ActualTech MEDIA

GUIDEBOOK

Moving Beyond Disaster Recovery to Achieve IT Resilience

How to Prepare for Disruptions to Ensure Continuous Availability

James Green & Tom Howarth

Zerto

Moving Beyond Disaster Recovery to Achieve IT Resilience

AUTHORS
James Green & Tom Howarth

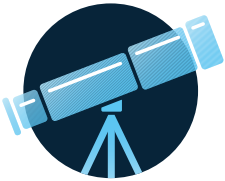
EDITOR
Keith Ward, ActualTech Media

LAYOUT AND DESIGN
Olivia Thomson, ActualTech Media

ACTUALTECH MEDIA
Okatie Village Ste 103-157
Bluffton, SC 29909
www.actualtechmedia.com

ENTERING THE JUNGLE

CALLOUTS USED IN THIS BOOK



THE 101

This is where we turn when we want to provide foundational knowledge for the subject at hand.

OFF THE BEATEN PATH

This is a special place where you go to discover insight into topics that may be outside the main subject but that are still important and relevant.

BRIGHT IDEA

When we have incredible thoughts (at least in our heads!), we express them through eloquent phrasing in the Bright Idea section.

DEEP DIVE

Takes you into the deep, dark depths of a particular topic.

EXECUTIVE CORNER

It's not all tech all the time! This is where we discuss items of strategic interest to business leaders.

ICONS USED IN THIS BOOK

DEFINITION

Defines a word, phrase, or concept.

KNOWLEDGE CHECK

Tests your knowledge of what you've read.

PAY ATTENTION

We want to make sure you see this!

GPS

We'll help you navigate your knowledge to the right place.

WATCH OUT!

Make sure you read this so you don't make a critical error!

CHAPTER 1

IT Resilience Keeps Business Moving

Maybe you've had this situation in your career: someone from higher up the corporate food chain calls and explains a project that has just been approved by the company's board of directors. The project will ultimately produce excellent results for the company, but since you've been working in IT for some time now, you can clearly see that the path between where you are today and where you'll end up is littered with late nights, planned downtime, and probably lots of frustrated co-workers as you commence a significant migration initiative.

And as this executive is wrapping up the call, they drop a bomb on you. "Oh, and we need to get started right away. We need to have this done before summer."

You instantly object in your head. "But...it's already spring, and this project will take at least nine months!" And yet, you know it doesn't matter. You're going to have to find a way. That's why they call you The Fixer. You take situations like this and combine your industry experience with the right tool for the job and make the impossible seem possible.

And so, with a sigh, you call a meeting with your best people, and you begin.

The Plot Thickens

Fast forward two months. The migration effort is well underway, and you're on track for a fat bonus thanks to your stellar leadership on this project.

Just as you're getting ready to head home from the office one late-spring Friday afternoon, you overhear some groans coming from the helpdesk area. The helpdesk team sits a few aisles up the cube farm from where your team sits, and you always know when something bad is happening because it starts to get suspiciously loud over there, and you can usually spot a bit of a crowd beginning to form.

You wander that way, which is also conveniently the same direction as the kitchen. You often use this "coincidence" to your advantage when a significant issue is brewing over in the helpdesk territory. You walk over to the kitchen to grab a glass of water, but as you pass the helpdesk team, you linger just long enough to get an idea of what's happening.

As you're strolling past, feigning disinterest like you should be on Broadway, you overhear a word that makes your heart sink like someone just unlatched a trap door and your heart fell right through it.

"...ransomware..."

And just like that, you can envision your weekend disappearing right before your eyes. Not only that, but your mind leaps to next week and the week after: all of the resources you need to complete your highly visible migration project are going to be tied up unwinding this ransomware mess.

Your mind begins to race. You're going to miss your bonus milestone, which means you're going to have to go home and tell your spouse that you have to cancel the vacation you had planned to take with the bonus money. In your mind's eye, you can vividly see your company logo on CNBC with a big skull and crossbones next to it. As

the anchor talks excitedly, the ticker across the bottom says, "Latest Ransomware Victim's Stock Price Falls at Record-Setting Pace." Are you even going to have a job in three months?

Now stop.

That knot you feel in your stomach and that shortness of breath... that's why this book exists. The drama you just read was fictional, but for many IT professionals and business leaders, parts of that story become all too real every day.

This book is about how to protect your business from outages – both planned and unplanned – so that you can focus your efforts where they matter: on completing those projects which mean the most to the progress of the business. The ones that help you increase market share, decrease costs, and innovate faster than your competitors.

This book will teach you how to create **IT Resilience**.

