

Making Everything Easier!™

2nd Stratus Special Edition

Availability

FOR
DUMMIES[®]
A Wiley Brand

Learn:

- How downtime can hurt your organization's bottom line
- What kinds of availability options your organization has
- How virtualization and the cloud can impact availability

Compliments of

Stratus[®]

Technologies For an Always-On World

Jessica Devine



Availability

FOR
DUMMIES[®]
A Wiley Brand

2nd Stratus Special Edition

by Jessica Devine

FOR
DUMMIES[®]
A Wiley Brand

Availability For Dummies® 2nd Stratus Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

Copyright © 2014 by John Wiley & Sons, Inc., Hoboken, New Jersey

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Trademarks: Wiley, For Dummies, the Dummies Man logo, The Dummies Way, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. Stratus, ftServer, and everRun Enterprise are registered trademarks of Stratus Technologies. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

ISBN 978-1-118-94174-4

ISBN 978-1-118-94221-5 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. Some of the people who helped bring this book to market include the following:

Development Editor and Project Editor:
Chad R. Sievers

Acquisitions Editor: Kyle Looper

Editorial Manager: Rev Mengle

Business Development Representative:
Susan Blessing

Custom Publishing Project Specialist:
Michael Sullivan

Project Coordinator: Melissa Cossell

Introduction

When was the last time you had an “out-of-service” experience and gladly continued to do business with the provider? If your bank’s ATM was temporarily out of service more than, say, twice a month, would you recommend it to a friend? Or would you trust the secured payment page of an online retailer whose server occasionally went down while you were making your purchase?

Admit it: Your expectations for technology to work as it should have skyrocketed over the past five or ten years, to the point that you have almost no tolerance for any sort of noticeable, regularly occurring outages.

Well, you’re not alone. Many consumers and most (if not, all) businesses have no patience for downtime, and they’re not afraid to let the offending party know it. In a global, Internet-driven economy, consumers have more choice than ever before. If your customers have a frustrating experience due to downtime even once, they’re likely to take their business elsewhere — and your competition will be ready and waiting.

To protect your business, you need to understand the true importance of availability, your options for protecting availability, and your business needs. That’s where *Availability For Dummies*, 2nd Stratus Special Edition, can help.

About This Book

Having your applications available most of the time might sound acceptable, or even great. However, availability is one of those instances when “most of the time” just isn’t good enough.

Imagine if your car randomly wouldn’t start 15 times a year. If you run the engine about four times a day, that’s 99 percent

availability. And on the days your car started and you got to work, what if, for a total of 87 hours each year, your office building randomly had no power? That too, would be 99 percent availability.

Suddenly 99 percent isn't looking that great, is it?

Availability For Dummies, 2nd Stratus Special Edition, aims to educate you on the true importance of availability. So many organizations assume that because their applications are up and running “most of the time,” they're protected from downtime catastrophes. Unfortunately this isn't true. *At all*. And it will become even less true in the coming years. This book explains why and exactly what you can do about it.

In this book, you can discover the true cost of downtime to your organization, from demanding and flighty customers to the true impact on your company's wallet and reputation. Then we delve into your availability options and how to pick the one that best matches your organization's needs. You uncover how virtualization is making everyone yearn for more availability and how the cloud fits into the picture. And finally, we quell your anxieties by explaining how solutions are available to meet your every availability (and virtualization) need — without busting your budget.

Icons Used in This Book

As you read this book, you'll notice a couple of eye-catching icons designed to highlight special information:



Flagging a tip is akin to kicking you under the table to warn you that if you're not careful, you may overlook something important that could end up costing you.



This icon reminds you of the really, really important things that can help you achieve your business continuity, revenue, cost, compliance, quality, and safety goals.

Chapter 1

Why You Need to Raise the Bar for Availability

.....

In This Chapter

- ▶ Shunning downtime in an always-on world
 - ▶ Acknowledging the needs and wants of those you serve
 - ▶ Steering clear of downtime consequences that may shock you
-

We live in an always-on world. Instant access is basically a given in our day-to-day lives. Young people in the workforce don't know of a world without high-speed Internet access. Even those individuals who remember the days of dial-up Internet won't stand for an unreliable company or service. Many good options are now available — so why would anyone want to be kept waiting?

