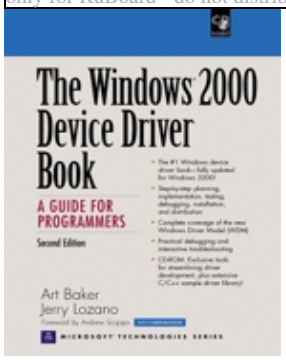





## The Windows 2000 Device Driver Book, A Guide for Programmers, Second Edition

Art Baker
Jerry Lozano
Publisher: Prentice Hall PTR

Second Edition November 20, 2000

ISBN: 0-13-020431-5, 480 pages

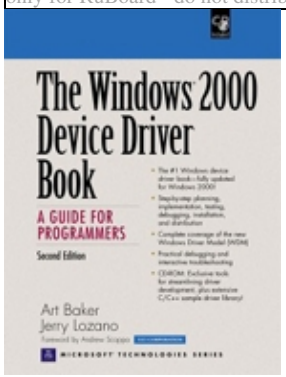
View
Buy Print Version

---

- The "100%" border="" cellspacing="0" cellpadding="0">

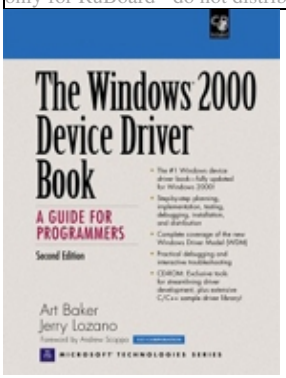







## The Windows 2000 Device Driver Book, A Guide for Programmers, Second Edition

only for RuBoard - do not distribute or recompile




The Windows 2000 Device Driver Book, A Guide for Programmers, Second Edition



Library of Congress Cataloging-in-Publication Data

The Windows 2000 device driver handbook / Art Baker, Jerry Lozano. p.cm.

Includes bibliographical references and index.

ISBN 0-13-020431-5

1. Device drivers (Computer programs) 2. Microsoft Windows (Computer file) I. Lozano, Jerry. II. Title.

QA76.76.D49 W56 2001

005.49'4769--dc21

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Prentice-Hall International (UK) Limited, *London*

Prentice-Hall of Australia Pty. Limited, *Sydney*

Prentice-Hall Canada Inc., Toronto

Prentice-Hall Hispanoamericana, S.A., *Mexico*

Prentice-Hall of India Private Limited, *New Delhi*

Prentice-Hall of Japan, Inc., *Tokyo*

Pearson Education Asia Pte Ltd.

Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*


only for RuBoard - do not distribute or recompile

# Foreword

Drivers are the most fundamental and technically difficult part of operating system development. As a reader of this book, you are probably well aware of the complexities involved. Even for the most seasoned software engineer the task can be daunting. Writing device drivers under Windows 2000 is a big challenge to learn. The most comprehensive, authoritative guide to Windows NT driver development, The *Windows NT Device Driver Book* by Art Baker is now a classic. I can not think of anyone better qualified to write the second edition of Art's outstanding book than Jerry Lozano. Jerry combines the qualities of strong technologist, excellent writer, and gifted educator. These qualities have translated into book form very well. Reading this book I felt I was taking one of Jerry's classes.

There are two kinds of books. Some books provide reference information that very much read like an encyclopedia. Such books are picked up occasionally to answer a specific question. Other books are tutorial in nature. They are designed to be read from front to back in order to transfer the knowledge and skill necessary to perform a task.

*The Windows 2000 Device Driver Book,* like its predecessor, falls clearly into the latter category. It is intended to be used as an instructional guide for device driver authors. Unlike other books on the subject, this book does not attempt to reproduce the DDK. The DDK stands as the definitive reference on the Windows 2000 device driver technology. Instead, *The Windows 2000 Device Driver Book* provides the guiding information needed to successfully master W2K driver development. This book gives developers the knowledge to design, write, and debug Windows 2000 devices, and is based on a course Jerry created and teaches for UCI. Based on feedback from the course, Jerry found that one of the biggest problems device driver and kernel-mode code

developers face is the lack of clear, concise technical information on driver models, kernel mode programming, and hardware interfaces. In this book Jerry has succeeded in solving this problem with detailed examples and informative coverage in all areas, and presenting it with exceptional clarity.

As the book went to press, it was clear that another chapter was highly desirable. The chapter concerns USB and IEEE 1394 driver specifics. The revision author has generously agreed to include this chapter on the book's companion web site: http://www.W2KDriverBook.com. Readers that need this information should visit this informative site.

Andrew Scoppa

President

UCI Software Technical Training



# Preface

This book explains how to write, install, and debug device drivers for Windows 2000. It is intended to be a companion to the Microsoft DDK documentation and software.

Windows 2000 represents a major improvement to previous versions of Windows NT. Device drivers for Windows 2000 may be designed for the new Windows Driver Model (WDM) architecture. If so, the driver will be *source* compatible with Windows 98. This book covers the new WDM specification.

This book will also prove useful to those studying the internals of Windows 2000, particularly the I/O subsystem and related components.



## What You Should Already Know

All instruction assumes a base knowledge level. First, the reader should be familiar with Windows 2000 administration  security and setup, for example. Since experimentation with kernel-mode code can (and will) cause system problems, the reader should be prepared and able to restore a chaotic OS.

Second, the reader should be competent in the C programming language and somewhat familiar with C++. Only a little C++ is used in this book, and then only for the purpose of simplifying tedious code.

Third, experience with Win32 user-mode programming is useful. Knowing how user-mode code *drives* I/O devices is useful in designing and testing device driver code. The test code for the examples in this book rely on the console subsystem model for Windows. To review this topic, the reader is referred to the *Win32 Programmers Reference,* particularly the chapters on I/O primitives (CreateFile, ReadFile, WriteFile, and DeviceIoControl). The bibliography lists other references for this topic.

