Overview and History of NTFS

In the early 1990s, Microsoft set out to create a high-quality, high-performance, reliable and secure operating system. The goal of this operating system was to allow Microsoft to get a foothold in the lucrative business and corporate market–at the time, Microsoft's operating systems were MS-DOS and Windows 3.x, neither of which had the power or features needed for Microsoft to take on UNIX or other "serious"operating systems. One of the biggest weaknesses of MS-DOS and Windows 3.x was that they relied on the FAT file system. FAT provided few of the features needed for data storage and management in a high-end, networked, corporate environment. To avoid crippling Windows NT, Microsoft had to create for it a new file system that was not based on FAT. The result was the New Technology File System or NTFS.

It is often said (and sometimes by me, I must admit) that NTFS was "built from the ground up". That's not strictly an accurate statement, however. NTFS is definitely "new"from the standpoint that it is not based on the old FAT file system. Microsoft did design it based on an analysis of the needs of its new operating system, and not based on something else that they were attempting to maintain compatibility with, for example. However, NTFS is not entirely new, because some of its concepts were based on another file system that Microsoft was involved with creating: HPFS.

Before there was Windows NT, there was OS/2. OS/2 was a joint project of Microsoft and IBM in the early 1990s; the two companies were trying to create the next big success in the world of graphical operating systems. They succeeded, to some degree, depending on how you are measuring success. :) OS/2 had some significant technical accomplishments, but suffered from marketing and support issues. Eventually, Microsoft and IBM began to quarrel, and Microsoft broke from the project and started to work on Windows NT. When they did this, they borrowed many key concepts from OS/2's native file system, HPFS, in creating NTFS.

NTFS was designed to meet a number of specific goals. In no particular order, the most important of these are:

- **Reliability**: One important characteristic of a "serious"file system is that it must be able to recover from problems without data loss resulting. NTFS implements specific features to allow important transactions to be completed as an integral whole, to avoid data loss, and to improve fault tolerance.
- **Security and Access Control**: A major weakness of the FAT file system is that it includes no built-in facilities for controlling access to folders or files on a hard disk. Without this control, it is nearly impossible to implement applications and networks that require security and the ability to manage who can read or write various data.
- **Breaking Size Barriers**: In the early 1990s, FAT was limited to the FAT16 version of the file system, which only allowed partitions up to 4 GiB in size. NTFS was designed to allow very large partition sizes, in anticipation of growing hard disk capacities, as well as the use of RAID arrays.
- **Storage Efficiency**: Again, at the time that NTFS was developed, most PCs used FAT16, which results in significant disk space due to slack. NTFS avoids this problem by using a very different method of allocating space to files than FAT does.
- **Long File Names**: NTFS allows file names to be up to 255 characters, instead of the 8+3 character limitation of conventional FAT.
- **Networking**: While networking is commonplace today, it was still in its relatively early stages in the PC world when Windows NT was developed. At around that time, businesses were just beginning to understand the importance of networking, and Windows NT was given some facilities to enable networking on a larger scale. (Some of the NT features that allow networking are not strictly related to the file system, though some are.)

Of course, there are also other advantages associated with NTFS; these are just some of the main design goals of the file system. There are also some *disadvantages* associated with NTFS, compared to FAT and other file systems-life is full of tradeoffs. :) In the other pages of this section we will fully explore the various attributes of the file system, to help you decide if NTFS is right for you.

For their part, Microsoft has not let NTFS lie stagnant. Over time, new features have been added to the file system. Most recently, NTFS 5.0 was introduced as part of Windows 2000. It is similar in most respects to the NTFS used in Windows NT, but adds several new features and capabilities. Microsoft has also corrected problems with NTFS over time, helping it to become more stable, and more respected as a "serious" file system. Today, NTFS is becoming the most popular file system for new high-end PC, workstation and server implementations. NTFS shares the stage with various UNIX file systems in the world of small to moderate-sized business systems, and is becoming more popular with individual "power" users as well.

NTFS Versions

Like most file systems, NTFS has gone through some evolutions. Over time, Microsoft has made changes to it, for many different reasons. These include correcting problems with the file system, adding support for newer hardware,

and enabling new operating system features. The biggest change to NTFS came with the introduction to the market of Windows 2000. NTFS changes enable some of the most important features of that operating system.

In this section I describe the two versions of NTFS that are commonly used on PCs. I first talk about NTFS 1.1, which is also called NTFS 4.0 (since it is most commonly used with Windows NT version 4.0). I then discuss the newest NTFS version, NTFS 5.0, which is an integral part of Windows 2000. I conclude with a brief discussion of compatibility between NTFS versions.

Note: The NTFS versions discussed here are used by almost all Windows NT and Windows 2000 installations. Apparently, there was also an earlier version of NTFS, called NTFS 1.0 or NTFS 3.1 (or so I believe), which was used on the first version of Windows NT: Windows NT 3.1. I have been unable to find any information about this version and how it differs from NTFS 1.1 / 4.0. Since Windows NT 3.1 is not really widely used, I am limiting my coverage of that older NTFS version to this note. :) Still, if you know anything about this first version of NTFS, and are willing to share that knowledge with me, I'd appreciate it!

