Brought to you by:

VEEAM

Data Protection by the Numbers

for dummies®
A Wiley Brand

Effectively establish RPO/RTO and set SLA's

Assess Risk and Business Impact for IT systems

Quantify the TCO and ROI of data protection

Jason Buffington

Veeam Special Edition

About Veeam

Today's modern enterprise faces the dual challenges of not only managing and mining the data they produce and use, but also ensuring that the digital experience is always-on for both internal and external customers.

With the hyper-growth and hyper-sprawl of today's data, it's not enough for data to be backed up, secure, and available. Data must travel across five key stages (http://bit.ly/5keystages) to a new state of intelligence, automatically able to anticipate need and meet demand, securely, across multi-cloud infrastructures in order to meet the expectations of the mobile, always-on world.

As the leader in Intelligent Data Management, Veeam® is uniquely positioned to help customers along their data management journey. Founded in 2006, we are industry recognized in customer satisfaction and trusted by 80% of the Fortune 500, with more than 350K customers and 60K channel partners worldwide.

Learn more **@ Veeam.com**

Data Protection by the Numbers

Veeam Special Edition

by Jason Buffington
Author of *Data Protection for Virtual Data Centers*

for dummies®
A Wiley Brand

Data Protection by the Numbers For Dummies®, Veeam Special Edition

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

ISBN 978-1-119-57742-3 (pbk); ISBN 978-1-119-57743-0 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

Publisher's Acknowledgments

We're proud of this book and of the people who worked on it. For details on how to create a custom *For Dummies* book for your business or organization, contact info@dummies.biz or visit www.wiley.com/go/custompub. For details on licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

Some of the people who helped bring this book to market include the following:

Development Editor: Paul Levesque
Project Editor: Vasanth Koilraj
Copy Editor: Becky Whitney
Editorial Manager: Rev Mengle

Table of Contents

Introduction

Numbers make everything equal.

Quantitative metrics allow you to compare a range of availability and protection alternatives objectively, without the bias of experience or preconceptions or vendor preference. In this book, we look at several metrics. We start by defining each one and then applying them to the discussion of what kind(s) of data protection, data retention, and data availability you may need for different scenarios.

Data protection, including backup, is often regarded as an unsolvable nuisance that adds complexity and budgetary burden while remaining unreliable and cumbersome. The goal of this book is not to persuade you to purchase anything — in fact, no vendors or products are listed except for Veeam's cover logo and back cover product information as the sponsor of this guidebook. Veeam feels it's critical for you to gain an understanding of and a formulaic approach to quantifying technical recovery capabilities, business impacts, and financial concerns related to data protection.

The Rosetta Stone of Data Protection

The famous *Rosetta stone* essentially has the same information carved into one side in three different languages. Because of this, people who could read in one language could derive the translations of the other languages, thereby enabling them to communicate. Each of the three "languages" of data protection has its own terms, priorities, and three-letter acronyms — the trick is simply to be able to translate between the native tongue of each "tribe":

- **IT professionals** deal with recovery point objectives (RPOs) and recovery time objectives (RTOs) as the basis of their measurements of certain data protection tools and technologies.
- **Operations professionals** deal with business impact analyses (BIAs) and risk assessments (RAs) to take stock of

the organization's preparedness for crises of all kinds (with and without IT considerations).

- **Financial professionals** are barely concerned with concepts such as backups and IT policies, but they do care about investments and returns. That means they tend to see everything through the lenses of total cost of ownership (TCO) and return on investment (ROI).

TIP

If you're an IT professional who can translate operational goals and business impact into IT mechanisms (and vice versa), you'll get the IT stuff you want — and eventually, a promotion from being "just" an IT professional.

Data Protection by the Numbers For Dummies is about helping you understand each of the three languages of data protection numbers — and helping you translate between them.

About This Book

To continue the Rosetta stone analogy, this book is meant to help IT professionals tasked with data protection responsibilities to empathize with, communicate with, and eventually collaborate with the operational and financial stakeholders within their organizations. Simply put, if you can translate operational requirements into metrics such as uptime and business impact (when systems aren't up), you can objectively choose the right data protection technologies. And, if you can translate downtime and data loss into financial impact, and then quantify the cost savings or new recognizable value, you can pay for it.

How This Book is Organized

Since we describe this book as a Rosetta stone, you can find chapters about each of the three data protection languages.

Chapter 1: Technical Metrics: RPO, RTO, and SLA

Chapter 1 starts where most backup administrators and other IT professionals responsible for data protection should start: with the understanding of downtime and data loss. RTOs and RPOs are

both the real metrics of data protection and the fodder for data protection marketing. We explain how to think of this concept in practical terms, and then we focus on establishing reasonable and attainable service level agreements (SLAs).

Chapter 2: Operational Metrics: RA and BIA

Chapter 2 covers the realm of business continuity and disaster recovery experts who serve operational groups, business units, and executive leadership. In this chapter, we cover BIAs, RAs, and "the art of worrying."

Chapter 3: Calculating the Cost of Downtime

The practical application of being able to translate between RTO/RPO (capabilities) and BIA/RA (needs) is the ability to quantify the cost of downtime. The reason that most organizations undervalue IT functions, like backup, is a lack of appreciation for what recoverability or nonrecoverability means to the organization. Chapter 3 breaks down the cost of downtime to four key variables. It then frames the dialogue on how IT and operational professionals can sit down to adapt the generic formula to one that accurately measures the organization.

Chapter 4: Financial Metrics: TCO and ROI

Most IT folks cannot simply choose to purchase technology in a vacuum; there has to be some level of justification to the business. Just as important, the costs of the solution(s) to be considered are typically far more than the prices of the components under consideration. Chapter 4 helps you work through quantifying the costs associated with data protection mechanisms (TCO) and how to effectively articulate the comparative benefits of the solution compared with the recently quantified cost of the problem explored in Chapters 2 and 3.

Chapter 5: Ten Keys to Better Conversations

Driving a conversation that aligns technical, operational, and financial goals while responding to objections to those technical,

operational, and financial goals is not trivial. Set those challenges against the jaded "burden" of backup and it can be daunting. Chapter 5 highlights ten facets of this process and the building of bridges to help you make these conversations happen, if for no other reason than to help you get the better data protection that you probably already know you want.

Icons Used in This Book

Icons are used throughout this book to call attention to material worth noting in a special way. Here's a list of those icons, along with a description of each one:

REMEMBER

Some points bear repeating, and others bear remembering. When you see this icon, take special note of what you're about to read.

WARNING

Watch out! This information tells you to steer clear of topics that may leave you vulnerable, cost you big bucks, or suck down your time.

TECHNICAL STUFF

This icon indicates technical information that is probably most interesting to technology planners and architects.

TIP

If you see the Tip icon, pay attention: You're about to find out how to save some aggravation and time.

Where to Go from Here

Seasoned IT professionals often "know" (by instinct) that they need better data protection than they're now using, but far too many don't know how to justify that purchase and make it happen. This book is intended to help you change the dialogue so that all three groups of stakeholders can agree on the imperative nature of the problem and the justification for a better solution.

IN THIS CHAPTER

- Recovery point objective (RPO)
- Recovery time objective (RTO)
- RPO and RTO combined as problem statement
- Service level agreement (SLA)

Chapter 1
Technical Metrics: RPO, RTO, and SLA

When comparing the wide range of data protection technologies and methodologies, the two technical metrics that provide a standard for comparison are the recovery point objective (RPO) and the recovery time objective (RTO). For an introduction to these terms, consider a typical nightly backup scenario where a full backup is made every weekend and an incremental backup is made every evening after users leave work.

Recovery Point Objective

The topic of RPO starts by asking, "How much data can we afford to lose?" and "How frequently should we protect the data?" Where RPO can really matter is as a method to objectively compare the diverse range of data protection and availability technologies.

RPO is often thought of as simply the amount of data that can be lost. That's not the whole story, but you can start there. If you're backing up every evening — and assuming that nothing goes

wrong during the backup or the recovery — the most you can lose due to an IT outage is one day's worth of business data. If your data is made up solely of documents from Word or Excel, you have lost only those documents that were updated that day. If your data consists of transactions such as financial records, the consequences could be worse. Imagine that you work in a financial institution and, in one day, most (if not all) of your accounts have some kind of activity, including not only deposits and withdrawals but also changes in value. If you lose a day's worth of those transactions, the entire data set is no longer valid.

