Sample Chapter

Inside Windows NT(R), Second Edition

David A. Solomon,
base on the original edition by Helen Custer


Chapter 2: System Architecture

- System Architecture
- Requirements and Design Goals
- Operating System Models
- Architecture Overview
  - Portability
  - Symmetric Multiprocessing
  - Windows NT Workstation vs. Windows NT Server
- Key System Components
  - Environment Subsystems and Subsystem DLLs
  - NTDLL.DLL
  - Executive
  - Kernel
  - Hardware Abstraction Layer (HAL)
  - Device Drivers
  - Peering into Undocumented Interfaces
  - System Processes
- Conclusion

System Architecture

Now that we’ve covered the terms, concepts, and tools you need to be familiar with, we’re ready to exploration of the internal design goals and structure of Microsoft Windows NT. This chapter expli overall architecture of the system—the key components, how they interact with each other, and the in which they run. To provide a framework for understanding the internals of Windows NT, let’s fi review the requirements and goals that shaped the original design and specification of the system.

Requirements and Design Goals
The following requirements drove the specification of Windows NT back in 1989:

- Provide a true 32-bit, preemptive, reentrant, virtual memory operating system
- Run on multiple hardware architectures and platforms
- Run and scale well on symmetric multiprocessing systems
- Be a great distributed computing platform, both as a network client and a server
- Run most existing 16-bit MS-DOS and Microsoft Windows 3.1 applications
- Meet government requirements for POSIX 1003.1 compliance
- Meet government and industry requirements for operating system security
- Be easily adaptable to the global market by supporting Unicode

To guide the thousands of decisions that had to be made to create a system that met these requirements, Windows NT design team adopted the following design goals at the beginning of the project:

- **Extensibility** The code must be written to comfortably grow and change as market requirements change.
- **Portability** The system must be able to run on multiple hardware architectures and must be move with relative ease to new ones as market demands dictate.
- **Reliability and robustness** The system should protect itself from both internal malfunction and external tampering. Applications should not be able to harm the operating system or other running applications.
- **Compatibility** Although Windows NT should extend existing technology, its user interface application programming interfaces (APIs) should be compatible with older versions of Windows well as older operating systems such as MS-DOS. It should also interoperate well with other such as UNIX, OS/2, and NetWare.
- **Performance** Within the constraints of the other design goals, the system should be as fast and responsive as possible on each hardware platform.

As we explore the details of the internal structure and operation of Windows NT, you’ll see how these design goals and market requirements were woven successfully into the construction of the system. Before we start that exploration, let’s examine the overall design model for Windows NT and compare it to other modern operating systems.

**Operating System Models**

In most operating systems, applications are separated from the operating system itself—the operating code runs in a privileged processor mode (referred to as *kernel mode* in this book), with access to system data and to the hardware; applications run in a nonprivileged processor mode (called *user mode*), with a limited set of interfaces available and with limited access to system data. When a user-mode program calls a system service, the processor traps the call and then switches the calling thread to kernel mode. When the system service completes, the operating system switches the thread context back to user mode and allows the caller to continue.

The design of the internal structure of the kernel-mode portion of such systems varies widely. For traditional operating systems were monolithic in nature, as illustrated in Figure 2-1. The system was constructed as a single, large software system with many dependencies among internal components. Interdependency meant that extensions to the system might require many changes across the entire base. Also, in a monolithic operating system, the bulk of the operating system code runs in the same memory space, which means that any operating system component could corrupt data being used by other components.
A different structuring approach divides the operating system into modules and layers them one on the other. Each module provides a set of functions that other modules can call. Code in any particular layer of code is given access only to the lower-level interfaces (and data structures) it requires, the layering (using multiple, hierarchical processor modes). One advantage of a layered operating system structure is that because code in any particular layer of code is given access only to the lower-level interfaces (and data structures) it requires, the amount of code that wields unlimited power is limited. This structure also allows the operating system to be debugged starting at the lowest layer, adding one layer at a time until the whole system works correctly. Layering also makes it easier to enhance the operating system because individual layers can be modified or replaced without affecting other parts of the system.