This is why application downtime is, was, and always will be the ultimate downer for any organization. Even the words used to describe downtime and its causes are riddled with pessimistic overtones: outage, offline, unavailable, failure, crash, bug, and error. That's why the preferred term is *availability*. Availability is the optimist's way to view the ups and downs of the applications running on your servers. It's defined as the percentage of time in a given time span (for instance, a year) that your applications are operational and accessible to users.

As a result, you have to take a good hard look at your total business IT environment to figure out just how much availability you really need, where you really need it, and how much you're willing and able to pay to get it. But first you need to understand exactly what downtime costs you — and not just in terms of quantifiable dollars. You need to understand the whole picture. First up: the very people you exist for.

Knowing What Happens When Your Customers Aren't Served

No matter whom your organization serves — whether they are businesses, consumers, patients, or the general public — they demand high levels of availability. And when the people you serve don't get what they want, your organization will feel their wrath. Potential consequences include the following:

✔ **Loss of business:** Customers have more choices for products and services than ever before. So why would a company put up with a production line that goes down? Why would a city and its residents tolerate unreliable emergency services? Basically, why would anyone put up with downtime when so many other options exist?

No one — individuals and organizations alike — has the time or patience to put up with an outage and will take their business elsewhere. And worse: They'll tell their friends and colleagues about it.

✔ **Reputation damage:** Bad publicity can cause major damage to an organization — and not just the big guys. Sure, the traditional press loves a good headline about bad news at a big company. But what can a complaint on Twitter or a negative post on Facebook cost you?

Bloomberg.com reported that one bad tweet can cost a company 30 customers. And another study by Echo Research Group found that people are twice as likely to talk about bad customer service experiences than they are to talk about good experiences.

Bad press can also devalue a company's stock and reduce its market capitalization. Especially in shaky economic times, the stock market reacts to negative press about a company, even more so if the news is about a significant sales loss — an event that is entirely possible when servers go down.

✔ **Health risks:** For certain organizations, the human impact is obviously the most serious potential consequence of downtime. Hospitals, public safety 911 call takers, and first responders all depend on applications that are managed by computer systems. Downtime of public safety answering point (PSAP) applications causes slower emergency response times and can tragically result in the worst type of loss: loss of life.





What availability really boils down to is managing the critical adverbs that define customers' needs:

- ✔ **Quickly:** Customers want your products and services — and they want them now.
- ✔ **Accurately:** Whether you're processing a financial transaction, shipping furniture, or routing a call, your customers expect pinpoint accuracy. If they don't get it, they'll let everyone know about it. And then they'll turn to your competitor.
- ✔ **Reliably:** Like Old Faithful, you have to deliver consistently and on time — morning, noon, and night.

Eyeing How Downtime Affects Employees (and Your Profit)

Downtime also takes down your employees and sneaks into your profit margin in these ways that you may not realize.

- ✔ **Indirect business costs:** *Direct costs*, such as lost wages, overtime, and remedial labor costs, all add up during an outage. Chapter 2 looks closer at these dollars.

But for now, consider the less-quantifiable indirect business costs, such as lost inventory, scrap of work in progress, and the potential legal penalties for not delivering on service-level agreements.
- ✔ **Productivity costs:** During an outage, employees can't perform their regular duties. The impact of this idle time varies by industry. For example, in an office environment, employees may not be able to access the Internet but can work on a desktop spreadsheet program, so perhaps their productivity would be cut in half. But in a manufacturing environment, if the line stops, employees may be 100 percent unproductive.
- ✔ **Recovery costs:** These costs include the price paid to repair the system that failed, IT staff overtime, and third-party consultants or technicians needed to restore services. You also have to consider the opportunity cost sacrificed when IT needs to focus on system recovery instead of working on other critical issues.

