Realtime publishers

The Shortcut Guide™ To

Eliminating Insecure and Unreliable File Transfer Methods

2012 Edition

sponsored by

Attachmate

Dan Sullivan

Introduction to Realtime Publishers

by Don Jones, Series Editor

For several years now, Realtime has produced dozens and dozens of high-quality books that just happen to be delivered in electronic format—at no cost to you, the reader. We've made this unique publishing model work through the generous support and cooperation of our sponsors, who agree to bear each book's production expenses for the benefit of our readers.

Although we've always offered our publications to you for free, don't think for a moment that quality is anything less than our top priority. My job is to make sure that our books are as good as—and in most cases better than—any printed book that would cost you $40 or more. Our electronic publishing model offers several advantages over printed books: You receive chapters literally as fast as our authors produce them (hence the "realtime" aspect of our model), and we can update chapters to reflect the latest changes in technology.

I want to point out that our books are by no means paid advertisements or white papers. We're an independent publishing company, and an important aspect of my job is to make sure that our authors are free to voice their expertise and opinions without reservation or restriction. We maintain complete editorial control of our publications, and I'm proud that we've produced so many quality books over the past years.

I want to extend an invitation to visit us at http://nexus.realtimepublishers.com, especially if you've received this publication from a friend or colleague. We have a wide variety of additional books on a range of topics, and you're sure to find something that's of interest to you—and it won't cost you a thing. We hope you'll continue to come to Realtime for your educational needs far into the future.

Until then, enjoy.

Don Jones

Copyright Statement

Chapter 2: Analyzing 5 Key Drivers for Improving File Transfer

IT professionals are constantly faced with a wide array of new technologies promising greater efficiencies, higher performance, and faster development times. Why should they concern themselves with file transfer, which is, after all, a solved problem? The truth is file transfer in many organizations is a partially solved problem. Ad hoc solutions, which are programs designed to meet a single file transfer requirement or a small number of requirements, typically do not fully address the range of business requirements. Certainly, ad hoc solutions can copy files from Point A to Point B but that is not enough.

File transfers have to meet certain levels of reliability, performance, security, cost effectiveness, and accessibility. When ad hoc solutions come up short on these requirements, businesses will likely find they have increased risk of security lapses, incurred hidden costs of debugging and maintenance, and been forced to work around performance problems. These unmet requirements underlie several key drivers to adopting a more managed file transfer solution, including the need for

- Compliance
- Flexibility and scalability
- Management
- Cost controls
- Workflow efficiency

This chapter will discuss how file transfer solutions affect each of these drivers and highlight ways managed file transfer solutions can help meet these requirements.

Compliance

Businesses are expected to conduct operations in ways that mitigate risk to the enterprise. Government and industry regulations are codified expectations of what kinds of information is to be protected and minimum standards for protecting that information. After large-scale corporate accounting scandals and well-publicized data breaches, it is not surprising that regulations have been established to protect the integrity and confidentiality of corporate and personal information.

Some regulations are broadly applicable and others are industry specific. Regulations often have common objectives with regard to data integrity and confidentiality and are therefore applicable to file transfer practices. We will consider five diverse regulations in order to form a sufficiently comprehensive picture of the types of requirements these regulations establish with regards to file transfers. The regulations we will consider are:

- Sarbanes-Oxley (SOX) Act
- Payment Card Industry Data Security Standard (PCI DSS)
- Health Insurance Portability and Accountability Act (HIPAA)
- Gramm-Leach-Bliley Act (GLBA)
- 21 CFR Part 11

These span the broadly applicable, such as SOX to the transaction specific (PCI DSS) to the industry specific, such as HIPAA, GLBA, and 21 CFR Part 11. In all these cases, file transfers can be part of data flows subject to some degree of regulation.

SOX

SOX is a set of regulations governing the financial reporting of publicly traded companies. The regulations were established in the wake of several accounting scandals that were deemed so potentially damaging to trust in the market that a federal law was passed to establish standards for protecting the integrity of financial reporting. At first glance, it might appear that SOX would only apply to file transfers involving financial reporting information, such as transferring a general ledger. This is not the case, as we shall see in a hypothetical scenario; but first, let's review some of the key requirements of SOX.