The key point here is that you must assess the potential for data loss in two ways:

- The time spent re-creating lost data
- The scope of data that will be lost or affected

To expand on this statement, assume that a reliable backup takes place every evening and that restore operations will always work, which is already a dangerous assumption. I explain, later in this chapter, why this assumption doesn't usually apply; but, for now, that supposition helps with the example. With that in mind, consider these two extreme scenarios:

- **If the server were to fail at the beginning of the business day,** almost no data would be lost since the last backup. The actual data loss would be measured at nearly zero because nearly nothing would have changed since the last recovery point or backup event.
- **If the server were to fail at the end of the business day,** that entire day's worth of data would be lost because no backups (recovery points) would have been created since the preceding midnight. The data loss would be measured as a full day's worth.


TIP

If you average these two extremes, *you can presume that the server will always fail at noon* — halfway into the business day. Statistically speaking, companies that rely on nightly backup alone will lose, on average, half of one day's worth of data.

A "SERVER" BY ANY OTHER NAME

For the purposes of this book, *server* refers to a physical, virtual, or cloud-instantiated platform that is hosting data in order to deliver IT services. According to industry analysts, 52 percent of IT outages within a data center are caused by hardware. In earlier days, this would equate to a single (physical) server being impacted, but today the vast majority of physical infrastructure is in support of virtualization, where 20 virtual machines (VMs) might reside within a single physical host, thereby compounding the impact of downtime exponentially.

So, although it would be easy to infer that individual VMs (on-premises or in a cloud) won't suffer individual hardware issues, the risk of their hosts experiencing a physical issue is gravely more impactful.

To learn the whole story when looking at RPO, it's the *O* that is most important. RPO is an objective (or goal). It specifies how much data you're willing to lose. In a nightly backup, the statistical probability is that you will lose a half-day of data. But if you establish your RPO at half-day and then your server fails in the afternoon, you have actually lost more data than you planned — and you fall short of your goal or objective. In fact, you'll likely miss that goal close to half the time. So, most would set an RPO as one-day, meaning that it's an acceptable business loss to lose an *entire day* of data, because of the recognition that backups are occurring only *nightly*.

In the case of nightly backup, the unit of measure for RPO is in days (for example, half-day or full day) because backup operations normally take place daily or, more specifically, nightly. To have an RPO (a goal) for how much data you can afford to lose using a measure *less than days*, you have to protect the data *more often than nightly*. That usually takes backup-alone — and by association, tape-only solutions — out of the equation. Disk-based solutions often replicate hourly or every few minutes or seconds or in real-time. RPO essentially becomes the measurement of data protection frequency, regardless of media choices.

REMEMBER

For the purposes of this book, the three most common data protection mechanisms are defined as

- **Backup:** A point-in-time copy of data that is stored on secondary disk, tapes, or within cloud-storage for the purposes of previous version retention. Backups are usually created on a daily or less frequent basis, have been compressed or deduplicated to save space over time, and are usually stored for months to years.
- **Snapshot:** A relatively recent point in time based on a series of blocks within the production storage system. Snapshots can be created every few minutes or hours, are composed of pointers to block-level changes, and are usually retained within the primary storage system for only days.
- **Replica:** A current or near-current copy of data that is in its original format but residing in an alternate location (offsite for BC/DR, for example). Replication can occur in real-time (synchronous), delayed (asynchronous), or on a recurring schedule, usually measured in minutes or hours. At each point, the secondary copy is overwritten with whatever has changed from the original source.

Note that these three data protection mechanisms complement each other — and are often (and should be) used together.

Recovery Time Objective

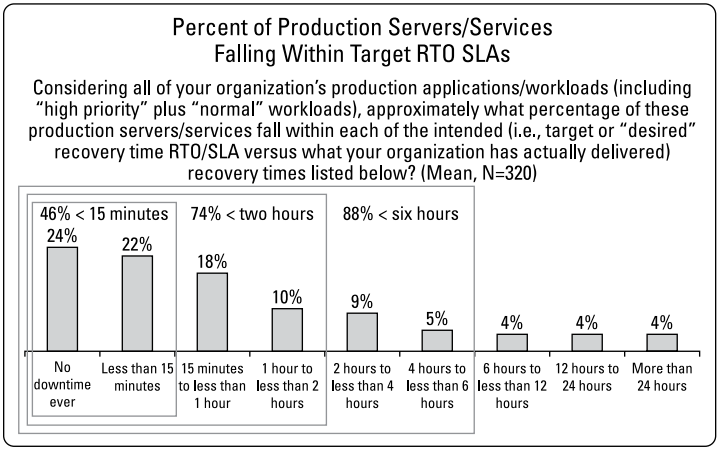
RTO starts by asking the question "How long can we afford to be without our data?" — which can also be asked this way: "How long can our services be out of operation?" Using the same nightly-backup scenario as in the earlier section "Recovery Point Objective", RTO is the goal (objective) for how long it takes to conduct a recovery. The question begins with "How long will the restore take?"

In the earlier example of nightly backup, RPO is measured in days or partial days because that's how often a data protection (backup) event is occurring — every night. But for this same example, RTO is measured in hours because it's a performance measure of the components in your solution itself. If your backup software and secondary storage can restore up to 2 terabytes (TB) per hour and

the production server has 6TB of data, time to restore the data is roughly 3 hours — or, more specifically, 3 hours *from when the restore function begins*.

If your largest server holds 10TB of data and your protection storage can restore 2TB per hour, and if you're confident that you could immediately locate the right tapes and restoration could begin soon after, then you might specify an RTO of 5 hours — or 6, to be cautious. But you'll also lose time between the moment the outage occurs and the moment you invoke the restore process (with or without diagnostics time on the original server). As such, you will likely round up and specify an RTO of one-day.

To help you appreciate what real-world RTO goals look like, industry research suggests that nearly half of all servers have an RTO goal of less than 15 minutes. (See Figure 1-1).



Source: ESG Real World SLAs and Availability Methods, December 2017

FIGURE 1-1: How recovery times stack up.

REMEMBER

Nearly seven of eight servers have an RTO of six hours or less, which, according to the earlier examples in this chapter, negates a backup-alone approach to data protection. Instead, roughly half of all server datasets supplement backups with replication, and a third supplement with snapshots.

Putting RPO and RTO Together

Let me now combine the examples I've used in this chapter.

Assume that the server being protected in this scenario failed on Wednesday at 4:00 p.m. It will take most of the next day to recover the server. If IT personnel are in the same office, you can optimistically identify what has failed and, if necessary, arrange for replacement parts to arrive early Thursday morning. On Thursday, the server will be repaired and the data restored. By Thursday evening, the server will be rebuilt and recovered. The recovery time will be one business day and (you hope) within the RTO.

The unfortunate part of this scenario is twofold for users:

- Thursday is a wasted day for users because they cannot get to their data while items are being repaired, replaced, and restored.
- Wednesday's data is likely lost because the server, when it's restored, will be restored to the latest successful backup (Tuesday night). Everything that was created on Wednesday (after the Tuesday backup) will likely have been lost when the server storage failed. The recovery point was within one day of lost data (Tuesday midnight through Wednesday at point of failure), which is again, you hope, within the set RPO.

REMEMBER

To improve RPO, you need to perform data protection more often than nightly. For this, you must turn to replication and snapshot mechanisms, in complement to backups (not replacement of).

REMEMBER

To improve RTO, you need a faster restore medium, which usually points to disk instead of tape or cloud for routine restore operations. In addition, frequency of protection can also shrink RTO because the amount of data to be restored shrinks with each iteration.

But just measuring RPO and RTO isn't enough. You have to translate your solution's RPO and RTO capabilities as something predictable that can be understood and agreed to by the business and operational stakeholders of the company. You have to set a service level agreement (SLA), as described in the following section.

Making RPO and RTO Real with SLAs

WARNING

When you recognize that the *O* in both *RPO* and *RTO* is *objective*, you run into one of the key problems in most plans for data protection and availability. An objective is a goal, not a promise. The promise comes when you describe your capabilities to the stakeholders in the business, for example, when you tell the management of the people who rely on the server that they will be "running again within one business day and will lose an average of a half-day of data, but potentially a full day of data." You might tell the management team that, with nightly backup, you could have an RPO of a half-day of lost data and that the RTO might be one business day to repair the server and restore the data. But those are goals based on ideal circumstances.