Avoiding Risky Business

The impact of *unavailability* (downtime) can vary greatly, depending on your organization. At the very least, unplanned downtime is an annoyance that reduces productivity as workers sit idle. Ill-timed outages can have lasting detrimental effects, causing temporary or permanent data loss, disgruntled customers, unproductive employees, and potentially dangerous interruptions. Consider these examples:

- ✔ The server that collects your manufacturing data goes kaput and takes an entire production batch with it.
- ✔ Your email server goes AWOL just as your biggest customer sends a request to double his recent widget order. You don't get the message on time, you lose the entire order, and your customer calls the competition.
- ✔ A police officer pulls over a car for speeding. An outage means she can't communicate whether or not the individual she just pulled over is a wanted criminal.
- ✔ Your company is part of a tightly integrated supply chain. For the second time this year, your server crashes and your production line grinds to a halt. Your company has become the weakest link in the chain — and therefore soon becomes the missing link.



You really want to avoid the pain and heartache of such events and the damage to your bottom line. Downtime is never good. In fact, your success often depends on making sure your superiors take the appropriate steps to avoid it. In worst cases, your superiors will have already heard about certain incidents from your customers. And you can bet that each customer will also tell two friends, who will tell two friends, and so on.

Despite lots of near-death experiences, many organizations still don't understand the level of downtime risk to which they are exposed and assume the coverage they have is good enough. Many never attempt to tally the exorbitant costs of their wishful thinking — until they either fade into oblivion or get wise to the fact that superior availability is a strategic asset that can help them shine among the competition.

Chapter 2

Grasping the Hard Costs of Downtime

In This Chapter

- ▶ Calculating the average cost of downtime
- ▶ Breaking down downtime costs
- ▶ Figuring out what downtime costs *you*

In June 2010, the Aberdeen Group found that a single hour of downtime cost the average organization \$110,000. By 2013, that number rose to \$163,674 per hour. Just imagine what the average hourly cost of downtime could be in 2016.

Published estimates of hourly downtime costs can be painful to see, with figures climbing into the millions across industries, such as energy, telecommunications, manufacturing, retail, and others. This chapter looks at how downtime affects some popular industries and then delves into how lack of availability can be affecting you, in particular.

Recognizing Downtime Dollars in Different Industries

Downtime can affect different types of organizations in different ways, so taking these additional factors into consideration is important when thinking about what downtime costs you.

✔ **Manufacturing:** Unexpected downtime for manufacturers can mean loss of revenue, a lower-quality product, and/or unsellable products. In some cases, a momentary

disruption in production can cause an entire run to be scrapped due to regulatory guidelines — a potentially devastating scenario and a harsh reality, especially for food and pharmaceutical manufacturers. Saying goodbye to 10,000 delicious chocolates that were sitting on the line for too long is painful for both the chocolate devotee and the company wallet.

- ✔ **Building security:** A security system works only if it's up and running all the time. Any downtime for video monitoring systems, access control, or other building security and automation systems can mean costly, dangerous, and potentially life-threatening consequences. No one wants the security systems of airports, stadiums, and nuclear power plants to suddenly go offline, because the consequences could be truly dire.
- ✔ **Retail:** The retail sector is hit hard by IT downtime. According to CA Technologies, the last available figures had losses at \$18.18 billion per year due to outages. A single downtime event for a retailer can be a huge blow to its financials, especially when such an event happens during a holiday shopping season.

When a server goes down for an online retailer, website performance is compromised, which frustrates customers and may cause them to abandon their online shopping carts and shop elsewhere. In a store setting, point-of-sale (POS) systems need to be up and running to process sales and maintain the flow of customers throughout the store.

