this, depending one what you prefer. If you want to run everything inside your Fedora VM, just type:

    ```
    ./autopsy
    ```

    The output of this command will give you instructions for connecting to the autopsy web client through Firefox. You will connect to autopsy by launching Firefox within your Fedora virtual machine and pasting in the url provided.

    Alternatively, if you would like to connect to autopsy from another machine on the same network (such as your host machine/laptop), determine the machine's IP on the network and launch autopsy as follows:

    ```
    ./autopsy <your machine's ip>
    ```

    This sets up the autopsy server to allow connections from a host with the specified IP address, to a special randomized URL that the above command generates for you. Start a web browser on your machine and browse to the autopsy URL. It should look something, though not exactly, like:

    http://server:9999/17476184371196035533/autopsy

    (Substitute the IP of your Fedora VM in place of “server.”)

6. Once you've connected to Autopsy using either of the two methods described above, follow the steps below to set up your image for analysis:

    a. Click on “New Case”, and provide the requested information
        i. Case Name: lab5
        ii. Description: lab5

iii. Investigators: your name without spaces
iv. Click “New Case”

b. You should get a message saying that the case is being created and prompting you to create a “host.” This represents the machine you are investigating.
   i. Make sure your name is selected
   ii. Click “Add Host”

c. Fill out the fields on the page that follows:
   i. HostName: victim
   ii. Description: leave blank
   iii. Time Zone: PST
   iv. Time Skew: leave at 0 (for this lab—you would actually set this in a real investigation)
   v. Paths to Alert/Ignore Hash Databases: leave blank
   vi. Click “Add Host”, then “OK”

d. You should get a message saying the host has been added to the case. Click “Add Image” to move on to the next step (adding your image to this host).
   i. Click “Add Image File”
   ii. Location: /root/lab5/image3 (or wherever you put it)
   iii. Look back at your notes from step 2 to determine whether you should select “Disk” or “Partition”
   iv. Import method: symlink
   v. Click “Next”
   vi. If you get a warning, go back and double check your choices...

e. Add some final details about the image:
   i. Check “Add…” a hash value and “Verify…” and enter the md5 hash of image3 that you generated in step 2c.
   ii. Indicate a mount point: /1/
   iii. Verify the file system type.
   iv. Click “Add”

f. Read the output to see if the verification step worked. If so, click “OK.”

g. Click on Analyze to start examining the image. There are many, many analytical tools available. The steps below are meant to give you a tour of the major functionality.

h. Click on ‘File Analysis” to examine the image at the “File” level
   i. Notice that you basically have a file system navigator in the main browser frame
   ii. Navigate to etc/ and click on the passwd hyperlink
   iii. The contents of the file are displayed in the bottom pane. In the middle pane, you have several other links, one of which is Export.
   iv. Click on the Export link to extract the passwd file from the image. You will turn this in.**

i. Click on the “All Deleted Files” button (on the lefthand side). Any deleted files should be listed.
   i. Files that autopsy can recover easily will have complete information displayed. Files that autopsy knows something (but maybe not

everything) about will be displayed but may not show an inode number or any mac times.

ii. This display will, at a minimum, give you an idea of what files have been deleted from a system. Files displayed here may or may not be fully recoverable.

j. Click the Meta Data button at the top of the page
   i. In this view, you can view any inode on the image. Inodes are individually numbered (and note that inode numbers are not the same as block numbers).
   ii. Click the "Allocation list" button on the left side to browse the list of allocated and unallocated inodes.
   iii. You will return to this section to answer Question 1.

k. Click the Data Unit button at the top of the display
   i. In this view, you can see the contents of individual allocation blocks
   ii. You can examine a specific block by entering the block number in the left hand window or clicking the "Allocation List" button on the left to browse blocks in order.
   iii. Click the Allocation List button then click the link for block 328. You can view its contents in either hex or ascii.
   iv. Clicking the "Find Meta Data Address" link above the content window will cause autopsy to attempt to determine what file the block belongs to. This is a sort of reverse lookup service answering the question: given a block on the disk, what file does it belong to? **NOTE: You may need to lower the sliding divider line between the white content window and the tan info window to see the "Find Meta Data Address" link.**
   v. Click the "Find Meta Data Address" link above the content window and make a note of which file this block belongs to (see Question 2).
   vi. Browse to block 1. Click "Find Meta Data Address" and note that autopsy does not generate a path to a file this time, and instead gives you the rather useless output "Meta Data." Ponder this (and see Question 3).
   vii. If you want to see a sequential list of blocks, enter the number of fragments to view in the field on the lefthand side.
   viii. If you come across something interesting, you can export the contents using the Export Contents button above the content window. It's possible to reconstruct deleted files in this way.

