Nervous System: SafeBack in the Day

BY DAVID KALAT

With the aggressive pace of technological change and the onslaught of news regarding data breaches, cyber-attacks, and technological threats to privacy and security, it is easy to assume these are fundamentally new threats. The pace of technological change is slower than it feels, and many seemingly new categories of threats have been with us longer than we remember.

Nervous System *is a monthly series that approaches issues of data privacy and cyber security from the context of history—to look to the past for clues about how to interpret the present and prepare for the future.*

In the early days of computer forensics and eDiscovery, no standardized tools existed for preserving electronically stored information. This led to problems in validating the integrity of electronic evidence.

In the early days of the profession, computer examiners often conducted their analyses on live systems. However, the examiner's actions could tamper with or taint evidence. Electronic data is inherently volatile. Even a computer that appears to be idle is making calls to its operating system and writing system operations to disk that the user may not even know about. Without a methodology to prevent the examiner from altering evidence, intentionally or accidentally, computer investigators were subject to various challenges, and their findings were questioned.

Programmer Chuck Guzis tackled the problem in 1991. Guzis' company, Sydex, was an Oregon-based technology company specializing in software utilities to directly access floppy diskettes, tape drives, and hard-disk drives. Guzis realized that the technologies that Sydex used to manage direct connections to computer media also could be used to provide secure connections for forensic purposes.

The idea behind SafeBack, and various digital forensic tools that have followed, was to mount the source evidence drive in a read-only state, such that the copying utility would be unable to write any data onto it. The tool then would read and copy each bit from the source evidence onto a different destination drive.

Copying is different from cloning. A cloned copy of a computer's hard drive could be substituted in that computer, and the user would not notice the difference. Cloning is not, however, a forensic collection. Cloning utilities do not typically make a true bit-for-bit duplicate: a clone can function as a substitute for another drive when it has reproduced the *active* data on that drive, but a forensic collection also preserves inactive and unallocated data no longer used by the system.

All electronic data exists in physical form, somewhere. A computer hard drive, for example, stores binary information in the form of magnetic charges on microscopic cells. The computer's operating system can interface with the drive to read those magnetic charges into live memory, but the magnetic charges persist even when the computer is powered down. The physical state of the computer system's storage media contains a wealth of information of possible investigative significance, including deleted files and file fragments in unallocated space and slack space; file metadata within the file system; operating system artifacts regarding user activity; and more.

Additionally, SafeBack and other forensic utilities encapsulate the data that they copy into an evidence file. This file can be very large (possibly as large, bit for bit, as the source drive) and contains all of the individual files and components of the source. Such a file is called an image; the process of creating this file is called imaging the source drive. The evidence image file is itself protected within a read-only wrapper and requires specialized forensic tools to be examined.

IT professionals usually think of disk images as something that moves data *onto* a target drive. A corporate IT department may have stock disk images for different devices; when a new computer needs to be deployed, the exact arrangement of the approved operating system with relevant enterprise software and network settings can be copied efficiently to that device's drive, rather than going through the process of multiple downloads, installations, updates, and configuration. The image is a sort of quick-press mold that can imprint a given arrangement of data onto a fresh drive.

By contrast, digital forensic investigators usually think of disk images as something that moves data *off of* a target drive. The subject computer's internal memory is copied to a disk image on an external medium, such as a standalone drive.

SafeBack's ability to create "mirror-image" bitstream copies of disks was ideally suited to law enforcement applications. Its implementation of mathematical hashing to verify the integrity of the duplicate copy has since become an industry standard.

Hash functions and hash values are essential features of computer forensic investigations. A hash function is an operation that takes as its input some binary data of any size and maps it onto a value of fixed size—the hash value. The function is a one-way street: a given input will produce only one hash value result, but that hash value cannot be used to determine the input value.

When a forensic investigator collects electronic evidence, the hash value of that evidence can be used to verify that the copy is mathematically identical to the source—and that any subsequent copies made from it are also identical. For this reason, hash values have become a widely accepted method of authenticating electronic evidence, often referred to colloquially as a "digital fingerprint."

SafeBack brought these essential features together to ensure the reliability and validity of electronically preserved evidence. SafeBack provided access to the source evidence in a read-only state; duplicated the source evidence bit by bit, including deleted data and unallocated space; encapsulated that copy in a read-only format; and verified that the copied image was mathematically identical to the source evidence. Collectively, these features provided a standardized utility for investigators to preserve electronic evidence in a way that ensured and documented its integrity.

SafeBack was originally designed for law enforcement use. It was quickly accepted by the US military, intelligence agencies, and thousands of law enforcement agencies worldwide. SafeBack also was made available commercially to civil and private investigators and was arguably the first commercial digital forensic tool.

Legaltech® news

This article was originally published in Legaltech News on July 6, 2021.