

Sparse Files

A file in which much of the data is zeros is said to contain a *sparse data set*. Files like these are typically very large—for example, a file that contains image data to be processed or a matrix within a high-speed database. The problem with files that contain sparse data sets is that the majority of the file does not contain useful data and, because of this, they are an inefficient use of disk space.

The file compression in the NTFS file system is a partial solution to the problem. All data in the file that is not explicitly written is explicitly set to zero. File compression compacts these ranges of zeros. However, a drawback of file compression is that access time may increase due to data compression and decompression.

Support for sparse files is introduced in the NTFS file system as another way to make disk space usage more efficient. When sparse file functionality is enabled, the system does not allocate hard disk drive space to a file except in regions where it contains nonzero data. When a write operation is attempted where a large amount of the data in the buffer is zeros, the zeros are not written to the file. Instead, the file system creates an internal list containing the locations of the zeros in the file, and this list is consulted during all read operations. When a read operation is performed in areas of the file where zeros were located, the file system returns the appropriate number of zeros in the buffer allocated for the read operation. In this way, maintenance of the sparse file is transparent to all processes that access it, and is more efficient than compression for this particular scenario.

The default data value of a sparse file is zero; however, it can be set to other values.

For more information about sparse files, see the following topics.

In this section

Topic	Description
Sparse File Operations	Determine whether a file system supports sparse files by calling the <code>GetVolumeInformation</code> function.
Obtaining the Size of a Sparse File	Get the allocated size or the total size for a file by using either the GetCompressedFileSize or the GetFileSize function.
Sparse Files and Disk Quotas	A sparse file affects user quotas by the nominal size of the file, not the actual allocated amount of disk space.

Community Additions