Finally, while no specific prior knowledge of hardware or device driver software design is assumed, it would be useful if the reader had experience with some aspect of low-level device interfacing. For example, knowledge of writing device drivers for a Unix system will prove quite useful when reading this book.

## What's Covered

The focus of this book is to first explain the *architecture* of the hardware, environment, and device driver, and then to explain the *details* of writing code.

Chapters are grouped within this book as follows:

**Chapter 1, Chapter 2, Chapter 3, Chapter 4, Chapter 5:**

The first five chapters of this book cover the foundation of what's needed to write a device driver. This includes coverage of the Windows 2000 architecture, hardware terminology and bus basics, and an in-depth view of the Windows 2000 I/O Manager and related services.

**Chapter 6, Chapter 7, Chapter 8, Chapter 9, Chapter 10, Chapter 11, Chapter 12, Chapter 13:**

The next eight chapters form the nucleus of this book. The chapters cover everything from the mechanics of building a driver to the specifics of instrumenting a driver to log errors and other events.

**Chapter 14, Chapter 15:**

These two chapters deal with somewhat more advanced topics within device driver construction. This includes the use of system threads, layering, filtering, and utilizing driver classes.

**Chapter 16, Chapter 17:**

The final chapters deal with the practical but necessary details of driver installation and debugging. The use of Windows 2000 INF files for "automatic" installation of a plug and play device driver is covered (as well as manual installation for legacy devices). The use of WinDbg is covered in sufficient detail so that the programmer can actually perform interactive debugging.

**Appendices:**

The appendices cover reference information needed for driver development. The mechanics of Windows 2000 symbol file installation, bugcheck codes, and so on are listed.


only for RuBoard - do not distribute or recompile

## What's Not

Since the purpose of this book is to cover driver development from "the ground up," some specific topics fall outside its scope. Specifically, the list of topics not covered includes

File system drivers: Currently, the construction of a full Windows 2000 Installable File System requires the acquisition of the Microsoft IFS kit. The bibliography of this book points to one source for more in- formation on this topic. Potential users of the IFS kit will benefit greatly from this book, as the material covered is essential prerequisite knowledge.

Device-specific driver information: The construction of NIC (Network Interface Card), SCSI, video (including capture devices), printers, and multimedia drivers is not specifically covered in this book. Chapter 1 discusses the architectural implications of such drivers, but even

individual chapters on each of these driver types would seriously shortchange the requisite knowledge.

Virtual DOS device drivers: The current wave of driver development is toward the WDM 32-bit model. Legacy 16-bit VDDs are no longer of interest.



## About the Sample Code

Most chapters in this book include one or more sample drivers. All code is included on the accompanying CD. Samples for each chapter are in separate subdirectories on the CD, so installation of individual projects is straightforward.

The CD also includes a device driver application wizard for Microsoft Visual C++ version 6. This wizard configures the build environment so that code can be written, compiled, and linked within Visual Studio.

**Platform dependencies:**

The sample code included with this book has been targeted and tested on Intel platforms only. Since it appears that the last non-Intel platform (Alpha) was dropped from the final release of Windows 2000, this should come as no surprise. Be advised, however, that Windows 2000 is intrinsically a platform-independent OS. It is a straightforward process to port the OS to many modern hardware sets. Driver writers should consider designs that take advantage of the Windows 2000 abstractions that permit source compatibility with non-Intel platforms.

**To build and run the examples:**

Besides the Microsoft DDK (Device Driver Kit) (which is available on an MSDN subscription or, at present, free for download from the Microsoft web site at http://www.microsoft.com/DDK), the sample code assumes that Microsoft Visual C++ is installed. The device driver application wizard was built for Visual Studio version 6. Obviously, with some effort the sample code can be built using other vendors' compilers.

Of course, an installed version of Windows 2000 (Professional, Server, or Enterprise) is required. For interactive debugging using WinDbg, a second host platform is required.



## History of this Book

The first version of this book was written by Art Baker, entitled *The Windows NT Device Driver Book.* By any account, the book was required reading for any NT driver author. The Microsoft driver model is a continuously moving target. As such, recently introduced books on this subject provided more and up-to-date information. The goal of this revision of the book is to carry forward the goals, style, and clarity of Art's original work while updating the material with the very latest information available from Microsoft.

If you are a previous reader of the original version of this book, I hope you will find this version just as useful. I have attempted to provide accurate, concise, and clear information on the subject of Windows 2000 device drivers. While I have relied heavily on Art's original work, any errors present in this book are entirely mine.

## Training and Consulting Services

The material in this book is based on training and consulting performed for various companies within the industry.

The subject matter of this book is presented exclusively by UCI in the format of a five-day instructor-lead lecture/lab course. The course is available as public or on site classes. UCI provides comprehensive training in high-end programming, web development and administration, databases, and system technologies.

For more information please visit the UCI web site at http://www.ucitraining.com or use the address information below:

UCI Corporation

4 Constitution Way

Suite G

Woburn, MA 01801

1-800-884-1772

The revision author, Jerry Lozano, provides seminars and workshops on the topic of device drivers and other related subjects. For more information visit the web site: http://www.StarJourney.com



# Acknowledgments

I am grateful to many people who helped me with this 2nd edition. First and foremost, I want to thank Art Baker for his original work. The structure and content of this revision is based on his initial efforts.

To my partner in life, Carol, who makes everything possible. I thank Carol for both her personal and professional support. Without your encouragement, I would never have started this project. Without your help, I would never have finished.

Thanks to Russ Hall, my development editor and friend, for making the book sound good.

Thanks to Patty Donovan and her staff at Pine Tree Composition for making the book look good.