What happens when the circumstances are not ideal? In the scenario I describe of the nightly backup of a failed server, I've made a few assumptions:

- You can react quickly to the server outage.

  If you have IT staff on-site, they can identify the issue almost immediately.

  If you don't have IT on-site, the entire restore-time window will be longer because nothing can happen until you get there (or remotely connect).

- The server is readily repairable. In the example, the server failed at 4 p.m. on Wednesday.

  If parts are available, repairs can begin immediately.

  If you happen to be on the US East Coast, you can expedite parts from a West Coast provider, where they can be sent overnight and repairs can begin the next morning.

  If you happen to be on the US West Coast, you may not be able to get parts for another whole business day — and everything else will be prolonged accordingly.

- Every recovery point (on disk, tape, or cloud) is readable.

  If the latest Tuesday evening backup is somehow corrupted, you will only be able to restore through Monday's copy. You will have lost another day of data (RPO), and you will likely lose time trying to restore

Tuesday's data before you can identify the failure (longer RTO).

If you're doing incremental backups (only nightly changes) instead of differential backups and Monday's data is bad, then Tuesday's data is mostly irrelevant. Tuesday's incremental contains the differences between Monday and Tuesday; without a successful restore of Monday's data, however, Tuesday's changes may not be substantive. This varies by the production workload (as well as by the backup software's tolerance for failed recovery points within the storage repository).

If one of the weekend full backups is bad, then Monday's and Tuesday's are irrelevant because everything is in the context of the last full backup (which is unusable).

It's not the best of both worlds, by any means, but I can describe two last-resort recovery scenarios:

- If the daily incremental data is not overwritten each week (that is, this Tuesday replaces last Tuesday), you can restore the full backup from a week earlier and then the incrementals or differentials from the previous week. In short, when the server is repaired on Thursday afternoon (and accessed Friday morning), the data will be as it was the Thursday from a week before — the last good copy.
- If the daily backups are overwritten, the data will be restored only to the full backup from a week ago. All data for the previous week, as well as the beginning of this week, is lost (10 days of data in the example).

These aren't niche cases or a dramatic calamity-of-errors. They're just examples of how easily the reality can fail to match the ideal RPO and RTO that are defined by the hardware and software of the data protection solution. It is for these reasons (where reality doesn't match the ideal RPO and RTO) that the SLA — your assurance to the business units of what your recovery capabilities are — needs to be broader than just stating the RPO and RTO of the technologies themselves.

REMEMBER

You need to consider the processes and potential pitfalls of whatever technical solution(s) and recovery processes you are considering:

- **Time to React:** Sites without IT staff should have longer RTO SLAs than sites with IT staff, because it will take time to get there, depending on the arrangement, connectivity, and type of failure. Perhaps IT staff can drive or fly from their primary location to the remote office. Perhaps a local integrator or channel reseller can be dispatched; in that case, a pre-negotiated contract may have to be put into place, including an SLA from them to you on their committed response rate to your issue.
- **Time to Repair:** Should spare parts or even complete cold-standby servers be acquired? Where can parts or servers be expedited from? Does a pre-negotiated agreement need to be signed between you and a vendor or distributor?
- **Technical RPO and RTO:** These relate to issues involving the RPO and RTO, as well as the perceived failure rate of the media.

Notice that technology isn't mentioned until the last item. The first aspects of a server recovery SLA relate to people and process, followed by materials and access. Once you get to the technology, you're likely more in the comfort zone of the IT professional, but there are still unknowns concerning the technology. With all of that addressed, after the server is ready to be restored, RTO still varies based on whether you're restoring from Monday or Friday:

- Presume that you start with a *full* backup, meaning a complete and autonomous copy of the data each weekend.
- With *differentials*, you start with the full backup and then will layer on the Thursday night differential (the changes since the last full). But a Thursday differential will have appreciably more data to restore than the Monday differential.
- With *incrementals*, you see a similarly linear increase in restore time, as each subsequent incremental is layered on top of the incremental before it: Monday-to-Tuesday, Tuesday-to-Wednesday, and so on. Unfortunately, in this imperfect world, if a recovery point (on tape, disk, or cloud) fails, such as the Tuesday incremental, then Wednesday's and Thursday's incrementals might be unusable.

All these overly dire and pessimistic examples are designed simply to prompt you to compare your data protection technology's presumed RPO and RTO to what you — as the IT professional responsible for data protection — can assure your management of being able to deliver.

TIP

Here's the secret to a successful SLA: Salespeople call it sandbagging, where what you forecast will sell is less than what you believe is likely. Others might call it underpromising and overdelivering. As a consultant and an IT implementer, I call it planning for Murphy's law.

TIP

When you're setting your own SLAs, don't believe the RPO and RTO printed on the outside of the box of whatever technology you're looking at. And certainly, don't repeat the RPO and RTO to the business managers. Test it. Assume that something will go wrong, and think about how you would address such issues. (Heck, you can even go to the extreme of thinking that *everything* will go wrong and then negotiate with the business managers back to a point of reality!)

SLAs ARE AS MUCH ART AS SCIENCE

SLAs can sometimes be more art than science because in order to have good SLAs that both the IT staff and business management can be satisfied with takes creative planning, usually by folks who have suffered an IT outage. Balance must be achieved despite two types of outlook:

- **If the IT team is overly conservative and cautious**, they may set the SLA performance bar so low that the business management team believes that the IT staff is unknowledgeable or low-performing.
- **If the IT team is overly optimistic or unrealistic,** the SLA performance bar may be so high that even well-executed recoveries may fail to meet the measure established in the SLA.

When negotiating your IT SLA with the business leaders, consider first leading a brainstorming session with your senior IT folks to map out the recovery plans: Identify the likely failure points and your mitigating actions when the plan does break down. Only after you have that workflow should you talk to the business managers about SLAs.

IN THIS CHAPTER

- Risk Analysis (RA)
- Business Impact Analysis (BIA)

Chapter 2
Operational Metrics: RA and BIA

Chapter 1 introduces some of the metrics you can use to assess what your data protection technologies are. This chapter looks at how well you can utilize these metrics within a business operations context. Here, you'll see how to apply them to the business as well as how to pay for the technologies that we believe you need.

Risk Analysis: The Science of Worrying

How likely is it that your particular backup solution will have a problem?

Perhaps more important, how likely is it that your production resource will suffer failure? Consider my house in Dallas, Texas. How likely is it that it will suffer a flood? Or a monsoon? Or an earthquake? Or a tornado?

Let's focus on this last question to take technology out of the process. It so happens that I live in Dallas, Texas, which is approximately 400 miles from the nearest large body of water, the Gulf of Mexico. Because of this, I have no fear of a monsoon. Statistically

speaking, the likelihood of Dallas being hit by a monsoon is effectively zero. Speaking of water, my home is in a 100-year flood plain, so, statistically, my land will be flooded once per century. I could buy flood insurance for my home, but the probability is low enough that I choose not to. If I lived in Houston, Texas, which is much closer to the coast, flooding is more likely and I might want flood insurance. But, because the probability is so much higher, I probably could not afford it — the insurance actuaries will have already calculated that fact into any possible premiums.

This has meaning for this discussion. Insurance is an entire industry built on consumers' presumptions that they pay a little every month to avoid a potentially significant and perhaps life-altering financial impact later. The amount of insurance that's paid is based predominantly on two factors:

- How likely is the crisis that you're anticipating?
- What is the financial impact that you're mitigating?


REMEMBER

Concepts such as data protection and data availability are similar to the idea of buying insurance. First, you assess what could go wrong and consider how much it will cost if it does, and then you purchase something that costs appreciably less than that to mitigate the crisis.

What could possibly go wrong?

The first step in planning your data protection and availability strategy is to look at each of the servers and applications in your environment and think about what could go wrong in what is formally referred to as a Risk Analysis (RA).

Go crazy: Think about *everything* that could possibly go wrong. The most important rule in this exercise is to simply list every single thing that could go wrong. Do not think (yet) about the probability of something occurring, but just *the potential* of its occurring. And let yourself think small *and* think big.


