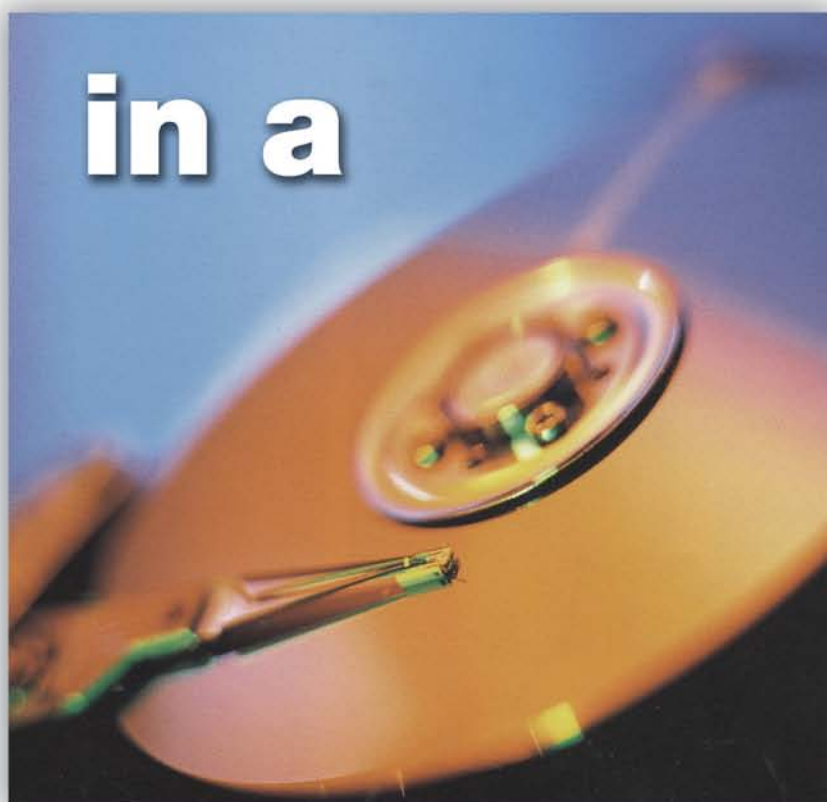




SERVER DISK MANAGEMENT



WINDOWS ENVIRONMENT

DREW ROBB

SERVER DISK MANAGEMENT in a WINDOWS ENVIRONMENT

OTHER AUERBACH PUBLICATIONS

The ABCs of IP Addressing

Gilbert Held

ISBN: 0-8493-1144-6

The ABCs of LDAP

Reinhard Voglmaier

ISBN: 0-8493-1346-5

The ABCs of TCP/IP

Gilbert Held

ISBN: 0-8493-1463-1

Building an Information Security Awareness Program

Mark B. Desman

ISBN: 0-8493-0116-5

Building a Wireless Office

Gilbert Held

ISBN: 0-8493-1271-X

The Complete Book of Middleware

Judith Myerson

ISBN: 0-8493-1272-8

Computer Telephony Integration, 2nd Edition

William A. Yarberry, Jr.

ISBN: 0-8493-1438-0

Electronic Bill Presentment and Payment

Kornel Terplan

ISBN: 0-8493-1452-6

Information Security Architecture

Jan Killmeyer Tudor

ISBN: 0-8493-9988-2

Information Security Management Handbook, 4th Edition, Volume 1

Harold F. Tipton and Micki Krause, Editors

ISBN: 0-8493-9829-0

Information Security Management Handbook, 4th Edition, Volume 2

Harold F. Tipton and Micki Krause, Editors

ISBN: 0-8493-0800-3

Information Security Management Handbook, 4th Edition, Volume 3

Harold F. Tipton and Micki Krause, Editors

ISBN: 0-8493-1127-6

Information Security Management Handbook, 4th Edition, Volume 4

Harold F. Tipton and Micki Krause, Editors

ISBN: 0-8493-1518-2

Information Security Policies, Procedures, and Standards: Guidelines for Effective Information Security Management

Thomas R. Peltier

ISBN: 0-8493-1137-3

Information Security Risk Analysis

Thomas R. Peltier

ISBN: 0-8493-0880-1

Interpreting the CMMI: A Process Improvement Approach

Margaret Kulpa and Kurt Johnson

ISBN: 0-8493-1654-5

IS Management Handbook, 8th Edition

Carol V. Brown and Heikki Topi

ISBN: 0-8493-1595-6

Managing a Network Vulnerability Assessment

Thomas R. Peltier and Justin Peltier

ISBN: 0-8493-1270-1

A Practical Guide to Security Engineering and Information Assurance

Debra Herrmann

ISBN: 0-8493-1163-2

The Privacy Papers: Managing Technology and Consumers, Employee, and Legislative Action

Rebecca Herold

ISBN: 0-8493-1248-5

Securing and Controlling Cisco Routers

Peter T. Davis

ISBN: 0-8493-1290-6

Six Sigma Software Development

Christine B. Tayntor

ISBN: 0-8493-1193-4

Software Engineering Measurement

John Munson

ISBN: 0-8493-1502-6

A Technical Guide to IPSec Virtual Private Networks

James S. Tiller

ISBN: 0-8493-0876-3

Telecommunications Cost Management

Brian DiMarsico, Thomas Phelps IV, and William A. Yarberry, Jr.