The staff of Prentice Hall PTR, especially Mike Meehan and Anne Trowbridge, deserve considerable credit for their patience and encouragement in leading me through the entire process.

I wish to thank Bryce Leach of Texas Instruments who tried to correct my misunderstandings of IEEE 1394. Your comments and suggestions for Chapter 2 were invaluable.

Thanks to Ron Reeves, for several great technical comments on several chapters.

And finally, thanks to the many people who attend my seminars, workshops, and classes for asking all those wonderful, thought-provoking questions.

# Chapter 1. Introduction to Windows 2000 Drivers

**CHAPTER OBJECTIVES**



***Device drivers on any operating system necessarily interact intimately with the underlying system code. This is especially true for Windows 2000. Before jumping into the world of Windows 2000 device drivers, this chapter presents the design philosophy and overall architecture of Windows 2000.***


only for RuBoard - do not distribute or recompile

## Overall System Architecture

Windows 2000 presents arguably the most aggressive attempt at operating system control in the history of computers. This section tours the Windows 2000 architecture, highlighting the features of significant interest to a device driver author.

### Design Goals for Windows 2000

The original goals for Microsoft's NT ("New Technology") operating system took form in early 1989. Interestingly, the original concept for NT did not even include a Windows operating environment. While the NT OS has indeed come a long way since 1989, the five fundamental goals remain intact.

- **Compatibility.**

  The OS should support as much existing software and hardware as possible.
- **Robustness and reliability.**

  The OS should be resilient to inadvertent or intentional misuse. A user's application should not be able to crash the system.
- **Portability.**

  The OS should run on as many present and future platforms as possible.
- **Extendibility.**

  Since the market requirements will change (grow) over time, the OS must make it easy to add new features and support new hardware with minimal impact on existing code.
- **Performance.**

  The OS should provide good performance for a given capability of the hardware platform which hosts it.

Of course, goals are not reality, and over time, serious compromise of one goal may be necessary to achieve another. NT is an operating system and, as such, is subject to the same sorts of compromises that affect all systems. The remainder of this section describes the delicate balance of solutions that Microsoft OS designers chose to implement their goals.

### Hardware Privilege Levels in Windows 2000

To achieve the robustness and reliability goal, the designers of NT chose a *client-server architecture* for its core implementation. A user application runs as a client of OS services.

The user application runs in a special mode of the hardware known generically as *user mode*. Within this mode, code is restricted to nonharmful operations. For example, through the magic of virtual memory mapping, code cannot touch the memory of other applications (except by mutual agreement with another application). Hardware I/O instructions cannot be executed. Indeed, an entire class of CPU instructions (designated *privileged*), such as a CPU Halt, cannot be executed. Should the application require the use of any of these prohibited operations, it must make a request of the operating system kernel. A hardware-provided *trap* mechanism is used to make these requests.

Operating system code runs in a mode of the hardware known as *kernel mode*. Kernel-mode code can perform any valid CPU instruction, notably including I/O operations. Memory from any application is exposed to kernel-mode code, providing, of course, that the application memory has not been *paged out* to disk.

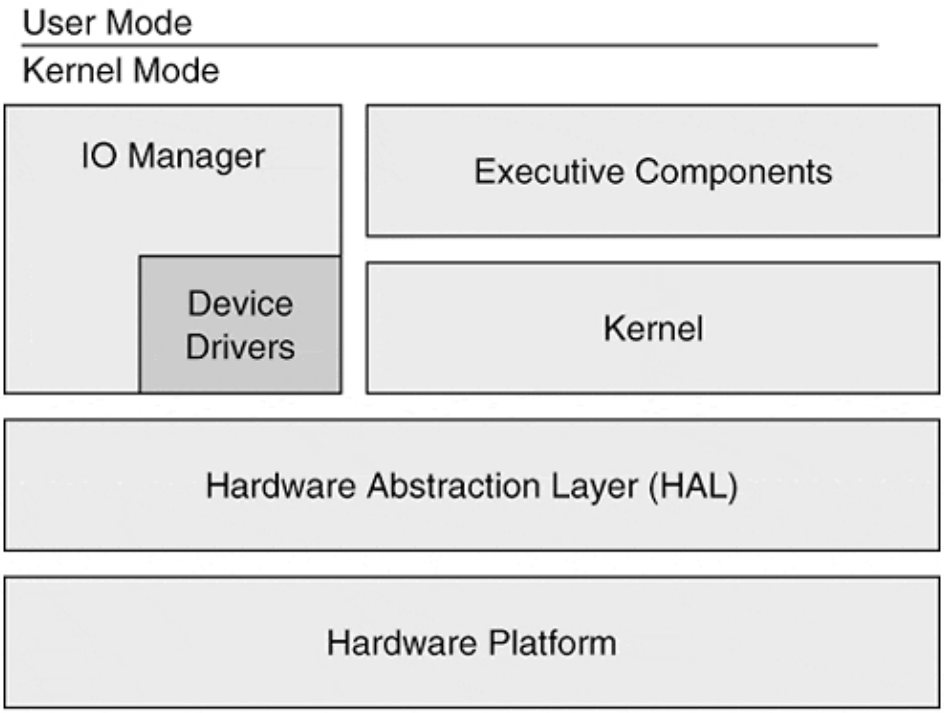
All modern processors implement some form of *privileged* vs. *nonprivileged* modes. Kernel-mode code executes in this privileged context, while user-mode code executes in the nonprivileged environment. Since different processors and platforms implement privileged modes differently, and to help achieve the goal of portability, the OS designers provided an abstraction for user and kernel modes. OS code always uses the abstraction to perform privileged context switches, and thus only the abstraction code itself need be ported to a new platform. On an Intel platform, user mode is implemented using Ring 3 of the instruction set, while kernel mode is implemented using Ring 0.