IT-Related SOX Requirements

Perhaps the best known, and most challenging, part of SOX that relates to IT responsibilities is Section 404. That part of the legislation describes requirements for management and IT to ensure proper internal controls with respect to financial reporting. The specific requirements include a number relevant to file transfers that may be involved in financial reporting:

- Assessing the adequacy of internal controls to mitigate the risk of misstatements in financial reports
- Identifying points in workflows where tampering could occur
- Understanding controls to prevent tampering and detect it if it occurs
- Assess controls on the financial report generation process

There are more requirements of information technology professionals but these are enough to demonstrate how poorly secured and insufficiently monitored file transfer procedures can undermine SOX compliance.

SOX and File Transfers

Given the requirements on information technology processes, it is clear that transferring files containing financial reports are subject to SOX. These are not the only times in which SOX is relevant, though. The financial reports themselves are based on data from potentially many sources. Any tampering with data in these source systems or when data is transferred from one system to another can result in compromised results. Just as a polluted tributary can contaminate other rivers, a compromised data source can undermine the integrity of other data management systems that use it as a source system.

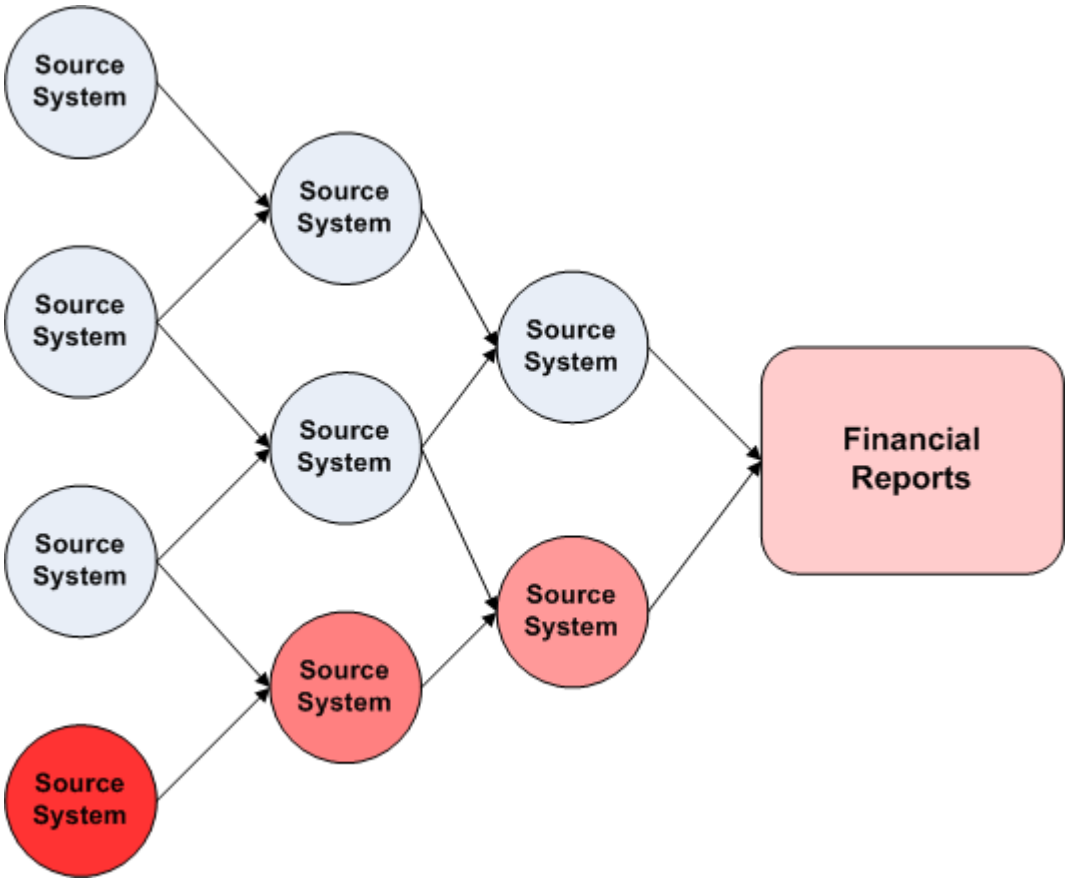

Figure 2.1: Financial reports are generated from source systems that are dependent on other systems through a chain of dependency. Lack of integrity at any point can compromise the integrity of the final product.

