LEARNING MADE EASY

Acronis 15th Anniversary Edition

Backup

for dummies®
A Wiley Brand

Answer common questions about backup

Obtain ten tips for easy backup and recovery

Address data protection concerns

Brought to you by

Acronis

Joel Berman
with Peter Hale, Rebecca Schimmoeller, and James Slaby

About Acronis

Acronis sets the standard for cyber protection and hybrid cloud storage through its innovative backup, anti-ransomware, disaster recovery, storage, and enterprise file sync and share solutions. Powered by the Acronis AnyData Engine and strengthened by its artificial intelligence-based ransomware defense and blockchain-based data certification, Acronis solutions deliver easy, reliable, efficient, secure, and private cyber protection for physical, virtual, cloud, mobile workloads, and applications.

Founded in Singapore in 2003 and incorporated in Switzerland in 2008, Acronis is currently celebrating its 15-year anniversary as a global leader in cyber protection. Its innovative technology is driven by 100+ patents, developed by 500 engineers, and supported by more than 1,000 employees worldwide. Acronis' products are used in over 150 countries in more than 20 languages, bringing complete protection for more than 5 million consumers and 500,000 businesses.

Acronis solutions are available worldwide through a global network of service providers, distributors, and cloud resellers. To learn more, visit **www.acronis.com**.

Backup

Acronis 15th Anniversary Edition

by Joel Berman
with Peter Hale, Rebecca Schimmoeller, and James Slaby

for dummies®
A Wiley Brand

Backup For Dummies®, Acronis 15th Anniversary Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN: 978-1-119-54114-1 (pbk); ISBN: 978-1-119-54117-2 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Project Editor: Carrie A. Burchfield
Acquisitions Editor: Katie Mohr
Editorial Manager: Rev Mengle
Business Development Representative: Sue Blessing
Production Editor: Magesh Elangovan

Table of Contents

INTRODUCTION 1
About This Book 1
Foolish Assumptions 1
Icons Used in This Book 2
Where to Go from Here 2

CHAPTER 1: **Data Protection 101** 3
Defining Data 3
Protecting Data 4
Making Backups Easy, Complete, and Safe 6
Working with ease 6
Understanding completeness 6
Relying on safety 7

CHAPTER 2: **Capturing Data for Backup** 9
Understanding Backup Types 9
File backups 10
Image backups 11
Backing Up According to Plan 11
Choosing full, differential, or incremental backups 12
Setting the RPO 13
Taking a snapshot 14
Backing Up with and without Agents 15
Considering Backup Products 16
Bare-metal restore 17
Single-pass backup 17

CHAPTER 3: **Storing Backups Safely** 19
Creating a Backup Policy 19
Backup plan 20
Retention policy 20
Selecting Backup Software 22
Picking Backup Media 23
HDD and SSD 24
Tape 25
Cloud 25

Choosing Off-site Storage Locations 26
Online network 27
Dark site 27
Cloud backup 27
Considering Compression and Deduplication 28
Calculating the Costs 29

CHAPTER 4: **Recovering Data** 31
Recognizing Data Loss 32
Putting Your Recovery Plan in Motion 33
A Word about Backup and Ransomware 35

CHAPTER 5: **Managing Backup** 37
Keeping Current on Backup Products and Services 37
Setting the Backup Window 38
Creating and Checking a Backup Plan 39
Keeping it simple (or not) 39
Setting backup windows 40
Checking execution 40
Monitoring the plan 40

CHAPTER 6: **Ten Things to Know about Backup** 41
The Value of Your Data 41
The Cost of Downtime 42
Workload Priorities 42
Where Your Backups Are Stored 42
How Long to Keep Backups 43
What Recovery Tools to Use When 43
The Details of Your Backup Plan 43
What Data Is Excluded from Backup 44
How (and How Deeply) to Test Backups 44
How to Frame Backup Questions 44

Introduction

Many people think that backup is simply about copying data and then copying it back if needed. The process probably was that simple in 1960, when you could punch another set of computer cards and keep them in a safe place.

Technology has advanced rapidly in the past 50-plus years, of course, so backup processes have also had to advance. Backup now covers a variety of use cases — from restoring a single file to recovering an entire IT infrastructure.

About This Book

Backing up by simply copying a few files to another drive and then copying them back if needed satisfies only the most trivial cases. This book covers a variety of situations from the simplest single server to large and complex systems.

Data protection (DP) is an umbrella term covering backup and recovery for every kind of information, including data, applications, and system software. The term is often confused with another meaning of data protection, which has to do with encrypting personal information such as credit card and healthcare data. This book is about backup and recovery. Along with DP you may see disaster recovery (DR). DR means recovering an entire system after some sort of physical damage such as a flood or a massive cyberattack with various viruses being injected.

This book is your guide to putting these protection systems in place.

Foolish Assumptions

In writing this book, the following things were assumed about you, the reader:

- You're familiar with information technology (IT), but you're not a data protection expert.

- You're somewhat experienced in administering systems but aren't a skilled sysadmin.
- You want to begin or improve your company's backup processes and are looking for some help on the foundational concepts, as well as guidance on choosing backup products and setting up a data protection solution.

Icons Used in This Book

As in all *For Dummies* books, icons in the margins point out certain types of information.

Text marked with the Tip icon provides helpful hints on backup concepts or techniques.

The Remember icon flags important facts that you should keep in mind.

You don't have to read Technical Stuff text, but I hope you will, because it gives you a deeper understanding of backup.

Don't skip anything that has a Warning icon. Failing to heed the warning could cost you time, money, and/or data.

Where to Go from Here

Like all *For Dummies* books, this one can be read in whatever order is most helpful to you. Start with Chapter 1 and go straight through or skip around. It's up to you.

IN THIS CHAPTER

- Understanding why data needs protection
- Knowing what data to protect
- Seeing what goes into a backup system

Chapter 1
Data Protection 101

Companies have protected their data ever since carbon copies were stored in bonded warehouses, but times have changed. The techniques for protecting information and recovering from disastrous data loss have also changed. This chapter brings you up to speed on modern data protection.

REMEMBER

Throughout this book, I use the term *data protection* to mean storing data in such a way that it's easy to search and restore, no matter where it's located in the backup archives. I use the term *disaster recovery* to mean recovering a server, workstation, or entire data center quickly after a major problem.

Defining Data

What is data? The question may seem simple, but answering it isn't. Data may mean simple text files, or it may mean a vast range of types of complex information. In a modern computer system — and for the purposes of this book — data means programs, files, and metadata.

Because data takes many forms, it's easy to get confused about which data needs to be protected so it can be reproduced easily. In general, however, items that have value should be protected. Just

as cavemen were careful to keep their children away from saber-toothed tigers, you're probably careful about where you leave an expensive watch or park your car. Why do you protect your things? Because they have value to you. If these items are lost or stolen, you lose the value associated with them. Even insurance can't restore covered items; it can only compensate you for its financial expense.

Data is a bit different. If you take proper care of it, you should be able to recover it — and its value — with relative ease.

Protecting Data

Information technology (IT) experts protect data against loss and destruction, such as through theft, accidental deletion, or deliberate alteration. Fortunately, protecting data is easier than you may think. If you plan carefully and execute your plans well, it's possible to protect your data to any level you desire.

Consider protecting the following items:

- **Bootstrap data:** *Bootstrap data* is used to start a machine. It's the program that runs first when a system is powered on or restarted. Without a valid bootstrap, a system never becomes operational.
- **File-structure metadata:** *File-structure metadata* describes where all the files and folders are located, as well as where the bootstrap data, operating system, and drivers are. File-structure metadata records which blocks on the disk are being used and which are free, and it maps every directory and filename to specific locations on the disk drive. This type of data also contains permission and access lists that prevent unauthorized read or write operations by checking against those lists.