IT resilience is an emerging term that describes a design goal that allows business to accelerate transformation by adapting to change while protecting the business from disruption. IT resilience can save more than just your weekends and your job, too; it can potentially save your company. Some companies never recover from a major outage or from an incident that results in a substantial loss of data. And those businesses that do recover can *never* recover the opportunity cost of fighting fires instead of making progress.

As businesses undertake a journey of digital transformation, their data becomes even more valuable than it was before such efforts, and requires more comprehensive levels of protection from outages – both those that we're expecting and those which we could never have anticipated. IT resilience allows organizations to build an "always-on business" and deliver a superior customer experience to their clients.

Digital Transformation and the Growing Need for Resilience

There are myriad conflicting definitions for Digital Transformation. One of the best, most comprehensive, is by storage industry executive Bill Schmarzo, who said: "Digital Transformation is [the] application of digital capabilities to processes, products, and assets to improve efficiency, enhance customer value, manage risk, and uncover new monetization opportunities."[1]

Schmarzo defined "digital capabilities" as those that are "electronic, scientific, data-driven, quantified, instrumented, measured, mathematic, calculated and/or automated."

In layman's terms, digital transformation is the process by which businesses are using technology to meet their business objectives, as well as revolutionize and reinvent the industry itself.

There are innumerable ways technology can improve business. It could be by applying digital measurement and analytics to a manufacturing operation to reduce material and labor costs while increasing output. It could be a healthcare organization using Artificial Intelligence and Machine Learning to apply security policies to sensitive patient data so that they consistently meet compliance requirements.

But digital transformation, applied to business models, becomes digital *disruption*. It could be as dramatic as applying new technology to an old business model to create an entirely new business model. There are plenty of famous examples of this:

Taxis → Uber

Long distance phone calls → Skype

Bed and breakfasts → AirBnB

Movie rental stores → Netflix

[1] https://www.cio.com/article/3199030/analytics/what-is-digital-transformation.html

And there are plenty of less well-known examples that have completely changed lives and upended the way business is done (in a good way).

If you think about how full-blown digital transformation would change a business, you see how there's no going back and why digital transformation is not just a buzzword du jour, but is here to stay. Once a company comes to rely on the technologies applied, the processes created, and the employees empowered by digital transformation, those technologies, processes, and empowerment tools become business-critical.

Any future loss of service causes a significant disruption to business operations. A healthcare organization that fails to control access to patient data will pay harsh fines and lose the trust of their clients, for example. A manufacturer that relies on a totally automated production system will have to stop the production line if technology fails.

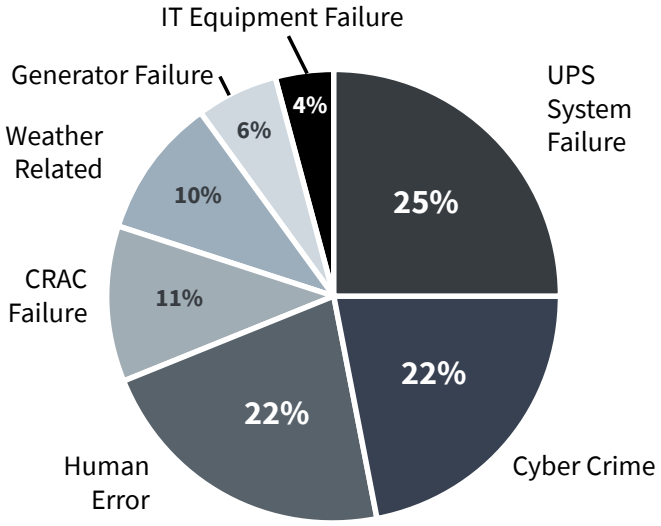


Figure 1: Root causes of unplanned outages. (Source: Ponemon Institute)

As a result of this new and enabling – yet potentially crippling – dependence on technology, businesses need to foster IT resilience. Resilience will afford them as little disruption as possible, even as IT systems encounter software bugs, as hardware fails, and administrators migrate, consolidate, and cloudify systems.

What Is IT Resilience?

In this book, we'll use the term *IT resilience* to refer to a framework for avoiding, mitigating, and remediating failures of information technology systems, and increasing availability during planned outages and migrations. The goal is to prevent disruption to the business. There are three main components of a robust resilience framework:

- Continuous availability
- Workload mobility
- Multi-cloud agility

Let's take a closer look at each.

Continuous Availability

When technology is the lifeblood of your business, downtime is unaffordable. Modern consumers expect availability at all times. Your internal consumers rely heavily on technology to do their jobs. Therefore, it's imperative that a modern IT organization provide services engineered to be immune to mistakes, hardware failures, security compromises, and data loss. There are a number of avenues available to you to help your company maintain high levels of availability.