Information Flows and Application-based Controls

Consider, for example, the flow of information that goes into generating a financial statement. Of course, there are the basic sections of a financial statement, such as the balance sheet, cash flow statement, and income statement, but where does the raw data come from? For example, the value of assets on hand could include the value of inventory that is stored in a business partner's warehouse. Let's consider how a seemingly simple file transfer process between business partners could undermine the integrity of financial reporting.

In this scenario, a business has contracted with a business partner in Asia to store inventory close to the manufacturer, which is also in Asia. Orders to Asian customers ship directly from those warehouses, while orders destined for Europe and North America are shipped first to warehouses on the respective continents. The European and North American warehouses share inventory levels with the warehouse managers who maintain sufficient levels of inventory to meet local demand.

The business partner uses its own enterprise resource planning (ERP) system to manage its warehouses but it reports to the owners of the inventory every week on current inventory levels in the Asian warehouse, details about what was shipped directly to Asian customers, and products staged at European and North American warehouses.

From a financial reporting perspective, one company owns this entire inventory and must report it regardless of the fact that it is distributed around the globe. Let's assume that each of the business' ERPs is well designed, properly configured, and meet best practices for complying with accounting standards. A CIO could be reasonably confident that these systems meet SOX requirements and would be willing to sign off on the compliance report. What about the file transfer process that is used to move inventory data from the warehouse managers to the inventory owner?

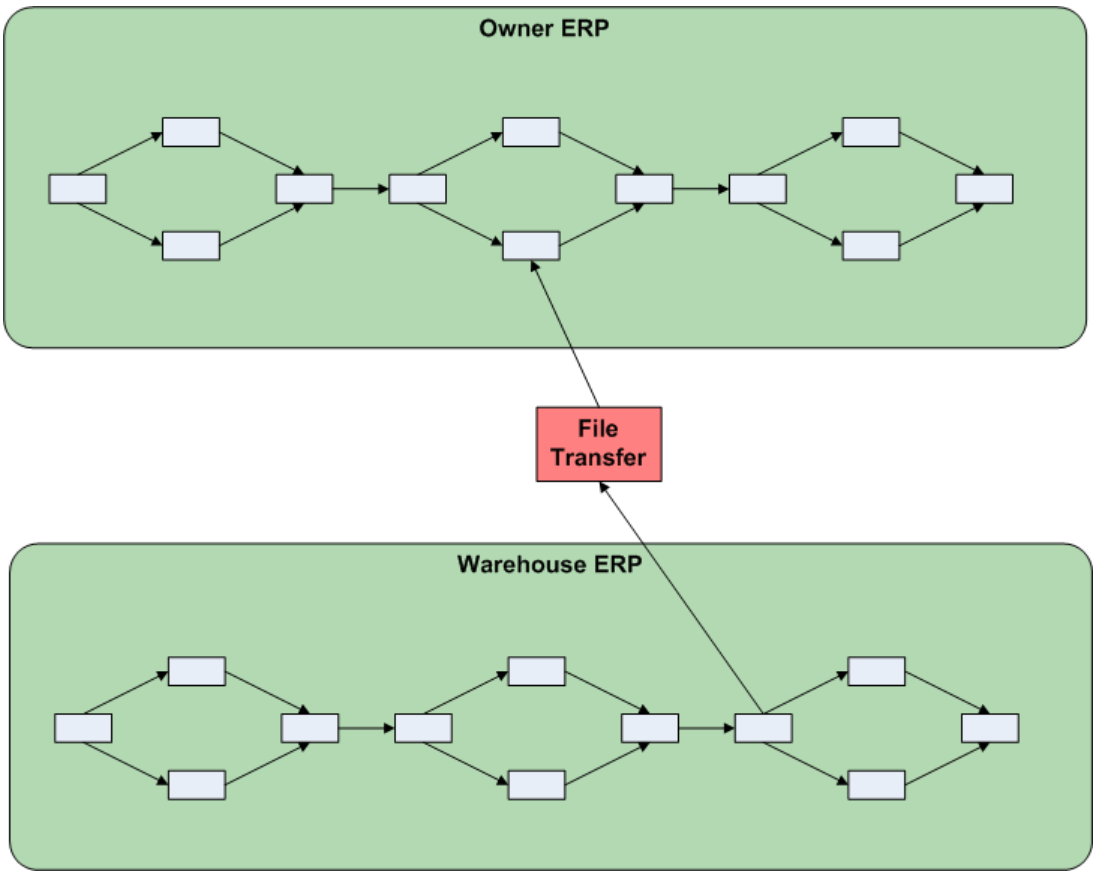

Figure 2.2: Source systems, such as ERPs, may be fully in compliance, but data exchange procedures between these compliant systems are not necessarily compliant.