  
  TECHNICAL STUFF

  In some systems, special logs keep historical records as an audit trail of changes. These logs are used to recover from power failures and other abrupt halts. All this metadata is important when you need to restore an entire system.
- **Driver binaries:** *Driver binaries* control devices that read from discs, tape, or the network. Drivers are often part of the operating system (OS; see the next bullet), and they must be

compatible with the rest of the system. They often come with devices that are purchased but not included with the OS.

- **Operating system (OS):** OS code is often shipped on a physical disc or is already installed on new hardware. The manufacturer often issues updates over the network, so the initial disc becomes obsolete.

  
  WARNING

  Make backups of the OS after updates are applied so you always have a current version for recovery purposes. If you don't save the updated version, you'll have to reapply all updates when you restore the system — an error-prone, time-consuming process.

- **Configuration files:** Configuration files are numerous. Some are as simple as a file containing the system's name or time zone; others are as complex as the Windows Registry, a large file containing thousands of pieces of information. Password files are configuration files, for example. Some software is so secure that if the password file is lost, no data can be recovered. Also, many applications have complex configuration files that store all sorts of information that's specific to the system.
- **Application programs:** Companies buy many of the applications they use from third parties, but some companies develop their own applications. One of the most tragic cases of data loss occurs when a program is developed in-house and then the original source code is lost. If a company loses these program files, it may be unable to use any of the saved data. Having the data, but not the programs that use that data, is tantamount to losing the data itself. Businesses must protect applications with the same care that they protect the application data itself.
- **Data files:** Data files are highly portable files used by many programs, such as spreadsheets and PDFs. Other data files are unique to specific programs, such as log and database files. These files may be short, permanent, large, or rapidly changing, and any company's server could contain hundreds of thousands or even millions of them.
- **Databases:** Databases could be considered another application with data files, but don't assume that any backup process also protects them. There are special considerations for protecting databases to ensure consistency.

Making Backups Easy, Complete, and Safe

The mantra for backup is "easy, complete, and safe." In this section, I break down these components.

Working with ease

Ease of use is probably the most important factor because it's the thing that prevents mistakes.

Companies that have evolved their backup solutions over years may be using incompatible products and backup media, which can cause problems. And if they have no standard recovery checklists or written procedures, they're almost always asking for trouble. More times than I can think of, people have recovered data from the wrong backups or even overwritten good data with bad data because they were trying to follow a complex, poorly documented process under pressure and in a hurry.

REMEMBER

Make sure to consider usability in choosing an easy, complete, and safe backup product. Arguably more important than clear checklists and plans are intuitive interfaces, with minimal setup and single-click recovery.

Understanding completeness

Completeness means having all the data necessary to recover the system itself after any failure as well as the proper tools and processes. Your backup software may help you create a bootable CD or DVD, which you can use to recover when you have no running system at all. If your system doesn't allow you to create such a disc, however, you have to reinstall the system — a lengthy process that requires configuration and OS updates (both discussed in the section "Protecting Data" earlier in this chapter).

Likewise, your backup program should be able to reproduce disk partitions and formats instead of expecting you to do the work manually; this task is tricky and must be done perfectly. Finally, your backup software should print recovery instructions for you to follow, just as an airline captain follows a checklist to land an

aircraft. Remember, if your system is completely down, you won't be able to log on to the Internet to download instructions. This can be easily summed up as "completeness means any data, any location, any environment."

Relying on safety

Safety has two components:

- **Reliability:** Backup data must be captured reliably in order to be safe. One missing piece of data can make the entire system unrecoverable. That means the data must be readable, and if the backup depends on any other data, that data must also be readable. Backup software may use a variety of techniques to capture data safely, so each element should be coordinated so the data can be read and the system can be recovered.
- **Security:** After data is captured, it needs to be protected against changes and theft. If your backup system is not secure, a bad actor could sneak into it, steal data, and then do something to harm your main system with no one being the wiser.

REMEMBER

The only cure for data loss is prevention. Keep your backup process simple and well documented. Start at the beginning by capturing the data properly (see Chapter 2). If you capture it incorrectly, all hope is lost.

Because of these rapid advances, including virtual machines in your backup and recovery plans is important. Data protection/disaster recovery (DP/DR) software that works across physical, virtual, and popular hypervisors is essential. This allows backup archives to be preserved and useful no matter where the workloads reside and also provides added value with the capability of migrating operating systems, applications, and data between physical, virtual, and cloud environments using any hypervisor. Choosing a backup product that works with any data, any environment, any location, and any device gives you the greatest flexibility in taking advantage of the full range of computing offered today and tomorrow.

ANY DATA, ANY ENVIRONMENT, ANY LOCATION, ANY DEVICE

While virtualization has been around for over 30 years, it has exploded over the last 15 years as VMware introduced a software hypervisor. Since then, a number of hypervisors have appeared, and hardware manufacturers have migrated many software functions into hardware for performance reasons. Additionally the emergence of multi-core processors and increasingly dense memory now allows more and more workloads previously only possible on mainframes and minicomputers. The result is that workloads are migrating from physical to virtual and back, and a number of hypervisor and cloud technologies are available.

IN THIS CHAPTER

- Understanding image and file backups
- Setting a backup plan
- Using agents (or not)
- Selecting backup products

Chapter 2
Capturing Data for Backup

For most of us, backup means copying just a few critical files to a USB stick or separate disc. But large systems and servers house many files, and the data in those files changes constantly. A systematic means of capturing system data, applications, and metadata even while files are open and being changed is mandatory if you want to recover from lost or damaged data and applications. That process requires backup software.

In this chapter, I walk you through the main types of backups, discuss the scope of backups, and introduce a couple of special backup situations.

Understanding Backup Types

Excluding tape, most storage devices look like disk drives to the low-level software. CD/DVD devices, for example, have special burning requirements but still appear as disk drives — a sequence of blocks. The hardware presents a sequence of blocks to the operating system (OS), each block being separately readable and writeable. Within some of these blocks is *metadata,* which

holds directories, lists of used and free blocks, bootstraps, partition information, bad-block lists, and remapping information, as well as those blocks used for holding data.

TECHNICAL STUFF

By the way, the storage space used by the metadata takes away from the free space on a freshly formatted disk. That's why the free space you see on a disk is less than the raw capacity specified by the vendor.

REMEMBER

Generally, you have two major ways to capture data for backup: image backups and file backups. Although both methods allow you to search for specific files and data, the image backup is a superset of a file backup because it also contains system metadata.

File backups

The original type of backup was a file backup, which is still a popular method. A *file backup* copies all the files and folders from the current data to backup media. The process is similar to copying personal files to a USB stick, a USB drive, or another directory.

Because file systems keep track of when a file is created and modified, the file backup may only copy the files and folders that have changed since the last backup. It's simple for a backup program to copy files, because the OS provides all the necessary functions to look up and copy files. Copying, however, creates a lot of overhead for the backup system, because for every file, the system has to do the following:

1. **Find the blocks where the folders are located.**
2. **Read the folders.**
3. **Look for filenames.**
4. **Determine where those files are located.**
5. **Read and copy those blocks.**

This process can take a great deal of time. Also, if the system is operational at the time, there can be a lot of contention for resources, which increases backup time and/or reduces performance.

