Testing Controls Associated With Data Transfers

Tommie W. Singleton, Ph.D., CISA, CGEIT, CITP, CPA, is an associate professor of information systems (IS) at Columbus State University (Columbus, Georgia, USA). Prior to obtaining his doctorate in accountancy from the University of Mississippi (USA) in 1995, Singleton was president of a small, value-added dealer of accounting using microcomputers. Singleton is also a scholar-in-residence for IT audit and forensic accounting at Carr Riggs & Ingram, a large regional public accounting firm in the southeastern US. In 1999, the Alabama Society of CPAs awarded Singleton the 1998–1999 Innovative User of Technology Award. His articles on fraud, IT/IS, IT auditing and IT governance have appeared in numerous publications.

There are several IT-related functions that inherently have a rather high degree of risk. Some of those include custom application development, logical access (especially where the Internet is involved) and data transfers. The latter has been growing in volume and risk recently.

In the last few years, there are more and more occasions in which entities find it necessary or beneficial to transfer data from one repository to another, or to a tool for analysis. There are a number of reasons why this has become fairly common. Business analytics (i.e., business intelligence) generally requires data to be transferred from online transaction processing (OLTP) systems to the analytic tool (e.g., data mining, data warehouse, even a spreadsheet). Then there are a number of occasions in which data are transferred internally from different accounting systems to a downstream general ledger or financial reporting system. Sometimes that system is also a spreadsheet. With systems such as electronic data interchange (EDI), there is a need to transmit/transfer data to an external entity (e.g., vendor). The same might be true for government agencies and other parties. Add to that the increased use of data transfers in the banking industry (e.g., ACH, wire transfers, electronic funds transfer [EFT]) and credit card payments for online transactions.

According to experts, most companies do not have a centralized methodology for tracking and managing data transfers, which puts them at risk for both data security/error problems and the lack of documentation and audit trails for relevant government regulations.

Gartner describes the risk this way: "FTP [File Transfer Protocol] alone is not a viable option to give organizations the insight, security, performance, and ultimately, the risk mitigation necessary to responsibly conduct business."[1]

The most popular term for this IT issue is managed file transfer (MFT). TechTarget defines MFT as "a type of software used to provide secure internal, external, and *ad hoc* data transfers through a network."[2]

This article addresses some of the IT audit issues associated with data transfers.

DATA TRANSFER TYPES

Generally, data transfers can be categorized into three types: system-to-system, partner-to-partner and person-to-person.[3]

System-to-system is a transfer of data between two systems. That transfer could be internal and involve computers of the entity, or it could be between the entity and some external party. The previous example of transferring data into a downstream financial reporting system is fairly common these days. For instance, if the entity is using SAP, Oracle or even Microsoft Dynamics, and has chosen to download the data into a spreadsheet for year-end or fiscal-year journal entries, generation of a trial balance, and possibly other financial reporting process functions, then that is a system-to-system transfer, even though one of the systems is a spreadsheet. System-to-system transfers can be systematic/regular, occasional or *ad hoc*. For instance, transfers from OLTP to a data warehouse system tend to be regular, whereas financial reporting transfers tend to be occasional.

A partner-to-partner transfer was described in the introduction regarding EDI. In this case, two partners are continuously transferring accounting data back and forth across agreed-upon systems. These transfers are generally done on a regular basis.

A person-to-person transfer is the type that probably goes unnoticed and unmanaged most often. It could be as simple as attaching a data file to an e-mail and sending from one person in the entity to another via the corporate e-mail system. Person-to-person transfers tend to be *ad hoc* and, thus, are more difficult to observe, manage, secure and control than the other types of data transfers.

An added complexity is the fact that most *ad hoc* transfers, be they system-to-system or person-to-person, tend to need a user interface that is simple to use to minimize transfer risks.

Suffice it to say that transfers using e-mail as the medium are a risky (i.e., horrible) option. But the idea that restricting the size of e-mail attachments will somehow reduce this risk is not sound logic. Users will either "zip" the data or use an online service (e.g., drop.io) to accomplish the same purpose, and work around that imposed restriction.

IMPLICATIONS FOR DATA SECURITY

Security of data being transferred is a critical component of the risk associated with data transfers. The primary objective here is to ensure that the data intended to be extracted from the originating system are *exactly* the same data as that recorded/downloaded in the recipient system, i.e., that the data were protected and secured throughout the transfer process. The secondary objective is to prevent unauthorized access to the data via interception, malicious activities and other means.

Verification

Controls need to be in place to ensure that the data received are exactly the same data set as what was sent, or the same data set as that residing on the sending system. There are a multitude of risks associated with data transfers. First, the encoder could have a technical glitch/error and improperly gather data from the sending system at the initiation point. Second, the transmitting channel could adversely affect the data and cause errors (e.g., transmission across telephone lines when an electrical storm is taking place, scrambling data). The receiving technology could improperly download the data. And most of all, there is the element of human error throughout the transfer activities.