The Weakest Link: File Transfers

ERPs and other enterprise-scale applications are designed from the start to provide an integrated flow of data through multiple processes. No matter how many ERP modules are licensed, there will be times when business requirements cannot be met within the ERP and data has to be transferred to some other system. In the case of the business and its warehousing partner, data exchange is needed to move data between ERPs.

Once the source ERP generates the data to transfer to the other ERP, the source system's controls no longer protect the data. The data is not in the target ERP, so that system's controls are not yet protecting the data either. The data has entered a veritable "no man's land" of integrity controls. To mitigate this situation, the workflow that transfers the data from one ERP to the other could employ several techniques:

- Writing the data to a file or set of files to a secure directory with sufficient access controls
- Encrypting the file to prevent unauthorized access to the information in the data files
- Calculating a message digest that is passed along with the file to ensure that there are no unauthorized changes to the files
- Logging any changes to the files and recording metadata about changes, such as the user ID making the change and time of change

The problem with this scenario is that developers writing ad hoc file transfer programs may not be aware of compliance issues or even the fact that the data being transferred is materially relevant to financial reporting in the first place. Line of business managers, for their part, may not be aware of the low-level technical details of file transfers and could be understandably unaware of the fact that a seemingly simple step in a complex information flow could leave essential data vulnerable to tampering. Although there are several ways to mitigate risks with ad hoc file transfers, there is no guarantee that those responsible for implementing them are aware of the need for them. This scenario depicts the limitations of ad hoc file transfer solutions with respect to SOX, but as the following sections will show, similar problems arise with other regulations as well.

PCI DSS

If you follow information security news, you have probably heard several times about the "largest data breach to date" involving either a retailer or a credit card processor. The credit card industry has stepped in with a bit of self-regulation to reduce the risk of credit card data breaches by establishing PCI DSS.

Resources

See the Privacy Rights Clearinghouse "Chronology of Data Breaches" at http://www.privacyrights.org/ar/ChronDataBreaches.htm for a sobering list of incidents; those that include credit card data as well as other types of data breaches.

The PCI DSS establishes several control objectives designed to minimize the risk to credit card information including:

- Building and maintaining a secure network
- Protecting cardholder data
- Maintaining a vulnerability management program
- Implementing access controls
- Monitoring networks
- Maintaining an information security policy

The second, third, fourth, and sixth control objectives are directly relevant to file transfers.

Protecting cardholder data includes encrypting that data anytime it is transferred outside of a secure network. Vulnerabilities can exist in workflows that include file transfers, but those vulnerabilities may be difficult to detect when custom scripts are used to transfer data files. Access controls may be the responsibility of systems administrators who manage directories where transferred data files are staged. The information security policy is highly relevant to the way file transfers are performed; however, that information may not be captured in design requirements provided to the programmer developing a custom transfer script. Unless there is a review and enforcement process in place, how can management be sure the policy is actually implemented?

HIPAA

HIPAA is legislation designed to protect personal health information. Part of the regulation, known as the Security Rule, specifies three types of safeguards:

- Administrative
- Technical
- Physical

The administrative safeguards describe policies that should be in place regarding governance and management oversight, access controls, training, and emergency response. The technical safeguards focus on access controls to data and protections of transmitted data. These include, for example, controls to prevent unauthorized changes to data, checks to verify the integrity of data, and the use of encryption to ensure confidentiality of private data. Physical safeguards address access concerns related to hardware and physical facilities.

With regards to file transfers, both technical and administrative safeguards must be in place. For example, if a file containing protected health information is transferred from one healthcare provider to another, several issues must be addressed:

- Is the file encrypted?
- Is a checksum, message digest, or other integrity check calculated for the file?
- Is the transfer partner authenticated using a digital signature or other means of authentication?
- Is a digital signature applied to the file so that the receiving party can verify the source of the transmission?
- Are access controls in place to prevent tampering with the file after it is generated but before it is transferred?

As with ERPs, data may be well protected when it resides within healthcare information management systems, but the process of exporting and transferring data, such as from a healthcare provider to an insurance company, can be the most vulnerable point in the process.