- ✔ **Public health and safety:** Hospitals, emergency call takers, dispatchers, and first responders all depend on applications and information across multiple touchpoints to protect lives and property. Downtime isn't just annoying. It causes slower access to critical electronic health records and can slow emergency response times, impacting the lives and health of patients and citizens. No one wants to think about what could have happened if an ambulance hadn't arrived on the scene quickly, so preventing downtime when it comes to saving lives is absolutely critical.
- ✔ **Financial services:** Financial services organizations rely on transactions. Customers want to complete their business quickly and securely, whether over the Internet, by telephone, at a local branch office, through an ATM, or via debit/credit card. If a bank's online system isn't

available 24/7, its customers have many, many other banks from which to choose. Therefore, when downtime occurs, financial institutions are hit hard on a company level. A CA Technologies report states that revenue loss due to IT downtime is \$224,297 for financial services organizations each year.

Knowing What Downtime Actually Costs You

Downtime cost estimates are downright scary. But of course, they're industry averages. What does downtime really mean to you and to your business? How can you calculate your specific business risk? Your potential downtime costs?



TIP

You can develop a ballpark estimate by looking at two major side effects of downtime: loss of productivity and loss of business. You incur productivity losses when some or all of your employees are twiddling their thumbs while your server is down, and business losses when transactions are disrupted and/or customers fly the coop. So, you need to calmly, coolly, and collectively sit down and massage the numbers.



REMEMBER

To estimate the cost of lost productivity, you need to know a few things:

- ✔ How many employees would be affected by a particular server outage?
- ✔ How much less productive would they be during an outage? 0 percent? 50 percent? 100 percent?
- ✔ What is the loaded hourly salary (base + benefits) of the average employee affected by the outage?
- ✔ How long (in hours) would a typical server outage last?

Multiply those numbers together and you get the cost of lost productivity due to a single server outage event. Then (brace yourself) multiply that cost by the number of times you think your server will take a dive this year to get an estimate of annual productivity losses.



Estimating the cost of lost business can be a bit trickier, but if you have a good understanding of how your profits or revenues break down by employee or transaction, you can do it. Here are two ways:

- ✓ Multiply the number of affected employees by the percent of productivity loss, by the average profit per employee hour, and by the average duration (in hours) of the outage.
- ✓ Multiply the number of business transactions per hour by the percent of affected transactions, by the average profit per transaction, and by the average outage duration (in hours).

Your annual business loss is the single outage loss you just calculated multiplied by the expected number of server outages per year.

By adding your annual productivity losses to your annual business losses, you can get a glimpse at the impact of downtime on your own economy.

Chapter 3

Choosing Your Favorite Flavor of Availability

In This Chapter

- ▶ Knowing your availability options
- ▶ Determining your tolerance for downtime
- ▶ Serving your needs with an availability solution

Being able to figure out how to protect your business is important. Ask yourself these questions: How much downtime can you tolerate? What are the availability options? And what solution is right for you? This chapter can give you some answers.

Spanning the Availability Spectrum

Generally speaking, availability comes in three different grades: good, better, and best. From a distance, all three appear to be pretty good choices. But upon closer examination you can see that the seemingly small distinctions between categories can be critically important when protecting your precious, more precious, and most precious applications. The three major categories are as follows:

- ✔ **Standard (or Conventional) Availability:** This is your average, run-of-the-mill option that comes with a regular, off-the-shelf server — approximately 99 percent availability. That may sound good, but what it really boils down to is an average of 87.5 hours of downtime per year, or

more than 90 minutes of uninvited downtime per standard work week (40 hours). And standard servers aren't designed to prevent downtime or data loss.

- ✓ **High Availability (HA):** HA systems, which include cluster solutions and software solutions, typically provide upwards of 99.9 percent availability — an order of magnitude (ten times) greater than standard systems. That one extra nine in the tenths place makes for a lot less downtime per year: only about 8.75 hours. Some HA systems can do better than that — 99.95 percent uptime (about 4½ hours per year) or even 99.99 percent uptime (52 minutes of downtime per year).



Some HA solutions are designed to get you back up and running as quickly as possible after a failure, while others are designed to prevent downtime from ever occurring in the first place. So how do you know the difference when looking at HA solutions?

You may hear the term *failover* in conversations about HA cluster solutions. It means if one server fails, another server will take over — ideally with as little interruption as possible. With a failover cluster solution, downtime isn't prevented. Advanced HA software is different from clusters. The prevention of downtime and no noticeable failures are the goals. The difference in downtime is noticeable: less than an hour of downtime for HA software in an entire year versus almost 9 hours each year for clusters.

