IBM data protection and recovery solutions
White paper

Tivoli software

Addressing the complexities of remote data protection.

July 2008

Contents

Introduction

Remote or branch offices are increasingly at the front lines of business – they have the closest contact with customers and business partners and therefore can have a dramatic impact on the success of the business. Analysts estimate that there are more than four million remote offices in the United States alone.[1] Many of these offices run autonomously from headquarters and are responsible for managing their own operations – including protecting and retaining the electronic information that they generate. Ignoring the protection and recovery needs of this remotely stored data is simply not an option.

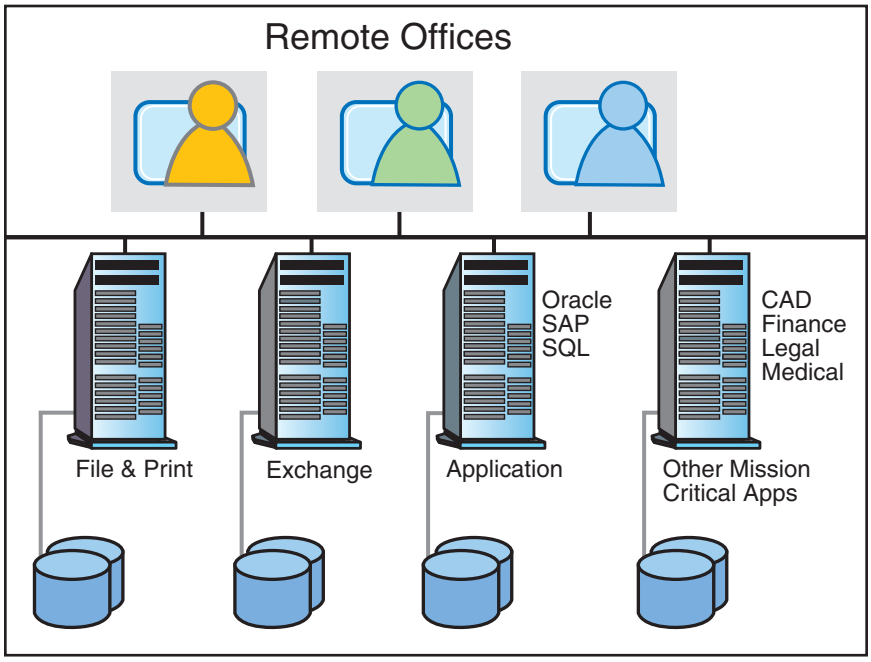
As companies expand operations into new markets, the percentage of total corporate data in remote offices is increasing—the industry average is now 31 percent.[2] However, many companies may not be adequately protecting these assets.

This white paper reviews the factors that can make data protection and recovery in remote offices a complicated problem and describes the range of technology options available, explaining the benefits and shortcomings of each in the context of remote offices. The paper concludes with a brief explanation of how the IBM Tivoli® Storage Manager FastBack™ product family delivers a full-featured adaptive data protection solution for the remote offices of enterprises of all sizes.

For the purpose of this paper, a "remote office" is a staffed office with one or more Microsoft® Windows® servers that support business activities (see Figure 1). By this definition, a remote office probably includes a file and print server and, depending on the operations, may also include:

- A Microsoft Exchange Server for e-mail management.
- An application server running a database program like Oracle, SAP or Microsoft SQL Server.
- Another application server to support business-specific, mission-critical applications.
- A backup server attached to a tape drive or autoloader.

Figure 1: Remote office infrastructure



Highlights

Factors to consider when protecting remote data are the types of applications to be protected, the potential risks, and your specific recovery objectives.

These systems can be the lifeline of your business, and losing access to them can result in lost productivity, lost opportunity and, ultimately, lost revenue.

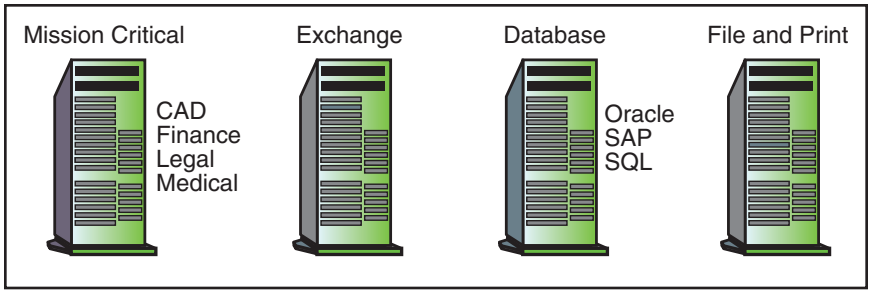
The complexities of protecting remote data

In designing an overall solution to the remote data protection problem, there are many factors to consider, including the types of applications to be protected, the potential risks, and your specific recovery objectives.