ISBN: 0-8493-1101-2

AUERBACH PUBLICATIONS

www.auerbach-publications.com

To Order Call: 1-800-272-7737 • Fax: 1-800-374-3401

E-mail: orders@crcpress.com

SERVER DISK MANAGEMENT in a WINDOWS ENVIRONMENT

Drew Robb



AUERBACH PUBLICATIONS

A CRC Press Company

Boca Raton London New York Washington, D.C.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2004 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20131031

International Standard Book Number-13: 978-0-203-49874-3 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Contents

Preface	xiii
Introduction	xxi
1 Hard Disks and How They Are Organized	1
Hardware Components of the Hard Drive	1
Platters	2
Spindle/Motor.....	3
Head-Actuator Assembly	4
Logic Board.....	4
Cache	5
Controller	5
SCSI versus ATA	5
ATA Drives	6
Small Computer Systems Interface Drives	6
What Is on a Disk	7
Disk Performance Definitions and Metrics	9
Disk Formatting	10
Grouping Disks: Stripe Sets and RAID	11
Evaluating Disks	12
Mean Time between Failures	13
2 Windows File Systems	15
Enterprise Windows	15
Magnetic Disk Systems.....	16
File Allocation Table (FAT).....	16
New Technology File System (NTFS).....	18
Additional NTFS Features	20
FAT versus NTFS.....	21
Converting FAT to NTFS	21
Optical Disk Formats	21
Partitioning Strategies	22
Summary.....	23
3 Hard Drive Reliability	25
MTBF Explained	25
Adding Up the Cost of Disk Failure	27

Do Not Rely on Backup	27
Do Not Rely on RAID	28
Predicting Failure	29
Self-Monitoring Analysis and Reporting Technology (SMART)	30
How Often Do Disks Fail?	32
Not So SMART	33
Products Using SMART	34
Compaq Insight Manager (CIM)	34
OEM Disk Monitoring Tools	34
SpinRite	35
Hot Swapability	35
Disks Will Fail	36
4 Disk Management Basics for Windows 2000	37
The Importance of Disk Management	37
Basic Tasks	38
Disk Initialization	38
Disk Formatting	38
Volume Creation	39
Basic Disks	39
Extended Partitions	39
Logical Drive	39
Dynamic Disks	40
Volumes	40
Simple Volume	40
Spanned Volume	40
Striped Volume	41
Creation of Fault-Tolerant Disk Systems	41
Mirrored Volumes	42
RAID-5 Volumes	42
Disk Management Console	42
Converting a Basic Disk to a Dynamic Disk	43
Basic and Dynamic Disks and Windows XP	43
Be Dynamic, Choose Dynamic	44
5 Backing Up Windows Networks	45
No More Simple Backups	45
Windows Backup: What It Includes	46
Using the Built-In Windows Backup Utility	48
Emergency Repair Disk and Restore Functions	49
Additional XP Backup Features	49
Beyond Basic Backups	50
Mirroring	50
Offsite Replication	50
Backup Cautions	51
Backing Up Mixed Windows/Linux Networks	51
Storage Portal Case Study	52
6 Disaster Recovery and Disk Management	55
Is Your Data Really Protected?	55
A Blended Approach to Data Protection	57
Non-Catastrophic Failure	58
Contingency Plan	58
Catastrophic Failure	59

Make the DRP Comprehensive	60
The Economics of Disaster	60
Recognition of Mission-Critical Elements	61
Drilling	61
Data Organization	62
Data Quantity	62
Data Priority	62
Data Medium	63
Hardware Issues	63
Networking Issues	64
Software Issues	64
Test and Refine Some More	64
7 Disk Performance and Fragmentation	65
History and Origins of Fragmentation	65
Does NTFS Suffer from Fragmentation?	66
How Much Fragmentation Occurs in NTFS?	66
Defining Fragmentation	67
Free Space Fragmentation	68
Fragmentation Adversely Affects Systems Stability	68
Fragmentation and Boot Time	69
Fragmentation and Slow Back Up Times/Aborted Backups	70
Fragmentation and File Corruption/Data Loss	71
Fragmentation and Crashes	71
Fragmentation and Errors in Programs	72
Fragmentation and RAM Use/Cache Problems	73
Fragmentation and Hard Drive Failure	73
Benefits of Defragmentation	73
Remedying Fragmentation	74
How Defragmenters Work	74
Preventing MFT Fragmentation	74
Converting FAT to NTFS	75
MFT and Full or Heavily Fragmented Disks or Partitions	76
Effect of MFT Fragmentation on Performance	76
Page File Fragmentation	76
Defragmentation Safety Issues	78
Defragmentation Performance Testing	78
Windows 2000 Tests	79
XP Performance Tests	80
Testing the Windows 2000 Disk Defragmenter	81
Running the Windows 2000 Disk Defragmenter	81
WDD Cautions	82
Speed	82
Scheduling	83
Priority Settings	83
Administrative Privileges	83
Free Space	84
Thoroughness	84
I/O Overload	85
Defragmenting a System	85
Hardware Upgrades versus Defragmentation	88
Total Cost of Ownership of Manual versus Network Defragmenters	90
Uptime Guarantee	90

8	Disk Optimization: Optimum or Not?	93
	Hard Drives Are Slow	93
	Enter Optimization	94
	Optimization Pros and Cons	96
	Optimization Shortcomings	97
	Logical Cluster Numbers	98
	Multiple Partitions	99
	Server Issues	99
	RAID Arrays	100
	Proven Disk Performance Solutions	100
	RAID	100
	Caching	101
	Defragmentation	101
	Conclusions	102
9	Disk Quotas	103
	What Is a Disk Quota?	103
	Quota Terminology	104
	Disk Quota Management	104
	Using APIs	104
	Operating System Integration	105
	I/O Filters	106
	Windows 2000 Built-In Quota Utility	106
	Strengths	106
	Limitations	106
	Third-Party Alternatives	107
	QuotaAdvisor	108
	Case Study: Premier, Inc.	108
	Conclusion: Upgrade to a Third-Party Disk Quota Tool	109
10	Hard Disk Forensics: The Hard Disk As a Source of Evidence	111
	The Rise of Computer Crime	111
	Melissa	113
	I Love You	113
	Definitions of Computer Crime and Computer Forensics	114
	Uses of Computer Forensics	115
	Recovering Evidence	115
	Printed versus Electronic Versions	115
	Microsoft Word Confessions	116
	Preserving Confidentiality in Microsoft Word	117
	Finding Missing Files	117
	Shredder Programs	118
	Forensics Software	119
11	Simplifying Disk Management for Storage Assets	121
	Snapshot Technology	121
	Copy-on-Write	122
	Split Mirror	122
	SQL Server Backups	123
	Breakthrough in Backup and Recovery Performance	123
	Storage Evolution	126
	Microsoft System Restore	128
	Shadow Sets	129
	Advantages of VSS	129