This discussion is relevant to device driver writers in that kernel-mode drivers execute in a privileged context. As such, poorly written device driver code can and does compromise the integrity of the Windows 2000 operating system. Driver writers must take extra care in handling all boundary conditions to ensure that the code does not bring down the entire OS.

### Portability

To achieve the portability goal, NT designers chose a layered architecture for the software, as shown in Figure 1.1.

**Figure 1.1.** ***The layers of the Windows 2000 operating system***




The Hardware Abstraction Layer (HAL) isolates processor and platform dependencies from the OS and device driver code. In general, when device driver code is ported to a new platform, only a recompile is necessary. How can this work since device driver code is inherently device-, processor-, and platform-specific? Clearly, the device driver code must rely on code (macros) within the HAL to reference hardware registers and buses. In some cases, the device driver code must rely on abstraction code provided in the I/O Manager (and elsewhere) to manipulate shared hardware resources (e.g., DMA channels). Subsequent chapters in this book will explain the proper use of the HAL and other OS services so that device driver code can be platform-independent.

### Extendibility

Figure 1.1 also shows an important design concept of Windows 2000  the kernel is separate from a layer known as the *Executive*.

The Windows 2000 kernel is primarily responsible for the scheduling of all thread activity. A thread is simply an independent path of execution through code. To remain independent of other thread activity, a unique thread *context* must be preserved for each thread. The thread context consists of the CPU register state (including a separate stack and Program Counter), an ID (sometimes called a Thread ID or TID, internally known as a Client ID), a priority value, storage locations local to the thread (Thread Local Storage), and other thread-relevant information.

The scheduler's responsibility is to manage which thread should execute at any given time. In a single processor environment, of course, only one thread may actually gain control of the processor at a time. In a multiprocessor environment, different threads may be executing on the different available processors, offering true parallel execution of code. The scheduler assigns a processor to a thread for, at most, a fixed period of time known as the *thread time quantum*. Processors are assigned to threads primarily based on the thread's priority value. Higher priority threads that become ready to run will preempt a running thread.

Since the kernel's prime role is to schedule thread activity, other OS components perform the necessary work of memory, process, security, and I/O management. These components are collectively known as the *Executive*. The Executive components have been designed (though the I/O Manager itself is a significant exception) as modular software. Over the years, Microsoft has added, deleted, merged, and separated these components as improvements and compromises deemed necessary. A good example would be the addition of

the Active Directory Services, which is relatively new to Windows 2000.

The notion of keeping the kernel itself small and clean, coupled with the modularization of Executive components, provides the basis for NT's claim to extendibility. The OS has now survived about ten years of revisions, maintenance, and significant feature improvement (a.k.a., *creeping elegance*).

## Performance

While the layered approach to software design is often characterized by lackluster performance, attention to fast layer interaction has been a continual effort with the NT design group. First, it should be noted that all the layers described so far execute within the same hardware mode, kernel mode. Therefore, interlayer calls often involve nothing more than a processor CALL instruction. Indeed, HAL usage is often implemented with macros, thus achieving inline performance.

Second, there has been a concentrated effort to parallelize as many tasks as possible by allocating threads to different units of work. The Executive components are all multithreaded. Helper routines seldom *block* or *busy-wait* while performing their work. This minimizes true idle time on the processor.

The performance goals of Windows 2000 impact device driver writers. As user and system threads request service from a device, it is vital that the driver code not block execution. If the request cannot be handled immediately, perhaps because the device is busy or slow, the request must be queued for subsequent handling. Fortunately, I/O Manager routines facilitate this process.

## Executive Components

Since the Executive components provide the base services for the Windows 2000 operating system (other than thread scheduling), their needs and responsibilities are fairly clear. These components are explained in the following sections.

#### SYSTEM SERVICE INTERFACE

This component provides the entry point from user mode to kernel mode. This allows user-mode code to cleanly and safely invoke services of the OS. Depending on the platform, the transition from user mode to kernel mode may be a simple CPU instruction or an elaborate Save and Restore context switch.

#### OBJECT MANAGER

Almost all services offered by the OS are modeled with an object. For example, a user-mode program that needs thread-to-thread synchronization might request a *mutex* service from the OS. The OS presents the mutex in the form of an OS-based object, referenced from user mode only through a *handle*. Files, processes, Threads, Events, Memory Sections, and even Registry Keys are modeled with OS-based objects. All objects are created and destroyed by a centralized Object Manager. This allows for uniform access, life spans, and security with all objects.

#### CONFIGURATION MANAGER

The Configuration Manager of Windows 2000 models the hardware and installed software of the machine. A database called the Registry is used to store this model. Device drivers utilize information in the Registry to discover many aspects of the environment in which they are executed. With the introduction of Plug and Play into Windows 2000, the role of the Registry for device drivers has been significantly reduced.

#### PROCESS MANAGER

A process is the environment in which threads execute in Windows 2000. Each process maintains a private address space and security identity. In Windows 2000, it is important to note that processes do not *run*;

instead, threads are the unit of execution and the process is a unit of ownership. A process owns one or more threads.

The Windows 2000 Process Manager is the Executive component that manages the process model and exposes the environment in which process threads run. The Process Manager relies heavily on other Executive components (e.g., the Object Manager and Virtual Memory Manager) to perform its work. As such, it could be said that the Process Manager simply exposes a higher level of abstraction for other lower-level system services.

Device drivers seldom interact with the Process Manager directly. Instead, drivers rely on other services of the OS to touch the process environment. For example, a driver must ensure that a buffer residing with the private address space of a process remains "locked down" during an I/O transfer. Routines within the OS allow a driver to perform this locking activity.