Image backups

An *image backup* bypasses most of the OS file-lookup overhead by simply copying blocks to back up in order from beginning to end, making a complete record of everything on the disk. Image-backup software is designed by specialists, so it's programmed to determine which blocks have been changed since the last backup and copy only those blocks. If a 2GB file has only a small change, for example, the file backup has to copy the entire 2GB, whereas the image backup copies only the changed block. This process results in extremely fast backups.

REMEMBER

The fastest backup software captures data only from blocks that are in use. It won't copy bad, temporary, unchanged, or unused blocks unless specifically requested to do so.

Image backups allow you to inspect the image so you can determine where individual files are and do fine-grained file recovery. Often, image backups can be mounted as full disk drives so an administrator can recover or compare data from different periods.

Both image backups and file backups can be full, differential, or incremental, and both types can exclude specified folders and file types. I cover backup types in the later section "Backing Up According to Plan."

TECHNICAL STUFF

Images used to be called *snapshots*, but that term has two different backup-related meanings, so it's better to say *image backup* than *snapshot backup*. I cover snapshots in "Taking a snapshot" later in this chapter. Also, both file and image backups should be able to exclude the copying of specific files so that temporary and other unnecessary files aren't copied.

Backing Up According to Plan

A backup plan describes the data to be backed up and the scope of the backup. In this section, I cover the basic decisions you need to make.

Choosing full, differential, or incremental backups

Backups are divided into three types:

- **Full:** The first backup of a system, capturing everything in it

  The upside of a full backup is that it's self-contained. The downside is that it takes up a lot of space, can take a long time to complete, and can be almost identical to a previous full backup.

- **Differential:** A backup that captures only the differences between the current state and the last full backup

  Recovering from a differential backup requires both the last full backup and the differential backup to be valid. The upside of a differential backup is that it's much faster than a full backup. The downside is that it takes up more space than an incremental backup (see the next bullet), and it requires at least two backup files to be read for recovery.

- **Incremental:** A backup that captures only the differences between the current state and the last differential, incremental, or full backup

  The upside of an incremental backup is that it's very small and very fast. The downside is that recovering with an incremental backup is time-consuming and requires all the data from the last full backup and every successive incremental backup until the recovery-point objective (see the section "Setting the RPO") to be valid and read.

  
  TECHNICAL STUFF

  Most backup software allows incremental backups to be consolidated offline, which greatly improves reliability and recovery time. Another backup type called *reversed incremental* automatically does the consolidation as the backup is captured. Yet another type, called *always incremental,* has different meanings depending on the vendor.


TIP

For most companies, the size of a daily backup — and thus the size of an incremental backup file — is 3 to 5 percent of a full backup. This figure varies widely, however, depending on the business. Also, right after upgrades or major changes, an incremental backup can be quite large. If you're planning a major

upgrade of systems, applications, or data, you should perform a full backup before the change and then another right afterward. If you intend to tinker with the system after the changes, you should make frequent backups.

Setting the RPO

The *recovery-point objective* (RPO) defines how often backups must be made. The amount of time the system is allowed to be paused for backup is the *backup window.* The goal is to eliminate the backup window by using techniques that allow a running system with changing data to be backed up without having to disable and pause the system. Generally, the backup window and the RPO conflict because frequent recovery points are desired, which requires frequent pauses or slowdowns during the backup.

The RPO determines how often recovery points must be created. If the RPO is 30 minutes, a recovery point must be established every 30 minutes. The backup window defines how much time is available for the backup process. For a 30-minute RPO, for example, the backup window should be much less than 30 minutes. Likewise, if the backups can be done only between midnight and 2 a.m., it's impossible to have an RPO shorter than a day because if the backup completes at 2 a.m., the most recent recovery point would be the previous day at 2 a.m.


REMEMBER

Many methods are available to shorten the backup window and allow more frequent recovery points. But they always involve more computing resources. The system must have enough power to run operations and complete a backup during the window. Good backup software allows you to limit the amount of resources consumed during a backup. This protects the running system response time, but adds to the total backup time.

Applying facts and judgment to the RPO

Determining the RPO is a business question that requires facts and judgment. The *facts* are the cost of the downtime experienced, the cost of lost work, and the cost of providing that RPO. The *judgment* is weighing the risks against the intangible losses, such as reputational loss. If your advertising and marketing campaigns position your company as being a low-cost provider, for example, you may be tempted to skimp on backup. But if you advertise as

never going down and never losing data, even a few minutes of downtime — or the loss of even one document or transaction — could damage your business.

If you want to restore to the exact recovery point time, you would have to retain an infinite number of backup files, so you may decide on a sliding RPO. For mission critical workloads, you may set the RPO as 5 minutes over the previous 24 hours, then every hour for the previous few days, daily for the previous month, and then monthly forever. For workloads that are less important or that don't change often, you may set the RPO daily for the first week and monthly thereafter. ***Remember:*** This is about deleting older recovery points to make space for new ones.

Weighing costs, benefits, and risks

Weighing the costs against the benefits, tempered by sound judgment of the risks, is difficult, but the process helps you understand what your RPO needs to be. It's easy to say that you can't afford to ever lose any data, but ensuring that you don't lose any data can be quite expensive. Backup doesn't provide fault-tolerant or zero-downtime operation, which requires second sites, redundant equipment, and specialized system design.

The time it takes to complete a backup job must fit into the backup window, which is also set by business requirements concerning the amount of planned downtime allowed. You have several ways to minimize the backup window without compromising the RPO, however. The easiest way to reduce the backup window is to use a snapshot (see the section "Taking a snapshot"). But that process carries some risk.

Taking a snapshot

One method to speed up a backup is to minimize the data copied. The system is paused for an instant to take a *snapshot* or copy the metadata. This process takes a fraction of the time that copying the data would take. Then the backup is performed by using the metadata to locate the files. If changes occur to the data during normal operations, the original metadata is updated, but the snapshot copy is not. So, the backup system won't back up any data added after the snapshot is taken. The snapshot has pointers to most of the data and only contains the actual data

when that data has changed. The alternative would be to pause the system while all the data (full backup) or the data that has changed since the last backup (incremental backup) is recorded. This alternative is safer but takes much longer, and it disrupts operations by pausing the running applications until the backup is complete.

Snapshots shorten the backup window considerably, and they're especially useful when you're doing a lot of updating because it's easy to revert the system to a snapshot. Also, snapshots are important in Storage Area Network (SAN) management because SANs are a widely shared resource that will disrupt much of the system if it's paused for more than a few seconds. Snapshots are safe only for short-term use, however, and managing snapshot deletion can take considerable resources.

REMEMBER

A snapshot isn't a complete copy of the data. If the original disk is damaged, the snapshot is also damaged. Therefore, snapshots may be safe in the short term but are no replacements for backups.

TECHNICAL STUFF

Sometimes, an application can help the snapshot process shorten the backup window. VMware uses one such type, called Changed Block Tracking (CBT), to enable backup software to reduce the time to capture. Another type is Microsoft Volume Shadow Copy Service (VSS). Both the applications and the backup software must work with these technologies; otherwise, a complete backup can't be made.

Backing Up with and without Agents

Backup programs may access data on your system in two ways:

- **With an agent:** A small backup program called an *agent* is installed on every physical and virtual machine (VM).
- **Without an agent:** In cloud and virtual environments, the number of VMs can grow quite large, so *agentless backup* comes into play.

Agentless backup still uses agents, but only a small number of them, so the process is easier to manage. Typically, one agent is installed for each virtual host, usually running in a VM itself. This agent can communicate with the host and can back up every VM on that host. Most systems have multiple hosts, and VMs can migrate among hosts, so the backup and capture systems must be aware of where each VM is at all times.