Storage Virtualization	130
Storage and RAID	131
Server-Attached Storage	131
Storage Area Networks	132
SAN Appliances	132
What Is True Storage Virtualization?	132
Virtualization Scenarios	134
Virtualization and Windows-Based Servers	135
12 Server Consolidation and Disk Management Simplification	137
Harsh Economic Realities	137
Assessing the Reliability, Availability, and Manageability of the Data Center	139
Data Center Cost Analysis	139
Leveraging New Technologies To Streamline Data Center Operations	140
Scaling Out	140
Scaling Up	141
Server Price and Performance	142
Wintel Servers Challenge Domination of Big Three	142
Windows 2000 Datacenter Server	143
Mid-Range Hardware for Windows 2000 Datacenter Server	144
Unisys ES7000	144
Hewlett-Packard ProLiant DL7602G	145
IBM x440	146
Dell PowerEdge 8450	146
Fujitsu Siemens PRIMERGY N850	146
Fewer Servers	146
Server Consolidation at La-Z-Boy	146
Limits to Consolidation	147
When Consolidation Makes Sense	148
13 Windows Past, Windows Present: Windows NT and	
 Windows 2000	149
Windows NT	149
Windows NT in Action	150
NT's Days Are Numbered	151
The Windows 2000 Server Family	151
Windows 2000 Server	151
Certified for Windows 2000 Logo	153
Windows 2000 Advanced Server	154
Windows 2000 Datacenter Server	154
Datacenter Server User Study	155
The Rigors of Windows 2000 Migration	158
The Three Active Directory States: Native Mode, Mixed Mode,	
and Parallel Mode	159
Windows 2000 Migrations Tools	161
Making the Transition to Native Mode	161
Consolidating Directories	162
Windows 2000 Migration Case Studies	163
Cincinnati State Technical and Community College	163
CVS Corporation	164
Pioneer Hi-Bred, Inc.	164
Mt. Sinai NYU Health: Managing AD in a Hybrid Network	165
No Universal AD Panacea	165

14	Windows Present: Windows XP	167
	No More DOS	167
	XP Fastboot	169
	Prefetching	170
	Layout Optimization	171
	Boot Optimization versus File Optimization	171
	XP Fast Resume	172
	Tips To Improve the Impact of Fastboot and Fast Resume	173
	Underlying XP Architecture	174
	Beware XP's Consumer Features	176
	The XP Windows Media Player	176
	The XP CD Burner	177
	The XP Defragger	177
	XP Security Features	178
	XP Task Manager	179
	XP Performance Tool	180
	XP Quotas	180
	XP Backup	181
	Backward Compatibility of XP	181
	Extending Hardware Lifespan	181
	Improved XP Management	182
	Only Room for So Much	182
15	Windows Present and Future: Windows Server 2003, 64-Bit Computing, and Beyond	185
	Microsoft's .NET	186
	What Is .NET?	186
	.NET and Web Services	187
	The .NET Starting Roster	187
	Case Study: Making the Case for Information Sharing among Police Departments	190
	Nothing but .NET: What Are the Alternatives to .NET?	192
	.NET Slam Dunk	194
	Windows Server 2003	194
	XML Web Services and Windows Server 2003	196
	Members of the Windows Server 2003 Family	196
	Standard Edition	196
	Enterprise Edition	196
	Datacenter Edition	196
	Web Edition	196
	64-Bit Windows	197
	64-Bit at a Glance	197
	The Next Few Years of Windows: Longhorn, BlackComb, and XP Second Edition	198
	XP Second Edition	199
	Longhorn and BlackComb	199
16	Asset Management for Server Hardware and Software	201
	The Asset Management Explosion	201
	Hot Spots	202
	Piracy and Asset Management	203
	Asset Management Strata	204
	System/Desktop Management Platforms	204

Asset Management Suites	206
Asset Management Utilities	206
Microsoft Systems Management Server	206
Simplifying Asset Management.....	208
17 Software Deployment and Distribution on Windows Platforms.....	211
Deployment Complexity	211
Security Breaches	212
Securing the Walls.....	213
XP Patches	213
Automated Software Deployment.....	214
Frameworks and Software Deployment.....	214
Microsoft Tools and Software Deployment.....	215
Site Licensing/Tracking Software	215
Software Deployment Case Study	217
Do You Feel Lucky?	219
18 The Future of Disks and Disk Management.....	221
The Storage Industry Is Healthy	221
Faster Data Transfer.....	222
Bigger and Better Drives	223
Recent Developments	223
Fujitsu Drives.....	224
2.3-GB MO Drive	224
2.5-Inch MHS Mobile Disk Drive.....	224
15-K MAS Series 3.5-Inch Disk Drive	225
Upcoming Fujitsu Read Head and Media Technologies.....	225
What the Future Holds	225
Conclusion.....	226
Appendix A: Glossary of Terms	229
Appendix B: Bibliography	237
Books.....	237
White Papers and Reports	238
About the Author	239
Index.....	241

Preface

Disks and disk health are given scant attention in the broad scheme of things. As you look over the computer press, you read plenty about CPUs, RAM, networks, integration, and the server wars. Hard disks barely seem to merit a mention, except perhaps once or twice a year whenever a bigger or faster drive comes out.

