Chapter

1

Storage Primer

TOPICS COVERED IN THIS CHAPTER:

- ✓ The importance of IT
- ✓ The role of storage
- ✓ Types of storage
- ✓ Where storage fits into the wider IT infrastructure
- ✓ The influence of flash memory
- ✓ Cloud storage
- ✓ CompTIA Storage+ Powered by SNIA exam

This introductory chapter outlines the increasingly important role of technology in the world and how this creates a demand for robust, high-performance, technology-based systems that need good, knowledgeable people to design and administer them. It explains the important place that storage occupies in the world of information technology infrastructure and how it interacts with other infrastructure components such as computing and networking. You will learn about the major forms of storage, including the mechanical disk drive and the new wave of solid-state storage technologies such as flash memory. You will also learn about the influence of the cloud.

The Importance of Information Technology

Information technology (IT) has never been more important than it is now. It is more important than it was yesterday and shows all the signs of being even more important tomorrow. From social media and your personal life, all the way up to major governments and corporations, information and technology are at the center of everything.

In our personal lives, we're increasingly reliant on our technology devices—smartphones, smart watches, tablets, laptops, and more. Think about how inconvenienced you were the last time your home Internet connection was down. Suddenly you were cut off from the world, isolated, severely limited in what you could do! Even when we're on the move, we often rely on services such as Google Maps when navigating around big cities or driving our cars. If we're running late from a meeting and need to know the time of the next train home, where do we turn? Our technology devices, of course. It's a fact that we're increasingly reliant on technology in our personal lives.

The same goes for businesses. Businesses are increasingly dependent on information and technology. Try cutting most businesses off from the Internet and see what happens to them! It's not just the Internet. Cut a business off—in any shape or form—from its information and data, and see how long it takes until that business goes bust!

It was not that long ago that high technology was a nice-to-have feature for most businesses. High-speed Internet connections were a luxury, and instant access to information was something that was nice to have. In today's world, businesses simply can't exist without instant access to information and the Internet. In today's world, technology *is* the business. No technology, no business!

As a result of this paramount importance and central nature of technology to businesses, it's vital that businesses have reliable, high-performance technology systems. A low-latency trading company needs super-fast IT systems that enable them to execute as quickly as possible on trades; if they trade too slowly, they stand to lose money. Service companies need reliable, high-performance IT systems that enable them to provide their services to customers; if those services are unavailable or slow, customers will go elsewhere. Online sales companies need high-performance, reliable IT systems to process customer orders; if the online purchasing system is down or too slow, customers won't be able to buy goods. Hospitals need high-speed, reliable access to patient information in order to make accurate clinical decisions; terrible things can happen to patients if doctors can't access up-to-date, accurate patient data in a timely manner. Transportation companies need high-speed, dependable access to IT systems in order to provide safe and reliable services such as air-traffic control and train-track signaling. Media companies need high-speed, reliable IT systems to create and deliver high-quality content to subscribers. The list could go on and on.

IT is important, and that fact is not about to change any time soon. In order for companies to have reliable, high-speed IT systems, they need good people to design and manage them, and that's where you come into the picture!

The Role of Storage in IT

Within the world of IT, *storage* is one of the three major IT infrastructure systems:

- Computing
- Networking
- Storage

At a high level, we often think of these as three layers: the computing layer, the network layer, and the storage layer. In the computing layer, applications such as web servers, databases, and apps live and run. The network layer provides the connectivity between computing nodes. For example, a common design approach has a computer running a web server service, which talks to a database service running on another computer, all via the network layer. Finally, the storage layer is where all the data resides. A simple analogy is as follows: computing nodes are like vehicles such as cars and trucks; the network layer is like the road infrastructure; and the storage layer is like the offices and warehouses that hold all the goods and stock. This analogy might be a bit crude, but it conveys the basic idea.