GLBA

In 1999, the United States repealed a long-standing law that kept commercial banks, investment banks, and insurance companies separate. The legislation, GLBA, included privacy safeguards to protect customer data that would be maintained by financial services companies. The most relevant part of the legislation to file transfer processes is the Safeguard Rule.

Under the Safeguard Rule, financial services firms are required to

- Evaluate the risks to private information maintained by the company
- Develop, implement, and monitor security practices to protect private information
- Establish management oversight for those security practices

Although the GLBA's safeguards may not be as precise as other regulations, they clearly dictate the need to protect customer information when stored and when transmitted. The GLBA opened the door to mergers of different types of financial services firms, which would clearly benefit from exchanging customer information for cross selling, marketing, and other business development purposes. Much of that exchange could be done using file transfers, thus creating the same potential security threats we have discussed in sections about other regulations.

Figure 2.3: Information sharing across financial service segments is allowed under GLBA, but private data must be protected, including during transmission.

21 CFR Part 11

21 CFR Part 11 is a section of the Code of Federal Regulations that specifies requirements for pharmaceuticals, medical device manufacturers, and others regulated by the Food and Drug Administration (FDA) with regard to electronic records. Clearly, the FDA has an interest in protecting the integrity of records related to the design and manufacture of drugs and medical devices. As with other regulations discussed in this chapter, the code specifies the need for access controls, audit trails, written policies, and so on. There is also special emphasis on electronic signatures in this regulation.

Data integrity is particularly important for the FDA. If, for example, there were a problem with a particular batch of a prescription drug, the FDA and manufacturer would want clear records of the components used to manufacture the drug, the location of the manufacturing, the distribution of the drug, and other information needed to recall the product. File transfer processes would once again be required to maintain sufficient information and enforce adequate controls to meet these requirements.

Compliance and File Transfer Requirements

Generalizing from the several regulations we have discussed, we can conclude that at minimum, regulations typically demand:

- Security procedures to protect the confidentiality of private data and the integrity of operational and financial data
- Effective controls to implement security policies
- Management and audit reporting sufficient to demonstrate compliance with the regulation

Repeatedly through this discussion, we have seen that file transfers can be a weak link in the overall security posture of a workflow. Enterprise applications, such as ERP and healthcare information management systems, may have sufficient controls in place to meet regulations; however, once data is exported for transfer to other systems, those controls are no longer protecting the exported data. At that point, the file transfer system should assume responsibility for securing data. Unfortunately for the developers of ad hoc file transfer solutions, that is a significant task to take on.

Developers should also keep in mind that security measures provided by a company's internal network may not be in place when working with public cloud providers. Consider the possibility that a custom file transfer program lacks sufficient security controls. For example, a custom program might have weak authentication, allowing someone to perform an unauthorized transfer. Fortunately, other controls are in place that log the fact that a file is created on the target server. If the transfer is made to a cloud storage provider instead of an internal server, no record of the transfer will appear in the log. Custom scripts that worked well with on-premise transfers might not be sufficient when working with cloud resources, including storage and software as a service (SaaS) providers.

Flexibility and Scalability

In addition to external drivers, like compliance, we have internal drivers that promote the adoption of managed file transfer methods. One of those is the need for flexibility and scalability. Flexibility is important because there are so many ways in which file transfers are used within an enterprise, between enterprises, and with external services, such as cloud providers. A file transfer method is ideally suited for multiple uses cases. Scalability has obvious implications for meeting service level requirements and continuing to meet those as demands change.

Three aspects of file transfer operations are relevant to understanding flexibility and scalability:

- Size of files transferred
- Volume of files transferred
- Number of transfer partners

These individually and collectively contribute to the need for managed file transfer solutions.

Size of Files Transferred

There is no "average" file size when it comes to managing file transfers. We could have small control files that pass between processes to indicate the state of one process or to pass a small number of parameters before initiating a process on another server. We can of course have quite large files. Some common large file exchanges include:

- Database dumps
- Multimedia files
- Extraction, transformation and load (ETL) files for data warehouses
- Input files to batch processing operations
- Network, server, and application log files

Difficulties can arise when transferring large files:

- Packets may be dropped and have to be re-transmitted
- Connections can be lost during a transfer leaving the target file partially copied
- Large file transfers can run the risk of running longer than the time window allotted to them

A robust file transfer solution will scale to large file transfers by incorporating a number of features, such as the ability to recover a disrupted transfer from the point of failure without starting over again and the use of efficient protocols that minimize overhead while preserving the integrity of the transfer process.

