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

Introduction to Realtime Publishers i

Chapter 4: Planning and Deploying a File Transfer Solution 50

Assessing Current State of File Transfer Methods 50

Identifying Homegrown File Transfer Solutions 51

Working Top Down with Formal Enterprise Software Management: The Best of All Possible Worlds 51

Distributed Management Models and the Bottom-Up Approach 52

Inventory Business Process Dependencies on Homegrown File Transfer Solutions 55

Identifying Business Requirements for File Transfer Solutions 55

Inventory Hardware and Assess Repurpose Potential 57

Multiple Dedicated File Transfer Servers 57

Countering Virtual Server Sprawl 59

Prioritizing Replacement of Existing Solutions with an Enterprise File Transfer Solution 61

Criticality of Business Process 61

Frequency and Volumes of File Transfers 62

Reliability of Existing Solutions 63

Security and Compliance Requirements 63

Establishing File Transfer Policies and Procedures 64

SOPs for Managed File Transfer Operations 65

Reporting and Monitoring Policies 65

Security Policies 66

Storage and Resource Management Policies 66

Rolling Out a Managed File Transfer Solution 67

Summary 67

Copyright Statement

Chapter 4: Planning and Deploying a File Transfer Solution

Throughout *The Shortcut Guide to Eliminating Insecure and Unreliable File Transfer Methods,* we have considered the limitations of traditional file transfer methods, analyzed key business drivers motivating improvements to file transfer methods, and assessed essential requirements for a secure and reliable file transfer solution. In the final chapter of this guide, we will turn our attention to planning and deploying a secure and reliable file transfer solution. We continue our methodical approach to understanding file transfer issues and solutions with a structured approach to deployment.

The chapter is organized into several sections, each describing a core step in the deployment process:

- Assessing the current state of file transfer methods used in an organization
- Identifying business requirements for a file transfer solution
- Inventorying hardware and assessing repurpose potential
- Prioritizing replacement of existing solutions with an enterprise file transfer solution
- Establishing policies and procedures for enterprise file transfer
- Rolling out a file transfer solution across the enterprise

By the end of this chapter, the reader will have concise description of critical questions that should be answered, methods for assessing the answer to those questions, and guidelines for constructing policies and procedures to promote the most efficient and effective use of an enterprise file transfer solution.

Assessing Current State of File Transfer Methods

Once you have determined that your organization would be better served with a secure, reliable file transfer solution, the next step is to catalog existing methods for file transfer and business processes dependent on those methods. The goal at this stage of the process is not to change any of the existing processes; we simply want to understand the extent to which ad hoc file transfer solutions are used and where they are used.

Identifying Homegrown File Transfer Solutions

The difficulty in identifying where in an organization homegrown file transfer systems are used can range from fairly simple to time consuming and complex. There are two basic strategies for identifying homegrown file transfer solutions: a top-down approach and bottom-up method. They are not mutually exclusive and often we can draw from both. The top-down approach is simple and direct; it takes advantage of work done in the process of formally managing software assets. The bottom-up method works for environments without centralized and detailed information about software deployments.

It is worth noting that even in organizations with formal software asset management, there might be custom file transfer solutions. Such is especially the case when departments or other small groups can use public cloud providers for special projects, like an analytics project or new business intelligence initiative. These kinds of projects can emerge outside of normal project management and software development procedures, and they can be particularly difficult to track down.

Working Top Down with Formal Enterprise Software Management: The Best of All Possible Worlds

At one end of the spectrum, we have organizations with formal software management policies that are strictly followed. In these ideal organizations, all production software is managed in code repositories. Approved copies of commercial software are maintained in a central repository along with tested and validated patches for that software. Copies of the software are deployed from the repository, information about the deployment is maintained, and one can readily report on where the software is deployed, which version is in use, and similar management-related data. Custom applications are similarly managed with the additional benefit of having copies of source code.

In this ideal situation, assessing the use of homegrown solutions becomes a matter of reviewing the contents of the software repository. Custom applications will likely have descriptions and other metadata that indicate whether file transfers are part of the services provided by the software. One should search for terms such as:

- ftp, ftps, and sftp
- extraction, transformation, and load (ETL)
- extract, dump, and backup
- replicate and synchronize
- file transfer, file copy, and so on

Although we are concerned here with "homegrown" solutions, we should not ignore commercial software. For example, a commercial ETL tool may have an ftp module that is used in ETL scripts to transfer files. The software repository may not have all scripts currently in production, but it will at least point to business processes and application users that might have developed such scripts.