REMEMBER

In the case of an application that the organization relies on, you can't just consider the application itself. An end user doesn't care if the reason he can't get to his data is that the application crashed or the OS hung or the hard drive failed or the DNS server isn't resolving correctly or Active Directory won't let anyone log on

or the browser isn't reading the page correctly. Users don't care because their data and their productivity are impacted, regardless of the cause.

That's from the user's perspective. Now think about the big problems. Is your company in a flood zone? If you're in southern California, are you near a forest that can catch on fire? If you're in northern California, what would an earthquake do? If you're in the Midwest, are you in tornado country? In the north, what would a blizzard do? On the East Coast, how likely is a hurricane? If you live in Florida, a hurricane is a *when*, not an *if*.


TIP

Here's a reality check: Several years ago, I was conducting a disaster recovery seminar in a town in Florida. My opening remark was this: "According to the National Weather Service, this city is in the eye of a hurricane every 2.83 years. It has been 3 years since you have actually been hit. You are past due." My advice to you? Be acutely aware of potential disaster situations you may confront. Don't assume that, just because it hasn't happened yet, it *won't* happen.

How likely is it?

I am not suggesting that every IT professional should turn into an *actuary*, someone who lives with the statistics of risk all day long. What I *am* suggesting is that when you are first imagining the entire realm of bad things that could happen to your data, servers, infrastructure, and even people, first just list them. Then, having done that, put on your practical hat and consider the reasonable probability of each one. The reason I don't buy flood insurance is because a flood, while possible, is not probable for me. I cannot buy hail insurance at a reasonable price, because it almost certainly will happen to me.

In technology, some calamities are certain to happen:

- You will lose hard drives.
- System boards will fail.
- Applications will crash.
- Databases will become corrupted.
- Users will accidentally overwrite last month's data with this month's data and then regret it.

In business, some crises are likely to happen:

- Someone may steal something — probably a laptop — from your company.
- Someone may maliciously delete data on their last day at work.
- The server room may catch fire or might be flooded from the bathroom immediately above it.

In life, natural disasters could affect your company facilities.

So, what is the likelihood of each event you listed? You may not have exact figures but stack the kinds of things you're protecting against in relative probability to each other. This risk analysis (RA) exercise is half of what you need in order to start planning your data protection and availability plan.

Business Impact Analysis: How Much Will It Cost?

REMEMBER

Data protection and availability aren't just about technology. In fact, both are mostly about reducing financial impact. To do that, you not only need to look at the technologies you can use and the calamities you fear but also to turn all of them into financial ramifications.

Let's look at the potential technology and business crises listed in the previous section. They're ordered approximately from most likely to least likely, with the exception of end users who accidentally overwrite precious data. (I *guarantee* that users will overwrite data.) Let's look at the two extremes of the list. Everything else falls between them, from a likelihood perspective as well as financial impact:

- **Scenario 1:** If I were to lose one hard drive within a production device, the physical costs are likely a few hundred dollars or less. Whether the lost hard drive was for a data drive or the operating system will determine the level of lost productivity, which is then compounded by the number of users affected. And, how long since my last backup will determine the amount of lost data that I may or may not have to re-create.

» **Scenario 2:** On the other end of the list, if my production facility were to be flooded, the entire server room — as well as many other production resources, desktops, infrastructure components, and even copy machines and coffee makers — would be destroyed. My business could be down for days and, in fact, might never reopen.


TIP

The goal of a business impact analysis (BIA) is to financially quantify what the cost of each type of crisis might be. Say you calculate that the total cost of a hard drive failure, lost productivity, and replacement is $1,000. You believe that it's a highly likely event, so you need to aggressively seek a data protection or availability solution that mitigates that $1,000 of impact to the business by finding a mitigation solution that costs less than $1,000 — hopefully, a lot less. Similarly, though you may believe that a flood would cost $3 million, it's admittedly far less likely than a hard drive failure. That statistical probability factors into determining what you might spend to mitigate that risk.

Always Turn Technologies into Dollars

Most often, the person who writes the checks — particularly the checks for buying new assets like software and hardware — doesn't care about RPO and RTO or about backups versus snapshots or about disk versus tape versus cloud. To move business-driven decision makers forward on data protection projects, always quantify the risk or the reward in dollars, not terabytes, minutes, or subjective assessments.

Data-protection and -availability projects are among the easiest topics to describe in financial terms. Availability — or, put another way, productivity, — can be calculated by looking at the cost of downtime. Protection and recoverability can be quantified based on the impact of lost data as it relates not only to lost productivity but also to the inability to comply with applicable regulatory compliance mandates.

In short, *if* you can objectively state that one hour of downtime equates to $10,000 and the solution to resolve it costs $800, *then* it is easy to justify new data protection or availability solutions.

Chapter 3 covers exactly that topic: How to calculate the cost of downtime.

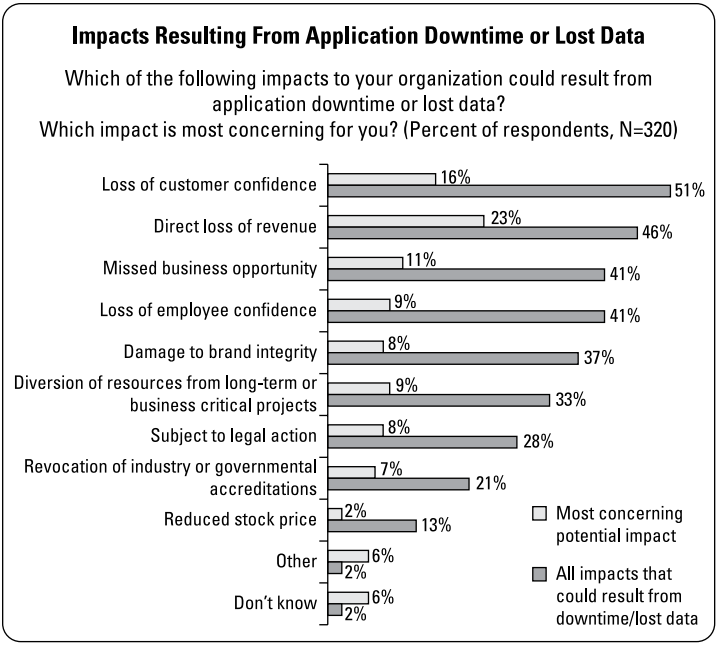
IN THIS CHAPTER

- Examining the intangible impacts of downtime and data loss
- Calculating the cost of downtime with four simple variables
- Applying the cost-of-downtime formula in the real world
- Adapting the cost-of-downtime formula to your organization

Chapter 3
Calculating the Cost of Downtime

Everyone "knows" that downtime and data loss are bad, though most don't know how to quantify it. Figure 3-1 shows industry research on what are the top concerns by senior leaders related to downtime and data loss.

Even so, you actually can quantify some of the impact in a way that helps bridge the dialogue between technical circumstances and operational realities.



Source: ESG Real World SLAs and Availability Methods, December 2017

FIGURE 3-1: The intangible impacts of downtime.

Taking a Look at Four Simple Variables

The main idea when quantifying downtime costs is how to turn technology problems into financial problems. If you can fiscally quantify the impact that something has on the business, then you can have a different kind of discussion with business (and budget) leaders on why you need to fix it. To start that conversation, you need to understand the cost of downtime. In other words, when a server breaks, how much does it cost the company?

REMEMBER

If a server fails, you have to look backward as well as forward in order to calculate the impact. Figure 3-2 shows a server failing at 2 p.m. on a Wednesday.

In this first example, you make these three assumptions:

- The business day is exactly 8 a.m. to 5 p.m.
- You have a successful backup from Tuesday night that is restorable.

» The server will be fixed by the end of the next business day (Thursday).

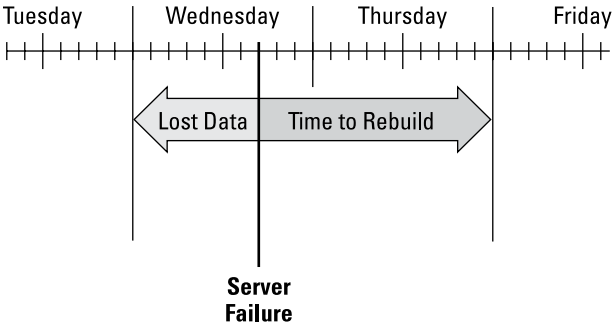


FIGURE 3-2: Downtime, forward and backward.