MITIGATE EXTENDED OUTAGES

Extended outages happen—and will always happen—as long as data centers continue to exist. But with the availability tools at your disposal today, extended outages don't have to happen to *you*. This

book is for you if you want to know how to protect yourself from outages and disruptions to your business.

REPEL RANSOMWARE

Ransomware infects a growing number of organizations every year, and recovery can take days, weeks or even months. And, unfortunately, in some cases, recovery is impossible. What are the practical implications for your business if you have to send everyone home for a week while the IT staff recovers systems after ransomware encrypted the file server? With the right level of protection, you can restore hijacked files back to normal in seconds; you'll learn about that later on.

REVERSE HUMAN ERROR

What do you think is the most common reason for a file to go mysteriously missing from a user's desktop? That's right—they accidentally deleted it! But you can be their hero by ensuring that your data protection strategy can quickly and easily restore a file—or an email, or a database entry, or whatever else they accidentally removed from existence—instead of waiting for an entire volume to restore. In Chapter 2, we're going to cover all of these availability concerns in detail and show how IT resilience can help you avoid the pain.

Workload Mobility

The gift of workload mobility was perhaps most famously bestowed upon the world by VMware with their then-revolutionary vMotion technology. vMotion impressed everybody with its ability to move a running workload from one ESX host to another, with no visible interruption of service. That same level of seamless transition is now desired throughout the data center, and the concept of workload mobility is the idea of creating a vMotion-like experience across all different types of resources and environments.

In a resilient IT environment, the workload moves from the old infrastructure to the new with no disruption of service as an organization refreshes infrastructure hardware. The same goes for migrations during company mergers and the like. All of this shuffling of workloads can and should happen with no impact to the end users.

Workloads aren't the static entities they're sometimes made out to be. They're always in flux, moving even when there aren't upgrades and maintenance underway. For example, data and workloads are continually shuffling through different parts of the software development lifecycle. Workloads that begin life in Dev might ultimately migrate to Test and then to QA and then to Production. And these workloads can exist across multiple hypervisors and clouds, further complicating the environment. In a resilient infrastructure, this migration between environments is seamless and doesn't hamper productivity. Workload mobility will be the sole focus of Chapter 3.

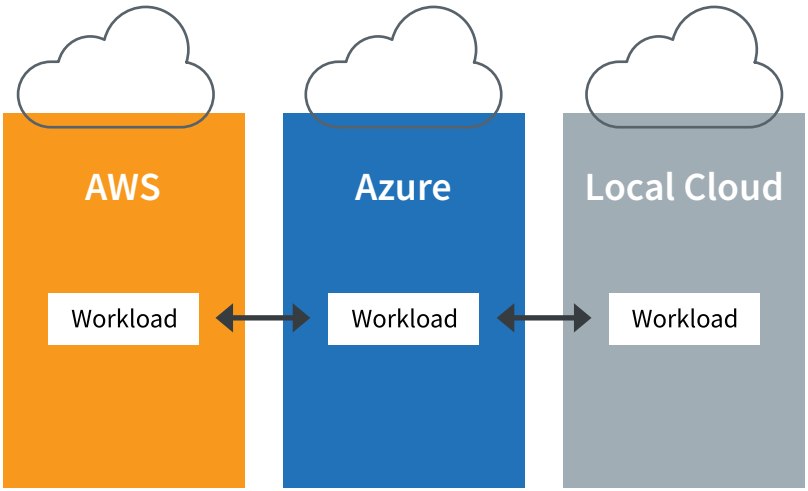


Figure 2: Workloads may move from cloud to cloud as requirements change.

Multi-Cloud Agility

In the cloud era, internal workload mobility is just the start. No longer are workloads confined to the data center perimeter. Now they move from one cloud to another, both on-premises and off; in fact, hybrid clouds are becoming the standard for many organizations.

Moreover, advanced IT organizations sometimes use multiple public cloud providers, and they want the ability to not only shift workloads physically within data centers or regions, but between cloud providers at will.

The ability to move workloads easily between clouds provides resilience against failures, but also helps to protect against changing cost models, feature support challenges, and service-level agreement/service-level objective (SLA/SLO) problems with the provider. If a resilient IT organization has problems with one cloud, they're prepared to swing their workloads over to a different provider without encountering downtime or technical complications. We're going to look at multi-cloud agility in depth in Chapter 4.

Data Protection Has Evolved

Data protection is not (all) about hurricanes, tornados, fires, floods, and squirrels running across power lines anymore! As we've seen, IT resilience goes far beyond the boundaries traditionally covered by data protection. It's no longer just about backup or replication, or even just about adding redundancy to the infrastructure layers or high availability into the logic layers.