Distributed Management Models and the Bottom-Up Approach

Not every organization has an enterprise-wide formal software management process. This is not necessarily a problem; different businesses have different drivers and requirements. A multinational bank might not hesitate to spend the time and resources required to maintain a formal software management system. A rapidly growing startup, in contrast, might have to focus on more immediate demands on its time. Many organizations will likely fall somewhere in between where the desire to adopt ideal software engineering practices are tempered by constraints of time and budget.

The bottom-up approach starts with low-level details of file transfer implementations. Let's assume a worst-case scenario in which we are working with an organization that has no structured information on the use of file transfer programs. The company has grown quickly and had to focus on delivering IT services on tight schedules. Corners were cut and documentation is limited. The organization's drive to deploy services was successful and IT management, systems administrators, and developers are all in agreement that the quick and dirty homegrown programs for file transfer need to be replaced.

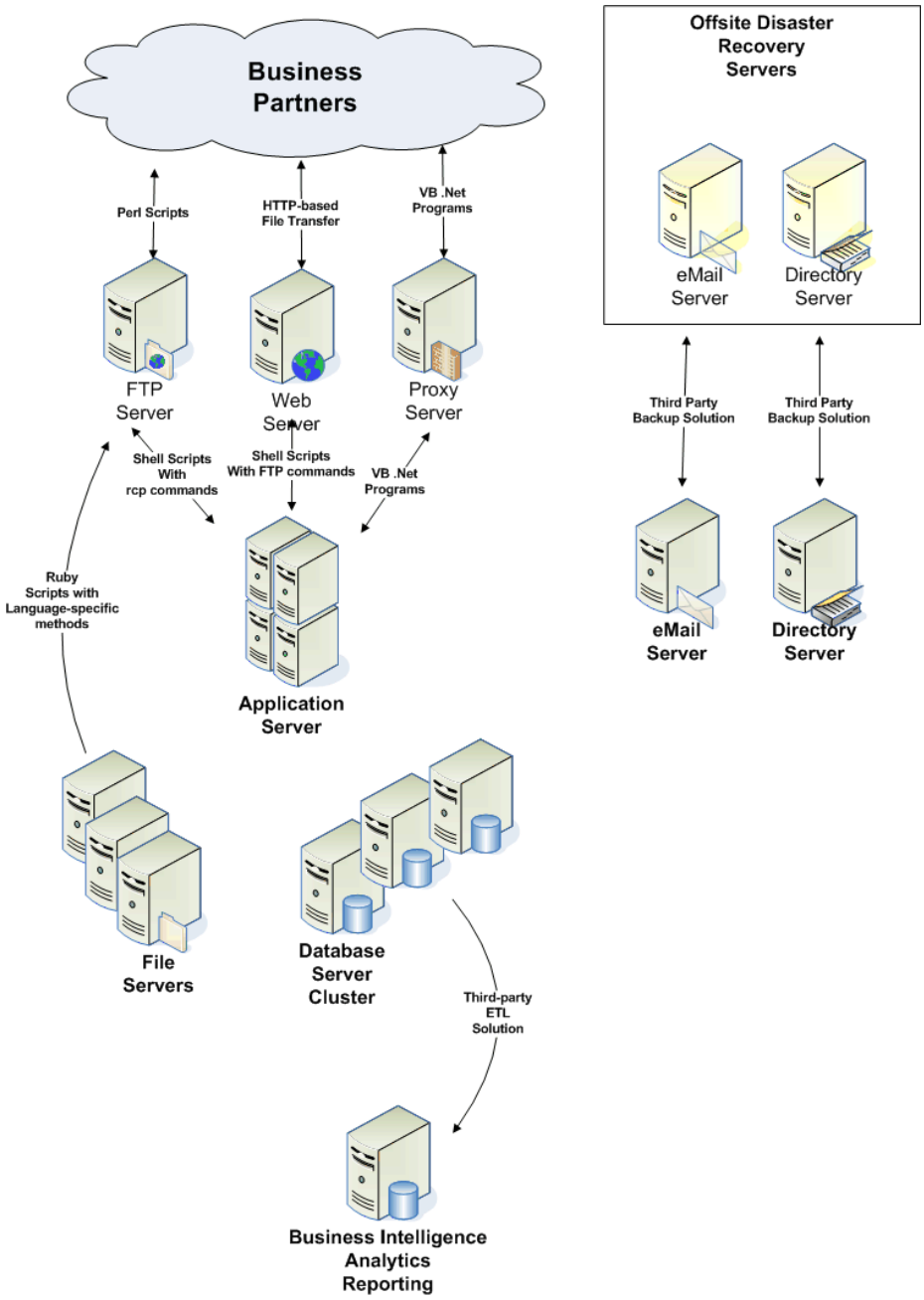

Figure 4.1: Different methods of file transfer may be used throughout an enterprise when each file transfer process is treated as a new requirement rather than an instance of a generalized problem.