- ✓ **Continuous Availability (CA):** If you think four nines (99.99 percent) of availability is impressive, you're right. But it can get even better. With a CA solution (also called an *always-on* or *fault-tolerant* solution), you get five or even six nines (99.999 or 99.9999 percent) of availability. Two completely redundant servers plus software that constantly monitors system components to identify, handle, and report faults before they impact the system accomplishes this feat. What does it get you? You have only a total of 1 to 5½ minutes of downtime in a year.

Figuring Out the Right Solution for You

You may be wondering why anyone in his or her right mind would settle for anything less than always-on availability. But then your left brain (the analytical, mathematical part) kicks

in and you realize that something as special as an always-on solution probably calls for a larger budget allocation (at least, upfront) than a standard availability solution.

So you come to terms with your cold, hard budget numbers, scrutinize your business applications, determine your cost of downtime (refer to Chapter 2), and decide just how much downtime you can tolerate.

If even a short server outage could sink your business, then you need an always-on solution. Emergency services, building access control and surveillance, stock trading, air traffic control, credit card validation, manufacturing, and e-commerce, scream out for continuous, real-time computing. Even HA systems aren't enough for these business-critical applications. Why? Because the 4½ hours of downtime per year might come as a few, lengthy outages spread over the year, or as weekly five-minute outages. What's your preferred dosage of misery?

HA systems are a good fit for business applications that can endure minor disruptions and minimal data loss — but not much more. The occasional hiccup in one business-critical application won't mean the end of your business, but if you try to get by with standard availability, you're likely to suffer the consequences.

For a few business applications, standard availability might be acceptable. Of course, you'll have to resort to manual labor or patiently wait until the system is back online, which will obviously impact productivity.



Keep all this in mind when you shop around for an availability solution and consider the following factors:

- ✔ **Uninterrupted processing:** How many nines (99.9999 percent and so on) does your business environment need? Can your applications withstand prolonged outages? Failover delays?
- ✔ **Data protection:** What is your risk tolerance for losing data? Can you handle it if multiple transactions disappear into the ether?
- ✔ **Ease of deployment:** Do you have the time and expertise to make applications cluster-aware and to develop and test failover scripts? Or do you need out-of-the-box or wizard-driven solution deployment? Are you willing to

add a layer of complexity to enable virtualization? (Refer to Chapter 4 for help in answering these questions.)

- ✔ **Administration and support:** Do you have skilled IT personnel onsite to oversee the operation of your solution and fix problems at a moment's notice? Or do you need a more automated solution with minimal human intervention? What about remote monitoring capabilities?
- ✔ **Recovery process:** Do you require automated recovery or can your staff initiate recovery procedures in a timely fashion? Are you prepared to test cluster failover scripts periodically to ensure they're properly designed?
- ✔ **Total cost of ownership (TCO):** Have you considered the cost of duplicate operating system licenses (required for clusters), solution design and implementation, a dedicated disk array or storage area network (required for clusters), and ongoing administration and maintenance? What about the costs attributed to downtime (such as penalties, recouping data, loss of business, and paying idle staff)?

Clearly, the level of availability you need is only part of the total equation. Staffing dependencies, proximity to resources, the long-term impact of outages, your capital and operational expense budgets, and your risk tolerance are all important considerations.

Protecting Your Servers, Data, and Reputation

Knowing your availability needs is great, but you also have to find (or build) solutions that address them. And, further complicating matters, appropriate availability levels may vary tremendously — throughout the functional units of a single company and from business to business.

You have these choices for guarding against downtime:

- ✔ **Robust Standalone Servers:** The latest x86-based servers includes features like redundant fans and power supplies, hot-plug PCI cards, and mirrored memory, offering improved reliability over unadorned servers. They can usually deliver about 99 percent uptime, but are susceptible to catastrophic failures because more than likely they

have only basic backup, data-replication, and failover procedures in place. In the event of a crash, the server stops all processing, and users lose access to their applications and information. Standard servers also don't provide protection for data in transit, which means if the server goes down, this data is also lost.