Volume of Files Transferred

In addition to the need to transfer large files, scalability requires the ability to handle a large number of individual files. As the number of files grows, so does the overhead associated with tracking each transfer. Even without the demands of regulations, it is a good practice to log information about each transfer; logging can include details about:

- The name of the file transferred
- File size
- File type
- Start and end of transfer
- Encryption status
- Authentication on the target server

As the number of transfers grows, so does the volume of logging data associated. Ideally, this information is centralized; that would ease management overhead.

It is especially important to ensure the integrity of file transfers when performing analysis on those files. For example, you might transfer a large number of application, server, and network log files to an on-premise cluster of servers or to a public cloud for analysis. The object of the analysis is to identify correlations between application performance and events on servers and the network. If data is missing because of a problem with file transfers, the analysis will not accurately reflect the actual performance of the application and correlated events.

Number of Transfer Partners

An increasing number of transfer partners, like the number of files transferred, can create scalability issues as well. Difficulties with increasing number of transfer partners tend to cluster around management issues:

- Scheduling transfers according to different partners' requirements
- Managing multiple authentication credentials
- Tracking different transfer requirements, such as whether to encrypt files before transferring them
- Providing custom reporting to trading partners

Flexibility and scalability are key drivers to adapting a managed file transfer solution. The demands for scalability manifest themselves in terms of file size, number of files, and the number of transfer partners.

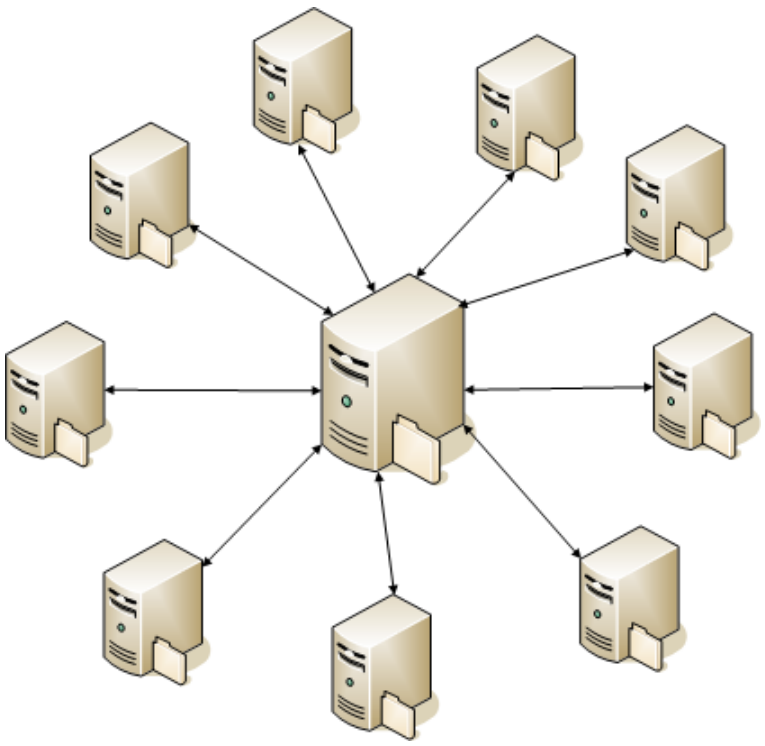

Figure 2.4: Increasing the number of trading partners increases the overhead of file transfer operations.

The need for scalability across these dimensions stems, in part, from the way we design applications and workflows today. They are highly distributed and can require large volumes of data transfers. Exchanging small amounts of data, for example, a single sales order, can be done with programmatic interfaces such as Web services; bulk transfers are still better done via file exchange.

Management Issues in File Transfer

As file transfer processes are so important to business operations, managing those processes has become a key driver to adopting managed file transfer practices and solutions. Three important issues in file transfer management are:

- Centralized control
- Monitoring and alerts
- Management reporting

Centralized Control

With so many applications of file transfers, from loading data warehouses and exchanging data sets for batch operations to sharing data with business partners and replicating data for disaster recovery protection, there is a growing need for centralized control.

Centralized management provides the ability to monitor:

- Status of file exchanges
- File exchange schedules
- Trends in growth of number of exchanges
- Trends in growth of data volumes
- Required procedures for exchanges, such as the use of encryption for exchange of confidential or private data

Of course, we could manage all these aspects in a more distributed fashion, but it would be less efficient and more prone to error.