The bottom-up approach begins with searching servers for scripts used to transfer files. There are many ways to transfer files (see Figure 4.1 for some examples), so there will be a variety of patterns that one should look for when searching code directories. Some example patterns indicative of file transfers include:

- Perl scripts that uses the module Net::FTP
- Python scripts that use the low-level sockets module
- Ruby scripts that use the FileUtils module, and the copy method in particular
- Visual Basic programs referencing the System.IO object, especially those that use the FileExists and FileCopy methods
- Unix/Linux shell scripts using the remote copy (rcp) or ftp commands

This is not an exhaustive list, but it does give an indication of the range of ways in which we can implement homegrown file transfer solutions. In addition, one should search for scripts or configuration files for ETL or backup programs that may be used to transfer files created specific to those operations.

> **Which Search Tool?**
>
> Window's file searching service provides basic search functionality and is suitable for simple search requirements. Linux and Unix users have the advantage of the grep command and its support for regular expressions. For a little bit of Unix/Linux in a Windows environment, use the Cygwin suite of tools (http://www.cygwin.com/), which offers a Linux-like command-line environment for Windows users.

Of course, this method assumes we have read access to all production code directories. In practice, these search scripts will likely be run by several systems administrators and application managers who would have appropriate access to the various servers.

In the case of small groups using public cloud providers, you might find an ally in the finance department. Cloud usage will leave a trail of transactions. If payments are made to cloud providers through invoicing, purchase orders, or other company payment processes, then they should be relatively easy to track down. If small payments are made on credit cards, you will need access to all the transactions on those cards to find all cloud users.

Gathering information about homegrown file transfer solutions, whether it is done in a top-down, bottom-up, or combination method is the first step in the assessment phase. The next step is assessing business processes dependent on homegrown solutions.

Inventory Business Process Dependencies on Homegrown File Transfer Solutions

Every file transfer program in production use is running to support one or more business processes. For each of these programs, we should document key characteristics of the business process, including:

- Whether it is an internal or external transfer
- How frequently the transfer is executed
- The number of files typically transferred
- The volume of data typically transferred
- Constraints on when the program can execute

We should also categorize the business processes in terms of how important the process is to overall business operations. For example, file transfers from an order entry system to an order fulfillment application are critical. File transfers from an enterprise resource planning (ERP) system to a data warehouse are important but less critical than customers' orders. File transfers to support updates to news stories on the internal employee portal fall into the least important category. This information is useful for planning migration from existing solutions to an enterprise file transfer system.

At this point, we have information about what homegrown solutions are in use and what business processes are supported by them. Next, we need to assess the actual business requirements of the business processes using file transfer. As noted earlier in this guide, homegrown solutions may not actually meet all the business requirements of business operations. It would be a mistake to analyze existing file transfer programs and assume we have a handle on the business requirements.

Identifying Business Requirements for File Transfer Solutions

It is safe to assume that even the most rudimentary homegrown file transfer solution meets some business requirements. At the very least, such programs copy files from one device to another. Usually there is more functionality provided as well. We can often find information about other requirements within existing solutions, specifically:

- Information about programs that invoke the file transfer program
- Scheduling details, such as cron job entries that define when the programs are run
- Information about programs that are run after the transfer is complete
- Details about file transfer processes that are recorded in log files
- Authentication requirements for running the file transfer program

There may be additional facts about file transfer operations that are not apparent in the scripts and supporting code that implement homegrown solutions. For these details, we need to research the underlying business processes that might be either internal or external processes.

Internal file exchanges are less complex than external ones with regards to management and oversight because all responsible parties are within a single organization. In many cases, the same authentication and authorization systems are used for both the file transfer source and target systems. There may be greater latitude with regards to access controls as well. For example, an internal file transfer process may use an ftp server configured to allow reading as well as writing, thus allowing users to list the contents of the ftp directory and confirm the file transfer completed successfully. Transfer partners that receive files from multiple external sources in the same ftp directory may be hesitant to give outsiders, even business partners, read access to a directory that is shared with others.