Another approach to structuring an operating system is the client/server microkernel model. The architecture in this approach divides the operating system into several server processes, each of which implements a single set of services--for example, memory management services, process creation services, or processor scheduling services. Each server runs in user mode, waiting for a client request for one of its services. The client, which can be either another operating system component or an application program, requests a service by sending a message to the server. An operating system microkernel running in kernel mode delivers the message to the server; the server performs the operation; and the kernel returns the results to the client in another message, as illustrated in Figure 2-2.

NOTE:
The client/server model of networking is distinctly different from the client/server model of processing. In client/server networking, a server provides resources (such as files, printer, and storage space) to the clients. Client/server processing is a method of distributing the processing load required by an application to best suit the capabilities of network, server, and client so that one part of an application is processed on a server machine while another is processed on the client.

In reality, client/server systems fall within a spectrum, some doing very little work in kernel mode and others doing more. For example, the Carnegie Mellon University Mach operating system, a contemporary example of the client/server microkernel architecture, implements a minimal kernel that comprises scheduling, message passing, virtual memory, and device drivers. Everything else, including various file systems, and networking, runs in user mode. However, commercial implementations of the Mach microkernel operating system typically run at least all file system, networking, and memory management code in kernel mode. The reason is simple: the pure microkernel design is commercially impractical because it is too computationally expensive--that is, it’s too slow.

So what model does Windows NT embody? It merges the attributes of a layered operating system with those of a client/server or microkernel operating system. Performance-sensitive operating system components run in kernel mode, where they can interact with the hardware and with each other without incurring the overhead of context switches and mode transitions. For example, the memory manager, object and security managers, network protocols, file systems (including network servers
redirectors), and all thread and process management run in kernel mode.

Of course, all of these components are fully protected from errant applications, because applications have direct access to the code and data of the privileged part of the operating system (though they can quickly call other kernel services). This protection is one of the reasons that Windows NT has the reputation for being both robust and stable as an application server and a workstation platform yet fast and nimble from the perspective of core operating system services, such as virtual memory management, file I/O, networking, and file and print sharing.

Does the fact that so much of Windows NT runs in kernel mode mean it is more susceptible to crashes than a true microkernel operating system? Not really. Consider the following scenario: suppose the file system code of an operating system has a bug that causes it to crash from time to time. In a traditional operating system or a modified microkernel operating system, a bug in kernel-mode code such as the memory manager or the file system would likely crash the entire operating system. In a pure microkernel operating system, such components run in user mode, so theoretically a bug would simply mean that the component’s process exits. But in practical terms, the failure of such a critical process would result in a system crash since recovery from the failure of such a component would likely be impossible.

The kernel-mode components of Windows NT also embody basic object-oriented design principles; for example, they don’t reach into one another’s data structures to access information maintained by in components. Instead, they use formal interfaces to pass parameters and access and/or modify data structures.

Despite its pervasive use of objects to represent shared system resources, however, Windows NT is object-oriented in the strict sense. Most of the operating system code is written in C for portability and because development tools are widely available. C does not directly support object-oriented constructs such as dynamic binding of data types, polymorphic functions, or class inheritance. Therefore, the implementation of objects in Windows NT borrows from, but does not depend on, esoteric features of particular object-oriented languages.

Architecture Overview

Now that you understand the basic model of Windows NT, let’s take a look at the key system components that comprise its architecture. A simplified version of this architecture is shown in Figure 2-3. Keep in mind that this diagram is basic—it doesn’t show everything. The various components of Windows NT are in detail later in the chapter.

In Figure 2-3, first notice the line dividing the user-mode and kernel-mode parts of the Windows NT operating system. The boxes above the line represent user-mode processes, and the components below the line are kernel-mode operating system services. As mentioned in Chapter 1, user-mode threads execute in a protected process address space (although while they are executing in kernel mode, they have access to system space). Thus, system processes, server processes (services), the environment subsystems, and applications each have their own private process address space.

Click to view graphic (8 KB)

Figure 2-3
Simplified Windows NT architecture

The four basic types of user processes are described in the following list:

- **Special system support processes**, such as the logon process and the session manager, that are not Windows NT services (that is, not started by the service controller).

- **Server processes** that are Windows NT services, such as the Event Log and Schedule service,
add-on server applications, such as Microsoft SQL Server and Microsoft Exchange Server, include components that run as Windows NT services.