- ✔ **Cold Standby:** Keeping a second x86 server close at hand provides a fallback option, albeit a somewhat iffy one. If the primary server fails, you need a skilled administrator to either connect the second server to a shared disk array or move disks from the primary server to the standby server. Don't expect much more than 99 percent availability with this option — and that's if your site is staffed with a capable administrator whenever an outage occurs.
- ✔ **Data Replication:** This off-the-shelf software option allows you to replicate data synchronously or asynchronously from one or more source servers to a target server. Should a source server fail, the target server takes over (either automatically or through manual intervention, depending on the product). The downside is that the target server doesn't instantly take over, so you lose any data transmitted between the last replication time and the failure time. Depending on the product and configuration, you may be able to achieve upwards of 99.9 percent uptime.
- ✔ **High Availability (HA) Clusters:** This solution aims to recover from downtime as quickly as possible (as opposed to preventing it). HA clusters are a custom-built system consisting of two or more nonspecialized servers (nodes) joined at the hip in a single network. Because there are two or more servers, you need to license software on each of them. Applications may need to be made *cluster-aware* so that when one node fails, the application automatically fails over to the surviving node using failover scripts, which require specialized skills to create. A con of clusters is that when they fail, you'll always experience a failover delay and loss of in-flight data.

HA clusters can usually get you at least 99.9 percent availability. If meticulously designed, configured, administered, and maintained by a highly trained cluster expert, you may be able to achieve 99.95 percent availability.
- ✔ **High Availability (HA) Solutions:** This type of software is designed to prevent downtime, data loss, and business interruption, with a fraction of the complexity and at a fraction of the cost of high-availability clusters. HA

solutions are equipped with predictive features that automatically identify, report, and handle faults *before* they become problems and cause downtime.

Two important features of advanced HA software are that it works with two standard x86 servers, and it doesn't require the skills of highly advanced IT staff to install or maintain it. Advanced HA software is designed to configure and manage its own operation, making the setup of application environments easier and more economical than clusters.

- ✔ **Fault-Tolerant Software:** This type of CA solution not only prevents downtime from occurring, but also provides the cost-saving benefits of fault tolerance on standard x86 servers. Each application lives on two virtual machines. If one machine fails, the applications continue to run on the other machine with no interruptions or data loss. If a component fails, the healthy component from the second system replaces it. Fault-tolerant software can also offer disaster recovery and split-site capabilities. It prevents data loss, is simple to configure and manage, requires no special IT skills, and delivers upwards of 99.999 percent availability — all on standard servers.
- ✔ **Fault-Tolerant Server Systems:** A fault-tolerant server system is truly turnkey in that its hardware, software, and services are all integrated for easy management. This type of CA solution delivers fault-tolerance through specialized servers that are purposely built to prevent failures from ever occurring.

You may think that this kind of always-on solution would be hard to manage, but surprise, it's not. FT servers are managed just like standard servers, so they don't require specialized IT personnel to manage them. Fault-tolerant server systems include redundancy of components and error-detection software. Automatic fault detection and correction is engineered into the design so that most errors are resolved without you even knowing they existed. Plus, if you utilize VMware for a virtualized environment, fault-tolerant servers can run it.

You can expect five nines or better (99.999 percent) availability performance for your applications if you choose this option.

Chapter 4

Virtualization and the Cloud: Their Impact on Availability

.....

In This Chapter

- ▶ Understanding what virtualization is and why it's so great
 - ▶ Recognizing the importance of availability in a virtualized environment
 - ▶ Assessing virtualization and the cloud: What's the difference?
-

Server virtualization is putting a new twist on availability planning. The benefits of virtualization are numerous, and that's why many organizations are moving to a virtualized environment.

But there's a slight catch: If you have more applications running on fewer physical servers, you're putting all your eggs in one virtual basket. If that basket falls, it takes out all those eggs. And that means you have to employ an ironclad availability solution to protect those precious eggs. This chapter explores why.