With these assumptions in mind, two kinds of time can be quantified: time of lost data and outage time. The next few sections look at the two kinds more closely.

Lost data

If a server fails, any new data *since the last recoverable backup* is potentially lost. Since the server failed at 2 p.m. on Wednesday and you're assuming a reliable restore from a successful backup on Tuesday night, you can presume that you will lose whatever data was changed between 8 a.m. and 2 p.m. on Wednesday.

In Figure 3-2, the arrow starts at the time of the server failure and points backward to the left to whenever the last successful backup can be reliably restored from — in this case, the previous evening.

If the last backup had failed or was not able to be restored, the arrow would point further to the left until a successful backup could be restored. But, for now, data loss is quantified at six business hours. Therefore, the formula would look like this:

Td = Time of lost data, *which is six hours in this case*

REMEMBER

You're quantifying lost data in a measurement of time because, in the basic example, you're assuming that if the end users took six hours to originally create the data, then they will likely consume roughly another six hours of business time to recreate the data.

Outage time

Because the server failed in the afternoon, you're assuming that the end users may be idle for the remainder of the day and (in the example) idle for the whole next business day.

In Figure 3-2, this arrow starts at the time of the server failure and points forward to the right until the server is completely back online, which in this case is the end of the next business day. If this is true, the outage time is the remaining three hours of Wednesday afternoon plus all nine business hours of Thursday. The formula here looks like this:

> ***To = Time of outage or lost productivity,*** *which is 12 hours in the example*

You can say that, added together, *To + Td = 12 + 6*. So, the total time impact of the server failure is 18 hours.

Now you need to decide how much those 18 hours are worth in dollars. Again, there are two kinds of $-per-hour costs to consider: human costs and profitability. Let me tackle the human costs first.

Human costs

If you presume that an end user is completely idle while the IT resources are offline, then the company is essentially paying the salary or hourly wage of that person for no benefit. That can be quantified this way:

> ***Hr = Hourly cost of impacted personnel ($ per hour)***

In a restaurant that is unable to do any business, you might reduce its losses by sending the waiters and cooks home for the day. But if 15 hourly staff members who each cost the company $8 per hour (along with two salaried managers paid $40,000 annually, or $20 per hour) were to sit idle, then $(15 \times \$8) + (2 \times \$20) = \$160$ per hour for idle time.

In an office, perhaps there are other activities the employees can do so that they aren't completely idle, but simply impacted. In that case, you might choose to divide the salaried costs in half to show that they're half-impacted as opposed to idle and nonproductive.

REMEMBER

Every business is different, but you should be able to assess a $-per-hour number for some percentage of your hard costs of paying people who are unable to perform their primary role due to an IT outage.

Profitability

When a team that creates revenue is affected, revenue is affected. So, if you know the weekly or monthly profitability of a team, you can quantify how much profit they're generating, or not generating, during an outage:

Pr = Hourly profitability or loss ($ per hour)

A team may produce $9,000 per day in profit, so their hourly profitability between 8 a.m. and 5 p.m. is $1,000 per hour.

REMEMBER

On a team that is subject to service contracts, you may be liable for fines or recouped losses if you aren't offering your service.

A shipping department may not lose any money for a few hours of downtime, but if an entire day is lost, expedited shipping charges may be incurred in order to make timely deliveries the next day.

REMEMBER

Again, every business is different, but you should be able to assess a $-per-hour for the business value that a team creates per day or per hour. Some of that productivity is lost or penalties incurred when the team is unable to perform their primary role due to an IT outage.

Adding together Hr and Pr gives you the total dollars per hour impact that an IT outage has on the team. Using the first example from each description, a team may cost $160 per hour by sitting idle (Hr) and also not create revenue (Pr) at $1,000 per hour. Thus, every hour is worth $1,160 to the company.

Applying the Formula in the Real World

In this section, I present a basic formula for measuring systems availability in financial terms. You can take the total time for data loss plus outage time and multiply it by how much an hour is

worth to the business or team in consideration of human costs as well as profitability or losses:

Cost of downtime = (To + Td) × (Hr + Pr)

To = Time, length of outage
Td = Time, length of data loss
Hr = Human cost $/hr (per person)
Pr = Profitability $/hr

In the examples, this would be

To = 12 hour outage
Td = 6 hours of lost data
Hr = $160/hour for the team to sit idle
Pr = $1,000/hour in lost revenue

Cost of downtime = (12 hours + 6 hours) × ($160/hour + $1,000/hour)

Cost of downtime = 18 hours × $1,160/hour
Cost of downtime = $20,880

This company will lose nearly $21,000 if a server fails, even if it's recoverable by the end of the next day.

So you now have a formula, but more work must be done. The idea here is simply to help identify the variables that you'll need in order to quantify the cost of downtime.

Calculating downtime costs for just one group or workload

It's time to explore a bit more how the (rather abstract) formula functions in the real world. You're still considering the same outage scenario of an environment that is using nightly backup that includes a full backup every weekend and incremental backups every night. Consistent with the scenario you've used in this chapter, the production server fails at 2 p.m. on Wednesday. As discussed earlier in this chapter, users are affected for the rest of Wednesday, and the recovery takes a good part of Thursday. By the end of Thursday, the server is running, the users are mostly happy, and business resumes. Two months from now, that

incident will be thought of from long-term memory as a minor blip. Yes, the server went down, but everything resumed within a day. Pretty good, right?

Wrong. Let's look at the numbers:

Cost of downtime = (To + Td) × (Hr + Pr)

To = Time, length of outage
Td = Time, length of data loss
Hr = Human cost $/hour (for team)
Pr = Profitability $/hour

Nightly backup = (1d + ½d) × (Hr + Pr) × hrs/day

To = RTO = 1 day recovery, including parts, shipping, and installation
Td = RPO = average ½ day (could fail early morning or late afternoon)
Hr = Human cost $/hour (per person)
Pr = Profitability $/hour

The time of outage (To) or recovery time objective (RTO) will likely be one business day, which includes identifying why the server failed, repairing those components, and restoring the data. If everything goes well, this should typically be one business day. If things don't go well, this might measure two or three days of downtime. In a perfect world, additional parts are already standing by, technicians are ready to go, and perhaps the server is up in just a few hours.

The time of data loss (Td) or recovery point objective (RPO) is statistically probable as one-half of a business day. As discussed earlier in the chapter, the server could fail at the beginning of a business day, resulting in near zero data loss since the last nightly backup, or it could fail at the end of the business day, resulting in a complete day of data loss. Splitting the difference, assume that data is lost from half of the day.

Assume that the outage impacts just one division within a larger organization, perhaps the inside sales team within the company. Various surveys presume that the average white-collar worker in the United States costs $36 per hour. The reality of a statistic this broad is that it's guaranteed to be wrong for your workforce, but it

serves as a placeholder for now. Now, presume a 10-hour workday and that this team generates $10 million in revenue annually:

Cost of downtime = (To + Td) × (Hr + Pr)

To = Time, length of outage
Td = Time, length of data loss
Hr = Human cost $/hour (for team)
Pr = Profitability $/hour

Inside Sales Team whose data is protected via nightly backup = (To + Td) × (Hr + Pr) × hrs/day

To = RTO = 1-day recovery, including parts, shipping, and install
Td = RPO = average ½ day (could fail early a.m. vs. late p.m.)
Hr = Human cost = $36/hour/person × 50 employees = $1,800/hour
Pr = Profitability = $10M annually = $3,850/hour (10-hour workday, 5-day workweek)

Inside Sales Team whose data is protected via nightly backup = (1d + ½d) × (1800 + 3850) × 10 hrs/day

Business Impact per Server Outage = $84,750

This is not a typographical error. If the primary server fails for an average group of 50 office employees that creates revenue (all of your earlier presumptions), then the business impact to that organization is $84,750.

Adapting the formula to your business

Crunching the numbers is the first step. To be fair, though, this isn't the whole story. The likely outcome of first evaluating the cost-of-downtime formula is that management will disagree with its validity. And that is fine, because they are probably correct.