Therefore, controls needs to be implemented with the objective of ensuring that the data residing on the sending system are precisely the same data that are recorded on the receiving system. For example, the entity could utilize a custom middleware package, or a commercial one, that has the ability to automatically generate control totals during the extraction that can be automatically used to reconcile the data once they are recorded at the receiving system. The full duplex[4] approach to transmission has this goal. An automated control has obvious advantages over the manual types.

Another control might be a manual reconciliation. For example, a report (such as an SQL query) could be run on the origination system of the data to be gathered and sent, and then compared to the same report run against the data recorded in the receiving system. This report mimics the old batch transmittal control sheet process used successfully for decades on legacy systems.

Another manual control might be a review of the data by a qualified person who has the ability to detect material differences in the data.

Secured Communications

One control to protect data during the actual communication is encryption. This control should prevent any unauthorized party from intercepting the data and using them for nefarious purposes (e.g., industrial espionage, identity theft, credit card data theft). Encryption is necessary when the risks are relatively high regarding unauthorized access or interception, but might be unnecessary for certain types of internal transfers. It also might be essential for transfers that occur via e-mail, where the attached file is encrypted.

Secured communications might include strong access and authentication controls. For example, data files might be password-protected during the transfer. Firewalls and transport systems are also concerns. For transmissions over the Internet, malware is a concern.

Auditability

Auditability is the ability to provide a cyber audit trail associated with a data transfer. It would capture important information, such as who sent the data, when they were sent, when they were received, what data structure (e.g., xls, csv, txt, xml) was used, how the data were sent (i.e., via what medium) and who received the data. For example, e-mail and header metadata bear this information.

That auditability could be associated with automated logs of servers along the path, if the transfer occurs over the entity's network, especially if it is transmitted to an external system or if it uses the Internet (e.g., an intranet transfer).

For internal transfers, the auditability factors might be less extensive due to restricted communications, but external transfers might touch a number of Internet hosts and be exposed to hackers and cybercriminals. The external type would require a more robust auditability to ensure proper information on how the data were handled and by what systems or persons.

In some cases (e.g., EDI), nonrepudiation is needed. Nonrepudiation refers to the recipient being unable to repudiate receipt of the data or, put another way, making sure that the intended recipient is the actual recipient of the data. Strong auditability provides some control over nonrepudiation.

Managed System/Methodology

Entities need to be able to track and manage *all* data transfers, including those internal and external to the regular information systems. The advantage of this is obvious: the entity can see all of the transfers that are being made, including the *ad hoc* ones.

Whether it is done by some commercial MFT software or by a custom proprietary system, an approach should be in place to centralize and manage all data transfers. The goal here is to ensure that all essential data transfers go through this system. There are server-based and Software as a Service (SaaS)-based commercial applications available.[5] In addition to managing all data transfers, the commercial applications generally allow for transfers to be scheduled (routinely at specified times), which reduces the transfer risk considerably.

This system should be able to provide the following features:[6]

- The ability to manage multiple file transfer mechanisms
- The ability to use multiple protocols (e.g., FTP, SFTP, FTP/S, SCP, HTTP, HTTPS, AS2, SMTP/POP3)
- The ability to automatically encrypt, decrypt and electronically sign data files
- The ability to compress/decompress data files (e.g., ZIP, TAR)
- The ability to connect to common database servers (e.g., DB2, SQL Server, Oracle, Informix, MySQL, Sybase)
- The ability to send and retrieve files via e-mail and secure e-mail (S/MIME encryption)
- The ability to schedule (automate) regular data transfers
- The ability to analyze, track and report any attributes of the data being transferred (refer to auditability mentioned previously)
- The ability to ensure compliance with regulatory mandates (e.g., the Payment Card Industry Data Security Standard, the US Health Insurance Portability and Accountability Act, the US Sarbanes-Oxley Act, the US Gramm-Leach-Bliley Act)
- The ability to support the automation of data transfers
- A checkpoint or restart capability for interruptions
- Integration with back-office applications to automate data transfers as much as feasible

CONCLUSION

The growth in frequency of data transfers makes this issue ubiquitous in today's business environment. Because the inherent risk is so high, there is a great need for controls to mitigate that risk. IT auditors need to understand the nature of the risk and some mitigating controls (or control objectives) that can be tested to provide assurance that the control risk is sufficiently low enough to reduce the inherent risk to a tolerable level. These controls were illustrated with verification, secured communications, auditability and a *managed* system.

ENDNOTES

[1] Kenney, L. Frank; James Lennard; "Magic Quadrant for Managed File Transfer," Gartner, 2008

[2] TechTarget, *www.techtarget.com*

[3] *Op cit*, Kenney

[4] Full duplex is a methodology that echoes the data received back to the sending computer, which then checks those data against the original to make sure all of the data received at the other end match the data that were sent, ensuring that no bits were distorted during transmission.

[5] For example, *www.easylink.com*

[6] Partially taken from GoAnyWhereMFT, a supplier of MFT, *www.goanywheremft.com/products/director*