Discovering What Virtualization Is and Why It's So Fantastic

The practice of virtualization enables you to run multiple applications on a single piece of hardware via an additional software layer called a *hypervisor*. The hypervisor allows each



application to interact with the server as if it has complete control of the server hardware.

The benefits of virtualization are huge:

- ✓ **Cost control:** Server virtualization allows your organization to save money because you can consolidate a number of applications on the same physical server. You can reduce significantly the capital costs and operating costs for a data center with virtualization. The energy savings alone can be dramatic.
- ✓ **Higher productivity:** With virtualization, you gain an IT environment that is easier for IT staff to manage. You don't need to maintain physical hardware, more applications are run on fewer servers, and you eliminate server sprawl (something that IT gets pretty excited about).
- ✓ **Better long-term planning:** Virtualization allows you to extend your software's longevity because legacy operating systems can be successfully run on today's computers. As a result, you can extend the life of a useful computer system and not be unexpectedly forced to upgrade when a vendor introduces a new version.

Altering Your Reality by Going Virtual

Most organizations are moving to virtualization because of the undeniable benefits (see the preceding section for those benefits). Most organizations start by virtualizing smaller or less important applications. What makes IT pros hesitant to virtualize everything is that virtualization means putting a lot of faith in your availability solution.



As soon as you depend upon virtualization, your server has the potential to be a single point of failure for all the applications it supports. So even when you take a bunch of noncritical business applications and pile them on a single server, you may be unwittingly lowering your tolerance for downtime.

Think about it: It's one thing when your HR system is offline for an hour or so, but add a few other ordinary applications to the same outage, and you may as well send everyone home for the day. And if you do put business- or mission-critical applications on a virtualized server, you had better have a continuous availability plan for them. Basically, when a single outage can pretty much cripple your operations, you have to raise your availability expectations.

Comparing Virtualization and the Cloud

Virtualization and the cloud have a lot of similarities, but they aren't the same thing. These sections examine the ways that virtualization and the cloud are related, and why that relationship really matters.

The cloud isn't happening without virtualization

Virtualization is software that allows you to run many applications on one piece of hardware, also known as a server. Cloud computing is actually a service that depends upon virtualization. A *cloud service provider* can offer you shared computing resources because it has enormous data centers full of virtualized servers. Alternatively, big enterprises can use virtualization to build their own private cloud.



But whether your own IT group depends on virtualization in-house or if you're thinking about moving to a cloud service for some or all of your applications, you're ultimately depending upon virtualized servers to run your applications. Do you trust your own servers enough to virtualize critical applications on them? You shouldn't, unless you have an always-on availability solution for them. Is it possible to move critical applications to a vendor's cloud or even to your own private cloud?

The cloud has high availability limitations

Many organizations are moving to cloud computing for at least some of their applications. Some have understandable hesitation about moving important applications to the cloud because the cloud usually can't offer high availability, let alone continuous availability. If a cloud goes down, you'll often get reimbursed for that downtime, but not for the value of the data you lost. So if your company has anything in a cloud right now, it's likely data and applications that can handle the downtime associated with 99.9 percent availability, which is what a typical cloud provider offers.

Clouds face many problems concerning availability because cloud infrastructures are usually built on large numbers of commodity systems for scalability to keep down hardware costs. In these environments, IT managers assume that components will fail. That's not exactly a place you want to keep business-critical applications, right?

Plus, delivering the right level of availability at the right time for each application isn't possible for existing clouds. Not all applications need continuous availability all the time, so assigning a fixed availability service level without factoring in the usage patterns of each application simply isn't cost-effective.

So can the cloud offer high availability or continuous availability? The answer for most clouds is no. But there is good news: There is definitely such a thing as an always-on cloud. Refer to Chapter 5 for some reasons to use an always-on cloud.