If your users cannot get to their primary application, they might catch up on email, or they might have some documents on their local workstations or laptops. In this case, assume that the employees aren't completely idle but are simply affected. If that were the case, you might add a multiplier to the formula to imply that the user base is operating at two-thirds efficiency. If so, a one-third multiplier against the formula results in a business impact of only $28,250.

In today's world of information workers, you might presume that the users have a variety of activities to perform. Between email and database applications (including contact management), as well as traditional office applications from a file server, perhaps you could presume only a minor inconvenience to a percentage of the users. Perhaps this results in a 10% impact to 10% of the employees. Literally, this would mean that only five of the 50 employees had any impact, and therefore the cost would be only $8,475.

But some departments don't have multiple functions that they can balance. In the example of inside sales, what if all of the data were within a single database application or the sales folks fundamentally could not operate without access to the database? In that case, they would suffer the whole (and what may have originally seemed extreme) business impact of $84,750.

REMEMBER

The business impact may be even higher when you recognize that a server doesn't usually go down just once. Although these minor inconveniences may fade in the memory of users, they typically don't fade from your system's event log. You might be surprised to find that a particular server fails twice per year, in which case you would double all the previous numbers (which still don't include hardware or services costs).

Statistically, 18% of servers will have at least one outage per year. But even at one failure per year, if you presume that a typical server asset is expected to have a three-year lifespan, then you should multiply the per-outage cost times the number of outages per year times the number of years the resource will be in service:

Total Cost per Server = Co × Opy × LS

Co = Cost per Outage = the result of the earlier formula
Opy = Number of Outages per year (presume only 18% per year and one outage per server)
LS = Expected Lifespan of Server (typically 3 years)

Total Cost per Server = $84,750 × 0.18/yr × 3 = $45,765

Here's the punchline: The server that has been recently purchased and deployed to service the inside sales team of the company is considered reliable and well managed, so it's presumed to have

an 18% chance of failing at least once each year. Note that by adding statistical likelihood to the equation, you're now dealing with a probability, not a possibility (that is to say *when*, not *if*). With that in mind, the company should plan on that server's impact to the business being $45,765 over its lifetime of service.

That is the business impact analysis (BIA) for this *one* server. It took a while in this chapter to break it down, but in real life, the BIA goes quicker than you might expect. Essentially, as you're looking at what kind of protection or availability solutions you might consider per server or application platform, you first need to understand what kind of risks you're protecting against as well as the financial impact if one such risk were to occur.

IN THIS CHAPTER

- Total cost of investment (TCO)
- Return on investment (ROI)
- You should hope that they argue with your formulas

Chapter 4
Financial Metrics: TCO and ROI

Sure, you can discuss data protection and availability technologies in terms of *cost of impact*, meaning the business cost of the status-quo. There is, of course, the factor of *price of solution*, which is different depending on whom you're talking to.

Total Cost of Ownership

In this section, I want to recognize the fact that the price is always more than what is printed on the invoice. In fact, in many backup-and-recovery scenarios, the greatest contributing cost is labor.

Let's consider a traditional nightly backup solution. The initial acquisition costs might include

Backup server (software)	$2,500
Backup agents (software)	$995 per production server
Backup server (hardware)	$2,500
Tape backup drive (hardware)	$2,000

If you assume a traditional midsize company network with 25 servers, then to purchase a nightly backup solution for this environment, you might be requesting $37,000, not including deployment services.

WARNING

Not including the labor for deployment is your first mistake because you will "pay" for deployment even if you do it yourself. If you contract a local reseller or backup specialist, there's likely a fixed cost for the deployment, which hopefully also results in a fast-and-reliable solution, because presumably the reseller has previous experience and close ties with your backup software vendor. If you go the DIY route because you don't want to pony up for the additional labor costs, you will pay for it in time — literally. (Anyone who has ever completed a significant home improvement project can attest to this fact.) Your own IT staff, who would otherwise be doing other projects, will be deploying this instead. The project will likely take longer if your staff has not deployed this particular technology before, and, if they're not following best practices, it's sure to result in additional labor at a later date.

In any event, presuming that everything is equal, assume eight hours for the server deployment plus 30 minutes per production server for agent installation and backup-scripts configuration. Splitting the difference between an in-house IT professional at $75 per hour and a local reseller, which might charge $250 per hour, this results in 20 hours, which you could equate at approximately $150 per hour, or an additional $3,000 total labor.

But you aren't done yet. You should also calculate the cost of media. If you assume that each of the servers has 5TB of storage, then you would have 125TB of active storage across the environment. At an average 60% utilization rate, you would need approximately 75TB of data to be protected. With an aggregate daily change rate of 5% (more for applications, less on file shares), you'll be writing about 4TB of new data per day. Backups can be stored on disk, in the cloud, or within tapes. Using tapes for this example, most backup software will use a different tape for each daily job, plus four weekly tapes and 12 annual tapes. Conservatively, this puts you at 20 tapes at $100 each for an additional $2,000 in tape media (not including additional costs, like offsite storage or services).

You still aren't done. There will also be ongoing costs, such as power, space, and cooling. Space would be associated with your facilities costs, but simply running the new backup software on a commodity server platform might use 500 watts (plus the tape drive's 200 watts). The monthly power cost for this server alone is

700 Watts × 24 hours per day = 16,800 Watt Hours (WH) or 16.8 Kilowatt Hours (KWH) per day

16.8 KWH × 30 days in a month = 520.8 KWH per month

At $0.06 per kilowatt-hour, this server will cost $31.20 per month, or $375 per year. You also need to add in the ongoing labor costs for these tasks:

- Rotating the tapes daily, which isn't a lot, but perhaps 10 minutes per business day
- Checking the backup jobs, 10 minutes per business day, plus one -hour error resolution every two weeks

Those aren't significant numbers when looked at that way, but when you add them up, you see 8,220 minutes, or 137 hours, or 3.4 working weeks per year, just managing backups (assuming that most things work correctly most of the time) and not including restores. The labor for managing backups in this environment will consume at least a month of every year with no productivity benefit and will cost $10,300.

This gives you the bigger picture, the total cost of ownership (TCO):

- **Initial purchase:** The initial purchase price of your backup solution might be $37,000, plus $3,000 to install it.
- **Additional costs:** The operational costs in the first year will be an additional $12,800.
- **Extended over its service life:** Assuming that most hardware and software assets have a presumed lifespan of three years, you can add software maintenance (15%), upgrade labor (half of deployment), and new annual plus daily tapes (five annually) for the second and third years. The ongoing costs for them are $6,100 annually.

Thus, the TCO for this backup solution would be $65,000, which is *nearly double* the initial purchase price and doesn't include the labor for any restores.

Return on Investment

If TCO is thought of as the bad number to consider in any financial assessment, then return on investment (ROI) would be the good one. Think back to the beginning of Chapter 2 on business impact analysis (BIA), where we ask, "How much does the problem cost?"

If a problem costs $150,000, you should consider that amount as quantifiably lost money. But if you solve the problem, the company gets $150,000 back. Think of it like an ante in poker or a coin dropped into a slot machine: That money is gone. If you make any money from poker or slots, then that is winnings — a positive. Of course, if you bet $5 and then later win $5, you haven't actually won — you've broken even. Similarly, if your technology problem or vulnerability costs $150,000 and you get it back by spending $150,000 on a protection solution, then you haven't actually solved the problem of losing the money for the business — you've just chosen to spend it in a different way. That may be okay to your CFO, based on accounting practices, but that's outside the scope of this book.

REMEMBER

If you spent $65,000 (TCO) to solve a problem that will cost the company $150,000 (BIA), then you have solved the problem. You literally added $85,000 to the company's bottom line profitability because it otherwise would have lost those dollars due to the outages you mitigated. This is where ROI comes into the picture — how much you saved or gained for the company, in comparison to the amount you had to spend to accomplish it.

Calculating ROI

You can quantify ROI in different ways:

- **Think about it as savings, where you save the company $85,000 versus its existing costs.** Taking servers completely out of the discussion, if you could show your accounting manager that she's used to spending $150,000 per year on something but you could save her $85,000 by doing it a different way, it's usually an easy business decision.