Agentless is a good way to operate. In special cases, when the host/hypervisor is unable to back up all the objects connected to the VM, you should install an agent to back up that machine directly. In most cases, however, the agent in the host takes care of everything.


REMEMBER

You need to ensure that you install the right number of agents, update them regularly, and maintain a license database if you're required to do so.

Considering Backup Products

Image backups used to be done by one application while file backups were performed by another. Today, good backup software is capable of doing both types of backups. High-quality backup products can perform image backups in full, differential, and incremental styles (discussed earlier in this chapter), as well as use snapshots to reduce backup windows (also discussed earlier in this chapter). In addition, they can capture data at the block level and then recover files from that image.


TIP

For some network drives, as well as for some SAN and Network Attached Storage (NAS) systems, it's impossible to perform an image backup because the backup agents don't have programmable access to the metadata. Generally, though, create image backups if possible and create file-level backups only when you have a good reason not to perform an image backup.

A full-featured backup application can recover the entire image to a new system. It can even adjust to different-size disks and possibly inject drivers, as well as change bootstraps to allow recovery to a different model of hardware with different device

controllers — even different CPU types and different storage and memory configurations on physical machines and VMs.

REMEMBER

The fastest backup software captures data only from blocks that are in use; it doesn't copy bad or unused blocks.

Bare-metal restore

Bare-metal restore is the recovery from backup to a system with absolutely no software installed in it. Because file backups, unlike image backups, don't have all the system metadata and bootstrapping, they can't restore to bare metal. One major benefit of an image backup is that it provides both the capability to recover files to a running system and a complete restore to bare metal even when the backup wasn't made on an identical system. File backup can only recover files to a running system, but image backup is able to both recover individual files to a running system and recover everything to bare metal. Additionally, the backup software should be able to convert a physical image to a virtual image, and vice versa, and the virtual image should be exportable to any common virtualization system. Ask your vendor whether it has a universal backup format that can be recovered to physical machines and VMs.

Single-pass backup

Single-pass backup means that only one pass through the data is required to capture and store the backup, and only one pass is required to recover the data. Single-pass backups are faster than multi-pass backups and therefore afford more frequent recovery points and a shorter backup window.

If image and application backups are combined in the same product, all the data required for complete recovery can be captured in a single pass. If, however, you have an image-backup product, a file-backup product, and an application-backup product, even if they're all single-pass products, you must run three passes. Data is stored in separate archives for each product and managed separately, creating additional complexity and opportunity to fail at recovery time.

TIP

DATA-CAPTURE REQUIREMENTS

Here are some things to think about when deciding your data-capture requirements:

- RPO by subsystem and application
- Backup window, or amount of downtime you're willing to accept for backup
- Type of backup (image, file, or both)
- Which applications are covered
- How many backup processes can be administered safely

If your image, physical, virtual, cloud, database, email, and user backups require different programs and management processes, the resulting backup files will be incompatible, and your data may be in danger because of this multiplicity and complexity. The rapid pace of innovation in computing means that significant value can be gained from adopting new architectures or new types of hardware. Often though, the difficulty, time, and cost of migrating all your data and applications to the new systems holds you back. A backup system that can capture systems, programs, and data from physical and virtual machines and any hypervisor means you have flexibility to integrate approaches. Choosing a backup system that has a long life and works with any data and any style of computing should always be a priority.

IN THIS CHAPTER

- Setting a storage policy
- Choosing software, hardware, and sites
- Understanding compression and deduplication
- Estimating storage costs

Chapter 3
Storing Backups Safely

The age-old maxim for backup has been the *3-2-1 rule:* Keep three copies of your data on two media types, with one copy stored at a separate location.

Traditional media types included tape and disk, but cloud has become a critical part of the media mix in businesses. Data size and the cost of networking make cloud storage economical because it provides both a distinct media type and a separate location. This isn't to say that tape is going away — only that cloud is being used more and more frequently as a primary backup medium as well as an alternative to tape. If you're starting off with small amounts of data, you don't need to make large capital investments to use cloud backup.

This chapter discusses the safe storage of your backed-up data.

Creating a Backup Policy

Your business may already have a backup policy, based on existing systems and practices. But if you're considering adding new technologies, you may need to update your old practices. This section gives you some pointers on setting an effective backup policy and plan.

Backup plan

A *backup plan* is the record of the data that is to be backed up on what schedule. It can be as simple or complex as you like:

- A simple backup plan would be to perform a full image backup at midnight every day.
- A more complex backup plan would be to perform a full backup weekly, with differentials every night and incrementals every four hours. The backups on machines A and B could start at 10 p.m., with a random delay of up to two hours; the backups on machines C and D could start at midnight, with a two-hour random delay; and the backups of machines E and F could start at 2 a.m., also with a two-hour random delay. (See Chapter 2 for more on full, differential, and incremental backups.)
- An even more complex plan would back up the system image weekly, the Windows Exchange system continuously, the Microsoft SharePoint system nightly, user files every other day on a random schedule, the system configuration data weekly, the virtual hosts weekly, and the Active Directory data every eight hours.


TIP

It's important to have considerable flexibility in a backup plan and to consider backup products that let you set up the kind of plans you want to use. If you're creating full image backups, for example, you should use software that allows you to create virtual machines (VMs) from image backups. (See the "Going virtual" sidebar later in this chapter.)

Retention policy

A key part of every backup plan is its *retention policy.* Unless you have infinite space for backups, you eventually have to delete recovery points to make room for new ones. The retention policy determines which recovery points you delete next.

The most primitive retention policy is to monitor space and delete some of the older backups to make room when available space runs low. The tricky part is determining which ones to delete.

In the following sections, I discuss two common types of retention policies.

GFS

Imagine that every day, you make a backup. At the end of a week, you have seven backups and are running low on space, so you take one of the daily backups, rename it as a weekly backup, and start making daily backups again. At the end of the second week, you take the latest daily backup, rename it as a weekly backup, and continue making daily backups. You always have an entire week's worth of backups, plus weekly backups. You're still running out of space, however, so every four weeks, you rename the weekly backup as a monthly backup and reuse the weekly backups.

This type of policy is called *Grandfather-Father-Son (GFS).* The daily backups are the son, the weekly backups are the father, and the monthly backups are the grandfather.

Follow this policy long enough, however, and you eventually run out of space, so you must delete something. What do you delete? If you can't decide, consider the following policy.

TOH

One usefully sophisticated retention plan is called *Towers of Hanoi (TOH),* named after a child's game in which discs must be moved from one tower to another, but only one disc can be moved at a time, and no disc can be placed on a smaller one. The sequence used to solve this puzzle is a binary pattern. This sequence as applied to a backup retention plan allows you to reuse space and place backups on various media. If all your daily, weekly, and monthly backups reside on one disk drive, and both the main system and that drive fail, you lose everything. For that reason, you must rotate your use of storage media so that you're not putting all your eggs in one basket.

TOH isn't used as frequently as it should be because it's complex and hard to manage. (The details, in fact, are too complex to discuss in this small book.) Some backup products do automate this retention plan, however, and I recommend that you look for a product that does.

Selecting Backup Software

TIP

When you choose your backup software, look for the following high-level points:

- **Sufficient recovery features:** If you back up a SQL database, for example, you should be able to mount the backup file as a SQL database and start using it immediately. If you suspect that you backed up a disk image containing virus code, your software should allow you to mount that image as a disk and run an antivirus scan on it first.
- **No reinstallation requirements:** Vendors of backup products that can't back up and recover VMware, Windows, or Linux virtualization hosts may tell you that it's very easy to install the hypervisors from scratch. Well, the process is easy if you have all the necessary software and skills, good guidelines, and all the configuration parameters you need. Easier still is performing image backups of the virtual host so you don't need to reinstall anything; just load and go.
- **Compatibility with existing hardware systems:** Servers used to be quite finicky about operating-system (OS) configuration; if the configuration was wrong, they wouldn't start. Standardization is more common today, and modern OSes do make some adjustments for hardware differences.