Big data analytics will drive additional requirements for file transfer solutions. Those kinds of projects can drive processes that work with a wide range of files, such as log files, data base extracts, and third-party data sets, such as demographic information. Be prepared for potentially large file transfers to support analytics efforts. Also consider how frequently the file transfer processes will need to run. If you are collecting data from a heavily used Web application, you might want to download application logs every few hours. When you are downloading third-party demographics data, you might find monthly updates are sufficient. Optimizing file transfers will depend, in part, on understanding the frequency with which various jobs need to be run.

When assessing requirements for both internal and external file transfers, consider the following:

- Number of files and volume of data transferred
- Bandwidth required to complete transfers in time window allotted
- Retention policies on files transferred
- Need for encryption, both during transfer and while the files reside in ftp sites or other staging areas

In the case of external file transfers, one must also look into authentication and authorization issues. In cases where transfer partners have close relationships, there may be trust agreements between the organizations that allow identities on one of the partner's systems to be trusted by the other partner.

> **Note**
>
> These are known as federated identity management systems; they come with their own host of complex management challenges.

In other cases, a trading partner may simply embed a username and password in a transfer script. Although this is clearly a security risk, it is the kind of shortcut that is sometimes taken in homegrown file transfer solutions. An important requirement for any enterprise-scale file transfer solution is the ability to securely manage user credentials, such as usernames and passwords or digital certificates.

Existing file transfer programs are a starting point for quickly assessing basic file transfer requirements; however, they should not be the sole source. These programs likely do not capture the full extent of business requirements, so it is important to consider the broader business processes that include the execution of these programs.

Inventory Hardware and Assess Repurpose Potential

A fragmented approach to file transfers can lead to a fragmented infrastructure when dedicated file transfer servers are used. It can also promote virtual server sprawl in a virtualized environment.

Multiple Dedicated File Transfer Servers

Consider a typical project involving a multi-tiered architecture. Servers are needed to host a Web server, another to support an application server, and yet another for database services. The project also requires file transfer between the application server and a business partner's server. The system architect on the project is concerned about security vulnerabilities that have been found in ftp software and decides to host the file transfer application on a separate server in the DMZ (see Figure 4.2).

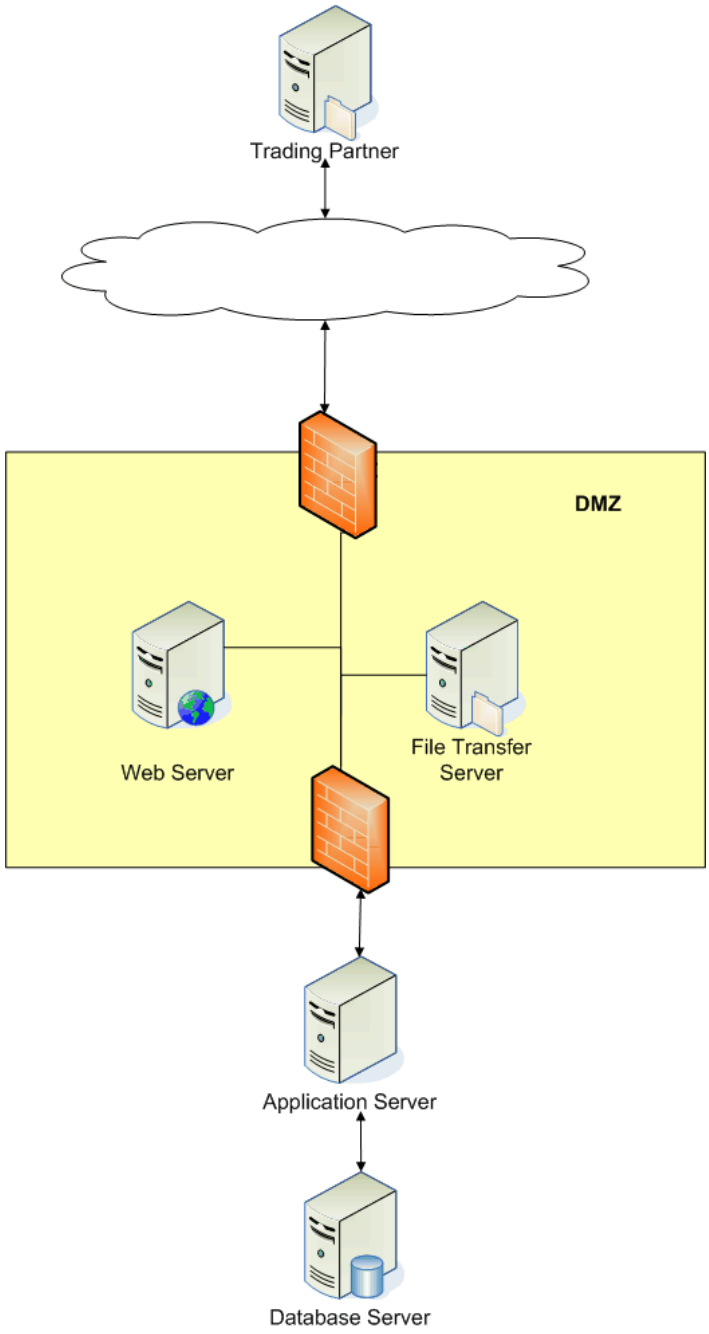

Figure 4.2: A single project solution to file transfer results in inefficient use of hardware.