Persistent and Nonpersistent Storage

There are several types of storage, but at a high level there are two main types:

- Persistent (sometimes called nonvolatile)
- Nonpersistent (sometimes called volatile)

As the names suggest, *persistent storage* is the standard choice for long-term storage of data. We call it *persistent,* or *nonvolatile,* because it doesn't lose its contents when the power is turned off. The spinning disk drive is still probably the most popular form of persistent storage, although flash memory is becoming increasingly popular. Just about every modern personal technology device—smartphone, smart watch, tablet, laptop, and so on—has internal flash memory. Because it's nonvolatile, when the battery runs out on that device (which happens all too often), your data is safe!

Probably the most common example of *nonpersistent,* or *volatile,* storage is random access memory (RAM). Just about all forms and variations of RAM lose their content when the power is removed. So why bother with RAM? Well, it's fast—quite often a lot faster than many forms of persistent storage. At certain types of operation, RAM can be a million times faster than disk and a thousand times faster than flash.

Generally speaking, when we refer to *storage*, we are referring to persistent, nonvolatile storage, which is what the majority of this book focuses on. Occasionally you may hear somebody refer to RAM as storage, but it's rare. Most of the time *storage* refers to persistent storage, whereas *memory* refers to nonpersistent technology such as RAM. An exception to this is flash memory, which is a form of persistent memory meaning that we consider it storage (it doesn't lose its contents when the power goes out) but behaves a lot like memory and is what we call *solid state*.

Although large chunks of this book are devoted to solid-state storage (SSS), we almost always use the term *solid state* to indicate that a device is semiconductor-based and has no moving parts. On the other hand, the disk drive is not a solid-state device because it has many complex moving parts and is therefore considered *mechanical*.

Performance and Availability

Both mechanical and solid-state storage devices have two important characteristics:

- Performance
- Availability

On the performance front, like everything else in IT, it's important that storage is fast. But performance can be a somewhat complicated subject in the storage world. For example, the disk drive is that last remaining mechanical component in the computer, and as such, will always struggle to keep pace with its silicon-based neighbors (CPU, memory, and so on). Because of all the mechanical components involved in every spinning disk drive, the disk drive just doesn't perform well at random workloads. However, it performs excellently at sequential workloads. Flash memory, on the other hand, is like greased lightning when it comes to random-read workloads, and the latest drives based on flash memory are getting faster and faster at sequential workloads, too (some rated at over 400 MBps vs. the latest 15K spinning drives rated at closer to 250 MBps). So storage performance isn't a simple matter and requires a good level of knowledge of the underlying components. Of course, this book addresses all of this, so if these topics seem complex and daunting right now, they should be as clear as crystal after you've read the relevant sections of the book.

On the topic of availability, because of the central and crucial nature of storage in the IT world, it's of paramount importance that storage solutions are designed to deal with component failures such as failed drives. It's a fact of life that things break, and as amazing a feat of modern engineering as the spinning disk drive is, spinning drives still break. And when they break, they tend to break in style, usually taking all their data with them. So it's vital that you understand the concepts that will allow you build highly available, resilient storage solutions. Again, this topic is covered throughout the book.

50-Something Years Young

The disk drive has been around for a long time. There's a good chance that it's older than you. Although it's advanced a lot over the years, it also hasn't changed too much.

Does it sound like the previous sentence contradicts itself? Let me explain. In 1956, IBM manufactured and shipped the IBM RAMAC 350 disk storage unit. This was essentially the first-ever disk drive and the father of all mechanical disk drives. By today's standards, it was a bulky beast. Weighing in at over a ton, it stood 5 feet 6 inches tall and sported no less than 24 50-inch platters (disks). Despite its size and weight, it stored less than 4 MB of data (quite a lot at the time). In terms of capacity and physical dimensions, things have advanced by leaps and bounds since then. But architecturally speaking, the physical design of platters, heads, actuators, and spindle motors has stayed pretty much the same.

