Chad Ballay

CYBR 420-342N

10/11/2020

**Week 5**

**Explain the process for cloning a hard drive.**

<https://www.pcmag.com/how-to/how-to-clone-a-hard-drive>

<https://www.avast.com/c-how-to-clone-hard-drive>

<https://www.avg.com/en/signal/how-to-clone-a-hard-drive>

Multiple third party solutions and open source solutions exists for cloning hard drives. In each of those instances the specifics differ. At the macro level you get access to the drive. (Usually by physically pulling it.) Then you attach a mechanism to read all data on the drive in a non-destructive manner. Finally, you write that data to a new device. (Usually, a matching instance of the drive you are cloning.)

**Briefly explain in your own words the process for collecting digital evidence.**

Examples:

<https://nij.ojp.gov/topics/articles/new-approaches-digital-evidence-acquisition-and-analysis>

<http://www.forensicsciencesimplified.org/digital/how.html>

<https://www.sans.org/blog/best-practices-in-digital-evidence-collection/>

The industry best practices seem to coalesce into methodically peeling the layers back one at a time and doing this in a very methodical way. I would start with documenting the situation first as it is. (Photos, apparent hardware, etc, etc…) This may be an unneeded step since the digital/physical divide breaks the connection between the two realms for some aspects. The next step would be to stabilize the existing system. (Power, inputs of data, outputs of data, etc…) You aren’t looking to start taking control of these things, just ensuring that they are working. Next, you slowly begin to control these external connections. I’ve got no idea how to prioritize which to start with, but you get to the point where you control the “signal” that these connections can provide. Finally, you begin to capture what’s the most fragile of data first. What’s in volatile memory? What is in the temporary systems? Those come before things like BIOS firmware or OS version.

This is all theoretical on my part though. In practice, I would focus on doing nothing that risked permanently destroying data or wasting an opportunity to get more detail without being absolutely sure that it was the only way forward.

**Provide the MD5 and the SHA256 hash values for the following phrases, "The quick brown fox jumped over the lazy dog." and ""The quick fox jumped over the brown lazy dog." "**

Using these two sites:

<http://onlinemd5.com/>

<https://emn178.github.io/online-tools/sha256.html>

The quick brown fox jumped over the lazy dog.

MD5: 5C6FFBDD40D9556B73A21E63C3E0E904

SHA256: 68b1282b91de2c054c36629cb8dd447f12f096d3e3c587978dc2248444633483

"The quick fox jumped over the brown lazy dog."

MD5: 5904B55E0BA529371876B12755590CC7

SHA256: c8239512459e4bcef235f5b992c16231ebd49cfd43c1575423556f1a9e0ffec0

**Why is the value for each different?**

The nature of what hashes do. Hashes take a data set, apply a mathematical function, and generate a unique result that is smaller than the data set, consistently repeatable, and importantly unique. So every variation of the data set to be worked on should itself reduce to a single unique answer or hash result. The difference in the data set doesn’t proportionally generate a difference in the hash. Any difference whatsoever generates a difference in the hash result with no regard to degree of difference.

**How are hash values used in cyber investigations?**

Since the results of the hash calculation are repeatable and are unique to the inputing data set hash functions server immense value. Demonstrating that seized data has not been altered can be done by securely storing the hash result and comparing the later instance of the data’s hash result. Another usage is as a checksum. Send the data through one channel and the checksum through another. On the other end of the communications use the data received and the expected checksum output to confirm the sanctity of the data transmitted.