#### VIRTUAL MEMORY MANAGER

Under Windows 2000, the address space of a process is a flat 4 gigabytes (4 GB) ($2^{32}$). Only the lower 2 GB is accessible in user mode. A program's code and data must reside in this lower half of the address space. If the program relies on shared library code (dynamic-link libraries or DLLs), the library code also must reside in the first 2 GB of address space.

The upper 2 GB of address space of every process contains code and data accessible only in kernel-mode. The upper 2 GB of address space is shared from process to process by kernel-mode code. Indeed, device driver code is mapped into address space above 2 GB.

The Virtual Memory Manager (VMM) performs memory management on behalf of the entire system. For normal user-mode programs, this means allocating and managing address space and physical memory below the 2 GB boundary. If the needed memory for a given process is not physically available, the VMM provides an illusion of memory by *virtualizing* the request. Needed memory is *paged* onto a disk file and retrieved into RAM when accessed by a process. In effect, RAM becomes a shared resource of all processes, with memory moving between files on the disk and the limited RAM available on a given system.

The VMM also acts a memory allocator in that it maintains heap areas for kernel-mode code. Device drivers can request the VMM to assign dedicated areas of pagable or nonpagable memory for its use. Further, devices that operate using DMA (direct memory access) can assign nonpagable memory as needed to perform data transfers between RAM and a device. Of course, these topics are covered in more detail in subsequent chapters.

#### LOCAL PROCEDURE CALL FACILITY

A Local Procedure Call (LPC) is a call mechanism between processes of a single machine. Since this *interprocess* call must pass between different address spaces, a kernel-mode Executive component is provided to make the action efficient (and possible). Device driver code has no need for the LPC facility.

#### I/O MANAGER

The I/O Manager is an Executive component that is implemented with a series of kernel-mode routines that present a uniform abstraction to user-mode processes for I/O operations. One goal of the I/O Manager is to make all I/O access from user mode device-independent. It should not matter (much) to a user process whether it is accessing a keyboard, a communication port, or a disk file.

The I/O Manager presents requests from user-mode processes to device drivers in the form of an I/O Request Packet (IRP). The IRP represents a work order, usually synthesized by the I/O Manager, that is presented to a device driver. It is the job of device drivers to carry out the requested work of an IRP. Much of the remainder of this book is devoted to the proper care and processing of IRPs by device driver code.

In effect, the I/O Manager serves as an interface layer between usermode code and device drivers. It is therefore the most important block of code that a device driver must interact with during operation.

#### ACTIVE DIRECTORY SERVICE

The Active Directory Service is somewhat new to Windows 2000. It provides a network-wide namespace for system resources. Previously, the internal names used to identify system resources (disk drive names, printer names, user names, file names) were managed within a restricted space of the OS. It was the responsibility of other OS components (e.g., the networking services) to *export* names across different protocols.

The Active Directory is now a uniform, secure, and standard way to identify system resources. It is based on a hierarchical scheme (strictly defined by a schema) whereby entities are categorized into organization units (OUs), trees, and forests.

#### EXTENSIONS TO THE BASE OPERATING SYSTEM

Although the Executive components of Windows 2000 define and implement core services of the OS, it might be interesting to note that these services are not directly exposed to user-mode applications. Instead, Microsoft defines several Application Programming Interfaces (APIs) that user-mode code treats as abstractions of OS services. These APIs form different *environmental subsystems* that application code live within. Currently, the following environmental subsystems are included with Windows 2000.

- The Win32 subsystem is the native-mode API of Windows 2000. All other environmental subsystems rely upon this subsystem to perform their work. All new Windows 2000 applications (and indeed, most ported ones as well) rely on the Win32 subsystem for their environment. Because of its importance (and interesting implementation), this subsystem is described in more detail in the next section.
- The Virtual DOS Machine (VDM) subsystem provides a 16-bit MSDOS environment for old-style DOS applications. Despite its promise of compatibility, many existing 16-bit DOS programs do not operate properly. This is due to Microsoft's conservative and safe approach that *emulates* device (and other system resources) access. Attempts to directly access these resources results in intervention from the OS that provides safe, but not always faithful, results.
- The Windows on Windows (WOW) subsystem supports an environment for old-style 16-bit Windows applications (i.e., Windows 3.X programs). Interestingly, each 16-bit program runs as a separate thread within the address space of a single WOW process. Multiple WOWs can be spawned, but 16-bit Windows applications are then prohibited from sharing resources.
- The POSIX subsystem provides API support for Unix-style applications that conform to the POSIX 1003.1 source code standard. Unfortunately, this subsystem has not proved workable for hosting the ports of many (most) Unix-style applications. As such, most Unix applications are ported by rewriting for the Win32 environment.
- The OS/2 subsystem creates the execution environment for 16-bit OS/2 applications  at least those that do not rely on the Presentation Manager (PM) services of OS/2. This subsystem is available only for the Intel (x86) version of Windows 2000.

A given application is tightly coupled to exactly one environmental subsystem. Applications cannot make API calls to other environments. Also, only the Win32 subsystem is native  other subsystems emulate their environments and therefore experience various degrees of performance degradation compared to native Win32. Their purpose is compatibility, not speed.

Environmental subsystems are generally implemented as separate user-mode processes. They launch as needed to support and host user-mode processes. The environmental subsystem becomes the *server* for the usermode *client*. Each request from a client is passed, using the Local Procedure Call Executive component, to the appropriate server process. The server process (i.e., the environmental subsystem) either performs the work to fulfill the request directly or it, in turn, makes a request of the appropriate Executive component.