  If you're buying hardware from different vendors and/or of different generations, you should ensure that your recovery software allows you to restore to dissimilar hardware.

WARNING

- **Compatibility with your virtualization systems:** If you're performing image backups or merely backing up the virtual disks, there's no reason why you can't recover to either virtual or physical machines.

  Look for backup software that, after completing a backup operation, can export a VM file and insert it into the virtual management program of your hypervisor. That way, if you need to restore to an earlier recovery point, the VM will already be there, ready to start.

TIP

Picking Backup Media

To choose which backup media you need, first you need to determine how much storage you need for backup. Here are some guidelines to get you going. The amount of storage needed is a function of how many copies you want to keep, how far back you need to maintain copies (this could be mandated by regulations in some industries), and how rapidly your data grows and changes. To start off, try three to five times your current data size, watch it carefully over a few months, and begin to project future needs.

If you want to be more precise, measure your own data to see if compression and deduplication can help. You will have to watch how much new data is backed up per day for a few weeks, and consider how many backups and how old they need to be.

After you figure out how much storage you need, decide which backup media to use. Your choices are hard disk drive (HDD), solid-state drive (SSD), tape, and cloud, each of which has pros and cons:

- HDD is relatively fast but relatively expensive.
- SSD is even faster than HDD, but the most expensive of commonly used storage media.
- Tape is also fast but more complex than disk to track and manage.

  
  WARNING

  Tape is arguably less reliable than disk because it can be damaged easily when handled. Tape, however, is by far the lowest-cost option, especially when larger amounts of data are involved.
- Cloud is great for remote end-points and for small servers.

  When backup is done locally and then staged to the cloud, it's an effective solution. The local backup is used for recovery of anything other than a major disaster, and in the event of a disaster, the backup is safe in a remote location. This dual-protection of local and cloud is consistent with the 3-2-1 rule and very economical.

  
  TIP

  Make sure your cloud provider allows initial seeding where you can send hard drives to the cloud location to be copied and the provider also can send drives to you for a large scale restore.

Which type of storage media do you choose? Often, the answer is "all of the above." All four media types, which I discuss in the following sections, are viable in terms of cost and space.

HDD and SSD

HDD and SSD have many advantages as backup devices:

- They're reliable.
- They're non-volatile, so they retain data even when the power is off.
- They're fast, so they give you the shortest recovery time of all four media types.

The largest-capacity HDDs as of early 2018 held 16 terabytes (TB; see the nearby sidebar) with manufacturers expecting to make 40TB HDDs by 2022. It's best to purchase drives that deliver the best cost per gigabyte. See the "Retention policy" section earlier in this chapter for tips on determining how much disk media you need.


TIP

If you're going to be transporting disk drives or tape cartridges — say, to off-site storage — it's best to make two copies of each, in case one is lost or damaged during transport.

HOW MUCH IS 16TB?

Probably the biggest change in data has been driven by the falling cost of storage. In 1980, a single gigabyte (GB) of disk storage cost $200,000, whereas today, 1GB of disk storage costs less than 4 cents. As a result, all the computing power used during World War II could fit into the sound chip in a musical birthday card, and the total computing power used for an Apollo moon shot would easily fit into a smartphone.

So how much data can 2018's largest HDD — 16TB — really hold? A gigabyte easily holds 100,000 emails, and 16TB = 16,000 GB, so 16TB = 16,000,000,000 emails.

WARNING

SSDs have also emerged as backup devices because they're even faster than disk drives and much more durable. SSDs holding 100TB were available in early 2018 and also expected to grow steadily in size. These drives, however, have some life-cycle issues that the backup software needs to manage. These issues are related to the way that data is written to the devices. Ask backup vendors whether their products take special care when writing to SSD. If not, the reliability of those products may be less than optimal.

Tape

Every year, conventional wisdom claims that tape is on the way out, yet tape vendors manage to keep tapes reliable and cost per gigabyte low by increasing capacity. The current tape version, LTO-8 (Linear Tape Open), stores 12TB per tape. Because you can keep 560 slots of mounted tape in one standard 19-inch rack, 6.72PB of online tape per rack is more capacity than all but the largest data centers need.

TIP

If you're storing tapes off site, you should make a copy of each tape you transport, due to the risk of damaging tapes during transport.

Cloud

The greatest advantage of cloud storage is convenience. You don't have to concern yourself with transporting media and making multiple copies in case the disks or tapes are damaged in transport. But cloud services also have a few disadvantages:

- **Security:** Make sure that you know exactly how safe each cloud facility that you're considering is. Ask vendors the following questions:
  - Is the facility fireproof?
  - Does it have emergency power generators and redundant network attachment points?
  - Who performs the actual backups?
  - Is stored data encrypted?
  - Who has access to data within the storage center?
  - Is the facility staffed around the clock, or is it fully automated?

- **Price:** The pricing of cloud storage is tricky. In addition to charging you for the storage you use, many providers charge extra for network usage, operations costs, and deletion and retrieval transactions. As of early 2018, typical cloud storage service prices ranged from $.01 to $.08 per month per gigabyte, or about $100 to $800 per year per terabyte.

  Also, cloud backup may not entirely eliminate your need for staff, which may reduce your effective savings. You may still need to employ some people to administer your storage, plan capacity upgrades, set backup schedules, monitor backup completions, and track data use.

- **Network bandwidth:** You should determine the network bandwidth you need to meet your recovery-time objective (RTO; see "Recovery in the cloud" later in this chapter). Your daily backups will be smaller than a full-scale recovery.

Along with the seeming negatives, cloud and network have many positives:

- The cloud provider is going to have excellent network connectivity, so you can access your data from anywhere.
- Using the network is very convenient.
- You don't need to worry about storing and testing media; the cloud provider ensures reliability for you and often makes redundant copies (although some companies charge more for highly available cloud storage).
- The cloud provider allows you to increase and decrease capacity very quickly.

Choosing Off-site Storage Locations

The 3-2-1 plan (see the introduction of this chapter) advises you to keep three copies of your data: the running system, local backups, and off-site backups. You can physically transport disk or tapes to an off-site location (or use a vendor that picks up the media for transport), or you can transfer data wirelessly to be written on a networked storage device. In this section, I cover a few options.

REMEMBER

When you keep one set of backups with your running machines and a second set at an off-site location, consider how far from your site the off-site location needs to be. Most people want to be able to drive to the off-site location and back within eight hours, so that's a good maximum distance. If your area is at high risk for natural disasters — earthquakes, floods, tornadoes, or hurricanes, for example — you may want to keep your data even farther away.

Online network

If you're considering using an online network as your second location, ask providers about data security and network bandwidth. Also look at how much of your data is changing. The industry norm is 5 percent, but your volume could be higher or lower. Databases change frequently, whereas application code doesn't.

Speed is yet another consideration. Suppose that you have a small business with ten users at 5GB each and two servers at 10GB each, or a total 70GB per day and 1.4TB for the initial backup. Also suppose that your network speed is 100Mbps. Your daily incremental backup takes 90 minutes plus any network overhead and delays in routing.