- **Environment subsystems**, which expose the native operating system services to user applications and thus provide an operating system environment, or personality. Windows NT ships with the following environment subsystems: Win32, POSIX, and OS/2 1.2.

- **User applications**, which can be one of five types: Win32, Windows 3.1, MS-DOS, POSIX, and OS/2 1.2.

In Figure 2-3, notice the "Subsystem DLLs" box below the "User applications" one. Under Windows NT, user applications do not call the native Windows NT operating system services directly; rather, they go through one or more **subsystem dynamic-link libraries (DLLs)**. The role of the subsystem DLLs is to translate a documented function into the appropriate undocumented Windows NT system service call. The translation might or might not involve sending a message to the environment subsystem process serving the user application.

The kernel mode of the operating system includes these components:

- The **Windows NT executive** contains the base operating system services, such as memory management, process and thread management, security, I/O, and interprocess communication.

- The **Windows NT kernel** performs low-level operating system functions, such as thread scheduling, interrupt and exception dispatching, and multiprocessor synchronization. It also provides a set of routines and basic objects that the rest of the executive uses to implement higher-level constructs.

- The **hardware abstraction layer (HAL)** is a layer of code that isolates the kernel, device drivers, and the rest of the Windows NT executive from platform-specific hardware differences.

- **Device drivers** include both file system and hardware device drivers that translate user I/O calls into specific hardware device I/O requests.

- The **windowing and graphics system** implements the graphical user interface (GUI) functions (better known as the Win32 USER and GDI functions), such as dealing with windows, controls, and drawing.

Each of these components is covered in greater detail both later in this chapter and in the chapters that follow.

Before we dig into the details of these system components, though, let’s review two key attributes of Windows NT architecture—portability and multiprocessing—and also examine the differences between Windows NT Workstation and Windows NT Server.

**Portability**

Windows NT was designed to run on a variety of hardware architectures, including Intel-based CISC systems as well as RISC systems. The initial release of Windows NT supported the x86 and MIPS architecture. Support for the DEC Alpha AXP was added shortly thereafter. Support for a fourth processor architecture, the Motorola PowerPC, was added in Windows NT 3.51. Because of changing market demands, however, support for both the MIPS and PowerPC was dropped after the release of Windows NT 4.0. Windows NT 5.0 will run only on x86 and Alpha machines. Eventually, Windows NT will also run on the Merced chip, the first implementation of the new 64-bit architecture family being jointly developed by Intel and Hewlett-Packard, called IA64 (for Intel Architecture 64). As Microsoft has stated publicly, Windows NT will be enhanced to support a true 64-bit programming interface on both IA64 and A architectures.

Windows NT achieves portability across hardware architectures and platforms in two primary ways:
Windows NT has a layered design, with low-level portions of the system that are processor-architecture-specific or platform-specific isolated into separate modules so that upper layers of the system can be shielded from the differences among hardware platforms. The two key components that provide operating system portability are the HAL and the kernel. Functions that are architecture-specific (such as thread context switching) are implemented in the kernel. Functions that can differ from machine to machine within the same architecture are implemented in the HAL.

The majority of Windows NT is written in a portable language—the operating system executive, utilities, and device drivers are written in C, and portions of the graphics subsystem and user interface are written in C++. Assembly language is used only for those parts of the operating system that must communicate directly with system hardware (such as the interrupt trap handler) or extremely performance-sensitive (such as context switching). Assembly language code exists only in the kernel and the HAL but also in a few places within the executive (such as the ex routines that implement interlocked instructions as well as one module in the local procedure facility), in the kernel-mode part of the Win32 subsystem, and even in some user-mode libraries such as the process startup code in NTDLL.DLL (explained later in this chapter).

**Symmetric Multiprocessing**

Multitasking is the operating system technique for sharing a single processor among multiple threads of execution. When a computer has more than one processor, however, it can execute two threads simultaneously. Thus, whereas a multitasking operating system only appears to execute multiple threads at the same time, a multiprocessing operating system actually does it, executing one thread on each of its processors.