Yes, data protection still means being prepared for the biblical flood that threatens to drown your data center. And being ready for the chaos that ensues when Ethel opens up an email promising a picture of Toonces the Driving Cat[2] and finds that it was actually a malware attack, and all her files are now encrypted thanks to the scourge that

[2] https://www.youtube.com/watch?v=5fvsltXYgzk

is ransomware. But to be *genuinely* protected today, redundancy alone isn't enough. Backups alone aren't enough. What's needed is a platform that offers resilience from beginning to end.

IT resilience takes the traditional notion of data protection—which is *reactive*—and adds to it the objective of accounting for planned outages, which is *proactive*. Even anticipated unavailability can cause a problem for the business: disruptive upgrades, workload relocation, and cloud migrations are all legitimate reasons for downtime, yet there is still a cost to the company for the outage. It would be better — and perhaps even mandatory in the future—for potentially disruptive activities like upgrades and migrations to be done without impact.

Moreover, traditional data protection focuses on preparing for the inevitable failure of data center systems. In a world where security breaches and malware infections are the norm rather than the exception, the scope of IT resilience is broader than just equipment failures and misconfigurations. Organizations need a reliable plan for recovering from a ransomware attack, for example, or they could find their business offline for days or weeks.

So how do you begin to build a resilient IT infrastructure? It all starts with perspective.

Identify Risk

Risk has been and will always be the enemy in IT, but that doesn't mean you should ignore it. The identification of risk is the first step in remediating potential system weaknesses and creating resilience. The rest of this Guide will focus on practical ways to reduce your risk exposure.

As you undertake this journey, however, it's important to remember that risk isn't limited to technical risk such as a single point of failure. Your risk identification activities also involve processes and people.

Consider this: Gary the Guru knows everything there is to know about your systems, but one day he leaves for his dream job with an underfunded startup. Has Gary correctly and sufficiently documented everything so that one of his peers can pick up where he left off? Or will the project be delayed because everyone else is entirely in the dark? In other words, is Gary the Guru's tribal knowledge a risk for your organization?

Control What Is Within Your Control, Design Around What Isn't

As you integrate IT resilience strategies, you'll have to account for the level of control you have relative to the concern. For example, your wide-area links are an external risk that you have little power to control; your telecommunications provider controls them. If something goes wrong on the telco network, you're powerless to fix it and are at the mercy of the provider.

Variables being outside your control doesn't mean that you're without mitigation options, though. You may not be able to control the telco network directly, but you can design around this issue. For example, because of this risk, are you utilizing diverse routing over disparate links owned by different providers? These sorts of decisions are part of IT resilience, too. If you don't have the reach to fix the real risk (e.g., the WAN link could go down), then you have to design around it (implement redundant WAN links).

The art of IT resilience is the orchestration of multiple different processes and technical solutions to protect what is effectively a company's crown jewels: the availability of its data and applications. The move, first to virtualization, and then to software-defined computing and all forms of cloud computing, has enabled the possibility of a heretofore unknown level of resilience. But with the luxuries of cloud computing and the ubiquity of the Internet come some

new challenges as well. The remainder of this guidebook will highlight the new challenges in the evolving landscape of IT, as well as some advice for dealing with them to create a resilient IT infrastructure.

CHAPTER 2

Provide Continuous Availability

Once upon a time, IT's internal customers didn't leverage technology at home the same way they did at the office. So those users had no expectations. They were happy with what IT gave them because it helped them do their job better.

How things have changed! With a smartphone in every pocket and the Netflix, Dropbox, and Gmail experiences as the measuring stick by which users judge IT, the expectations for functionality and availability are at an all-time high.

Since modern IT consumers are spoiled with hyperscale functionality and availability that's delivered at the click of a mouse and the swipe of a credit card, the goal of enterprise IT must be the same if there's any hope of subverting shadow IT activity within the organization. What consumers are asking IT for today looks something like this:

- Never go down; keep users working
- Never be crippled by malware
- Never be unable to recover data

In this chapter, we're going to take a look at some of the most prominent reasons productivity within a business can suffer, and discuss how to create resilience in those areas.

Outages and Disruptions

As much as we strive to avoid them, outages and disruptions are part of life for every IT department. The occurrence rate is still hovering around 100% over a long timescale. A service disruption or unplanned outage can have disastrous consequences for organizations, both financial and reputational. That's especially true in this age of Twitter, Facebook, and other social media outlets. The PR fallout of an outage can be nearly instant, merciless and severe. The social backlash and loss of business due to one of these interruptions can even spell the end of an organization.