Applications

The applications that are most critical to the success of your business should have the greatest degree of protection (see Figure 2). Applying the same level of protection to all the applications and services running in each office may seem simpler to implement, but it could cost too much and have too great an impact on your operations to be feasible. You should evaluate each application and the data that it generates or manages, and set relative priorities for data protection.

Figure 2: Remote office applications



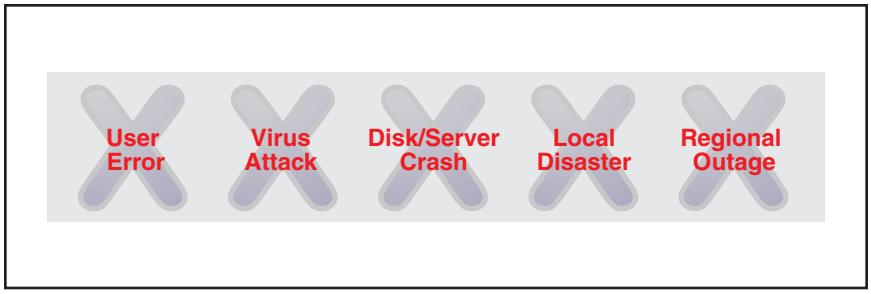
For example, many remote offices probably have a shared file and print server. Losing data from this system probably won't put the organization in jeopardy, although it may be inconvenient trying to recover or recreate that data. Running a backup on this system once per day may be sufficient. An Exchange e-mail server or an SQL database that contains real-time business transactions may need a higher level of protection due to the value of the data being generated and managed. Once an hour or a few times per day may be the right approach, as long as you don't impact the performance and availability of the applications to perform the backup.

Your mission-critical applications require the highest degree of protection and availability. For these systems, you should look for a solution that limits the amount of data you might lose, as well as the amount of time it takes to recover when something goes wrong.

Risks

There are many risks that can threaten your remote data (see Figure 3). First, there is the simple and all-too-often case of someone losing a file or e-mail, usually through human error; it may be through an accidental deletion or a deliberate act by a disgruntled employee.

Figure 3: Remote data risks



Another common risk is a malicious attack by a virus, worm or hacker. This can result in immediate loss of data, or the gradual and difficult-to-detect corruption of your data. To recover from this type of threat, you need to be able to roll back your systems to a point in time before the attack started.

Disk and server crashes pose another threat, requiring specialized tools, talent and time in order to rebuild a disk or a server quickly and avoid business losses.

Local disasters include fire, flood and other events that can make the office and its systems unusable for some period of time. If the data in the office is destroyed, you would need an off-site copy of the data—and the ability to restore it—in order to resume operations.

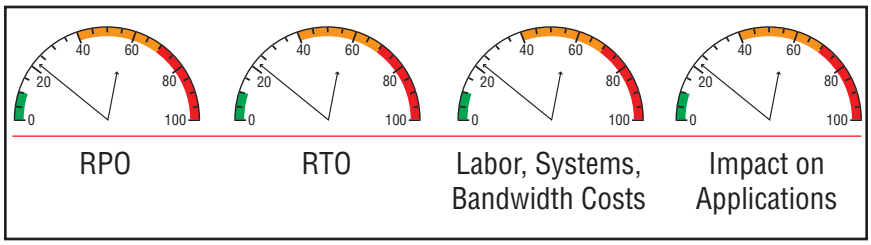
Finally, there are regional outages such as power blackouts and hurricanes. To recover from these types of disasters, you need backup facilities, systems, data and, in some cases, additional personnel in order to get back in business.

Now consider each of these data loss threats in the context of each of your prioritized applications. What is your planned response for each threat, based on each application or data type?

Recovery considerations

There are a number of variables involved in developing a data recovery strategy (see Figure 4). The first variable is the **recovery point objective** (RPO). RPO answers the question, "How much data am I willing to lose?" For example, running a tape backup every night would result in an RPO of 24 hours—you could lose all the data that was created or edited since the last backup was run, assuming the last backup was successful. As with failure types, you should apply your application priorities to RPO; maybe you can easily afford to lose 24 hours of typical office or e-mail data, while any loss at all from your business critical systems may be very painful.