As mentioned at the beginning of the chapter, a key Windows NT design goal from the start of the was to run well on multiprocessor computer systems. Windows NT supports *symmetric multiprocessing* (SMP). There is no master processor—the operating system as well as user threads can be scheduled on any processor. Also, all the processors share just one memory space. This model contrasts with *asymmetric multiprocessing* (ASMP), in which the operating system typically selects one processor to execute operating system code while other processors run only user code. The differences in the two multiprocessing models are illustrated in Figure 2-4.

*Figure 2-4*

Symmetric vs. asymmetric multiprocessing

Windows NT was architecturally designed to run on up to 32 processors. The number of licensed processors is stored in the registry at HKLM\System\CurrentControlSet\Control\Session Manager\LicensedProcessors. (Tampering with that data is a violation of the software license; and modifying Windows NT to use more processors is more complicated than just changing this value. The default value depends on the edition of Windows NT, as you can see in Table 2-1.

<table>
<thead>
<tr>
<th>Edition</th>
<th>Number of Licensed Processors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows NT Server, Enterprise Edition</td>
<td>8</td>
</tr>
<tr>
<td>Windows NT Server</td>
<td>4</td>
</tr>
<tr>
<td>Windows NT Workstation</td>
<td>2</td>
</tr>
</tbody>
</table>

System manufacturers that sell Windows NT Server systems that support more than eight processor
ship their own remastered Windows NT CD-ROM with a registry set to enable a higher number of processors. They might also need to provide their own HAL.

One of the key issues with multiprocessor systems is scalability. To run correctly on an SMP system, operating system code must adhere to strict guidelines and rules to ensure correct operation. Resource contention and other performance issues are more complicated in multiprocessing systems than in ordinary operating systems and must be accounted for in the system’s design. Windows NT incorporates several features that are crucial to its success as a multiprocessing operating system:

- The ability to run operating system code on any available processor and on multiple processors at the same time. With the exception of its kernel component, which handles thread scheduling and interrupts, all operating system code can be preempted (forced to give up a processor) when higher-priority thread needs attention.

- Multiple threads of execution within a single process, each of which can potentially execute simultaneously on different processors.

- Fine-grained synchronization within the kernel as well as within device drivers and server processes allow more components to run concurrently on multiple processors.

- Server processes that use multiple threads to process requests from more than one client simultaneously.

- Convenient mechanisms for sharing objects among processes and flexible interprocess communication capabilities, including shared memory and an optimized message-passing facility.

Chapter 4 describes how threads are scheduled in a multiprocessor system.

Are there two versions of Windows NT—one for uniprocessor systems and one for multiprocessor systems? Not really. Besides the HAL, which by its very nature is different for a uniprocessor system than for a multiprocessor system, of the more than 2000 files on the Windows NT CD-ROM, only one file is shipped in different uniprocessor and multiprocessor versions: the core operating system image that contains the executive and kernel, NTOSKRNL.EXE. The rest of the binary files that comprise Windows NT (including all utilities, libraries, and device drivers) are built to run properly on both uniprocessor and multiprocessor systems. For example, they handle multiprocessor synchronization issues correctly. You should use this approach on any software you build, whether it be a Win32 application or a device driver—build your code assuming it might run on a multiprocessor system so that if it does, it won’t break.

The Windows NT CD-ROM includes two versions of NTOSKRNL:

- NTOSKRNL.EXE is the executive and kernel for uniprocessor systems.

- NTKRNLMP.EXE is the executive and kernel for multiprocessor systems.

These two images are built from the same source files. They are built using compile-time condition so that multiprocessor-specific support is not included in the uniprocessor version of NTOSKRNL versus the uniprocessor version of NTOSKRNL. Because of this, single processor systems don’t have to pay for the overhead of multiprocessor synchronization at the operating system level.

At installation time, the appropriate file is selected and copied to the local \winnt\system32 directory. In either case, however, the file is named NTOSKRNL.EXE on the local hard drive.

You’ll notice that on the checked build CD-ROM (the special debug version of Windows NT, which is explained on page 22 in Chapter 1), both NTOSKRNL.EXE and NTKRNLMP.EXE are identical—both built for multiprocessor systems. In other words, there is no uniprocessor version of the check version of NTOSKRNL.
EXPERIMENT: Checking Which Version of NTOSKRNL You’re Running

You can tell which version of NTOSKRNL you’re running by running WINMSD.EXE. (From the Start menu, choose Programs, and then select Administrative Tools, Windows NT Diagnostics.) If you click the Version tab, you’ll see something like the following:

As you can see, the system is running the multiprocessor free build for x86 systems. (This screen shot was taken from the dual processor Pentium Pro workstation that Compaq so graciously loaned me for this book project.)