l. Try a keyword search by clicking the Keyword search button
   i. Search for the text "GIF89a". It might help you out later on.

m. File Type is a useful capability especially if you are looking for (visual) images. It can dump all of the images it finds to disk.
   i. Select File Type and click the link to sort files by type.
   ii. Follow the instructions to recover all deleted visual image files (jpg/bmp/gif…etc.). You will turn these in.**
   iii. **NOTE: In the results section of the File Type search, you should have links to some indexed files, but this functionality is missing from the version we are using. Try exploring the case output directory to see what was found: /home/student/evidence/lab5/victim/output/sorter-graphics-vol1/images.html**

n. Explore some of the other features of autopsy on your own.

Phase I Questions:

1. (3 pts) What type of file system is each image formatted with? How did you identify this?

2. (5 pts) Is debugfs useful in determining the file system type? If not, is there any super block information that debugfs presents that might assist you in distinguishing between the two images?

3. (2 pts) Which image did you choose to recover deleted files from in step 6?

4. (5 pts) Discuss how, if at all, you were able to determine the exact size of each of the recovered files.

5. (5 pts) Compare and contrast the performance of scalpel and foremost. Is one or the other faster? Easier to configure? More powerful? More accurate?

6. (5 pts) Compare the hashes of the files recovered via debugfs with the hashes of the files recovered using foremost and scalpel. Discuss any differences in the hashes.

7. (2 pts) Can any of the recovered files be viewed/used as originally intended? In other words, is there any corruption in the recovered files?

8. (5 pts) Is it possible to determine the original file names of the deleted files? If so how? If not why not?

9. (8 pts) Discuss your efforts to recover files from the image used in step 11. Was it easier or more difficult to recover the files from this image? Compare and contrast the performance of debugfs, foremost, and scalpel on this image with their respective performance on the image used in step 6.

10. (3 pts) One of the deleted files contains a photo of a person. Who is this person?

11. (2 pts) Who took the photo discussed in Question 10, above?

Phase II Questions:

1. Use the autopsy Meta Data section tools to answer the following 2 questions:
   a) (2 pts) How many inodes have been allocated out of the first 499 inodes?
   b) (2 pts) What file does inode 18460 refer to?

2. (2 pts) Which file does block 328 belong to (refer to step 6k-v).

3. (5 pts) Even though block 1 is reported as allocated, autopsy is unable to tell you what file it belongs to, why?

4. (2 pts) What file does block 8508 belong to?

5. (6 pts) What is/are the inode number(s) of the file(s) you recovered in step 6m?

6. (2 pts) You may find pictures of an airplane in the visual images you recover. What type of airplane is depicted?

7. (4 pts) Of the images you recovered, which, if any, were deleted from the partition before imaging? Give the time of deletion for each. If none were deleted, say “none.”

8. (5 pts) Describe one feature of autopsy that you discovered in step 6n and a possible use-case for it.

9. Extra Credit (2 pts): Name a person in the CS department who used this image and provide forensic evidence to back up your assertion.

Phase I Deliverables:

1. (5 pts) The sha1 hashes you created in step 7 (for the deleted files that you recovered in step 6).

2. (5 pts) The hashes you created in step 10 (i.e. the sha1 hash for the first 1024 bytes of **each** file recovered by foremost and scalpel in step 9).

3. (5 pts) The hashes you created in step 12 (i.e. the sha1 hash for the first 1024 bytes of each file recovered by foremost and scalpel in step 11).

Phase II Deliverables:

1. (5 pts) The passwd file that you recovered in step 6h.

2. (5 pts) The image file(s) you recovered in step 6m.

Post your files and answers to the questions to Sakai NLT than midnight on the due date.

Appendix: Troubleshooting Instructions for Building and Installing Software

"If you have any problems compiling the tools, you may see me or send me an email that contains the error messages you are receiving **and describe what you think the problem is**. (As discussed in class, "it won't compile" is not sufficient; you need to read the output and think about what it is telling you.) This is a real world problem when making use of open source tools. There are many occasions when the tools will fail to build properly and you will be on your own to come up with a solution. Occasionally tools like Google may be helpful, other times they may not. Feel free to discuss what works and what doesn't in the discussion forums on Sakai."