- **Measure it as the percentage of BIA/TCO.** In this case, $150,000 divided by $65,000 yields 2.3 — or a 230 percent yield. Others invert the percentage (TCO/BIA) as the percentage of the problem you're spending to solve it. In this case, you can spend 43% of the problem to resolve it. That also means that you save 57% of your projected losses.
- **Think in terms of payback windows (time).** If a problem costs $150,000 over the three-year lifespan of the asset, then consider how long into that window before the solution pays for itself. In this case, with an average of $50,000 costs annually, the first-year cost of $52,800 is basically breaking even, but the second and third years go from $50,000 to $6,000 annually, saving almost everything.


TIP

The actual calculation for ROI is to take the net gain ($150,000 minus the costs of $65,000) of $85,000 and then divide it by the costs, after which you can multiply it by 100 to arrive at a percentage:

(Total Gain – Costs) ÷ Costs

($150,000 – $65,000) ÷ $65,000

$85,000 ÷ $65,000 = 1.31, which is 131% ROI

TIME TO VALUE

Somewhat related to the ROI of a solution is how quickly you will start to see the benefits of the solution you're deploying. When considering that you will see *x* dollars over the lifespan of the project, look also at when you will see those dollars.

Compare when the costs are to be incurred to when the savings will start to be realized. Will you just break even for the first year and then see gains in the second and third years (such as when you deploy a new component that will solve an ongoing problem)? Or will you see gains the first year but fewer gains in later years, as you postpone a problem or take on incremental costs throughout the project?

How else you might use (and grow) the earlier money can also affect the overall costs for the project.

Any positive ROI is a relatively good decision, and any negative ROI is a relatively poor decision. Consider a $10 problem:

- **Spending $6 to solve a $10 problem** is good because ($10 – $6) ÷ $6 = $4 ÷ $6 = 0.66, or 66% ROI. Said another way, for every $1 that you spend in this way, you would get it back as well as an additional 66 cents.
- **Spending $9 to solve a $10 problem** is not as good because ($10 – $9) ÷ $9 = $1 ÷ $9 = 0.11, or 11% ROI. Said another way, for every $1 you spend in this way, you would gain only an additional 11 cents. There are likely other ways that the business could invest that dollar and gain more than 11 cents in return.
- **Spending $12 to solve a $10 problem** is obviously not a good idea: ($10 – $12) ÷ $12 = a negative 12% ROI. Said another way, for every $1 you spend, you lose 12 more cents than what the original problem was already costing. It would (obviously) be cheaper to live with the $10 problem than to solve it for $12.

The third example may be overly obvious, but sometimes IT administrators do solve $10 backup or availability problems with $12 solutions (improperly utilized cloud services, for example) because they don't understand the BIA or TCO well enough or they aren't aware of the $6 alternative solutions.

Determining which ROI method is the most accurate

This book talks about converting technology issues into quantitative — and specifically financial — assessments. After you have converted your protection or availability problem and potential solution(s) into this financial language, you can convert it from one denomination (ROI metric) to another as easily as converting the denominator of a fraction by multiplying or dividing it by a common number.

One of the reasons that I prefer to deal in actual dollars is because CFOs and other accounting types can often crunch the numbers to their own liking, once you present two key numeric facts (though they must be defensible facts and not subjective opinions):

- The problem is currently costing the company $XX,XXX. (BIA)
- I can solve the problem by spending $YY,YYY. (TCO)

From there, you might subtract one from the other for savings, or you might find a ratio that helps you appreciate it. However, based on some anecdotal findings from surveys and the experience of many years in supporting sales efforts, there is a credibility concern to be aware of. The next section explores that concern.

Examining the credibility challenge of ROI

Notwithstanding the recognition that every technology vendor (or other sales organization) always preaches how wonderful its widget is and how amazing its ROI (often unfounded) could be, ROI can have a credibility challenge.

Using the percentage ROI method, assume that the ROI of a solution is 43 percent: You're spending $70 to solve a problem that costs $100. The challenge is that the solution is costing over half of what the problem costs. That means that if your assessment of the cost of the problem is perceived as too high (qualitatively, not necessarily quantitatively) or that you may have underestimated something in your TCO, then the ROI goes down from 43% as costs start getting closer to what the problem itself costs. If your CFO is willing to wager that a problem won't happen as often as you project, she might actually save money (or at least break even) by just allowing the problem to happen — hopefully less often than you expect it to. The project doesn't have enough ROI to warrant the initial expenditure.

On the other hand, what if you need to spend only $5 to save $100, resulting in a 1,900% ROI? This situation presents the opposite challenge: It sounds too good to be true. If you have a good amount of credibility with the financial decision maker, then you will be seen as a hero and your project will be approved (although with that much credibility between you and your CFO, you may not have calculated a specific ROI to begin with). For the rest of us in reality-land, if it sounds too good to be true, some financial decision makers will assume that it is not true (or isn't viable as a "legitimate" solution). There must be some significant cost factor that is either drastically inflated in the problem or underestimated

in the solution. Either way, the solution is not perceived as credible. After all, how likely is it that you can purchase a mere toy to solve a real problem?


TIP

Here is a rule of thumb that at first glance looks goofy:

25 percent ROI may be better than 60 percent ROI.

Based on anecdotes, experience, and a few old surveys, it appears that 20% to 25% ROI is the best way to justify a solution. The gain is enough that the solution is likely worth pursuing, though the investment is substantive enough that the solution can be considered reasonable for addressing the issue. Using this approach, you might consider the following ROI boundaries:

- Over 40% may lack credibility or legitimacy.
- Under 15% may not offer enough potential gain.


TIP

One of the most interesting pieces of advice that I ever heard related to ROI was from someone at a CFO conference who attended a session on ROI. They heard that if a significant proposal was submitted for review and it had a TCO projection and ROI analysis on its first submission, it would be approved over 40% more often than those that did not have those calculations. If the same type of proposal were pushed back down to get the TCO/ROI analysis and it was resubmitted, it had only a 15% greater likelihood of approval over similar projects without one. Here's the first ROI success tip: Present the TCO and ROI assessment with the initial proposal because it not only clarifies the legitimacy of the project to you but also proactively clears a big hurdle for you with those who guard the dollars.

You Should Hope They Argue with Your Formulas

The best thing that can happen when you present your methodology and resulting BIA/TCO/ROI justifications for a project is that the business/operational/financial stakeholder challenges your formula (in a constructive way). When working with your

business leaders and establishing the formula you will use in your process, here are a few key ideas to frame the conversation:

- Working backward, ROI is just a comparison of BIA to TCO.
- TCO is simply a prospective invoice, along with some simple assumptions of fixed costs. It is likely that any challenges here will result in minor tweaks to the fixed values, not wholesale changes to the math.
- BIA is where challenges are most likely to occur — your business stakeholders and dependents don't agree with how you calculated the cost of downtime. (See the example in Chapter 3.) This is great news because then the group gets to decide why the formula doesn't apply to a particular business unit or technology resource.

If your discussion circle can collectively agree that when the database server is down for as long as a day, employees can catch up on email, or vice versa (and thereby reduce some variable by half), then the collective team has turned your formula into their formula.

If the HR person can provide more specific hourly dollar values across a large department (though you're unlikely to get a list of individual salaries), your team now has much more accurate fixed values that both the IT management and the operational management will agree on.

In short, every pushback that can be discussed or refined brings buy-in and agreement by the other parties. When you have five variables to work with, the formula may seem academic. But if you get more accurate modifiers and the dollar variables are filled in with real numbers, you're left with only the technology numbers, such as these:

- How often does the server go down?
- What is the cost of replacement hardware?
- How much do tapes cost?

These numbers are usually easily accessible by IT management, so completing the equation should pose no problem. From there, you now have a new BIA that is even more defensible and that now has credibility in the eyes of the other stakeholders.

REMEMBER

TCO comes from the invoice and projections. ROI is simply the mathematical comparison of the BIA and the TCO.

But now, because everyone has weighed in on the financial values and the relational impact of the formula, everyone believes the ROI, no matter how big or small. Going back to the concern you had around presumed credibility of the ROI formula:

- If the ROI is less appealing (for example, TCO is 50 percent of BIA), at least everyone was involved in understanding the legitimacy of the numbers, and you have a greater likelihood of their agreeing to the project.
- If the ROI is too appealing (not emotionally credible), you have the simpler problem of working with the vendor through side meetings to educate your stakeholder peers as to the legitimacy of the solution and the higher potential of being that hero by spending $10 to save $100.