Windows NT Workstation vs. Windows NT Server

Many people wonder what exactly the differences are between Windows NT Workstation, Windows Server, and Windows NT Server, Enterprise Edition. First, Windows NT Server behaves differently than Windows NT Workstation does--Windows NT Server is optimized to be a high-performance network server platform, whereas Windows NT Workstation, although it has server capabilities, is optimized for interactive desktop use.

Second, Windows NT Server, Enterprise Edition, is a superset of Windows NT Server, which in turn is a superset of Windows NT Workstation. For example, the following optionally installable network server components come with Windows NT Server but are not available for Windows NT Workstation:

- Enterprise network management and directory services through the formation of domains (Windows NT systems treated as a single security perimeter)
- Disk fault-tolerance features (striping with parity and mirroring)
- Services for Macintosh: file and printer sharing, user administration
- Gateway Service for NetWare, which permits a number of Windows NT clients to access a server using the Windows NT Server as a gateway
- TCP/IP server addressing management, such as a complete Domain Name System (DNS) or Dynamic Host Configuration Protocol (DHCP)
- Remote boot server for diskless MS-DOS, Windows 3.1, and Windows 95 PCs

Windows NT Server, Enterprise Edition, contains additional components and features beyond those of Windows NT Server, such as Microsoft Cluster Server, Microsoft Message Queue Server, and Microsoft Transaction Server. (The Windows NT 4.0 Option Pack, which installs on both Windows NT Server and Windows NT Server, Enterprise Edition, includes the latter two components in addition to Microsoft Internet Information Server 4.0 and Internet Connection Services for Microsoft RAS.) Also, on x86 systems, Windows NT Server, Enterprise Edition, can allow certain applications to have a 3-GB user address space (as opposed to 2 GB on the other editions). This capability is explained in further detail in Chapter 5.

There are also licensing differences between Windows NT Workstation and Windows NT Server:

- The Windows NT Workstation license permits only 10 unique IP connections in a 10-minute period (though the code doesn’t enforce this connection limit). Windows NT Server has no such re-
- Windows NT Server supports an unlimited number of clients (assuming that you have licenses for all of them) accessing the built-in file and print-sharing services, whereas Windows NT Workstation permits only up to 10 simultaneous inbound connections to shared files or printers.


Although Windows NT Server and Windows NT Server, Enterprise Edition, contain significant added functionality over Windows NT Workstation, the majority of the files in all three products are identical including such core components as the executive, kernel, device drivers, utilities, and libraries. However, number of these components operate differently depending on which edition is running.

How does Windows NT know which product is running? At boot time, the registry is queried and this information is stored in the system global variable `MmProductType`. One element of this information is in the registry key HKLM\System\CurrentControlSet\Control\ProductOptions. Changing this information is a violation of the software license. Table 2-2 shows the values for this key as they correspond to the different editions of Windows NT.

**Table 2-2. Product Type Registry Values**

<table>
<thead>
<tr>
<th>Edition of Windows NT</th>
<th>Value of Product Options*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows NT Workstation</td>
<td>WinNT</td>
</tr>
<tr>
<td>Windows NT Server (domain controller)</td>
<td>LanmanNT</td>
</tr>
<tr>
<td>Windows NT Server (server only)</td>
<td>ServerNT</td>
</tr>
</tbody>
</table>


If user programs need to determine which Windows NT product is running, they can query for this information. (For sample code to do this, see the article Q124305 "Which Windows NT (Server or Workstation) Is Running?" in the MSDN Knowledge Base.) Device drivers running in kernel mode can call the internal executive routine used by Windows NT itself, `MmIsThisAnNtasSystem`, documented in the Windows NT Device Driver Kit (DDK).

