Data recovery

is the process of restoring data that has been lost, accidentally deleted, corrupted or made inaccessible.

In enterprise IT, data recovery typically refers to the restoration of data to a desktop, laptop, server or external storage system from a backup.

Causes of data loss

Most data loss is caused by human error, rather than malicious attacks, according to U.K. statistics released in 2016. In fact, human error accounted for almost two-thirds of the incidents reported to the U.K. Information Commissioner's Office. The most common type of breach occurred when someone sent data to the wrong person.

Other common causes of data loss include power outages, natural disasters, equipment failures or malfunctions, accidental deletion of data, unintentionally formatting a hard drive, damaged hard drive read/write heads, software crashes, logical errors, firmware corruption, continued use of a computer after signs of failure, physical damage to hard drives, laptop theft, and spilling coffee or water on a computer.

How data recovery works

The data recovery process varies, depending on the circumstances of the data loss, the data recovery software used to create the backup and the backup target media. For example, many desktop and laptop backup software platforms allow users to restore lost files themselves, while restoration of a corrupted database from a tape backup is a more complicated process that requires IT intervention. Data recovery services can also be used to retrieve files that were not backed up and accidentally deleted from a computer's file system, but still remain on the hard disk in fragments.

Data recovery is possible because a file and the information about that file are stored in different places. For example, the Windows operating system uses a file allocation table to track which

files are on the hard drive and where they are stored. The allocation table is like a book's table of contents, while the actual files on the hard drive are like the pages in the book.

When data needs to be recovered, it's usually only the file allocation table that's not working properly. The actual file to be recovered may still be on the hard drive in flawless condition. If the file still exists -- and it is not damaged or encrypted -- it can be recovered. If the file is damaged, missing or encrypted, there are other ways of recovering it. If the file is physically damaged, it can still be reconstructed. Many applications, such as Microsoft Office, put uniform headers at the beginning of files to designate that they belong to that application. Some utilities can be used to reconstruct the file headers manually, so at least some of the file can be recovered.

Most data recovery processes combine technologies, so organizations aren't solely recovering data by tape. Recovering core applications and data from tape takes time, and you may need to access your data immediately after a disaster. There are also risks involved with transporting tapes.

In addition, not all production data at a remote location may be needed to resume operations. Therefore, it's wise to identify what can be left behind and what data must be recovered.

Data recovery techniques

Instant recovery, also known as recovery in place, tries to eliminate the recovery window by redirecting user workloads to the backup server. A snapshot is created so the backup remains in a pristine state and all users write operations are redirected to that snapshot; users then work off the backup virtual machine (VM) and the recovery process begins in the background. Users have no idea the recovery is taking place, and once the recovery is complete; the user workload is redirected back to the original VM.

One way to avoid the time-consuming and costly process of data recovery is to prevent the data loss from ever taking place. Data loss prevention (DLP) products help companies identify and stop data leaks, and come in two versions: stand-alone and integrated.

- Stand-alone DLP products can reside on specialized appliances or be sold as software.
- Integrated DLP products are usually found on perimeter security gateways and are useful for detecting sensitive data at rest and in motion.

Unlike stand-alone data loss prevention products, integrated DLP products usually do not share the same management consoles, policy management engines and data storage.

Integrating data recovery into a DR plan

An organization's disaster recovery plan should identify the people in the organization responsible for recovering data, provide a strategy for how data will be recovered, and document acceptable recovery point and recovery time objectives. It should also include the steps to take in recovering data.

For example, if a building is inoperable, affected business units must be advised to prepare to relocate to an alternate location. If hardware systems have been damaged or destroyed, processes must be activated to recover damaged hardware. Processes to recover damaged software should also be part of the DR plan.

Some resources worth reviewing are the National Institute for Standards and Technology SP 800-34 standard, as well as ISO 24762 and 27031 standards.

A business impact analysis can help an organization understand its data requirements and identify the minimum amount of time needed to recover data to its previous state. One challenge to data loss and data recovery is getting a handle on the unstructured data stored on various devices.

But there are steps that can mitigate the damage. Start by classifying data based on its sensitivity and determine which classifications must be secured. Then, determine how much data would have to be compromised to affect the organization. Undertake a risk assessment to determine what controls are needed to protect sensitive data. Finally, put systems in place to store and protect that content.