With the demand for 24/7, always-on availability rising, it's not surprising that unplanned outages become front-page news. In early 2018, Australia's largest airport caused considerable disruption to passengers because of an outage that they blamed on a technical issue. The failure was immediately and widely visible. Because of the modern news cycle, the public often measures failures by the speed and sure-footedness of the recovery. And beyond the shame, there's usually a real financial impact, too. According to Gartner, enterprise downtime costs more than $5,000 per minute, on average.[3]

Outages Then and Now

In the early days of IT, we protected our files by saving them to floppy disks. Fortunately, we've improved on that: we're not typically worried about a SCSI drive failure or a DIMM corruption anymore. The rise of virtualization has had much to do with that.

It's meant, among other things, that the failure of a physical server is not generally a business-crushing event, and unplanned outages due to failures at the micro level are now rare. Such failures still take place, but they don't have a negative impact because organizations have designed around them. Clustered resources have meant that

[3] https://blogs.gartner.com/andrew-lerner/2014/07/16/the-cost-of-downtime/

EXECUTIVE CORNER

The Staggering Cost of Outages

Consider that the cost of outages for the company can be much higher than just a lost weekend for you. The estimates of financial losses when a data center goes down are downright frightening:

- 98% of enterprise organizations say a single hour of downtime costs more than $100,000
- 81% of enterprise companies indicated that 60 minutes of downtime costs their business more than $300,000
- 33% of those enterprises reported that one hour of downtime costs their firms between $1 million - $5 million

virtual machines (VMs) can either be moved pre-emptively (vMotion/Live Migration) or automatically restarted on a second node (high availability) or never fail at all (e.g., vSphere's Fault Tolerance).

Despite those impressive availability gains, there are still issues at the operating system layer. So, for some mission-critical services such as databases, clustered services at the OS layer are still widely used.

As we move into the age of cloud-native applications, the rise of stateless application stacks will reduce the scope of machines that need to be protected by traditional backup processes. Stateless systems can be protected once as a golden master, and the individual configuration files can reside in a source code repository; any failed workload can be restored to service with little or no outage to the production services that rely on it by simply instantiating a new one.

However, the buck has to stop somewhere, and there will *always* be a need to protect data. While there may be fewer individual machine

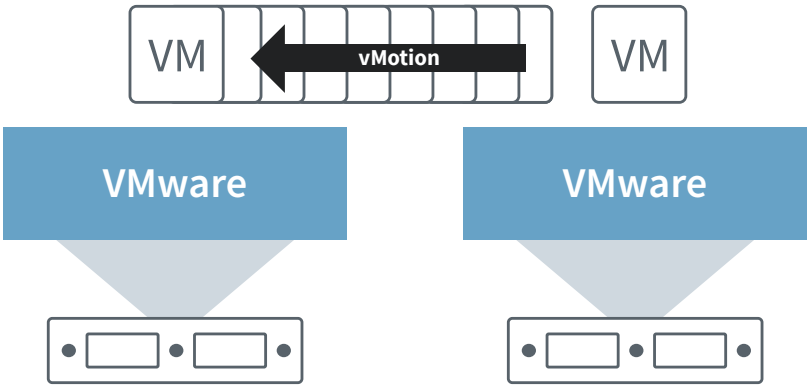


Figure 3: Clustered hypervisors allow VMs to move between physical servers.

instances to protect in the future, there will certainly be more data to protect - a lot more. And the recovery of stateless apps needs to be orchestrated; doing that alongside the persistent data is still a serious challenge. As businesses rely on their data more and more, even a performance degradation is viewed as an outage, regardless of whether or not the system is technically available. How we go about implementing data protection and recovery orchestration will determine our level of IT resilience into the future.

Mitigation of Outages or Disruptions

Thankfully, widespread outages caused by natural disasters are rare in comparison with outages caused by humans. The vast majority of outages and disruptions are caused during periods of change, patching or upgrades when something goes wrong.

With the rise of machine virtualization, IT pros started to use snapshot technology to return to a pre-change or pre-upgrade state to rapidly recover from a failure; this dramatically lowered disruption time. Snapshot technology also provides better, lower recovery point objectives (RPO) and recovery time objectives (RTO) by having a regular snapshot policy to shrink the data loss window.

Snapshots, however, come with serious downsides, including substantial disk space utilization, increased read and write overhead, and added processing costs when reverting or deleting a snapshot. Storage-level snapshots and array-offload with VAAI or ODX enable more efficient snapshotting, but even using storage snapshots comes with overhead.