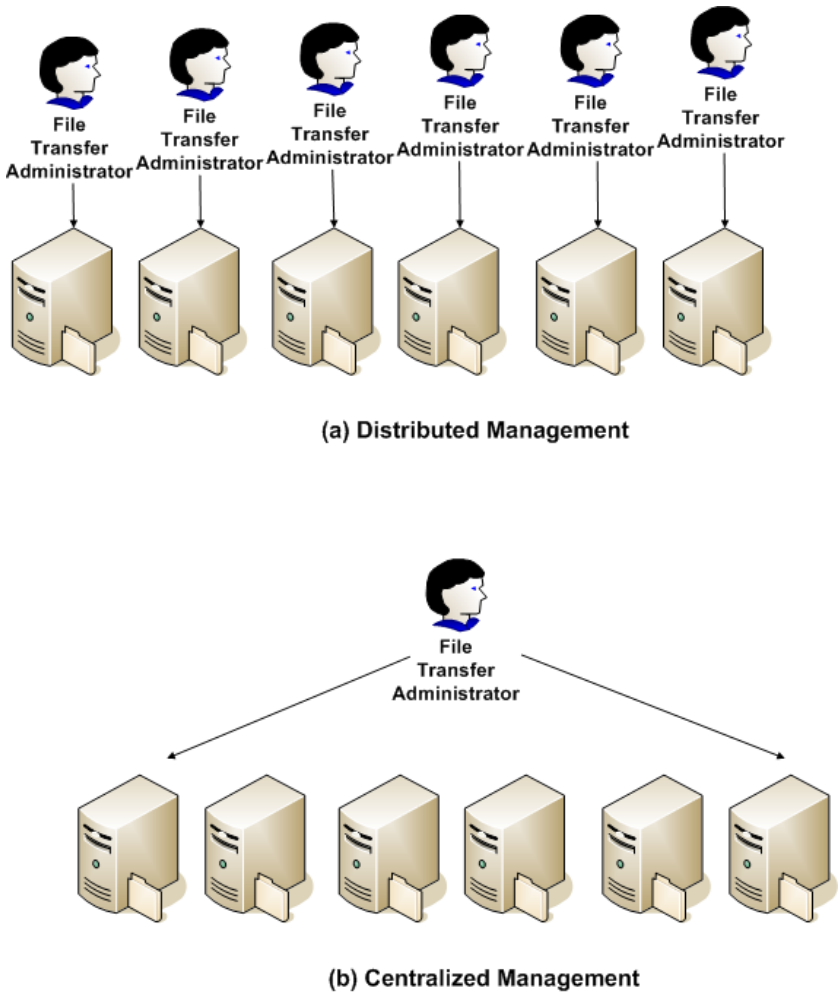

Figure 2.5: Distributed management is less efficient than centralized management.

Monitoring and Alerts

In addition to the normal, expected processes we need to manage within file transfer operations, we have to plan for the unexpected. Monitoring is the routine process of reviewing essential performance indicators:

- Time required to transfer files
- Average time to transfer per file size unit
- Number of encrypted/unencrypted transfers
- Number of transfer partners sending and receiving files

Monitoring helps to establish baselines and detect trends and changes in those baselines. This is especially useful when planning for future requirements.

Alerts keep us aware of unexpected events that need more immediate responses. Alerts may be triggered, for example, when:

- File transfers fail due to insufficient space on the target server
- Authentication fails on the target server
- Target server is unreachable
- Connection is refused by target server

With proper logging, alerts can also be tracked over the long term along with other areas covered by management reporting.

Management Reporting

Management reporting differs from monitoring and alerts in that the former provides a more aggregate view of the status of file transfer operations. Management reporting is especially difficult when ad hoc, homegrown solutions are used because each program may have its own method for structuring data, log entries, and management reports (if they exist at all). In a centrally managed system, a single log of events can be used to generate data about operations across multiple jobs, users, and even departments to provide a comprehensive overview of file transfer operation.

Cost Control

The discussion of centralized management and monitoring demonstrates the impact of distributed management on cost of implementing and maintaining multiple file transfer solutions. Certainly, different users will have different requirements, but a properly designed, centralized file transfer solution will be able to adapt to multiple requirements.