Hard Drives: The Biggest Source of Failure

Yet, disk drives account for more server failures and repairs than any other factor. According to the 2002 *PC Magazine* reader survey, hard drive repairs are by far the number one reason for server downtime among Windows-based servers (this holds true for UNIX servers, too, by the way). When several thousand server users were asked to name a specific cause of failure in servers, the results were surprisingly consistent (Exhibit 1). As you can see, hard drives account for half the reported failures. Other surveys have recorded similar findings, so it seems that more needs to be written about hard drives and disk management in general.

Why Windows Servers?

When it came to writing this text, I decided to stick to Windows-based systems. Why? Although Microsoft has about a 95 percent share of the desktop market, its server dominance has sneaked up on people quietly but forcefully. Currently, some surveys show Microsoft servers outnumbering other server operating systems by about 6 to 1. And, as Windows servers are still maturing, a disk management book in this area seemed to have more value.

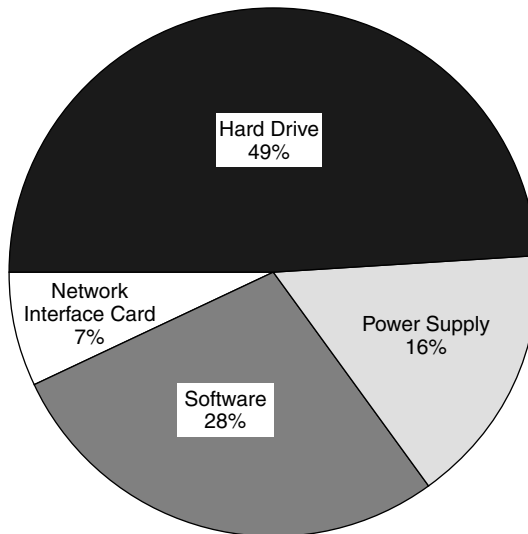


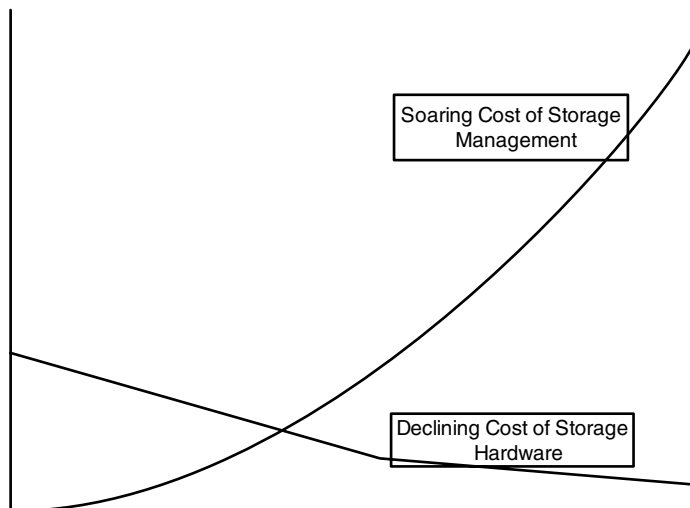
Exhibit 1 Causes of Server Failure

Exhibit 2 Storage Management Costs

Storage Management Costs

Another factor that drove the writing of this text was the terrific rise in storage management total cost of ownership (TCO) over the past decade. Ten years back, storage hardware accounted for the bulk of total costs, with management taking up only a small percentage of the total. The lines crossed around 1996 (Exhibit 2). Nowadays, the hardware accounts for a relatively small percentage of the costs, with management being by far the dominant cost. Think about

it for a moment: Management costs include people, software, security, facilities, energy, hardware, and many other expenses. As the size and complexity of distributed computing environments have multiplied, so has the amount of management required. So, it appeared to be time to give the subject of disk management and, in particular, server disk management a little more attention.

What This Book Is *Not*

That said, there are a great many things that this book is *not*. An abundance of fine books are out that cover all sorts of related technical subjects in tremendous detail, and that is not the purpose of this text. Let us look, then, at what else this work is not:

- It is not a hardware treasure trove of details concerning hard disk hardware and configuration.
- It is not a detailed rundown of every command, GUI, and tab involved in every conceivable action an administrator needs to manage disks on Windows 2000.
- It does not go into intricate detail about all the Windows internal type of material that only the highest level technical staff can understand.
- It is not a compendium of storage management and storage networking information.

In reality, dozens of works in existence already cover these subjects adequately. In the bibliography at the end (Appendix B), I have endeavored to list a few of these texts. They go into far more detail, and some demonstrate far more knowledge of these areas than I will probably ever come to possess.

What This Book *Is*

So, rather than try to compete with some of the most brilliant minds in the field, in this book I instead decided to cover areas that appear neglected in modern texts. This work, therefore, stays focused on disks and disk management, particularly with regard to servers. What are the basic elements of a disk and various disk architectures? How do you manage the disks you have or plan to have? How do you keep them running well with a minimum of fuss? The book also discusses such key areas as:

- Windows file systems and operating systems
- Backup
- Disaster recovery
- Basics of disk management on Windows 2000

Perhaps even more importantly, though, it takes up areas that are paid scant attention among the books available on the market — vital areas such as:

- Disk quotas
- Fragmentation
- Optimization
- Hard drive reliability
- Asset management and software deployment
- Forensics

Vital areas, yet they are given a page or two at best in current literature. This book fills this gap with chapters addressing each of these areas, written in such a way as to tie them all together as essential ingredients of sound disk management.