Many companies also use storage-based replication technologies to provide two or more copies of data, replicating them across storage arrays either asynchronously or synchronously, depending on availability requirements. Storage-based replication is excellent for site failures, but suffers when dealing with data corruption like ransomware. Because a committed write at the active site is also committed at the replication site, a problem on the active data set is almost immediately propagated to the DR site, which renders the replicas useless. That's bad.

Finally, whether VM-based or storage-based, as the number of snapshots on hand grows, the DR strategy becomes cumbersome, and keeping track of the snapshots becomes a tall order. Large numbers of snapshots can be unwieldy. But there is a better way.

Journals Are Better Than Snapshots

In the modern age, *any* loss of data is generally deemed unacceptable by the business. IT organizations strive to provide the lowest RTO and RPO that their budgets will allow. A unique way of protecting data more efficiently than the snapshot-based method is to leverage Zerto's Continuous Data Protection (CDP) engine. Zerto CDP is hypervisor-based and doesn't utilize snapshots indefinitely to protect the VM; therefore, it suffers none of the usual performance penalty associated with snapshot-based data protection.

Instead, Zerto CDP utilizes a journaling technology to keep a log of all changes that have occurred during a customizable period — configurable down to the second. This technology significantly

reduces the period of potential data loss (the RPO), and thus the potential financial impact or reputational cost of failure.

The use of journaling technology to keep track of changes provides multiple benefits over traditional snapshot protection:

- **Zero performance impact.** Journaling provides continuous block-level replication with zero impact on application performance, as opposed to snapshots which have a substantial performance impact as the snapshot tree grows.
- **Storage capacity savings.** A journal allows placement of protection data on any datastore, and administrators can specify maximum file size limits and warnings. With snapshot-based technology, there's no way to control the total space used for snapshots. Therefore, you run the risk of eating up all available storage capacity.
- **Increased storage utilization.** Journaling technology uses no extra space on the local datastores, since no snapshots are being created. Storage-based replication often sets aside significant capacity on the arrays for replication reserves; it can be as high as 20% - 30% on both source and target storage. With journaling, significantly less storage at the target site is used for storing changed data, thus freeing up significant amounts of additional disk space.

Journaling can bring storage capacity savings, because they can dynamically reclaim unused space; they use only the capacity they require. If the journal fills up, it will simply start to reduce the number of recovery points available, starting with the oldest; alternatively, the system may dynamically increase the journal space usage depending on the configuration parameters. This behavior is in contrast to storage-based snapshots; when they run out of replica reserve, you can plan on having a bad day.

Journaling is a powerful tool in developing IT resilience, and combining it with Zerto's automation and orchestration capability allows the development of complicated recovery processes with zero manual intervention. It also provides simple fail over and fail back with as little as three mouse clicks.

It's so easy to fail over and fail back, in fact, that organizations can regularly perform failover testing (in a sandbox, of course) to confirm that their recovery will be successful in the event of a real catastrophe. The fear of crippling, extended outages can finally be put to rest with a CDP-based availability strategy from Zerto.

Ransomware

Ransomware. It's a term that strikes fear into many a systems admin's heart. But what exactly is it?

A successful ransomware attack results in data on the infected machine being encrypted with a key that the user doesn't have access to. The result is that all data on the computer is effectively inaccessible to its owner. At this point, payment—the "ransom" in ransomware—is demanded before the data will be decrypted and access restored. One distinctive feature of a ransomware attack is that the victim is informed of the hijacking, and instructions are given on how to recover from the attack. This is a unique characteristic compared to other types of common malware, which wreak havoc without any suggestion as to how to recover.

Payment is usually requested in the form of bitcoin (or some other cryptocurrency where anonymity is a feature) to avoid creating a digital paper trail that leads straight to the attackers. Even if the unique wallet address of the attacker is known, it's impossible to deduce who's controlling that particular cryptocurrency wallet.

Although ransomware incidents have grown significantly over the past decade, it's not an entirely new concept. One of the first ransomware

attacks was called AIDS, and was first observed way back in 1989. But ransomware has gained prominence and mass public awareness in the last couple of years with variants like CryptoLocker[5], which is now one of the most iconic examples of how fast-spreading and ruthless it is. More recently, advanced evolutions such as WannaCry and Petya have increased the fear and trepidation about the possibility of a ransomware infection. The unfortunate truth is that ransomware has moved well beyond a simple financial transaction. Recent attacks have had actual life and death impact that's impossible to quantify in financial terms alone.