NTFS 1.1 / 4.0

The most widely-implemented version of the NTFS file system has two different names. It is "officially" called NTFS version 1.1. However, it is also commonly called NTFS 4.0, in reference to its use by the most popular version of the operating system, Windows NT 4.0. NTFS 1.1 is also used natively by the prior version of Windows NT, Windows NT 3.51. It is also supported by several other operating systems, for compatibility purposes, though generally only in "read only"mode.

All of the fundametal characteristics and features explained in the other sections in our coverage of NTFS are supported by NTFS 1.1. When people talk about NTFS, they typically consider NTFS 1.1 to be the "default"version of the file system. Reflecting this, I have not specifically identified each NTFS feature as being supported by NTFS 1.1. Instead, I have labeled the few features particular to NTFS 5.0 appropriately, so you can watch for these notes if you want to be sure to identify where the differences in NTFS versions are.

Windows 2000 is designed specifically to use NTFS version 5.0. Because of this requirement, and because Windows 2000 (and its successors) will likely replace Windows NT over time, it seems likely that NTFS 1.1 will eventually be relegated to a "background"role. However, the vast popularity of Windows NT, and its millions of current installtions, mean that NTFS 1.1 partitions will be around for a long time to come.

NTFS 5.0

When Microsoft finally shipped its long-awaited new Windows NT operating system, they surprised a few people by naming it not Windows NT 5.0, as had been expected, but Windows 2000. However, the version of NTFS that shipped with Windows 2000 provided a hint at the original name of the operating system: NTFS 5.0. The fact that Microsoft created a new version of the file system to go along with its new operating system demonstrates just how important the NTFS file system is to Windows NT and Windows 2000. Several of the new features in Windows 2000 actually depend on features built into NTFS 5.0, such as the Active Directory service.

The following are the most important of the new features and capabilities that were added to NTFS with version 5.0. Each has been linked to the page where you can read about it in more detail:

- **Reparse Points**: Files and directories within the file system can have actions associated with them, so that when the file system object is accessed in a particular way, the action is carried out.
- **Improved Security and Permissions**: The mechanisms for managing file system security and assigning permissions were improved.
- **Change Journals**: Disk volumes can be set to keep track of all operations performed on the files and directories they contain.
- **Encryption**: NTFS 5.0 allows you to encrypt files and automatically decrypt them as they are read.
- **Disk Quotas**: Administrators can track how much disk space is being used by users or groups of users, and even limit disk space use if necessary.
- **Sparse File Support**: To save space on the disk, support was added for the more efficient storage of sparse files, which are large files that are mostly empty.
- **Disk Defragmenter**: Windows 2000 includes a disk defragmentation program, where Windows NT did not. (Arguably, this is an operating system feature of Windows 2000, and not a file system enhancement, but I thought I'd mention it anyway, since it is obviously file system related.)

Of course, this list isn't comprehensive; there were also some other, minor improvements made in a variety of areas within the file system. This includes fixing some bugs that existed in the older version of the file system, though the new features undoubtedly mean that new ones were included as well. Even the items above represent a rather substantial set of enhancements to what was already a very powerful file system. Of course, NTFS 5.0 also supports all of the features included in older versions of the file system.

As I mentioned above, NTFS 5.0 is required by Windows 2000. In fact, if you install Windows 2000 on top of Windows NT, Windows 2000 will convert any older NTFS partitions to NTFS 5.0.

NTFS Version Compatibility

Microsoft's decision to greatly enhance the NTFS file system under Windows 2000 resulted in a number of new features that most users consider advantageous. However, in creating the new 5.0 version of NTFS, compatibility issues became a concern under some circumstances. In particular, file system compatibility becomes an issue when mixing disk volumes between systems that have different versions of Windows installed. Multiple operating system PCs that have both Windows NT and Windows 2000 installed also may run into difficulties.

There are in fact several different compatibility issues here, which are related. In no particular order:

- **Windows 2000 Automatic Conversion**: Windows 2000 will automatically convert to NTFS 5.0 any NTFS 1.1 file systems it sees when it boots. Even well after the operating system has been installed, if you add an NTFS 1.1 partiton to a Windows 2000 system, it will be converted to NTFS 5.0. This can cause problems, as mentioned above, on systems that boot to both Windows NT and Windows 2000. In some circumstances it may be better to avoid using NTFS under Windows 2000 to avoid this situation.
- **Automatic Conversion of Removable Disks**: Apparently, the behavior above also applies to removable media that has been formatted using the older versions of NTFS! This means that those who move files between Windows NT and Windows 2000 machines may need to pay attention to how they use their media.
- **Windows NT Compatibility with NTFS 5.0**: In order for Windows NT to be able to read or write NTFS 5.0 partitions, Service Pack #4 (SP4) or higher must be installed on the system. This patch contains a new version of the NTFS.SYS driver file. However, even though this gives Windows NT access to NTFS 5.0 partitions, the operating system components needed to enable the new features in NTFS 5.0 are not installed.
- **Non-Microsoft Operating System Compatibility**: As shown on this cross-reference chart, certain non-Microsoft operating systems can access both NTFS 1.1 and NTFS 5.0 partitions, in read-only fashion. Being newer, NTFS 5.0 support generally requires a newer version or build number than NTFS 1.1 support.