In addition, I have pulled together a couple of more speculative chapters towards the end, going over disk and disk management trends that are likely to affect Windows-based systems in the near future. In addition to attempting to predict where hard drives themselves are headed, the book looks at the future for storage and how storage issues will affect those managing the ever-increasing number of disks at work in the modern enterprise, as well as taking a look at where the Windows platform is going with a chapter on .NET and 64-bit computing. Also included is a discussion of what Microsoft plans for the next few years on the server and desktop side.

Intended Audience

This book has been prepared in such a way as to provide value for IT veterans as well as less technically savvy individuals. For students, business executives, or less experienced IT staff, it gives plenty of definitions and explanations of the basics, including chapters on disk organization, file systems, and basic disk maintenance actions. In order to make it more accessible, technical words are defined and key concepts explained (and a glossary is provided in Appendix A). Veteran IT staff can skip those explanations and still find plenty of useful information on keeping disks rolling (or spinning) in the most efficient way and for maximum performance. But, be warned, IT vets. It never hurts to go over material again, and you should be wary of the dangers of thinking you already know everything. At the very least, the chapters on hard disk reliability, fragmentation, and optimization, as well as the trend chapters at the end, should provide plenty of food for thought and, more importantly, some material you can use to do an even better job.

Overall Direction

The overall direction that this material takes, then, is determining how best to manage and optimize the performance of the multitude of hard drives that exist within the enterprises of today. Important issues include server disk management, managing storage assets, backing up, managing free space, managing fragmentation, maintaining high disk performance, guarding against

and preventing disk failure, and future developments with regard to hard disks. It is also important to note that this text concentrates entirely on Windows-based operating systems and file systems. Within that context, it speaks primarily about Windows 2000, as that is now the primary Windows server platform; however, the book also covers Windows 9x, NT, and XP in brief, as well as Windows Server 2003. In terms of file systems, NTFS is emphasized, but FAT and FAT 32 are also covered in some detail.

Software, Not Hardware

To further define this work, it takes a software approach and gives little attention to the hardware maintenance side (again, hardware is aptly covered in many other places, as listed in the bibliography section in Appendix B). I may reference some of these books but do not try to replace them. That said, the book does cover basic definitions and explanations of hard drive hardware in more than enough detail for the purposes of this text. In fact, an entire chapter is devoted to hard drive hardware and the different types available. In summary, then, this book has been written for those who want to know more about how to maintain their disks in optimum health, monitor them, and get the best out of them. It also highlights how IT managers can improve the efficiency of hard drives in the enterprise and how to manage them better.

Chapter Organization and Content

It all begins with an overview of the subject of server disks and their management in the Introduction. This section talks about the explosion in drive size and overall storage capacity and why that makes it more imperative than ever to take charge of server disks and disk management.

Chapter 1 covers the fundamentals of what a hard disk is and the basic definitions of terms such as cluster, tracks, platters, head, arm, spindle, controller, seek, caching, rotational latency, SCSI, IDE, and more. It includes plenty of charts and diagrams to make things clear. Though some readers will know all of this, some may be surprised to find at least a couple of basics concepts that have so far evaded them. By understanding *all* the basics, more understanding results. This translates into a greater ability to manage disks throughout the enterprise.

Chapter 2 goes over the various file systems for Windows, including FAT, what it is, and how it works, as well as NTFS and what it is, how it works, and how it compares with FAT. The chapter then discusses how data is arranged in these file systems and the various problems that can arise. I also address special file types such as the Master File Table (MFT).

Chapter 3 gets into an interesting area. Unexpected disk failure is an expensive proposition. It is not simply the cost of the drives themselves, but rather the cost of technical personnel time to replace and reload the disk that should be considered. Add in the user downtime, value of lost information, and, in the case of externally facing applications, lost customers and the

amount of damage skyrockets. This chapter investigates the frequency of disk failure, finding it to be a far more common experience than many realize. It also discusses one of the most common measures of disk failures — MTBF — and debunks the myth that today's high MTBF ratings equate to disks that will last for half a century. Finally, it gives some tips, strategies, and tools that can be used to guard against disk failure.

Chapter 4 investigates the basic principles of disk management, including the difference between basic and dynamic disks, the difference between the volume types, and the various types of fault-tolerance schemes.

Chapter 5 encompasses the importance of backup. It gives a brief rundown on how backups work, how backups should be organized, the media involved, tools to use to speed backups, etc. The chapter takes a look at strategies for doing backups and restoring the data as needed and goes into depth regarding the backup functions that are preinstalled with Windows, as well as the use of third-party software, mirroring, and off-site replication. The chapter also addresses the special needs of backing up Linux in heterogeneous environments.

Chapter 6 explains the importance of disaster recovery planning (DRP) and how this ties into responsible disk management. DRP actually goes way beyond RAID and backup, and this point must be thoroughly grasped. Further, the chapter covers a subset of DRP, known as the contingency plan (CP), which addresses non-catastrophic failures due to accidental deletions, disk failures, and other potential data loss scenarios. By creating and implementing a CP, organizations are more able to eliminate significant amounts of data loss smoothly and relatively effortlessly.

Chapter 7 covers a tool — defragmentation — vital to keeping systems stable, improving disk performance, and maintaining them properly. It investigates what fragmentation is, the impact it has, safety considerations with regard to it, and the best ways to defragment networks. It also discusses manual defragmenters (such as the one built into Windows 2000), comparing them with third party products and showing benchmarks of their performance.

