

Data Exfiltration: How Data Gets Out

Most attention goes to keeping hackers out. But once they're inside, how do they extract data from your organization? Research from Trustwave's SpiderLabs shows the answer is often surprisingly simple.

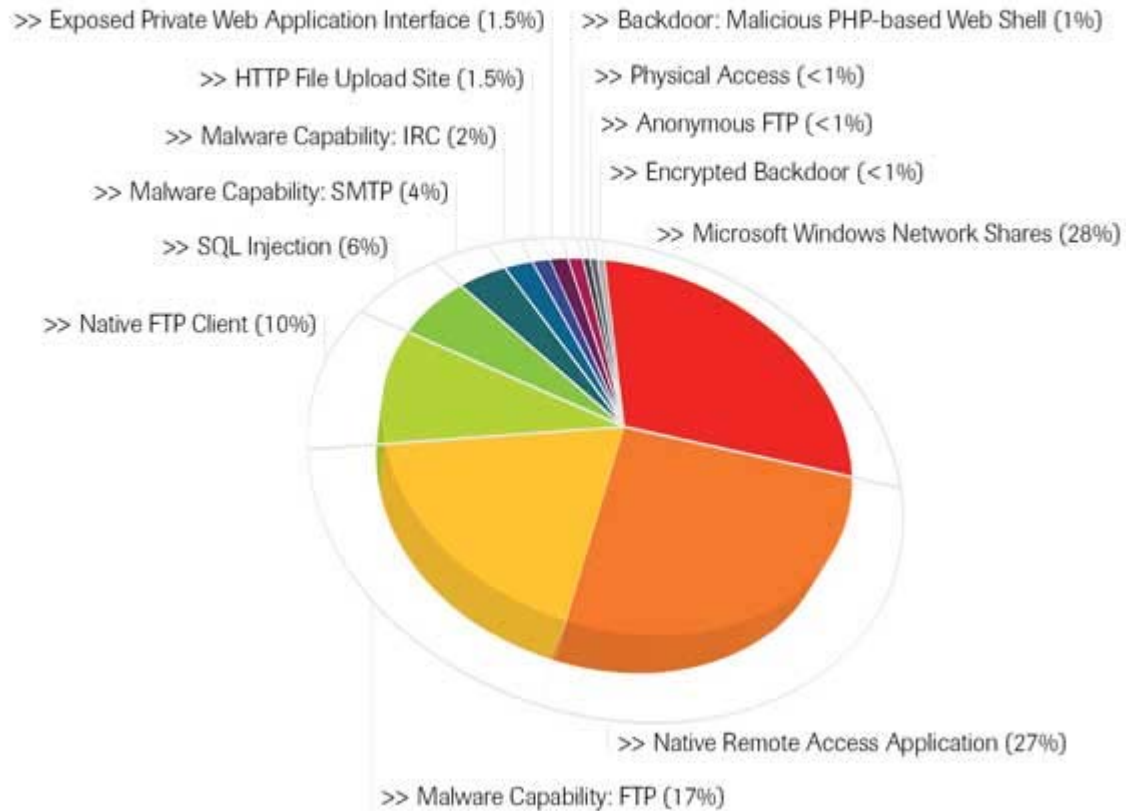
Nicholas J. Percoco, SpiderLabs, CSO
March 12, 2010

Cyber criminals are increasingly becoming more sophisticated in their methods of attack. Often we can equate this to the methods of data exfiltration as well. Exfiltration, or exportation, of data is usually accomplished by copying the data from the system via a network channel, although removable media or physical theft can also be utilized.

In 2009, the SpiderLabs team at Trustwave investigated over 200 data breaches in 24 different countries. While the methods used by cyber criminals to exfiltrate data from a compromised environment varied, the method of entry into an environment was often via the remote access application being utilized by the target organization. In the SpiderLabs investigations, 45 percent of compromises occurred by attackers gaining access to a system through a remote access application. These were not zero-day exploits or complex application flaws, and the attacks looked no different to the IT staff than, for example, the CEO connecting from London while on a business trip. The attackers also didn't need to brute-force the accounts they used. SpiderLabs found that 90% of these attacks were successful because of vendor-default or easily guessed passwords, like "temp:temp" or "admin:nimda."

Once a foothold is established, attackers often launch network enumeration tools. Network enumeration tools are often used by the attacker to discover additional targets within the environment and retrieve system information, such as usernames, group privileges, network shares, and available services. The noise generated by enumeration tools can indicate a prelude to an attack. Unfortunately, we've found that most entities are not properly monitoring their systems and therefore fail to observe these indicators.

Percentage of Methods Used to Exfiltrate Data



It was these types of tools that led attackers to the systems of additional hotel properties through trusted private circuits. The internal connections were subsequently exploited, resulting in the breach of data from physically dispersed sites. Without the existence of these connections, breaches within the hospitality sector would likely have been contained to only a few properties.

Once attackers gained access to the target environment, they harvested data using either manual or automated methods. Using manual processes, potentially valuable databases and documents were located, and searches of the operating system were conducted using specific keywords to further identify data.

The automated method was custom written malware that took advantage of a flaw found in the security controls of the applications being used to process confidential data. Generally, many application security designs do not apply more controls and alerting capabilities over components that process data in the clear. A target system that receives data encrypted and stores data encrypted but transmits data to an upstream host is susceptible to a data breach while the data is processed by the target system. This occurs because data processed by a system must be decrypted in RAM for the application to understand it. During this process, cyber criminals in 2009 frequently employed RAM parsers—67 percent of SpiderLabs' investigations involving malware concluded that automated tools were used to harvest data out of RAM while the system was using the data in some capacity.

The average time the cyber criminals were able to access the target systems and data was 156 days. For that period of time, attackers entered the environment, set up their tools to remove data and also harvested the data before a single IT or security department reacted to their activities. Some 2009 investigations showed recurring activity from the same cyber criminals over the course of three years. Long times to detection were typical in 2009 and, seemingly armed with this knowledge, cyber criminals are not practicing stealth in their activities.

In 38 cases, cyber criminals used the remote access application previously utilized for initial entry to extract data. Other existing services, such as native FTP and HTTP client functionality, were also frequently leveraged for data extraction. Specifically, when malware was utilized for data extraction, FTP, SMTP and IRC functionality were regularly observed. (In reverse analysis of custom malware, binaries would disclose the existence of FTP functionality including hardcoded IP addresses and credentials.) With off-the-shelf malware, such as keystroke loggers, attackers most often used built-in FTP and e-mail capabilities to exfiltrate data. When e-mail services were employed for extraction, the attackers often opted to install a malicious SMTP server directly on the compromised system to ensure the data was properly routed.

Only a single case contained the use of an encrypted channel for data extraction, further suggesting that criminals are rarely concerned with raising alarm. Due to natively available network services, lack of proper egress filtering and poor system monitoring practices, criminals are using available network services or choosing to install their own basic services.

It is clear that in all of these cases sensitive data was sent out of a target environment. During this time, the IT security teams did not detect the loss.

When looking for signs of an attack, IT Security teams seem to expect something complex. However, attacks are usually very simple, not "noisy" and likely appear benign in a routine log review. It is not until the data has left the target environment and shows up in some other capacity is the breach detected. Paying close attention to the behaviors of "normal" activity against "standard" systems is the key to identifying a problem before it is too late. Every anomaly should be viewed with a degree of suspicion and addressed through internal investigation or, if necessary, review by an outside expert.