The PC Guide is not a site that focuses specifically on operating systems, so I do not come even close to discussing all the nuances of Windows NT or Windows 2000 installations. If you are going to be working extensively with these operating systems, I would recommend that you consult more comprehensive documentation on the operating systems, and in particular, issues involved with file system installation and support. This applies even more to those who will be setting up complex systems, such as those that boot more than one operating system.

NTFS Architecture and Structures

In order to achieve the many goals that Microsoft had for its new file system, it was necessary to be very diligent in the way that NTFS was designed, starting with the fundamentals. The term architecture is often used to refer to the basic structures and concepts used to organize a complex system, be it a file system, operating system, microprocessor or what have you. To ensure that NTFS meets the needs of its many demanding users, it incorporates a very specific set of structures and techniques.

In this section, I provide a description of the architecture of NTFS volumes, and the key structures that make up an NTFS partition. This includes an overview of NTFS architecture and an explanation of how NTFS uses certain key file system structures to store information. I also discuss NTFS partition size and cluster size considerations, and contrast these to the way FAT partitions work.

NTFS Architecture Overview

Most of the weaknesses of the FAT file system result directly from the rather simplistic and outdated architecture it uses. No provisions were made in the internal structures of FAT partitions to allow for security and reliability features, making it very difficult to add such capabilities later on. In contrast, NTFS has a special architecture that not only allows for these advanced abilities, but also uses a simple conceptual scheme that makes it easier for more features to be added in the future with a minimum of changes. (In fact, this characteristic was employed when NTFS 5.0 came out with several new options.)

The elegance of the NTFS architecture can be seen in how information is stored in an NTFS partition. Virtually every structure in NTFS is a file, including the structures used to manage the partition and maintain statistics and control information about the partition itself. The control information is stored in a set of special files that are created when an NTFS partition is first created; these are called metadata files and include such items as lists of files on the partition, volume information, cluster allocations, and so forth. One exception to the "everything is a file" rule is the partition boot sector, which precedes the metadata files on an NTFS partition and controls the most basic of NTFS operations, such as loading the operating system.

The same simple conceptual model used for files and control structures is extended to the internal level of files as well. Every file in an NTFS partition is a collection of attributes; this even includes the data that the file contains, which is just considered one of many attributes. Other attributes include items such as the file's name and size. This arrangement really a database-like setup-the operating system view files as being objects with various characteristics, and manages them accordingly. This makes it easy to manage files and add attributes if needed in the future.

Internally, NTFS stores all files (including metadata files) using a cluster system-each file is broken into clusters, each of which contain a binary number of 512-byte sectors. On the surface, this is somewhat similar to how FAT stores data, but the implementation of clusters in NTFS is somewhat different.

NTFS Volume Boot Sector

When an NTFS partition is first created, the first block of information created on the partition is the volume boot sector. This fundamental structure, which is part of a very small block of NTFS management information that is not stored within the master file table on an NTFS volume, is very much the analog of the volume boot sector under the FAT file system. Like volume boot sectors in FAT, the NTFS volume boot sector is sometimes called the partition boot sector, volume boot record or other similar names.

Note: Despite the use of the word "sector"in the name of this structure, the volume boot sector can in fact be up to 16 sectors (8 kiB) in length.

The NTFS volume boot sector begins in the first sector of the partition, and consists of two different primary structures. Again, these are similar to the structures in a FAT volume boot sector:

- **BIOS Parameter Block**: This is a block of data that contains fundamental information about the volume itself. This block identifies the volume as an NTFS partition, and includes such information as the volume label and its size. In addition, NTFS provides for an extended BIOS parameter block, which contains additional information about the volume such as the location of the key metadata files.

- **Volume Boot Code**: This is a small block of program code that instructs the system on how to load the operating system. With an NTFS volume, this code will be specific to Windows NT or 2000, whichever is installed on the system. It will generally load NTLDR, the NT loader program, and then transfer control to it to load the rest of the operating system. Note that this code is also present in the partition as a system (metadata) file.

The volume boot code on a FAT file system partition is a favorite target of virus writers, as changing this code can allow a virus to automatically load whenever the system has started. The higher-security design of Windows NT and 2000, however, makes it more difficult for viruses in the boot sector code to spread, due to the fact that the operating system maintains tighter control over disk access routines once it has loaded. (Viruses are still possible on Windows NT/2000 systems, of course.)