Either way, having the initial formulas and variables challenged turns the project from yours to theirs and will help you pay for what you already know you want.

IN THIS CHAPTER:

- Talking more
- Keeping the conversation going

Chapter 5
Ten Keys to Better Conversations

We finish the book in the same way we started it — by recognizing that metrics on downtime and data loss may be expressed in terms of either IT technologies, operational concerns, or financial impacts. Keeping all three balls in the air may seem daunting, but the trick is simply to understand the audience(s) involved and to translate the numbers from one set of metrics to another.

Start with *More* Conversations

For any given business process, there are several different stakeholders. The conversations should start with the three key groups: IT technologists, business unit leaders, and financial leaders. After you've translated the business processes into both financial impacts and required IT dependencies, then there are more contributors to be invited: application owners, infrastructure architects, legal/compliance, and executive sponsors. Yes, this sounds like a lot of work, but getting them talking upfront will rapidly accelerate solution selection and the purchasing workflow as well as increasing overall organizational satisfaction with the resulting outcome.

Get Your SLAs Right the First Time

Starting SLA discussions with an assessment of current IT technology inevitably limits your imagination and may reduce your openness to consider what the business units need for assured uptime or data protection. Instead, first consider the business requirements and then assess your IT capabilities as they relate to satisfying them.

Realize Backup Admins Rarely Make Good Disaster Recovery Planners

Every business continuity/disaster recovery strategy includes an assured capability to recover data. But it would be a mistake to assume that a backup administrator has the breadth of skills necessary to assess the business processes and financial impacts necessary to frame out IT's portion of such a strategy. But if you turn that argument around, by learning how to have discussions on operational requirements and financial considerations, you're no longer "just" a backup admin.

Quantify IT Challenges in Operational and Financial Impacts

The only people who care about RPO and RTO for their own sakes are backup admins and data protection vendors. For everyone else, downtime should be measured in capabilities and impacts instead of speeds and feeds. This doesn't diminish the importance of measuring the uptime and frequency-of-protection of IT systems. But it should change the vocabulary and taxonomy that you as an IT professional use to articulate the imperatives associated with modernizing your data protection strategy.

Recognize That There is a Huge Cost in Doing Nothing

If you haven't gotten buy-in in all three languages on the importance of a robust data protection strategy, you'll likely hear

comments such as, "We can't afford to do anything more right now." But that's like complaining that you can't focus on the water bill while staring at a broken faucet with water gushing out. Server outages are occurring in your environment today, which means that you're experiencing some level of downtime and data loss today. And, thanks to Chapter 3, you can calculate what that data loss is costing the business today. Investing in better data protection and recoverability is far cheaper than continuing to pay for lost productivity caused by downtime and data loss.

Acknowledge That Downtime Costs More Than You Think

Most people think about downtime as an inconvenience because they haven't put numbers to it. So put numbers to it. If you start the conversation with the conceptual view of the formula in Chapter 3, you can reduce the four variables to two, by converting Hr (human costs) and Pr (profitability), plus any relevant fines or other ad hoc charges. At this point, the remaining two variables can be owned by IT, whereby you can check systems or management logs to count outages, lengths of each outage, and practical recovery times to return to service.

Make Sure Your Stakeholders Argue with Your Formulas

The best part of any discussion on the formula from Chapter 3 is when an operations or finance person says, "That doesn't really compute because of X." For example, if a CRM system goes down, maybe the sales teams aren't sitting idle (completely unproductive, in other words) but are instead encumbered. At that point, ask "How encumbered are they? Are they half as productive? One-third as productive? What number can you give me?" That number then becomes a multiplier of Hr or Pr. The great news is that now *your formula* becomes *the team's formula*. They're now bought in on the formula's validity in quantifying downtime costs.

Don't Forget the Soft Costs

Not every impact to downtime or data loss has a defensibly calculable numeric value. And yet, impacts such as brand dilution due to bad reputations, or employee morale due to unreliable systems, absolutely matter. For some organizations, the "soft" costs are the tipping point between an uncomfortable hard cost calculation and a decision to make improvements. Admittedly, the higher up the organization leadership you go, the more you may find that the soft costs are the actual motivation and that the hard calculations are needed to help justify the purchase. Either way, don't forget to capture the soft cost impacts while you're gaining consensus on the hard cost formulas and calculations.

Embrace the Fact That Regulations Are Your Friend

Regulations are often filled with the minutia of jargon, unrealistic mandates, and random hype that gets everyone excited, even though few people, if any, actually take the time to read the documents themselves. That said, regulations often have the side benefit of raising exposure among senior leadership: After the midlevel managers (or higher) have discussed the technical, operational, and financial considerations, they will very likely find executive sponsors who are more motivated to champion change if those changes also help with compliance to the regulation(s) that the organization is being held to.

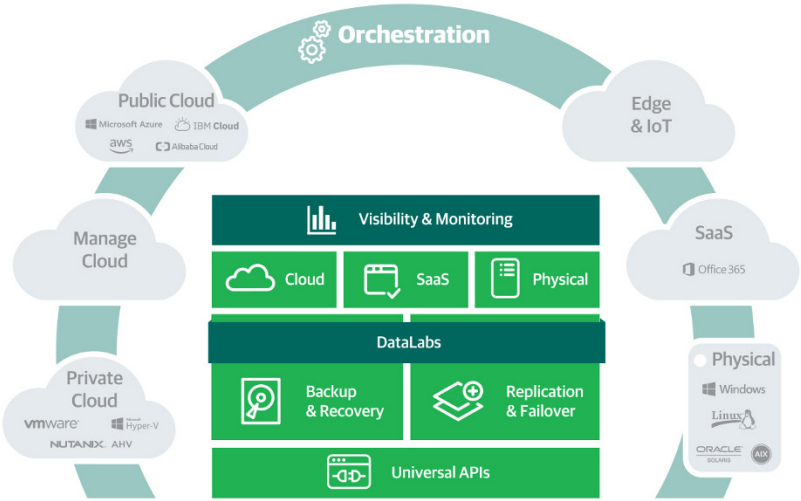
Get Your Data Out of the Building

Too many organizations begin their processes with meetings, surveys, and other administrative tasks, but their data remains insufficiently protected throughout the process. Later, they have a disastrous outage, and nothing survived except a big binder of plans that hadn't been acted on (yet). Instead, get your data out of the building. Ensure that your data is survivable, even if the current solution isn't particularly elegant or doesn't match how you'll likely deploy at some future date within a more modern data protection architecture. But at least this way you'll have a chance of getting what you need if disaster strikes while you're still in the early stages. So *get your data out of the building*.

Veeam Availability Platform

The most complete solution to help you evolve the way you manage data

- From policy-based to behavior-based
- Making it "smarter" and self-governing
- Ensuring availability across any application and any cloud infrastructure



Learn more: http://vee.am/vap

Downtime & Data Loss are Business Problems, not IT Issues

Data protection initiatives should start by understanding the business processes that rely on IT systems. Only after you've quantified the business impact of downtime and data loss can you effectively determine a data protection and availability strategy — and choose the appropriate tool(s) to mitigate those potential business impacts. Lastly, nothing will change unless you can justify the economics of investing in better-than-legacy mechanisms. This book addresses all of these concepts: RPO/RTO, SLA, RA/BIA, and ROI/TCO — as well as step-by-step instructions on how to calculate the real costs of downtime.

Inside...

- Define strategies that engage technical and non-technical stakeholders
- Identify key IT systems that business operations rely on
- Assess business impacts and risks due to system failures
- Calculate the real costs of downtime
- Map out strategies that ensure data protection & availability

VEEAM

Jason Buffington @JBuff has been working with backup, replication, and BC/DR solutions in the IT industry for over 30 years. Outside of IT, Jason has been married to Anita almost as long as he has been backing up data. He has three amazing kids, Joshua, Jaden, and Jordan, and is an active volunteer in Scouting. He lives by two credos: 2TIM1:7 and *"When you modernize production, you must modernize protection."*

ISBN: 978-1-119-57742-3
Not For Resale

for dummies®
A Wiley Brand

Also available as an e-book

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.