One of the most important aspects of a database is the storage system it is housed on and how it is managed. Advances in CPU, memory, and disk systems have allowed us to extend database technology to meet the needs of myriad data-intensive workloads. Indeed, advances in hardware continue to move at a fast pace. CPUs have for many years progressed according to Moore’s Law. Memory density and memory capacity follow a similar path, and we are now seeing solid-state disk drives that will eventually allow us to host an entire database in memory. Although technology advances allow us to do more, they also allow us to ignore the fundamental principles of good storage and database design. We have seen poorly designed databases run twice as fast on new hardware, and we have also witnessed well-designed databases see little or no performance improvement on the latest hardware. In this chapter, we will cover the various storage options that are available, how best to configure your database on the storage selected, I/O best practices, and finally high availability options.

Storage Systems

Database administration is a complex role. There are many different aspects of the database environment that must be cared for. With every new release of SQL Server, both the management and the maintenance of the database environment become easier. We find ourselves using DBCC less and less in favor of SQL Server Management Studio. But even with the help of Management Studio, the workload for the typical DBA has grown in complexity and volume. For example, the number of databases that you are expected to manage may have grown from 20 to 200; or, the 20 databases that you currently manage now account for terabytes of storage. Interestingly, we have customers who are looking at petabyte data warehouses in the not too distant future. Yes, that’s around a quadrillion bytes.

To the novice, setting up SQL Server is a simple matter of clicking through a wizard and answering a series of questions during the installation process. Accepting the defaults is perfectly fine for a simple development environment, but for a serious data management environment, you need to address a number of fundamentals and best practices. One of the most important among these is the storage system configuration and how it’s used with SQL Server. Your goal as a DBA is to ensure the best performance of the storage system while at the same time maximizing the availability of the system.

You will have to consider many factors as you select your storage system and as you configure SQL Server with the storage system. For example, you will need to factor in cost, expected workload, required uptime, data size, and expected growth rate, just to name a few. In this section, we will start at the lowest level, the disk subsystem.

As shown in Figure 7-1, there are three major categories of storage: direct attached storage (DAS), network attached storage (NAS), and storage area network (SAN). In the following sections, we will
describe each. Keep in mind that these storage types are not mutually exclusive. That is, you can combine them to achieve the storage system that is appropriate for your needs.

**Figure 7-1. The three major categories of storage (DAS, SAN, and NAS)**

Before we discuss the different storage types, we'll cover the different connection technologies that will allow you to attach your disk subsystem to your server (and database). Figure 7-1 illustrates how application software, in our case the database, connects via the operating system to the storage subsystem.