For example, one process may require a single file to be transferred from one local server to another at the same time every day. That is a pretty straightforward requirement. Another process may require multiple files to be transferred to a third party only after a particular event has been triggered. In addition, the transfer contains confidential information, so the files must be encrypted before transfer. These may sound at first like they need different programs to manage the transfer, but these tasks are easily accommodated by file transfer solutions that support job-specific configurations. A job configuration may include settings such as:

- Single or multiple files to transfer
- Transfer based on time or event
- Encrypt files prior to transfer
- Authentication required for external server

By using a single platform to perform multiple transfers and meet multiple use cases, we can reduce the cost of in-house deployment and maintenance.

Consolidating servers is an especially effective way to reduce costs. Consider the need for multiple security protocols. A business with multiple file transfer partners might have to support several security protocols, such as SFTP, FTPS, FTP with PGP, and HTTPS. Managed file transfer solutions can support all these protocols on a single server, reducing the need to procure and manage multiple servers to accommodate the varying requirements of different transfer partners.

Cost considerations play a major role in how you use cloud computing services. Many providers charge for servers by the hour. If you have 20 servers running but not enough data to keep them busy, you are paying for unused resources. It is important to have file transfer solutions that can transfer data to the cloud reliably and quickly. It is also important to have file transfer solutions that can recover from problematic transfer.

Some cloud providers charge for data transfers. How your file transfer solution handles errors can significantly affect your data transfer costs. Consider a large file download that fails after 80% of the file is downloaded. If the download must be restarted, you will pay for the 80% download as well as the next attempt to download the file. Assuming the second attempt is successful, you would pay 180% of the cost you would have paid had the file transfer not failed. Alternatively, if the file transfer solution detected a problem with the transfer and restarted the transfer from the point of failure instead of from the beginning, additional charges would be minimized.

A centralized system can eliminate the need for developing scripts targeted to single jobs that are managed by different administrators and debugged by different developers and that generate their own reports and log that are not easily consolidated for management reporting. Centralizing file transfers drives down the cost of developing, debugging, and maintaining file transfer solutions. The result is more efficient and cost-effective file transfer that also avoids the opportunity costs of having developers work on file transfer solutions when they could be working on other pressing business needs.

Workflow Efficiency

Another driver for improving file transfer efficiency is improving workflow efficiency. File transfers are typically one step in a larger workflow. If we ask the question “Why are we transferring these files?” we will likely find an answer that involves a complex series of steps implemented to support a larger business process. In a sense, no file transfer is an island unto itself. With this in mind, we can see two opportunities for improving workflow efficiency through improved file transfer.

Integration with Existing Workflows

The first is integration with existing workflows. Businesses are constantly executing workflows, from simple transaction-based workflows to complex multi-department operations. Often we find that workflows can be optimized and made more efficient by studying the global requirements of a process. For example, if we looked at just the first few steps in a workflow, we may miss the fact that later steps duplicate some of the same processing. By considering the global workflow, we may find new opportunities to optimize the process.

Similarly, if one workflow ends with the generation of a file to be transferred and another workflow begins after the file has been transferred, we may have the opportunity to optimize the overall flow by integrating the two workflows. The file transfer process becomes the lynchpin in this case; rather than having two workflows, and the separate management overhead, we can combine into a single workflow with consolidated management, monitoring, and development.

Support for Multiple Trading Partners

We can extend the idea of integrated workflows across organizational boundaries when we support the needs of multiple trading partners. File transfers are parts of larger business processes. If we were to custom design a file transfer solution for every trading partner (as some of us have), we reduce the efficiency of the file transfer process, introduce a weak link into the process, and make management reporting and operations management more difficult. A single file transfer solution for multiple trading partners gives us many of the benefits we have seen for other applications, including: centralized management, improved monitoring, alerts, and the ability to better monitor trends.

Summary

There is no single reason to improve file transfer, there are several. Compliance with regulations is a major concern and ad hoc, homegrown file transfer solutions may not meet all the requirements of regulations, and if they do, they are costly to build and maintain. Business is dynamic and workflows have to be dynamic too. Flexibility and scalability are essential characteristics of an enterprise-scale file transfer solution, especially with respect to the size of files transferred, the number of files transferred, and the number of trading partners. Management needs both tactical reporting, such as alerts, and longer-term management reports on trends. Here again, a single centralized file transfer solution is better able to provide that information efficiently than a large number of custom solutions. Cost controls and workflow efficiencies are also key considerations driving the adoption of managed file transfer solutions.