If we multiply this scenario by the number of homegrown solutions in use within an organization, we can see the potential for reclaiming underutilized servers. Repurposing can directly benefit the business processes that use these homegrown solutions by allowing dedicated file transfer servers to be used for a number of alternative purposes:

- An additional server in an application server cluster
- An additional server in a federated database system
- An additional server in a Web server load-balance configuration
- A failover server in a disaster recovery situation

Consolidation can also benefit virtualized environments.

Countering Virtual Server Sprawl

Virtual servers offer many advantages of physical servers without the dedicated hardware. This makes virtual servers an ideal solution for isolating file transfer operations. The security concerns with ftp described in the previous section could be addressed by deploying a virtual server. In fact, one way to deal with the need for multiple file transfer methods (for example, ftp, sft, ftps, and https) is to host multiple virtual servers.

Another potential advantage of virtual servers is that a single file transfer server could be deployed with multiple virtual servers each of which is dedicated to a different business process. This configuration has several advantages:

- Different file transfer protocols may be configured on each virtual machine
- Each virtual machine can be administered by a different business process owner
- Access controls can be configured according to the needs of each business process
- Virtual machines can be configured with only the application stack needed by each business process, such as Perl, Python, and Ruby interpreters

The drawback of this approach is the additional management overhead that comes with deploying multiple virtual machines.

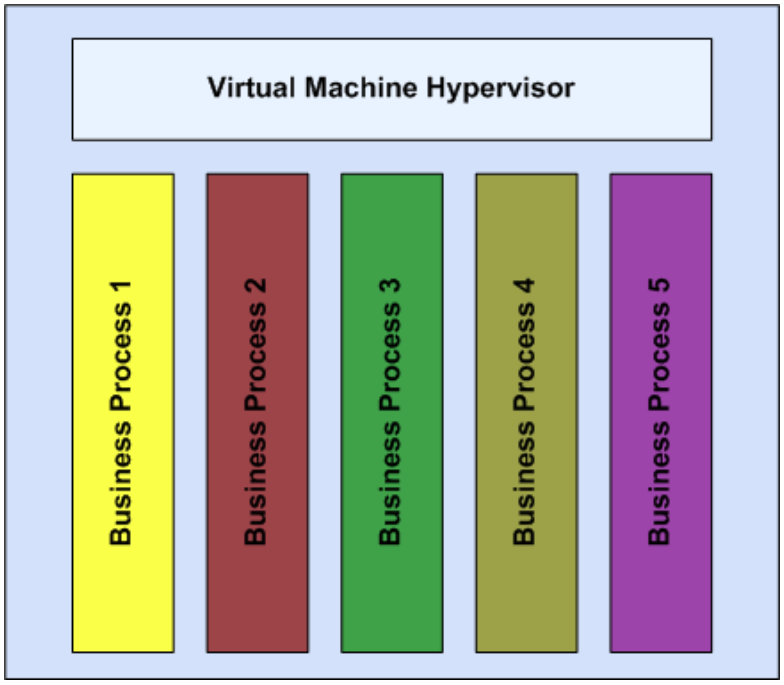

Figure 4.3: Deploying dedicated file transfer servers on virtual machines reduces the number of dedicated servers but maintains the management overhead of multiple servers.

Enterprise file transfer solutions can dramatically improve on these two scenarios. Rather than force architects to implement and enforce different file transfer policies and management schemes at the server level, enterprise file transfer solutions allow system designers to define logical policies and workflows that execute within a single file transfer application. There is no need for dedicated multiple servers or virtual server sprawl because file transfer processes can be implemented directly as a manageable, logical entity. The first step in the move away from fragmented architecture and collections of homegrown solutions is to prioritize the replacement of existing programs.