During the past 50+ years, the mechanical disk drive has entrenched itself as the de facto medium of choice for high-performance, high-capacity persistent storage requirements. Up until around 2009/2010, it was relatively unchallenged in this category. However, significant changes are ripping their way through the storage world, predominantly because of the following two technologies:

- Solid-state storage/flash memory
- The cloud

Solid-state storage is challenging the supremacy of the spinning disk drive. It is muscling its way into not only personal electronics such as phones and tablets, where it has eradicated the use of the mechanical disk drive, but also the corporate data center. It's not uncommon for laptops and PCs to ship with solid-state drives (SSDs) rather than disk drives, and many storage arrays in the data center are either all-flash (no disk drives) or a combination of spinning disk and SSD.

The cloud is both interesting and extremely disruptive, albeit more disruptive to corporate and personal IT in general, rather than the disk drive per se. However, whereas a user or organization would once naturally have had data stored on its own spinning disk in its own data centers and equipment, this data is increasingly being stored in the cloud, where technology choices such as disk technologies are somebody else's responsibility.

Although the venerable disk drive is over half a century old and still going strong, its dominance in the data center and IT devices is being threatened more than ever before. These are exciting times!

The Influence of the Cloud

The cloud is here and cannot be ignored. In fact, the cloud is exciting and is changing the way we design, deploy, and interact with IT. These aren't just empty promises. The cloud is already here and is delivering on many of its promises.

In a traditional IT infrastructure, a company's IT assets—their servers, storage, and network equipment—were owned and managed by the company's IT department. The equipment was typically housed in a computer room or data center and accessed remotely via network links leased by the company. A typical example is a sales company with an office in New York City that is connected to its servers and applications hosted in the company's own data center somewhere else, such as New Jersey, Boston, or Chicago. In a cloud model, things are different. The company may still have its office in New York, but instead of owning and managing all of its own IT equipment in its own data centers, that equipment is now owned and managed by a cloud provider and sits in the cloud provider's data centers. That cloud provider could be a public cloud company such as Amazon or Rackspace, a private cloud provider, or a software-as-a-service provider such as Salesforce.com or Google Apps.

In the cloud model, it's the responsibility of the cloud provider to purchase IT equipment, manage it on a day-to-day basis, upgrade it, deal with failures, manage the data center, and perform all the traditional IT functions. This allows the sales company to get on with its core line of business and not have to worry about managing IT equipment. Instead of the sales company having to buy equipment up front to deal with anticipated growth and demand, it can pay for the use of cloud-based IT resources on a month-by-month basis. For example, imagine that the sales company has moved its IT infrastructure to the cloud and does a steady trade for most of the year but has huge peaks in sales around Black Friday, Cyber Monday, and the Christmas holiday season. If the company's normal level of trade requires 100 virtual servers and 500 TB of storage, but that requirement peaks for the holiday periods to 250 virtual servers and 750 TB of storage, the cloud is a great choice. This is because for most of the year, the company uses and pays for 100 virtual servers and 500 TB of storage, but then for the holiday periods, it increases the number of virtual servers to 250 and the amount of storage to 750 TB. When the holiday periods are over, the company turns off the additional servers and storage and stops paying for them until the same time next year.

This is just one example of how the cloud is changing the way we consume and pay for IT services. This topic is covered in more detail later in the book.

CompTIA Storage+ Powered by SNIA (SGO-001) Exam

This book covers the topics required for the CompTIA Storage+ Powered by SNIA (SGO-001) exam. However, the book is absolutely far more than an exam prep guide. It covers a lot more material than on the exam and is focused on preparing you for working in the real world. That said, I've personally taken, and passed, the exam while writing this book, and I am confident that the book does justice to the materials covered by the exam. But the book also covers more, a lot more.

Chapter Essentials

The Importance of Information Technology IT has never been more important in our personal lives and our businesses. As individuals, we increasingly feel isolated and unconnected without our technology devices. Likewise, businesses are increasingly dependent on technology. The concept of no technology = no business has never been more true.

Persistent Storage Persistent storage, or a persistent storage medium, is one that doesn't lose the data stored in it when the power is turned off. The disk drive is a prime example. You can turn the power off to a disk drive and power the drive up again a year later, and the data will still be there. The majority of this book focuses on forms of persistent storage and solutions built around persistent storage devices.