The impact on businesses of a successful ransomware invasion can be devastating. More than the loss of a few thousand dollars' worth of bitcoin or a few hours of productivity, when ransomware is deployed

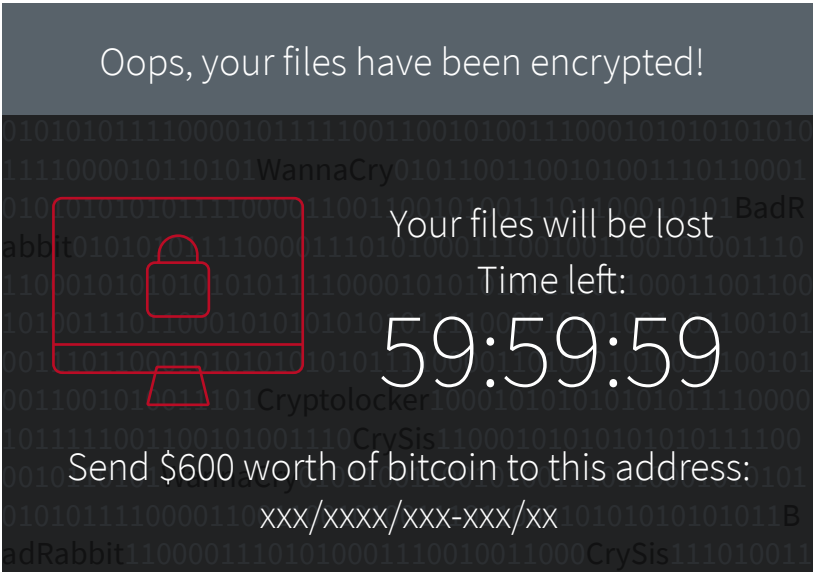


Figure 4: A typical ransomware demand is to send bitcoin in exchange for decryption.

[5] https://blog.malwarebytes.com/101/2013/10/cryptolocker-ransomware-what-you-need-to-know/

in certain places, it can have widespread ramifications which directly threaten human lives. Here are a few real-world examples:

- **Nuclear power plant monitoring failure.** Systems used to monitor the area around the Chernobyl nuclear power station – which is still an active and dangerous area following the disaster in 1986 – were rendered unavailable because of the cyber-attack known as Petya. This dangerous situation had to be closely monitored manually until systems could be restored.[6]

- **Hospital forced to rebuild from scratch.** Following another Petya infection, Princeton Community Hospital in rural West Virginia reportedly had to scrap and completely rebuild their entire IT infrastructure from scratch. Doctors were unable to access patients' medical history, transmit lab orders and results, or place pharmacy orders.[7]

- **Boeing production plant grinds to a halt.** An airplane production facility in Charleston, S.C. found itself prey to the WannaCry attack, warranting a company-wide memo calling for "all hands on deck." Although the footprint was ultimately found to be small, the panic was hard to contain and surely disrupted productivity for the entire IT staff.[8]

- These are the sorts of situations that IT leaders hope to avoid for their organization. With a strong IT resilience strategy, recovering from a ransomware attack can be done in minutes and without panic.

[6] https://ind.pn/2lox2uH

[7] https://www.wsj.com/articles/cyberattack-forces-west-virginia-hospital-to-scrap-its-computer-systems-1498769889

[8] https://www.seattletimes.com/business/boeing-aerospace/boeing-hit-by-wannacry-virus-fears-it-could-cripple-some-jet-production/

Mitigation of a Ransomware Attack

The overwhelming majority of ransomware attacks are delivered via email attachments, and rely on human curiosity to infect their intended victims. Some are obvious: an email comes into their inbox from a company they've never dealt with before, including an invoice for payment or a notice of a return shipment.

Others appear to be from people the victim is familiar with and contain either a threat ("I just saw these pictures of you") or make an offer that's hard to refuse, coupled with tales of how their lives will be immeasurably improved. The attachments are usually compressed to avoid detection by scanners, or they masquerade as PDFs and other common files. The file instead contains the destructive payload.

So, given the very human intricacies involved in these attacks, how can they be defended against? The best weapon for ransomware defense is education—teach your users to never open unsolicited emails. Common sense is the key here; if an offer appears to be too good to be true, it probably is, and that should send up a red flag.

But even with education and care, compromises will still happen. Sometimes people are distracted, they're in a hurry, or they're just technically unskilled. Count on it: sooner or later, your organization's machines *will* be infected.

This highlights the importance of regular backups, which allows the recovery of data to a known point in time. However, traditional backup methodologies will usually result in the loss of at least a day's worth of data; and this is assuming that your backup data is recoverable at all. In many cases, a ransomware infection will have also encrypted its victim's backups, rendering them inaccessible. A better option is to have multiple copies of data, in multiple locations, with more frequent updates (a shorter RPO). This allows for a more granular recoverability window.