Prioritizing Replacement of Existing Solutions with an Enterprise File Transfer Solution

When faced with a wide array of existing homegrown solutions, it can be difficult to know where to begin with the migration to an enterprise file transfer solution. In such cases, it helps to consider several factors when prioritizing the move to a centralized solution. These factors include:

- Criticality of business process
- Frequency of transfer
- Volume of transfers
- Reliability of existing transfer solution
- Security requirements
- Compliance requirements

When considering these factors, the foremost consideration should be how well the existing file transfer solutions meet the business needs embodied in the factor.

Criticality of Business Process

File transfer processes are not created equal. The impact of a failed file transfer can range from a minor inconvenience to a costly one-time mistake to a long-term setback. Consider a few examples.

A business provides an employee portal for self-service human resources administration. Employees can perform routine tasks, such as submitting time cards, changing payroll deductions, and keeping up with company news. The portal administrator has established a simple mechanism for department managers to submit news to the employee portal. Managers email content to a designated email address, and a custom scripts extracts those emails, restructures them into an RSS feed format, and transfers them to the portal server. It's not difficult to imagine how such a script could fail: error message formats, changes to access controls on target directories, servers down for maintenance, and so on. Assuming the complete failure of the transfer script, there is still minimal impact on business. Customers are not adversely affected, business partners would not know of the minor glitch, and the business would not lose money due to the failure.

Now consider a more costly example. US businesses with sufficiently large payrolls are required to submit payroll taxes on a weekly basis. Companies that grow through mergers and acquisitions may find themselves continuing to operate multiple payroll systems for some parts of the payroll process while using a consolidated reporting system for tax purposes. Information about payroll taxes are collected from division-level payroll systems and a single payment is made based on the aggregated information. If a file transfer process fails or succeeds in transferring only partial information, there could be an error in tax payment. Penalties for such errors can be significant.

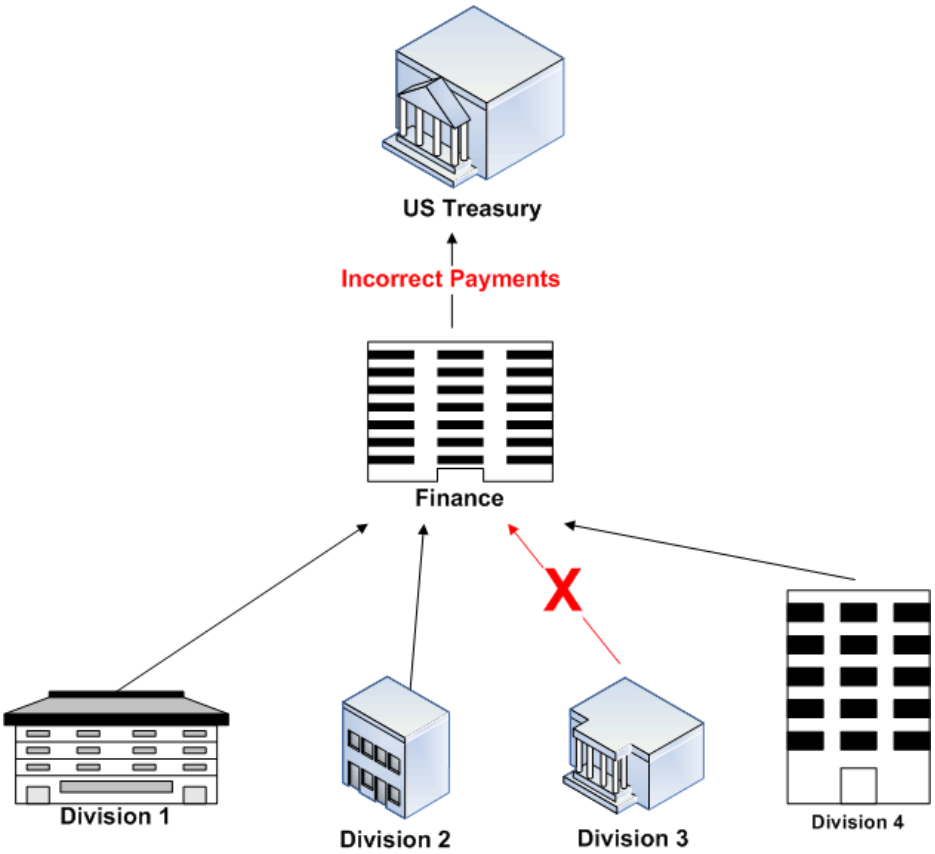

Figure 4.4: A failed file transfer can adversely affect business processes downstream of the transfer.

Frequency and Volumes of File Transfers

Another consideration when prioritizing new implementations of file transfers is the frequency with which the transfers are made. Frequent transfers are often found in core business operations, such as order fulfillment. As a general rule, if a file transfer is important enough to be done frequently, it is probably important enough to do it right all the time. This makes it a high priority to migrate to a reliable, secure enterprise solution.