TIP

Compression can make a big difference. If the data is compressed to half its size, only half the data has to travel on the wire, yielding a major improvement in speed. (I discuss compression in more detail in "Considering Compression and Deduplication" later in this chapter.)

Dark site

If you don't need or can't afford to have two running sites, the next option is a *dark site:* a remote computer room, usually not operational, with a minimal amount of equipment. Generally, all backups are sent over the network to storage devices running there. Periodically, a company's dark site is lit up and run to make sure that the disaster-recovery plan actually works.

Cloud backup

Cloud backup has gained great popularity among companies that want to have on-site data centers and are willing to use a cloud for temporary needs and disaster recovery. It has become easier to use the cloud as a second site, store backups in the cloud, and recover them to servers in that cloud.

Recovery in the cloud

If you decide to use a network/cloud solution, ensure that you have enough bandwidth available to meet your RTO goals. The first backup to the cloud can take quite a while because everything is being sent on the wire, so your cloud facility should let you ship initial seeding disks to reduce the time for the first backup. That way, if you ever need a full recovery, your cloud-storage vendor can ship your data back on disk drives.

RTO specifies how long you're willing to be down. If your recovery-point objective (RPO; see Chapter 2) is four hours and your RTO is two hours, you'll be up and running two hours after the failure, but you may have lost as much as four hours' worth of data, depending on the age of your last recovery point. If you want to be operational and up to date in two hours, your RPO and RTO together must add up to two hours. RTO generally applies to a catastrophic failure of a subsystem or the entire system. Your order-intake department may have an RTO of five minutes, for example, but payroll may have an RTO of two days.

Public or private cloud

One big decision is whether to use a public cloud facility. Many backup vendors offer cloud storage that's optimized for backups, and major public clouds also offer storage online. A big advantage is that capacity can be increased easily when needed.

The fee structure for public cloud services can be confusing. Some services charge a fee based on capacity per year; others charge for capacity per month, plus a fee for data transferred from storage to the Internet. Cloud companies may have other charges, such as data-deletion fees. Finally, be sure to check out the security features.

Considering Compression and Deduplication

Compression — the process of making files smaller by using various algorithms to substitute abbreviations for repeated information — is useful for local copies of backups, and most storage specs that you see assume compression. Compression

works when the data contains a lot of predictability. If your company name is Acronis, for example, and that name appears often in text, compression notices that fact and creates an abbreviation for the name.

WARNING

Some data, such as pictures and video, is already compressed, so it can't be further compressed by backup. Encrypted data by its very nature shouldn't be predictable, so you can't compress it either. In fact, if you *are* able to compress encrypted data, you should be very suspicious of your encryption system.

Deduplication (also called *dedupe*) is similar to compression. If you back up 1,000 system images of standard corporate laptops, for example, you find the same OS files over and over. High amounts of duplication require a lot of space that can be saved through the use of deduplication. Dedupe works by keeping one copy of the original data and inserting pointers into each set of backup data that contains the duplicate data. If deduped data contains thousands of characters and the pointer is only 20 characters long, the savings can be huge.

Deduplication works remarkably well when you have a lot of duplicate data. You can use multiple methods — one being deduplicating on the fly, which reduces the data that needs to be written to the backup archive. The other is a post process that deduplicates the archives after they've been created. You need to consider a number of performance and space tradeoffs so if you are counting on deduplication to save a lot of time and space, make sure you test it well and compare the various options. It may be better to just reduce the amount of duplicate data you're creating.

Calculating the Costs

Calculating an accurate cost for backup storage can be tricky, but a few general rules may help:

- At the raw storage level, the cost of HDD is $.03 or $.04 per 1GB, with SSD anywhere from 10 to 20 times as expensive, while tape costs under $.01 per 1GB.

- Cloud vendors charge a monthly price for raw storage and may add charges for network usage, operations, and retrieval and deletion transactions. At this writing, for example, Amazon.com charges anywhere from $50 to $300 per TB per year, depending on performance metrics like storage durability and latency.
- Cloud prices include the costs of space, air conditioning, controllers, racks, and power. When you add those items yourself for local storage, the total cost can be 5 to 20 times the raw storage cost, depending on your location and how much storage you have.


TIP

Over a three-year period, tape is the cheapest backup method, followed by HDD, then SSD, with cloud the most expensive. But from a cash-outlay perspective, you can start using the cloud very inexpensively and pay as you go, whereas with disk and tape, you need to purchase most of the infrastructure right away. Ensure that your backup software works with HDD, SSD, tape, and cloud, as well as with physical, virtual, and cloud systems, and then choose the most economical system for you while obeying the 3-2-1 rule. As long as your backup works with all the popular storage media, including the cloud, you can adjust media strategy when it makes business sense to do so.

IN THIS CHAPTER

- Knowing when you've lost data
- Putting your recovery plan to work
- Striving for simplicity

Chapter 4
Recovering Data

Data needs to be recovered for many reasons. The most common is human error, such as accidentally deleting an email or copying an old version of a file over the new version. Other reasons include viruses and malware; sabotage by disgruntled employees; hardware issues in disk drives, controllers, and networks; and software bugs in applications and operating systems (OSes). Sometimes, data is lost simply because no one can remember where it was filed.

Data recovery is how you get your data back, whether the lost data is as small as a phone number or as large as a complete business destroyed by a flood.

In this chapter, I show you how to create a data recovery plan that will serve you well if the worst happens. With proper backups, you can be up and running again quickly.

WARNING

If you've come to this chapter because you lost everything and don't have a backup, I can give you little advice. Some companies can try to pull the data from your disks if the loss is due to an electronic failure inside the drive and if the disk and heads aren't physically damaged by rain, heat, or dirt. Also, some programs available on the web may be able to rescue your data if you deleted it accidentally or reformatted a disk before backing up. But if you securely erased the disk, if the disk is encrypted and

you don't know the decryption key, or if you wrote on the disk quite some time after the accidental erasure occurred, all I can offer is my sympathy.

Recognizing Data Loss

Data loss isn't always obvious. Sometimes, it's confused with hardware failures, software bugs, low memory, or insufficient storage. Here are a couple of scenarios:

- **The system boots after a crash, but the applications crash.** You can look at the applications' logs and error messages, and maybe you can consult system monitoring tools, but you probably have a deadline for getting the system back up. If you're having problems with several applications, full recovery may take time. So you need to determine whether the problem is, say, a recent patch that caused your applications to malfunction or whether the data that your applications are using is corrupt.

  
  REMEMBER

  Often, it's easier to recover the system into a virtual machine (VM) — or to mount the backup as a virtual disk and do some quick compares — than it is to diagnose the problem. But after everything is running, you should try to diagnose what happened to prevent a recurrence.

- **Everything is running after a crash, but you're getting data-corruption error messages.** You may not be sure whether you have data loss and need to recover. Because the system is running, you can perform some queries to see what happens. If you notice that the corruptions fall within a certain date range or concern only one type of operation, for example, your database may have a corrupted table. Again, it's often easiest to just recover everything as long as you have an up-to-date backup.

  
  WARNING

  One common mistake is trying to repair the damage, thereby using up a lot of time that could have been better used in recovering to the last known good point.

You will suffer data loss. It's not an "if" but a "when" kind of thing. While many cases of data loss are minor and contained, they occur more frequently than most people realize. Smaller

companies report needing to recover data once or twice a week. Maybe some user accidentally deleted an email and needs it back, or another user can't find a presentation from a few months ago. If you can't recover the data, then you need to redo work that can take hours. Or maybe decisions get made without all the data.

Recovery is more complicated if an entire system becomes corrupted during an update, but if you have a tested recovery plan in place, the task of recovery is far from impossible. And unfortunately, there are more problems with updates than you realize.