Based on the product type, several resource allocation decisions are made differently at system boot time such as the size and number of operating system heaps (or pools), the number of internal system worker threads, and the size of the system data cache. Also, run-time policy decisions, such as the way the memory manager trades off system and process memory demands, differ between Windows NT Server and Windows NT Workstation. Even some thread-scheduling details are handled differently in the two editions. Where there are significant operational differences in the two products, these are highlighted in the pertinent chapters throughout the rest of the book. Thus, unless otherwise noted, everything in this book applies to both Windows NT Server and Windows NT Workstation.

---

**Windows NT vs. Windows 95 and Windows 98**

Windows NT and Windows 95 (and its follow-on release, Windows 98) are part of the "Windows family of operating systems," sharing a common subset API (Win32 and COM), device driver model (WDM), and in some cases shared operating system code. Although Windows NT 4.0 doesn’t have some of the features that Windows 95 has today, Microsoft has always made it clear that Windows NT was to be the strategic operating system platform for the future--not just for servers and business desktops but eventually for consumers as well. Following are some of the architectural differences and advantages that Windows NT has over Windows 95. (These comparisons also apply to Windows 98.)
- Windows NT supports multiprocessor systems--Windows 95 doesn’t.
- Windows NT runs on a variety of machine architectures--Windows 95 is limited to x86 systems.
- Windows 95 doesn’t have a file system that supports security (such as discretionary access control).
- Windows NT is a fully 32-bit operating system--it contains no 16-bit code. Windows 95 contains a large amount of old 16-bit code from its predecessors, Windows 3.1 and MS-DOS.
- Windows NT is fully reentrant--significant parts of Windows 95 are nonreentrant (mainly the older 16-bit code taken from Windows 3.1). This nonreentrant code includes the majority of the graphics and window management functions (USER and GDI). When a 32-bit application on Windows 95 attempts to call a system service implemented in nonreentrant 16-bit code, it must first obtain a systemwide lock (or mutex) to block other threads from entering the nonreentrant code base. And even worse, a 16-bit application holds this lock while running. Thus, although the core of Windows 95 contains a preemptive 32-bit multithreaded scheduler, because so much of the system is still implemented in nonreentrant code, applications many times run single threaded.
- Windows NT provides an option to run 16-bit Windows applications in their own address space--Windows 95 always runs 16-bit Windows applications in a shared address space, in which they can corrupt (and hang) each other.
- Shared memory on Windows NT is visible only to the processes that have the same shared memory section (called file mapping objects in the Win32 API) open. On Windows 95, all shared memory is visible and writable from all processes. Thus, any process can write to any file mapping object.
- Windows 95 has some critical operating system pages that are writable from user mode, thus allowing a user application to crash the system.

What does Windows 95 have that Windows NT 4.0 doesn’t? Full Plug and Play, power management, infrared support, and support for the FAT32 file system. However, all of these features will be a part of Windows NT 5.0, making it the first release of Windows NT to be a true superset of the Windows platform.

The one thing both Windows 95 and Windows 98 can do that Windows NT will never do is run all older MS-DOS and Windows 3.1 applications (notably ones that require direct hardware access) as well as 16-bit MS-DOS device drivers. Whereas 100 percent compatibility with MS-DOS and Windows 3.1 was a mandatory goal for Windows 95, the goal for Windows NT was to run most existing 16-bit applications.
Inside Windows NT Disk Defragmenting. Copyright © 1997 Mark Russinovich Note: The information presented here is the result of my own study. No source code was used. Introduction. The first defragmentation support for Windows NT was introduced by Executive Software. Their Diskeeper product was initially released for Windows This continued all the way until Windows NT 3.51. We now knew that the AND operation was implemented at some point during Windows NT 4.0™s development. Thanks to friends that had beta builds of NT 4 saved locally, we were quickly able to find the very first publicly available build with this “feature” build 1264. To test our theory, we modified the kernel of build 1264 to identify itself as build 16384. This showed us the two expected faces of the build / the code “the raw data and the handled version info.
Inside Windows NT was the first book ever published about Windows NT and provided key insights into the architecture and design of the system. Inside Windows NT, Second Edition (Microsoft Press, 1998) was written by David Solomon. It updated the original book to cover Windows NT 4.0 and had a greatly increased level of technical depth. Inside Windows 2000, Third Edition (Microsoft Press, 2000) was authored by David Solomon and Mark Russinovich.