The volume of data transferred is another good indicator of the relative priority of a transfer operation. Large volume transfers may be required for backup and replication processes. For example, full backups may be transferred to offsite disaster recovery centers on a weekly basis followed by smaller incremental backups performed during the week. In other cases, a large amount of data may be transferred as part of an ETL process for updating a data warehouse. Business intelligence reporting may be less critical than ensuring up-to-date disaster recovery capabilities; thus, as these examples show, volume alone is not enough to establish a priority order.

Reliability of Existing Solutions

File transfer programs that frequently fail drive up business operation costs. Failures require the time and attention of systems administrators and developers to debug and correct problems. Such failures entail opportunity costs as well: When IT professionals are chasing down problems with file transfers, they are not attending to other pressing business needs.

Reliability problems are best avoided with a combination of robust software and adequate management reporting. Dedicated file transfer solutions are more likely to be rigorously tested and provide error handling to respond to common problems, such as insufficient disk space. Reporting and alerts can be used to keep application administrators aware of changing conditions, such as low storage space or changes to access controls. This information is useful for preemptively correcting conditions before they adversely affect a file transfer operation.

Security and Compliance Requirements

Insecure homegrown solutions should be considered high-priority targets for migration to a more secure, enterprise file transfer application. Practices, such as embedding user names and passwords in clear text within a script, present security risks that should be eliminated. Quick and dirty techniques like this can compromise both internal security and that of file transfer partners.

Compliance is also a factor to consider with file transfers. Regulations governing the integrity of business data often require we demonstrate that we are employing sufficient controls to prevent tampering. That is a tough requirement to meet if we are using ftp sites to which multiple users can write. A host of compliance problems can arise with custom solutions:

- What is to prevent one user from maliciously or accidently overwriting another user's data on the same server?
- Do our homegrown solutions perform basic checks, such as checksums, to ensure files transfer correctly?
- If so, how are errors reported?
- Are logs that record details of transfers tamper-proof?

Security and compliance are critical considerations in file transfer operations that are easily overlooked with potentially costly consequences.

Prioritizing business process migration to a managed file transfer solution is just the first step to comprehensive management of file transfer operations. Many of the factors we consider when prioritizing migrations, such as volume of transfers and security considerations, also play an important role in file transfer policies and procedures.

Establishing File Transfer Policies and Procedures

One of the benefits of a consolidated file transfer solution is that the same system can be applied to a wide variety of business requirements: internal transfers, file exchange with business partners, encrypted or unencrypted transfers, and so on. Mature managed file transfer applications have many features built-in and ready for use. Our job is to use them most efficiently and effectively to do what we need within established policies and procedures.

Policies and procedures should be in place to guide the implementation of individual transfer operations as well as the management and governance of the enterprise service. In particular, we should consider:

- Standard operating procedures (SOPs) for managed file transfer operations
- Policies governing reporting requirements
- Process monitoring procedures
- Security policies
- Storage and other resource management policies

Together, these constitute the core management subject areas of enterprise managed file transfers.

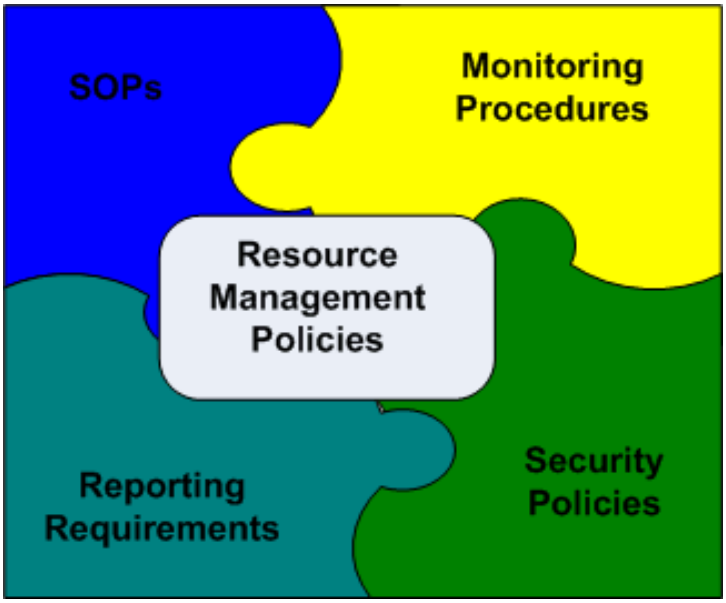

Figure 4.5: Managed file transfer requires a set of policies and procedures that defines how to operate and maintain the enterprise application.