Also, although most of your recovery actions are related to user mistakes that involve only a file or two, you need to be ready to call for a full system restore when doing so makes sense. It's better to have good backup procedures that allow you to recover easily than to spend hours trying — and failing — to get nonexistent data back from the system. Albert Einstein defined insanity as "doing the same thing over and over again and expecting different results." We've all been in denial that data is lost and have tried over and over to restart the system.

REMEMBER

The best advice for recovery is to practice recovery procedures frequently. Practice proves that you know how to recover, that you have the materials and software you need, and that your backup plans are working. Practice also gives you the confidence you need to recover quickly and effectively.

WARNING

Be careful that you don't hurt your backups. Sometimes, when a system is corrupting disk drives, users claim that they don't have time to restore, so they boot from a backup image. A few minutes later, however, they discover that their backup image becomes corrupted by the same problem. In other cases, users mount backups and decide to reformat the system drive before restoring, just to make sure, but they're so nervous that they accidentally reformat the backup drive.

Putting Your Recovery Plan in Motion

I hope that you have a written set of recovery procedures that you practice on a regular basis. This plan will make it easy to recover in the case of data loss.

TIP

Some backup software can print a recovery procedure for you. It's very easy to forget which tapes or disks contain which files, especially when you're rushed and under pressure, and printed backup instructions help keep you on track.

If you lose data, recovering that data is a fairly simple process when you have a recovery plan. Follow these steps:

1. **Boot up the host system that you want to place in operation with the backup program.**
2. **Recover the hypervisor and possibly the VMs to the disk on the host.**
3. **Boot the host.**
4. **Start the VMs, or recover them from another backup set and then start them.**

If you've practiced these steps, they shouldn't take long to complete and should be successful.

REMEMBER

You should appoint a recovery leader to take charge of any recovery efforts. If too many people are trying to help and taking conflicting steps, the recovery is likely to fail. You may also choose a vendor that provides an active restore capability that allows the system to run as soon as enough of the data has been recovered. Then the recovery completes while the system is running. This is a very useful feature.

WARNING

KEEP THINGS SIMPLE

I want to warn you about complexity. The simplest thing to do is choose one vendor that can supply backup and recovery for physical, virtual, and cloud storage, as well as for Windows, Linux, and granular application recovery, and do it all with image backup. You'll be better off than if you choose one product to recover Microsoft Exchange, another for SQL, a third for bare-metal recovery, yet another for file backup and recovery, and so on. Working with multiple vendors and systems can get confusing and can provide less interoperability, and you'll have to update and change your procedures much more frequently.

Try to keep your backup environment comprehensive and complete but as simple as possible by using a minimal number of vendors.

A Word about Backup and Ransomware

As of this writing, one of the gravest threats to user data is known as *ransomware,* a type of malware that infects a user's server, desktop, laptop, or mobile device, proceeds to quietly encrypt the user's data, then demands a ransom payment of several hundred or several thousand dollars to be paid in Bitcoin. The user must pay the ransom in order to receive the decryption keys necessary to unlock the data, then pray that the criminals deliver the keys and that the keys actually work. ***Note:*** The odds are poor here: Fewer than half of victims who pay successfully recover their data.

Backup is the single most foolproof defense against ransomware attacks because a victim can effectively turn back the clock to a time before the system was infected by restoring from backup. But depending on how recently this backup was performed, the user faces the loss of some amount of data, photos, and other files created in the interval between the attack and the prior backup. The problem gets tougher if the ransomware is sophisticated enough to encrypt any backup copies it finds, which is a common capability.

The widespread and growing nature of the ransomware threat, which is projected to inflict over $11 billion in damage worldwide by 2019, points to the usefulness of data protection solutions that can do more than simply help users restore their systems from the most recent pre-infection backup. Rather, it can be extremely helpful to deploy a backup solution that's also capable of detecting ransomware attacks in progress, terminating them, and repairing any files that were damaged before the attack was shut down. A solution that's properly hardened against ransomware attacks will also defend its backup archives against malicious encryption.


TIP

One example of this capability is Active Protection, a built-in feature of the business and consumer backup products offered by Acronis. End of commercial. But seriously: Take extra precautions to defend yourself against ransomware.

IN THIS CHAPTER

- Staying up to date on technology
- Knowing what and when to back up
- Planning your work and working your plan

Chapter 5
Managing Backup

The secrets to successful backup and recovery are developing the proper habits and being meticulous in your work. When things fall apart, you're the last resort, so you train and practice for a day when everything fails at the same time. Your job is simple and twofold:

- Create a great backup plan and track its execution.
- Put a great recovery plan in place and verify its efficacy.

Keeping Current on Backup Products and Services

The first task is staying up to date with the backup products you use, as well as with the hardware and software being used in your company. This task itself is a lot of work, not only because the products themselves gather new features as new versions are released, but also because the way that information technology (IT) is used constantly changes.

Virtualization, for example, has grown from a small part of a system that allowed the consolidation of lightly loaded servers to a major (and growing) component of data centers that allows much

greater resource sharing and use. The growth of virtualization alone has led to many new features and capabilities in backup products. Things to watch for are the expanding use of software-defined storage, the Internet of Things, and the increasing merging of development and operations (DevOps). These emerging technologies demand new capabilities from backup software.

TIP

You should have a long-term backup strategy that fits your company's IT strategy. Increased use of software as a service (SaaS), increased use of virtualization, or expansion to multiple locations may be the basis of an IT strategy, and the backup paths chosen should be able to move to new IT structure as well.

In addition to operational features, new technology is changing the security features of backup software. Cybercriminals know that backups help users recover from malware attacks, so new malware strains target backup files, software, and agents to prevent recovery. For example, ransomware encrypts data files and any backup archives it can find and then demands an online payment from the victim for the decryption key. New generation backup and recovery solutions ensure the integrity of their software and backup files by including self-defense mechanisms against attacks by using technologies like artificial intelligence and blockchain.

Setting the Backup Window

The next task is determining the recovery-point objective (RPO), fitting it into the backup window, and meeting the recovery-time objective (RTO). (I cover RPO in Chapter 2 and RTO in Chapter 3.) To do this, you must know

- The major applications and workloads you're protecting
- The amount of data (current and projected) associated with those applications and workloads
- The RPOs and RTOs needed for the applications and workloads

This information allows you to determine how much time you can afford to give backup. Unfortunately, some short pauses may occur during backup, and applications may run more slowly. You may decide to offset these problems by pushing backup to

less-busy times, but the busy times are when the data changes most frequently and needs protection most.

Creating and Checking a Backup Plan

TIP

If you want short recovery times, frequent recovery points, and long retention, you need to be clever. Here are a few ideas:

- Hybrid local and cloud storage (see Chapter 3) can help limit the data that has to be moved to another location.
- Deduplication and compression can minimize storage requirements (see Chapter 3).
- Image backups, along with appropriate use of incremental backups and consolidation, can reduce storage requirements without sacrificing RTO (see Chapter 2).

Also, you should have one backup plan. Your backup vendor should offer a single management console that can take a master backup plan and customize it for each system, install it on the system, monitor its progress, and check for errors.

Keeping it simple (or not)

A backup plan can be as simple as "Take a full backup of everything every night at midnight." It can also be as complex as

> For order intake, take a full backup weekly and an incremental backup every hour. For inventory control, take a backup of just the database every 15 minutes. For the manufacturing controllers, take a full image backup every 4 hours. For all the user endpoints, take a full backup monthly, dedupe at the source, and take incremental backups with encryption and compression of each user directory every 12 hours, but randomize the times so they don't all occur at once.