Chapter 8 explains the subject of disk optimization, which is said to reduce the time it takes to recover data from disks via intelligent placement of files in order to minimize head movement. But, how effective is disk optimization in the real world? Some say it produces great gain, but I am not so sure. For the average consumer on a low-end PC, optimization appears to produce some benefit. Where it may fall short, however, is in the enterprise. This chapter explains why.

Chapter 9 is all about another aspect of disk management that receives little attention in computer texts — disk quotas. Even though disk space seems to be nearly infinite (i.e., the amount of disk space and the sizes of disks seem to be getting bigger and bigger), the fact is that disk quotas are more important than ever. Pay no attention to disk quotas and in a short time your enterprise will be chock full of MP3s and other miscellany, and users will be screaming for more space.

Chapter 10 is all about the amount of hidden data that exists on disks and in documents. The area of computer forensics (recovering data from computers

for use as evidence in a court of law) has grown in recent years. This chapter explains some of the basics of computer forensics, some of the main sources of hidden data, ways to retrieve it, and methods to get rid of it.

Chapter 11 covers changes being made by Microsoft and some of its partners that will make disk management in general much easier, particularly with regard to the management of storage assets. In particular, the chapter focuses on shadow sets and storage virtualization for Windows-based servers.

Chapter 12 addresses the trend toward consolidating multiple commodity servers into fewer large Windows boxes. It includes a discussion of scaling out the number of servers versus scaling up, which strategy to employ when, and the value of each approach. It also covers the ways to consolidate Windows-based servers, the hardware and software available, and the disk management advantages that can be realized; for example, instead of managing thousands of small disks, it is possible to streamline server operations into a few dozen large disks.

Chapter 13 begins with a brief discussion of the early days of Windows, then gets into Microsoft's first foray into server systems — Windows NT. Within a year or two, NT will begin to disappear as the vendor no longer supports it. Windows 2000 already dominates the Windows server landscape, and that domination will last another couple of years. So, lengthy coverage of Windows 2000 was essential to this book. This chapter goes into the basics of the Windows 2000 Server operating system, its various flavors (such as Windows 2000 Advanced Server and Windows 2000 Datacenter Server), and some basics on the Active Directory. Without an understanding of the Active Directory, your disk management skills are not going to get you very far on Windows 2000.

Chapter 14 is all about the main Windows desktop system of today — XP. Why have a chapter on XP if this book is more about server systems and server disks? Quite simply, it is impossible to talk about disk management without also getting into the desktop side of the equation. System administrators dealing with server disks must know all about XP and XP disk management, as they are going to encounter a lot of machines running the XP operating system. This chapter covers the essentials of the XP operating system, what is new, what is good about it and not so good, and how to manage it.

Chapter 15 covers upcoming developments. By the time you read this text, many of the items covered will already be in general release. Windows Server 2003 promises to add significantly to the gains made by Windows 2000 and certainly offers greater disk management capabilities than ever before. 64-bit computing will take a few years to become commonplace, but eventually 32-bit systems will become a thing of the past, as 16-bit systems are today. The chapter also investigates where Microsoft is going over the next few years in terms of server and desktop operating systems.

Chapter 16 covers an essential area of disk management that is given scant attention — asset management. Many organizations today have tens of thousands of hard drives and hundreds of applications running, yet they have no idea what machine runs what type of disk, what software is where, what

version is running, or how all of this ties in to current licensing purchases. This subject, then, is not only an essential element of disk management, but it is also a vital organization basic when it comes to protecting an enterprise against attack from policing agencies that are increasingly on the prowl. Just last year, the Business Software Alliance levied big fines amounting to tens of millions of dollars.

Chapter 17 covers the ins and outs of distributing software in heterogeneous environments. Software deployment, or the lack of it, is actually one of the biggest sources of weakness in today's networks. A threat emerges and a patch comes out, but only a small percentage of companies actually distribute the patch in a timely fashion. This chapter, then, covers the threat and the steps that must be taken to avert it.

Chapter 18 asks such questions as: How prepared are you to meet the disk needs of the future in order to keep up with storage demands? Where exactly is the industry going and how big can disks get? Will disks be around forever? This chapter investigates what is on the disk management horizon and perhaps just beyond.

How To Read This Book

The chapters have been laid out in a general sequence that should make sense for most readers; however, they have also been written in such a way that if you want to find out about a specific subject, you can go into the relevant chapter and read it without much reference to earlier material. In whatever way you read it, though, heed this point: Make sure you know the meaning of the basic words being used. Contrary to common belief, it is bad practice to try to figure out word meanings from the general context. You will get it wrong more than half the time and end up with an incorrect idea. So, use the glossary at the back and crack open a dictionary or computer dictionary when necessary to clear up the definition of terms being used. That is the surest way to guarantee full comprehension.

Good reading!

Introduction

This section discusses the explosion in drive size, overall storage capacity, and why it is more imperative than ever to take charge of server disks and disk management.

Taking Hard Disks for Granted

Stripping away the tremendous complexity that can exist in regard to disks, the basic purpose of disks is very simple — to store data. That is exactly what hard disks have been doing since the 1950s when IBM worked out how to store 100 K on a 12-inch platter. This apparently tame event was actually a serious breakthrough, as it marked a move away from the cumbersome punch cards that had previously been the standard medium for data storage. Yet, despite the enormous expansion in disk capacity over the past fifty years, disks largely tend to be taken for granted. For many they are sort of a “black box” item. They spin away, buried inside little beige boxes that are usually gathering dust under desks or hidden in server closets. Most system managers hope they keep on going until it is time to replace all the hardware for the latest batch of bigger and better machines. In most people’s opinion, disks are the last thing they want to worry about. And, anyway, if something does go wrong, you can just put in a new one ... right?