SOPs for Managed File Transfer Operations

SOPs for managed file transfers are procedures that should be followed for all, or most, file transfer operations. The first procedure is creating a managed file transfer process. Doing so establishes jobs that run to execute the transfer. For each such process, we should collect information about:

- The business process supported by the transfer
- The business owner of the process
- The security requirements for the process, such as the need for encryption
- The frequency of the transfer
- The expected initial number and size of files transferred
- The expected growth rate of the number and size of files
- The time window in which transfers are to occur
- An indication of the criticality of the process

This type of data is important for day-to-day management as well as long-term planning. By storing this data in a centralized repository, we can have easy access to a comprehensive picture of managed file transfer processes. This, in turn, will support other management efforts, such as audits and other compliance reviews.

Reporting and Monitoring Policies

With a centralized system, we have the ability to report on the status of all file transfer operations. This is especially useful for logging details about transfer operations and notifying administrators about failures or other issues with transfers. Policies should be defined to:

- Establish rules for notifying administrators based on severity of issues
- Define types of information to routinely report on, such as number of files transferred, volume of data transferred, number of failed or interrupted transfers, and so on
- Establish response procedures based on types of errors and business importance of file transfer processes

The objective of these policies is to ensure that transfers are executing as expected and problems are detected rapidly and addressed in a standard manner. There should be no need for ad hoc responses, which are sometimes required with homegrown transfer applications.

Security Policies

It is especially important to have security policies in place to define acceptable use of file transfer services. Managed file transfer applications can automate the movement of large amounts of information within an organization and across organizational boundaries. This situation is itself a potential security risk if these services are not used according to established rules.

Security policies should address the need for:

- Patch management to ensure the managed file transfer application code is kept up to date and known code vulnerabilities are corrected.
- Vulnerability assessment because managed file transfer applications have access to significant amounts of data, making them targets for attack. Vulnerability assessments should examine configurations to detect improper settings that could be exploited by an attacker.
- Audit controls to securely log significant events, such as changes to system parameters, privilege elevation, or other events that alter the security stance of the application.
- Separation of duties with regard to supporting and maintaining file transfer operations. For example, administrators with the ability to create or delete transfer jobs should not have privilege to overwrite the event log.

Security policies such as these are designed to ensure the confidentiality, integrity, and availability of file transfer services.

Storage and Resource Management Policies

Preserving the availability of file transfer services also requires attention to limited resources, such as storage. Another set of policies should be in place to establish the use of quotas on bandwidth usage, the use of staging area storage, and rules governing the purging of data left in file transfer areas. Security policies are focused on preventing malicious disruption of services while these resource management policies are designed to avoid unintended loss of service due to resource limitations. Once policies and procedures are in place, a business can begin to effectively and efficiently use managed file transfer applications across the enterprise.

Rolling Out a Managed File Transfer Solution

Throughout, this chapter has outlined the steps required to migrate from homegrown file transfer programs to a managed file transfer solution. To deploy a managed solution, we need to:

- Assess the current use of file transfer methods
- Identify business requirements, especially those that are not addressed by existing file transfer programs
- Inventory hardware and assess how best to repurpose redundant servers
- Prioritize the order of replacement of existing solutions with an enterprise file transfer solution
- Establish policies and procedures for enterprise file transfer

As we replace existing custom solutions, we should verify several aspects of the new implementations. First, we need to able to account for all file transfer operations. They should be linked to business processes and business owners. As part of verifying the link to business processes, we need to verify that the new file transfer functions properly within larger workflows. For example, once a file transfer is complete, are follow-on events properly triggered and executed?

Second, we need to verify that the security requirements are met. This is one area where existing homegrown solutions are not necessarily good guides for judging the completeness of the new system. Authentication, authorization, and the use of encryption should all be verified.

Finally, we need to ensure that external file transfers are tested to ensure transfer partner's requirements are met. Ongoing management practices should be implemented according to the policies and procedure defined earlier.

Summary

File transfers are commonplace in today's enterprises. Data is constantly moving within and across organizational boundaries. Managed file transfer applications have taken their place alongside other enterprise applications as core business infrastructure. Moving away from homegrown, ad hoc solutions can and should be done in a methodical manner. The benefits of the approach outlined here include the ability to detect and implement business requirements, efficient use of hardware resources, a policy and governance structure for long-term efficient use of manage file transfer applications, and security practices that promote the confidentiality, integrity, and availability of file transfer services.