Chapter 5

Ten Advantages of a Stratus Always-On Solution

In This Chapter

- ▶ Achieving continuous availability through software
- ▶ Taking advantage of the always-on cloud

Here are ten reasons you can rely on Stratus Technologies, a leading provider of availability solutions:

- ✔ **Have close to 100 percent availability.** Whether you want to keep your standard servers or move to turnkey fault-tolerant ones, keep everything local or move to the cloud, you can affordably achieve nearly 100 percent availability for your applications with Stratus solutions.
- ✔ **Prevent downtime, not (hopefully) recover from it.** Stratus bets its business on a unique design philosophy: finding and fixing faults before they can hold you hostage. That's how Stratus achieves cream-of-the-crop availability for its always-on solutions.
- ✔ **Require no special skills to manage.** A Stratus ftServer system is built to prevent failures from occurring, yet it's managed exactly like a standard server, making the system easy to install, use, and maintain. The sophisticated back-end technology runs in the background, invisible to anyone administering the system.
- ✔ **Eliminate cluster headaches.** Because redundant components in a Stratus ftServer system are working all the time, if you have a failure, you won't suffer from it. You'll

experience no downtime, no degradation of performance, no loss of in-flight data, and no failover delay, because you don't need a failover process like clusters must have.

- ✔ **Fix errors before they happen.** The Stratus ftServer system is amazing at sniffing out both randomly occurring transient errors and hardware errors, and fixing them before they can wreak havoc. And, in the unlikely event that a hardware error can't be fixed by the system itself, on-board diagnostics figure out which hardware component is guilty, disable it before it can cause harm, and notify a human being to attend to the problem.
- ✔ **Still be able to use standard servers.** Stratus has a great option for you. Its everRun Enterprise software makes continuous availability achievable with standard x86 servers. Like the ftServer systems, everRun's design focuses on failure prevention rather than failure recovery, and it automates processes and leverages its twin hardware architecture to allow for continued operation during repairs and upgrades.
- ✔ **Protect against disaster.** The everRun Enterprise solution also protects against localized disaster with SplitSite, which provides application fault tolerance across physically separated sites. You can also mitigate disaster impact with everRun's disaster recovery capabilities, which utilize built-in asynchronous replication between sites over a wide-area network connection.
- ✔ **Get built-in virtualization.** With everRun Enterprise, virtualization is embedded so you can pile up several applications on your server pair and rest easy knowing you'll achieve the level of availability you need. Plus, everRun Enterprise gives you selectable levels of availability for each virtual machine, enabling you to prioritize based on need versus performance.
- ✔ **The always-on cloud is real with Stratus.** Want to get your legacy applications in the cloud? How about your business-critical ones? Now you can with Stratus Cloud Solutions. In fact, you can get *all* your applications in the cloud with the management platform that delivers the right level of availability at the right time to applications quickly, easily, and cost effectively.
- ✔ **You can get to nearly 100 percent availability quickly.** When you're ready to make an availability move, Stratus Technologies can get you there, quickly and cost effectively. You can find out more at www.stratus.com.

Keep connected with your customers so they keep coming back for more!

Availability For Dummies, 2nd Stratus Special Edition, explains why avoiding downtime is important so your customers have uninterrupted access to you, identifies your different availability options, and helps you pick the one that best matches your business needs. It also discusses how server fashion trends — virtualization and the cloud — are making every business year for more availability. Finally, it quells your anxieties by explaining how Stratus Technologies has a solution that meets your every availability (and virtualization) need — without busting your budget.

- ***Downtime drags dollars*** — when customers can't access your website, they will go elsewhere, so eliminating or significantly decreasing downtime will keep customers coming back
- ***Find your availability match*** — discover what your needs are to better know what type of availability can help you reduce downtime
- ***Up in the cloud?*** — cloud availability has unique traits that may be able to assist your organization with your availability needs



Open the book and find:

- **Make sense out of downtime dollars**
- **Evaluate your availability options**
- **Bridge the availability gap**
- **Examine virtualization and the cloud**
- **Identify your availability solutions**

Go to **Dummies.com**[®]
for videos, step-by-step examples,
how-to articles, or to shop!