The Disk Is the Bottleneck

This out-of-sight, out-of-mind attitude toward disks, however, may be misguided. Consider the mechanics of disks with regard to one very important factor — system performance. A 2-GHz CPU means 2 billion operations per second. RAM operates at 20 to 30 million operations per second, sometimes faster. Now factor in the hard disk. Good ones take about 5 milliseconds to access data — that is only 200 operations per second (Exhibit 1). Some can

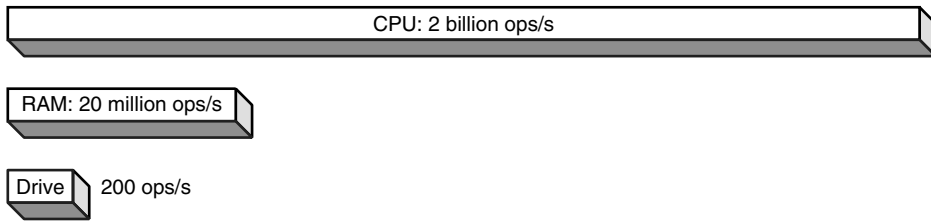


Exhibit 1 Relative Speed of Components

go faster, but not much faster. The point is that poor disk maintenance and disk management can mean that big powerful systems are held up badly by the sluggardly nature of the hard disk. So, it makes sense to know about disks and to be able to take action to avoid disks becoming a bigger bottleneck than they already are.

Disks Do Fail

After all, disks do fail. When an important one fails, IT has to scramble to replace it, while trying desperately to recover some, if not all, of the data, and this scenario plays out far more often than some manufacturers would like you to believe. A survey by San Jose-based Survey.com of 1293 IT staff and business executives, for example, revealed that the majority had experienced computer downtime in the previous year due to disk drive failure, and 30.3 percent of the time the computer was down for more than 24 hours. In addition to impacting IT staff and employees, one third of respondents said that these failures shut out customers. Thus, as well as the cost of the disk and labor, one must also factor in lost user time and the amount of business lost when customers are impacted by the failure. This survey discovered that most system administrators managed up to 500 disks on their Windows NT or Windows 2000 servers. The main actions performed to manage and protect data on the hard drive are backup, defragmentation, and hard drive replacement. When asked for the top problem with disks, most system administrators said drive failure and crashes, followed by running out of disk space.

Storage Explosion

This situation is made even more serious by the explosion in demand for disk space. According to META Group, Inc. (Stamford, Connecticut), storage capacity experienced 60 to 80 percent growth in 2002 and was expected to show annual growth rate continuing through 2006. Over the past year, in fact, more data has been recorded by mankind than in the entire rest of our history. To meet this demand, technology is utilizing two approaches: designing larger and larger disks and making it far easier to use more of them. The capacity of individual disk drives has already surpassed 180 GB, and IT research firm

Gartner Group (Stamford, Connecticut) foresees it hitting 1.44 TB in two years. Recent developments support the belief that bigger drives and more capacity are trends that will continue for quite some time. An industry group consisting of the likes of Maxtor, Microsoft, and Hewlett-Packard has announced a new ATA interface standard called “Big Drives” that allows drives to scale up to 144 petabytes. Seagate Technologies has also developed technology that permits a single 3.5-inch platter to hold up to 125 GB, three times the current level. And, the release of Microsoft’s 64-bit Windows operating system ups the ante by making it possible to manage up to 16 TB of addressable virtual memory.

In addition to more and bigger disks, they are also being compressed into smaller and smaller spaces in terms of the number of disks per square server rack. The big original equipment manufacturers (OEMs), including Dell, Hewlett-Packard, IBM, and Sun, have followed the lead of, for example, nStor Technologies, Inc. (San Diego, California) in producing rack-dense servers. Take the case of the Compaq QuickBlade, which allows enterprises to fit 200 to 300 servers in a single rack, rather than the usual 42 servers.

Desktop Disks

It is important to note, though, that these trends are not limited to just the server market. A friend still uses my old desktop from five years ago, complete with its 500-MB drive, and it works just fine for what she uses it for — e-mail and word processing. These days, though, some desktop drives can exceed 100 GB, and multiple drives are becoming more and more common, so what is the upshot of this hard drive explosion? Manufacturing quality is hard pressed to keep up with the explosion in the sheer number of hard disks deployed and the quantity of data they hold. A recent white paper from a hard drive vendor was frank enough to say that the increase in the number of drives has resulted in a dilution of overall system reliability. It goes on to state that given the growth in disk sizes, each potential drive failure now puts more data at risk. So, ensuring the reliability of each drive has become more important than ever.

Neglect disks and disk management, then, at your peril. In today’s disk-intensive world, it is vital for organizations to take care of the basics of managing space quotas, monitoring disk performance, organizing disks and partitions properly, planning for future storage needs, taking care of backups, having an adequate Disaster Recovery Plan, and keeping disk I/O low by keeping volumes free of fragmentation. And that is what this book is all about. It will not turn you into an expert in every aspect and activity of disk management or hardware maintenance, but it will certainly give you a good overview of the subject and enough familiarity to make more qualified decisions when it comes to the overall health of your server disks and how you should best go about managing them.

Chapter 1

Hard Disks and How They Are Organized

While disk drives have been around since the 1950s, like anything else the technology keeps evolving. As a result, the assortment of options on the market available for data storage is expanding. Selecting the best disk drive for any particular operation and keeping it running at its best require an understanding of the hardware components, as well as how the data is stored and the terminology used by the manufacturers in describing their products. Only then can a correct comparison be made. This chapter provides details about the basic components and architecture of the hard disk and methods to improve its performance.