### THE WIN32 SUBSYSTEM

As the native API for Windows 2000, the Win32 environmental subsystem is responsible for

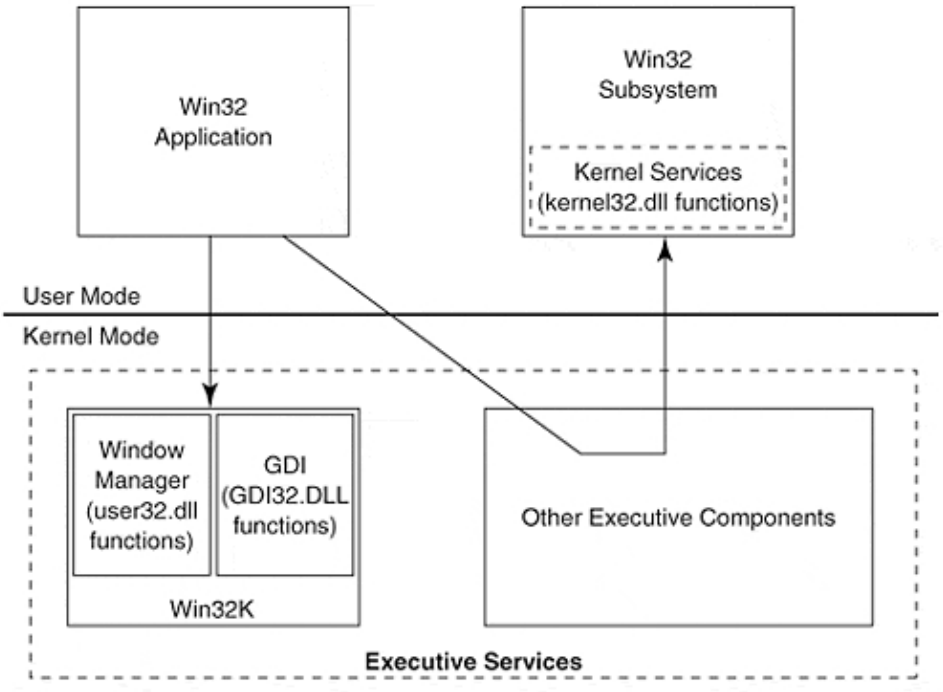
- The Graphical User Interface (GUI) seen by users of the system. It implements and exposes viewable windows, dialogs, controls, and an overall style for the system.
- Console I/O including keyboard, mouse, and display for the entire system, including other subsystems.
- Implementation of the Win32 API, which is what applications and other subsystems use to interact with the Executive.

Because the Win32 Subsystem holds special status within the system and because of its inherent requirement for high performance, this subsystem is implemented differently from any of the other subsystems. In particular, the Win32 subsystem is split into some components that execute in user mode and some that execute in kernel mode. In general, the Win32 function can be divided into three categories.

- USER functions that manage windows, menus, dialogs, and controls.
- GDI functions that perform drawing operations on physical devices (e.g., screens and printers).
- KERNEL functions, which manage non-GUI resources such as processes, threads, files, and synchronization services. KERNEL functions map closely to system services of the Executive.

Since NT 4.0, USER and GDI functions have been moved to kernel mode. User processes that request GUI services are therefore sent directly to kernel-mode using the System Service Interface, an efficient process. Kernel-mode code that implements USER and GDI functions resides in a module called WIN32K.SYS. The USER and GDI kernel components are illustrated in Figure 1.2.

**Figure 1.2.** ***USER and GDI kernel components.***




Conversely, KERNEL functions rely on a standard server process, CSRSS.exe (Client-Server Runtime Subsystem), to respond to user process requests. In turn, CSRSS traps into Executive code to complete the request for such functions.

#### INTEGRAL SUBSYSTEMS

In addition to the Environmental Subsystems, there are also key system components that are implemented as user mode processes. These include

- The Security Subsystem, which manages local and remote security using a variety of processes and dynamic libraries. Part of the Active Directory work also resides within this logical subsystem.
- The Service Control Manager (affectionately called the *scum*, or SCM) manages services (daemon processes) and device drivers.
- The RPC Locator and Service processes give support to applications distributed across the network. Through the use of remote procedure calls, an application can distribute its workload across several networked machines.


only for RuBoard - do not distribute or recompile

## Kernel-Mode I/O Components

The purpose of this section is to describe the goals and architecture of the Windows 2000 I/O subsystem. Since different kinds of drivers perform wildly different kinds of service, the I/O Manager's categorization of drivers is also discussed.

### Design Goals for the I/O Subsystem

The I/O subsystem of Windows 2000 added to the overall design goals of the operating system by including

- Portability, platform to platform.
- Configurability in terms of both hardware and software. For Windows 2000 drivers, this would include full support for Plug and Play buses and devices.
- Preemptable and interruptable. I/O code should never block and should always be written thread-safe.
- Multiprocessor-safe. The same I/O code should run on both uniprocessor and multiprocessor configurations.
- Object-based. The services provided by I/O code should be offered in encapsulated data structures with well-defined allowable operations.
- Packet-driven. Requests made of the I/O subsystem should be submitted and tracked using a distinct "work order" format, known as an *I/O Request Packet* (IRP).
- Asynchronous I/O support. Requests made of the I/O subsystem should be allowed to complete in parallel with the requestor's execution. When the request ultimately completes, a mechanism must exist to notify the caller of completion.

Besides these published goals, there is also strong emphasis placed on code reusability. This translates to heavy structuring of I/O code (including drivers) into logical layers. For example, bus-driving code should be layered separately from specific device code to allow for reuse of the bus code across multiple devices. In many cases, different vendors supply code for different layers. Only through careful *modularization* can this goal be achieved.

### Kinds of Drivers in Windows 2000

There once was a time when a device driver author could understand the intricacies of the new hardware, learn the OS device driver interface, scope the work, and "just write the code." For better or worse, the days of monolithic device driver code have passed. Today, an author must understand the *architectures* of both complex hardware buses and heavily layered I/O subsystems just to scope the work statement. Deciding what *kind* of driver to write for Windows 2000 is itself an interesting challenge. Deciding whether to implement or to reuse a layer is yet another challenge. The purpose of this section is to describe where different kinds of drivers fit within the hardware world and the OS.