Figure 4: Recovery considerations



Your protection and recovery systems should be tuned to meet the needs of each application.

The second consideration is **recovery time objective** (RTO). To define your RTO requirements, ask yourself, "How long can I afford to be without access to this application or data?" Again, this can vary depending on the critical nature of each application to your operations.

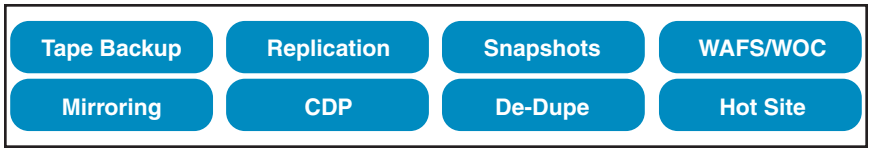
Costs are always a major factor in identifying data protection requirements. In the case of remote offices, you should consider more than just hardware and software costs. If you don't have experienced IT people in your remote offices, you should understand the impact of your data protection decisions on the central IT staff or local service contractors. And since you have to have a copy of at least some of the data off-site for disaster recovery, you should assess the security risks of physically moving the data, or the impact it could have on your wide area network (WAN) or Internet bandwidth to replicate the data electronically.

Lastly, you should assess the **impact** that the protection strategy could have on each of your applications. If you are running a standard tape backup, for example, you may have to place each application off-line for the required "backup window" each time you run the backup, remembering that the frequency of backups is defined by the RPO. For 24x7 applications, downtime may or may not be acceptable.

Backup and recovery solutions and technologies

You can see from the range of recovery considerations that it may not be feasible to apply a one-size-fits-all solution to protecting all of the applications and data in the remote office. There are a number of solutions and technologies available to provide remote office data protection, each offering its own approach and its own limitations (see Figure 5).

Figure 5: Solutions and technologies



Tape backup is a common data protection and recovery strategy, usually combined with off-site transport and storage of copies of the tapes. Tape backup has a number of limitations when used in remote offices, discussed in detail in the next section.

Mirroring captures all of the changes going to disk and copies them to another, usually identical, system. This method can be rather expensive, and real-time synchronous mirroring often doesn't work well over long distances, since the latency in the network can slow the application as it waits for each write operation to complete.

Asynchronous replication can be less expensive than mirroring. It may offer a superior solution over long distances, and allows for many-to-one consolidation of data from remote offices to a central site. However, both mirroring and replication can be ineffective solutions for individual data loss or corruption. Any loss would be replicated to the backup system, leaving you with two bad copies of your data. These solutions are best suited to system-level recovery.

Continuous data protection, or CDP, takes the benefits of replication and adds point-in-time recovery options. Each change is recorded and applied to the backup repository. This enables you to dial back to any previous point in time to recover from a data loss incident. However, most CDP, replication and mirroring solutions are susceptible to being unable to recover any files that are open at the time of the data loss incident, such as in the case of a system crash, although there are utilities available to help mitigate this issue.

Snapshots address the open-file issue by creating application-consistent, time-based recovery points. While it is necessary to pause or "quiesce" the application for a short period of time to accomplish a snapshot, it is typically far faster than a tape backup because it captures only incremental changes over a much shorter period of time. Snapshots can be run very frequently, often many times per hour, to meet aggressive recovery point objectives. However, hardware-based snapshot technologies are not always "application-aware," so they may not be able to fully recover open files consistently.

Data de-duplication can effectively reduce the amount of data that needs to be backed up or sent over the WAN. It works by eliminating any blocks of data that are redundant to a block already stored, and replacing them with a pointer to the stored block. While this approach can be very effective when protecting a large number of desktop-type systems, the amount of data redundancy in business application servers is usually much less. Also, you should consult your attorney or compliance officer before deploying data de-duplication technologies on applications that require data to be retained and retrieved in its original form.

Wide area file services (WAFS), combined with WAN optimization controllers (WOC), is a technology that allows organizations to pull servers, applications, and data out of remote offices, centralize them, and serve the data back to remote users over specialized WAN links. By itself, however, WAFS is not a data protection technology because it doesn't create copies of data.

Highlights

A **hot site** is the ultimate in data protection—everything in the remote office is replicated exactly in a standby site where business operations can continue following a major event. This can be a very expensive option, but may be required by some applications in some industries.