REMEMBER

The more complex a company's RPO and RTO are, the more complicated the backup plan is, so it's important to use the smallest possible number of backup vendors.

Setting backup windows

The *backup window* is the period of time during which the system can be down or degraded to allow backups to complete. If your company runs two shifts, an eight-hour period can be considered to be the backup window. If it operates 24/7, backups have to be done while the system is running.


TIP

Backup technology is improving, but enacting perfect zero-backup-window processes is difficult. Sometimes, using several short backup windows for different workloads can solve the problem. At other times, a strong centralized management console can automate and optimize backup windows.


TIP

Today's businesses require continuous data availability, which means being able to back up and recover data at all times. A mobile device-enabled management console ensures you can manage your backups whenever needed from wherever you are.

Checking execution

Tracking a backup plan involves checking that no backups failed and determining the cause of any failure. Running out of disk space is the most common backup failure, with network-connectivity issues running close behind.

Monitoring the plan

When the backup plan is created, daily backups are working, capacity management is in place, and network performance is established, you have only two more key tasks:

- **Watch for changes.** Today's virtual and cloud data centers let users split workloads easily. New virtual machines (VMs) may pop up at any time. Backing them up and understanding their file or synchronization constraints are major responsibilities of backup management.
- **Maintain a census.** Keep a census of systems, disks, and archives to ensure that everything that needs to be backed up is backed up.

IN THIS CHAPTER

- Understanding costs
- Setting priorities
- Knowing what to do when

Chapter 6
Ten Things to Know about Backup

When your backups are consistent and verified, and you've practiced recovery procedures, even a 2 a.m. call shouldn't throw you for a loop. You'll be prepared and confident. This chapter lists ten things you should know to make backup and recovery easier.

The Value of Your Data

Your company has many forms of data. Some data changes slowly; some changes rapidly. Some is tied to a sale; some is tied to a product or service; some is tied to financial reporting, marketing, or human resources. Knowing how important each type of data is and how often it changes helps you determine the recovery-point objective (RPO) for that data.

TIP

Usually, companies enumerate the RPO by workload or application. For more information on the RPO, see Chapter 2.

The Cost of Downtime

Sometimes, this calculation is easy. If a manufacturing system isn't running, you can compute the cost of the idle workers and the value of the products that aren't being made. The story is different in the case of a failed airline-reservation system, however. If the system can't sell seats, the airline may lose existing customers, its reputation, and potential customers, all of which represent value.

Downtime costs are business-dependent, of course, but in every business, data has value, and so does uptime. That's the reason to have backup processes in place.

According to a 2016 study by ITIC, 98 percent of companies surveyed reported that a single hour of unplanned downtime costs an average of more than $100,000.

Workload Priorities

In the event of a total loss, you should be able to prioritize recovery. Consider the following items:

- The order in which workloads should be brought up
- Which workloads should have redundancy and failover
- Which workloads can wait a few days and which can't
- Which workloads must be stopped to give their capacity to workloads that have failed

Where Your Backups Are Stored

The best practice for storage locations is the 3-2-1 rule: three copies, two media types, and one copy stored remotely (see Chapter 3). Ideally, you should have the running system, an offline copy stored locally, and a copy in a remote, off-site location such as the cloud. Be sure to consider safety and security.

How Long to Keep Backups

Storage space is limited, so at some point, you'll need to delete backups according to the company's stated retention policy (see Chapter 3). Consider three things:

- **Legal requirements:** Some companies are required to maintain certain records for a specified period. These requirements may be legislative (in regulated industries) or contractual (between the company and its customers).
- **How often files are needed:** Your business may typically exchange email and files with a customer for three months; then the job is over. In that case, you may want to keep backups for only four to six months.
- **Versioning:** Files may go through many revisions, and it may not be important to keep each revision. When the project is over, the final versions are kept in backup, but not all the intermediate versions are needed.

What Recovery Tools to Use When

To recover, you need a combination of hardware; operating-system (OS) backups, patches, and updates; applications; configuration data; and (of course) the data needed by the application. If you're performing full image backups, all this data is in the image, but if you're using file backups, you must collect and update them.

The Details of Your Backup Plan

You need documented recovery procedures for several reasons:

- If the CEO calls in the middle of the night because he accidentally deleted a file, he needs to know whom to call.
- The person whom the CEO calls needs to know how to recover the file.
- In the event of a more extensive failure, the technical staff needs to know where to find the backup and what servers need to be recovered, as well as how.

What Data Is Excluded from Backup

When space gets tight, administrators get clever with excluding unnecessary data. Generally, you don't find out until too late. It's fine to exclude files that can be re-created easily, but if your list of files excluded from backup is very long, make sure that you're not saving a few dollars of disk space by setting yourself up for days of installation work.

How (and How Deeply) to Test Backups

If your company has a good test plan for verifying performance and patches, you can use a similar test to verify that backups have been restored properly. If you have spare servers or spare capacity on a virtual host, it's a good idea to automate the testing of backups. At a minimum, you should recover the backups, run disk verification routines, and compare file sizes.

How to Frame Backup Questions

To answer any question about backup, turn it into a question about recovery. Instead of asking yourself what media to use or which retention plan makes the most sense, for example, ask one of these questions:

- "Which media allows me to recover fastest?"
- "Which media recovers most reliably?"
- "Will a GFS retention scheme allow me to recover from old backups?" (I cover Grandfather-Father-Son policies in Chapter 3.)
- "Which retention plan affords the fastest recovery?"

REMEMBER

There's no point in optimizing backup if doing so will cause problems with recovery.

Acronis

Cyber Protection & Hybrid Cloud Storage

PRODUCTS AND SOLUTIONS FOR ANY BUSINESS ENVIRONMENT

Powered by the Acronis AnyData Engine and strengthened by its artificial intelligence-based ransomware defense and blockchain-based data certification, Acronis solutions deliver easy, reliable, efficient, secure, and private cyber protection. Our business solutions deliver:

Proactive Ransomware Protection
helps avoid downtime by using artificial intelligence to actively stop attacks and unauthorized encryption.

Complete Protection
secures infrastructure and data for more than 20 platforms, including physical, virtual, cloud and mobile.

Unmatched Simplicity
allows IT pros to focus on other tasks, because our easy-to-learn solutions protect data with less effort.

Instant Restore
keeps disruptions to mere seconds, meeting recovery time objectives (RTO) and maintaining productivity.

Blockchain Notarization
ensures the integrity of data via blockchain-based technology that prevents tampering with your files.

Hybrid Cloud Architecture
lets you back up to any kind of storage and recover any piece of data via a single control interface.

Acronis

Download the Acronis Backup and Recovery Software Free Trial today.
Visit https://www.acronis.com/en-us/business/backup/

Never lose a file again!

Today, data is used and stored on a host of new platforms, including physical servers, desktops, laptops, virtual machines, and the cloud. Combine those new technologies with modern threats to your data like ransomware, and it becomes clear that the techniques used to protect your data must change, too. In *Backup For Dummies*, Acronis 15th Anniversary Edition, we discuss the strategies, tactics, and technologies that can keep your information safe from data loss in today's high-tech world.

Inside...

- Overview of data protection 101
- How to capture data for backup
- How to store your backups safely
- Protecting data from ransomware
- Ways to recover your data
- Creating a backup plan

Go to Dummies.com®
for videos, step-by-step photos, how-to articles, or to shop!

ISBN: 978-1-119-54114-1
Not For Resale

for dummies®
A Wiley Brand

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.