At the highest level, Windows 2000 supports two kinds of drivers, user-mode and kernel-mode. User-mode drivers, as the name implies, is system-level code running in user mode. Examples include a simulated, or *virtualized,* driver for imaginary hardware or perhaps a new environmental subsystem. Since Windows 2000 user mode does not allow direct access to hardware, a virtualized driver necessarily relies upon real driver code running in kernel mode. This book does not describe user-mode drivers. The purpose of this book is to describe *real* drivers, which in Windows 2000 are known as *kernel-mode drivers.*

Kernel-mode drivers consist of system-level code running in kernel mode. Since kernel mode allows direct hardware access, such drivers are used to control hardware directly. Of course, nothing prevents a kernel-mode driver from virtualizing real hardware  the choice between user and kernel mode is largely an implementer's choice. Again, however, the purpose of this book is to present the strategies for implementing true kernel-mode drivers for real hardware.

Moving down a level, kernel-mode drivers can be further decomposed into two general categories, legacy and Windows Driver Model (WDM). Legacy drivers were fully described in the first edition of this book. The techniques needed to discover hardware and interface with the I/O subsystem are well documented. Thankfully, most of the knowledge gained by understanding legacy Windows NT drivers is transportable to the Windows 2000 (and Windows 98) WDM world.

WDM drivers are Plug and Play compliant. They support power management, autoconfiguration, and hot plugability. A correctly written WDM driver is usable on both Windows 2000 and Windows 98, though at present, Microsoft does not guarantee binary compatibility. At most, a rebuild of the driver source is necessary using the Windows 98 DDK (Device Driver Kit).

Moving down yet another level, legacy and WDM drivers can be further decomposed into three categories, high-level, intermediate, and low-level. As the names imply, a high-level driver depends on intermediate and low-level drivers to complete its work. An intermediate driver depends on a low-level driver to complete its work.

High-level drivers include file system drivers (FSDs). These drivers present a nonphysical abstraction to requestors that, in turn, is translated into specific device requests. The need to write a high-level driver is apparent when the underlying hardware services are already provided by lower levels  only a new abstraction is required for presentation to requestors.

Microsoft supplies an Installable File System (IFS) kit, sold separately from MSDN or any other product. The IFS kit requires the DDK (and other products) for successful file system development. There are numerous restrictions on the types of file systems that can be developed using this kit. For pricing and ordering information, you can visit the HWDEV virtual site of Microsoft's Internet site. This book does not address file system development.

Intermediate drivers include such drivers as *disk mirrors*, *class drivers*, *mini drivers*, and *filter drivers*. These drivers insert themselves between the higher-level abstractions and the lower-level physical support. For example, a disk mirror receiving the request from the high-level FSD to write to a file translates such a request into two requests of two different low-level disk drivers. Neither the higher nor lower levels need to be aware that mirroring is, in fact, occurring.

Class drivers are an elegant attempt at code reuse within the driver model. Since many drivers of a particular type have much in common, the common code can be placed in a generic *class* driver separate from the physical, device-specific code. For example, all IDE disk drivers share considerable similarity. It is possible to write the common code once, placing it in a generic class driver that loads as an intermediate driver. Vendor and device specific IDE drivers would then be written as *mini drivers* that interact with the generic class driver.

Filter drivers are intermediate drivers that intercept requests to an existing driver. They are given the

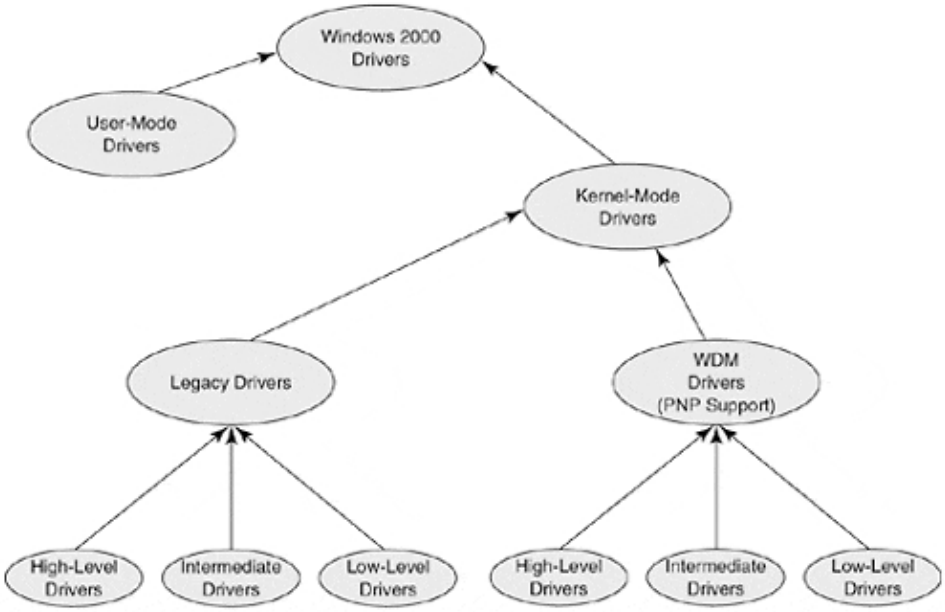
opportunity to modify requests before presentation to the existing driver.

Finally, within the WDM world, intermediate drivers can also consist of *Functional Drivers*. These drivers can be either class or mini drivers, but they always act as an interface between an abstract I/O request and the low-level physical driver code. Within the DDK documentation, the term *Functional Driver* is sometimes interchanged with Class or Mini Driver. The context determines the meaning.