Depending on the decisions you've made regarding protection policies, one, several or many of these technologies may be needed to completely meet the data protection needs of your remote offices.

Backup tapes can present a number of challenges to remote offices, including time-intensive manual processes and potentially disruptive backup windows.

The limitations of tape backup in remote offices

Most companies rely on tape backups for data protection and recovery, and enterprise-class tape drives and media are very reliable products when used in automated and controlled environments such as corporate data centers. But in remote offices, the manual processes used in operating the tape backup system may not be as reliable. Tape backup usually involves a number of manual processes to label, load, unload, tension, ship off-site, reuse and erase tapes, in addition to running backup and restore operations. These processes are often performed by non-technical office personnel who have other responsibilities and who may have little or no training in backup and recovery procedures. In addition, tape backups can also require application downtime for backup windows—a requirement that may impact productivity in some offices, or require staff to be available to run (or re-run) backups after normal business hours and on weekends.

Trying to recover data in remote offices from tape backups can also be problematic and often requires the help of central IT staff or outside contractors. If you use incremental backups to reduce the time it takes to run the backup job, you'll have to restore the last full backup and then each sequential incremental tape, all of which can increase your recovery time. In many cases, only the most essential data can be recovered. Successful recovery from tape backups also assumes that all backup operations were completed successfully, with no tape errors, labeling errors or tape loss.

Highlights

The IBM Tivoli Storage Manager FastBack family of products offers an extensive and cost-effective data protection and recovery solution for remote offices.

Most tape backup failures are due to tape errors, which can be caused by mishandling, excessive heat and humidity, and failing to periodically clean the write heads in the tape drive. The second leading cause of tape backup failures in remote offices is human error, such as forgetting to run the backup job, not changing the tape, and mislabeling.

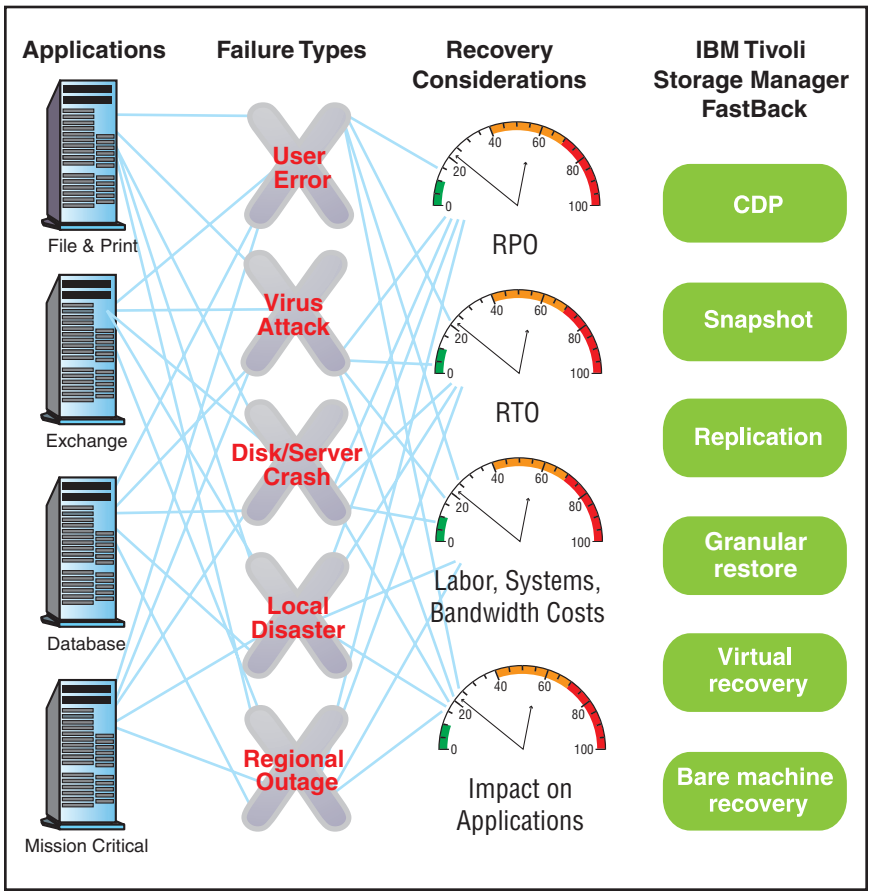
Tape is a point-in-time technology, so it can work well for recovering from a variety of data losses. However, the time it takes to perform a recovery makes it a poor choice for remote offices. And given the infrequency of backup jobs, tape probably doesn't address your more mission-critical protection and recovery needs.