An understanding of the basic components of the hard disk and how they operate together — as well as the different paradigms for storing data and factors affecting disk performance — makes it easier to select the optimum storage solution for a particular project or installation. Following is a discussion of the basics of hard disks, including their physical components, how the data is stored, and factors that should be considered when comparing disk performance. Also examined are several tried and tested strategies that will help to improve overall drive performance.

Hardware Components of the Hard Drive

Any hard drive has several major components. Below, we briefly outline the core components; other items, such as cabling and connectors, are not covered here.

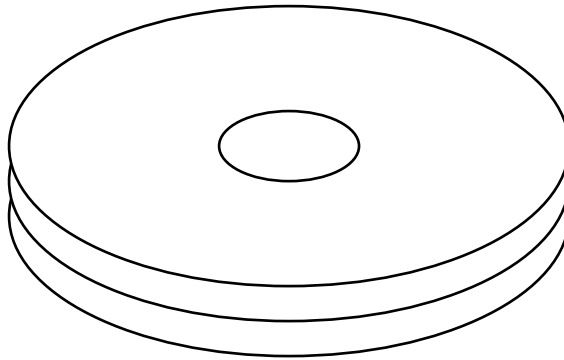


Exhibit 1 Platters

Platters

While the terms *hard drive* and *hard disk* are used interchangeably, hard drives typically contain multiple disks, or platters, stacked on top of each other (Exhibit 1). Personal computer (PC) hard drives generally have one to five platters, while servers will have up to a dozen. (The first disk drive, IBM's RAMAC, had a stack of 50 24-inch platters, each holding 100 K.)

The platter is the element that actually stores the data. It consists of a substrate coated with magnetic media. The substrate, made out of a nonmagnetic material, is there to act as a rigid support for the magnetic media. Traditionally, alloys made of aluminum have been used for hard drives, but to meet the demands of higher performance drives the industry is gradually transitioning toward the use of glass, magnesium, and other substances that provide greater rigidity, smoothness, and overall thermal stability. Switching to these newer materials allows manufacturers to create thinner disks and, perhaps more importantly in the long run, to operate them at higher speeds.

On top of the substrate goes a coating of magnetic media. Originally, this was a thin layer (less than 1 micron) of iron oxide. More recently, however, manufacturers have been replacing this by applying an even thinner coating of magnetic material using vacuum deposition, the same process used in manufacturing computer chips. The magnetic layer is then covered by a protective layer of carbon and a lubrication layer to prevent damage in case the head comes in contact with the platter surface. In most cases, both sides of the platter are coated with magnetic media. Why? The goal in coating both sides is to increase the “areal density” of the platter — the amount of data that can be stored in a given amount of space. Increasing areal density is what has continuously increased the capacity and brought down the prices of storage. A 3.5-inch floppy, for example, is the same size as the typical hard disk platter, but hard disk platters typically hold about 30,000 times as much data — 40 gigabytes (GB) — due to greater areal density, and this capacity will certainly increase even further in the coming years.

In 2001, for example, Seagate Technology demonstrated a drive with an areal density of 100 gigabits (Gb) per square inch. Such disks would make it

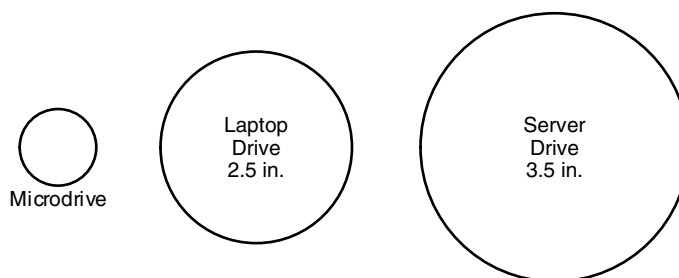


Exhibit 2 Disk Sizes

possible to store 125 GB of capacity on a single 3.5-inch platter. While drives built using that technology were not commercially available at the time of this writing, in June 2003 Western Digital released a 250-GB drive with over 80 GB per platter. It is only a matter of time before 125-GB disks are widely available on the market.

Interestingly, as areal density has increased, the size of platters has dropped. The typical server and workstation now have drives with a nominal 3.5-inch form factor, and the platters themselves are about 3.74 inches. This configuration has several advantages. Smaller platters are more rigid than larger ones, allowing the read/write heads to float closer to the surface. Also, the disks can spin faster, resulting in more rapid data access. High-performance drives, with speeds of 10,000 to 15,000 rotations per minute (rpm), achieve this speed by using even smaller disks (2.5 to 3.0 inches), thus trading lower storage capacity for quicker data access. Laptops, for example, typically utilize 2.5-inch platters, and the platter in IBM's Microdrive, which is used in devices such as digital cameras, is a mere 1 inch yet holds up to 1 GB (Exhibit 2).

Spindle/Motor

The platters are attached at the center to a spindle that is directly attached to the shaft of the motor. No belts or gears are used, as they could result in slippage and rotational speed variability; instead, the motors use a closed-loop feedback system to precisely control the speed.

Shock and vibration can cause the read/write heads to slap the disk surface, resulting in lost data and other damage, which is why spindle motors are specifically designed for minimal vibration. Even so, vibration is still a problem for manufacturers, one that worsens as rotational speeds increase. Newer rack-dense server installations, for instance, are beginning to experience vibration-related hard drive problems.

Fortunately, vibration is less of an issue in server installations thanks to some recent technological breakthroughs. In July 2002, for example, IBM released a new technology called Rotational Vibration Safeguard (RVS), which is designed to improve performance in large disk arrays by sensing the direction and intensity of the vibration on a disk and canceling it out. The first product to use RVS is IBM's Ultrastar 146Z10, a 148-GB drive running at 10,000 rpm.