Low-level drivers include controllers for the hardware buses. For example, the SCSI Host Bus Adapter is one such low-level driver. Such drivers interact with Windows 2000 HAL layer and/or the hardware directly. In the WDM world, low-level drivers include the notion of a *Physical Driver*. These Physical Drivers interact with one or more Functional Drivers.

Figure 1.3 shows the driver classifications in Windows 2000.

**Figure 1.3.** ***Driver classifications in Windows 2000.***





only for RuBoard - do not distribute or recompile

## Special Driver Architectures

Building upon the intermediate driver strategy described in the last section, Microsoft supplies driver architectures for several types or classes of devices.

- Video drivers
- Printer drivers
- Multimedia drivers
- Network drivers

These architectures conform to the spirit, if not to the letter, of the classifications of the last section. Each architecture is described in more detail in the following sections.

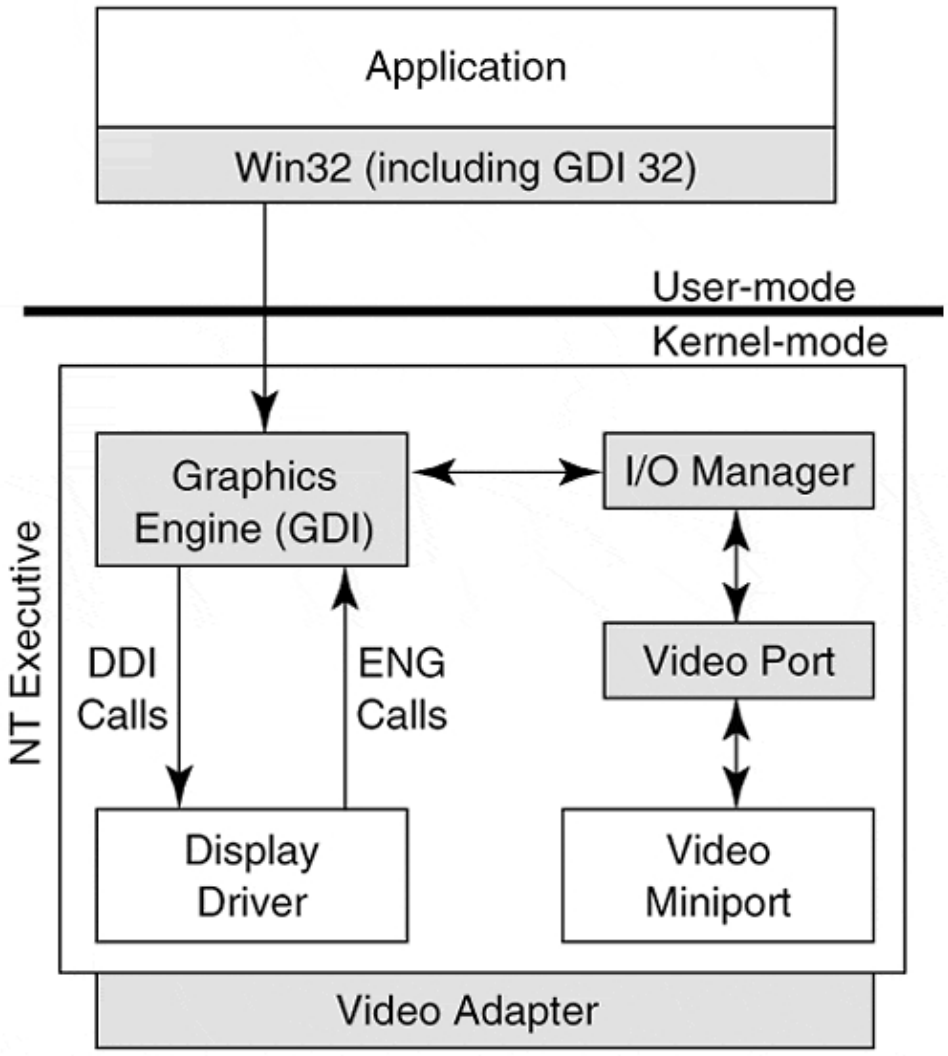
### Video Drivers

Video drivers in Windows 2000 present special requirements to the I/O subsystem. Because the graphical interface is constantly exposed to users, the apparent overall speed of the system is judged (often incorrectly) by the performance of this component. Competition among video adaptor hardware vendors has forced

aggressive graphic accelerators to be included on high-end boards. The video driver architecture must exploit such hardware when it exists, yet provide full compatibility and capability when it does not. Finally, video drivers have evolved since the 16-bit world of Windows. There is a need to provide as much compatibility as possible with legacy drivers.

The video driver architecture of Windows 2000 is shown in Figure 1. 4. The shaded components are provided with Windows 2000. Vendors of specific display adaptors supply the display driver. Since many display adaptors are designed using common chip sets, the chip set manufacturer supplies the video miniport *class driver* for its adaptor-manufacturing customers. For example, an ET4000 Miniport driver exists for all adaptors that utilize the ET4000 chip set. The extra hardware surrounding the chip set is driven with adaptor-specific display driver code.

**Figure 1.4.** ***Video driver architecture.***




Fundamentally, the video driver architecture differs from the standard I/O architecture in that user applications do not communicate directly with the I/O Manager when requesting drawing services. Instead, user-mode code interacts with a *Graphics Device Interface* (GDI) component of the kernel.

The GDI implements functions that allow the drawing of lines, shapes, and text in selected fonts. The GDI, therefore, is similar to a high-level driver. In turn, the GDI relies upon the services of the display driver and the I/O Manager to complete its work. Communication between the GDI and display driver is bidirectional. Where speed is paramount, the GDI can invoke functions in the display driver directly, bypassing the I/O Manager altogether. The display driver implements an interface known as the *Device Driver Interface* (DDI),