Data protection and recovery solutions from IBM

Through the IBM Tivoli Storage Manager FastBack family of products, IBM offers an extensive and cost-effective data protection and recovery solution specifically designed to help remote offices maintain operations, regardless of the type of data loss (see Figure 6). This solution provides:

- Adaptable data protection policies for remote applications.
- Fast block-level data acquisition with minimal impact on application performance.
- Virtual recovery capabilities, providing near-instant access to lost data.
- End-to-end automation to help eliminate manual protection processes.
- Selective replication for off-site data protection with minimal impact on WANs.
- Central management and control of remote office protection.

Figure 6: Data protection and recovery with IBM Tivoli Storage Manager FastBack



Highlights

Tivoli Storage Manager FastBack combines the best features of CDP, replication and snapshot technology.

Tivoli Storage Manager FastBack combines the best features of CDP, replication and snapshot technology to provide continuous but non-disruptive protection that enables point-in-time recovery. The solution can be extended through optional components that provide advanced recovery of Microsoft Exchange data objects and bare machine recovery capabilities. The Tivoli Storage Manager FastBack solution can be used to help organizations meet a broad range of remote data protection and recovery requirements (see Table 1).

Table 1: Addressing remote data protection problems

Problem	Consequence	Tivoli Storage Manager FastBack solutions
Failed or improper backup	Risk of data losses until next scheduled backup	Automated, continuous, non-disruptive data protection
Accidental or malicious deletion of files	Painstaking reloading of tapes; risk of overwriting newer data	Simple drag-and-drop process to restore files
Corruption of critical database	Days of downtime until an expert can correct the problem	Virtual restore capabilities for on-demand point-in-time access
Loss of important e-mail or contact information	Permanent loss of the data	Simple drag-and-drop process to recover any data object
SQL Server disk crash	Hours or days of downtime	Virtual restore capabilities for near-instant recovery
Flood, fire or other disaster	Business losses until restart of operations in a new facility	Recovery of workload on another server within an hour

Highlights

Tivoli Storage Manager FastBack makes it easier for organizations with remote offices to add remote data to their core storage-related processes.

Integrating remote and central operations

The ability to recover data from remote offices is becoming increasingly crucial as the number of remote offices grows and the amount of data in them increases. Because Tivoli Storage Manager FastBack can be easily integrated with IBM Tivoli Storage Manager, it makes it easier for organizations with remote offices to add remote data to their core storage-related processes such as archive, retention and compliance. The software's adaptive disaster recovery capabilities also make it easier to send copies of remote data to a central site or to a disaster recovery site while helping to minimize the risk of data loss during transport.

Summary

Remote offices have unique backup and recovery requirements in order to support a wide range of applications and data and to protect against a wide range of risk factors. While there are a variety of technologies available to help organizations protect remote data, Tivoli Storage Manager FastBack offers an extensive data protection and recovery solution for these offices.

For more information

To learn more about IBM Tivoli Storage Manager FastBack, contact your IBM represen-tative or IBM Business Partner, or visit **ibm.com**/tivoli

About Tivoli software from IBM

Tivoli software offers a service management platform for organizations to deliver quality service by providing visibility, control and automation—visibility to see and understand the workings of their business; control to effectively manage their business, help minimize risk and protect their brand; and automation to help optimize their business, reduce the cost of operations and deliver new services more rapidly. Unlike IT-centric service management, Tivoli software delivers a common foundation for managing, integrating and aligning both business and technology requirements. Tivoli software is designed to quickly address an organization's most pressing service management needs and help proactively respond to changing business demands. The Tivoli portfolio is backed by world-class IBM Services, IBM Support and an active ecosystem of IBM Business Partners. Tivoli clients and Business Partners can also leverage each other's best practices by participating in independently run IBM Tivoli User Groups around the world—visit www.tivoli-ug.org

IBM Corporation Software Group
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
July 2008

1 Aberdeen Group. “Remote Data Backup: A Fundamental Practice for Most Companies,” October 2006.

2 Aberdeen Group. “Remote Office Backup: Preparing for the Unavoidable,” September 2007.

TIW14010-USEN-00