Organizations also can, and should, implement access controls for users that limit access to only what's required for their roles, then implement a secondary personal administrative account with no access to email. This is considered a basic security practice but is often ignored.

Ransomware infections are fully recoverable when corrective measures are in place. You always have the option of paying the ransom; reports indicate that the majority of ransomware perpetrators do in fact decrypt files after payment. You can't be *sure* they will, however, and you're at the attacker's mercy with regard to how long they take to decrypt your files.

An infinitely better option is to restore the affected files, then find and delete the offending email from the email server and local client. If you haven't properly backed up your data and find yourself without a second offline copy stored somewhere, you're in trouble. But if you've got journaled replication, as discussed before, recovering from ransomware is simply a matter of rolling back to the journal checkpoint just before the ransomware infection and then taking a different course: in this case, deleting the e-mail instead of opening the attachment. It's almost like going back in time to correct a historical fumble.

Limiting, Or Even Eliminating, Ransomware Damage

Zerto's IT Resilience Platform™ uses the company's CDP engine to provide incremental, block-level replication. Coupled with the journaling discussed earlier and Virtual Protection Groups, it provides an excellent mitigation tool when the infected machine is running on a virtual host.

EXECUTIVE CORNER

The NIST Cybersecurity Framework (CSF)

The Framework for Improving Critical Infrastructure Cybersecurity, better known as the Cybersecurity Framework (CSF), defines five functions: Identify, Protect, Detect, Respond, and Recover. These functions are all critical for a complete defense. At a more fundamental level, the capabilities in the Recover function have a significant effect across the organization by providing realistic data for improving other capabilities.

Recovery can be described in two phases, focused on separate tactical and strategic outcomes:

- The immediate tactical recovery phase is largely achieved through the execution of the recovery playbook planned prior to the incident (with input from Detect and other CSF functions as required).
- The second phase is more strategic, and it focuses on the continuous improvement of all the CSF functions to mitigate the likelihood and impact of future incidents (based on the lessons learned from past incidents as well as from other organization and industry practices).

In the context of this Guide, it's interesting to consider how enhanced Recovery capabilities serve to fortify cybersecurity posture on the whole.

The Guide for Cybersecurity Event Recovery can be downloaded here: nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST SP.800-184.pdf

VIRTUAL PROTECTION GROUPS

With Zerto's Virtual Protection Groups (VPGs), you can group VMs and virtual disks to be protected and recovered together. This provides complete application consistency, regardless of the physical location on servers and storage. Protecting workloads with VPGs allows you to use an application-level granularity as opposed to a LUN-level granularity.

The ability to roll back to a point in time only *seconds* before the infection can save a company's hide, not to mention their data. In a traditional backup and snapshot protection scenario, RPO time is likely hours at a minimum; this is a significant amount of data loss for a company to suffer or recover from.

CDP, however, allows recovery points in increments as small as five seconds. The ability to roll back to a known-good point in time, seconds before the infection, means that a company can effectively turn back time. Further, simply mounting relevant infected VMs at a checkpoint that's known to be good enables the ability to restore individual files.

With the amount of granularity Zerto provides, data loss can be minimized to a level of near triviality.

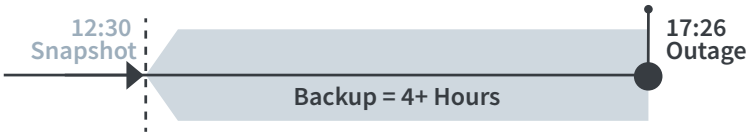


Figure 5: Traditional backup recovery lag can be severe, impacting an organization's bottom line.

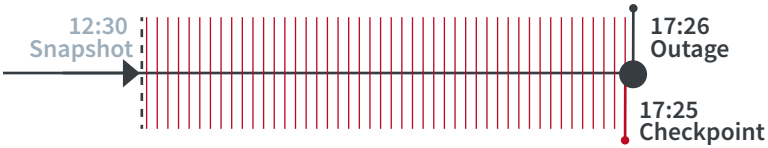


Figure 6: Minimal data loss with journaling. This CDP checkpoint can allow a much smaller RPO than a traditional snapshot or backup can.

BRIGHT IDEA

Ransomware Recovery Served Two Ways

Textiles technology group Royal TenCate (TenCate) is a multinational, global company with more than 4,000 employees. The company develops and produces solutions for the protection of people and their working and living environment.

TenCate has experienced multiple ransomware infections and responded both with and without the aid of Zerto technology. You may find their results surprising. The table below illustrates the difference. The Zerto advantage is clear.

	Without Zerto	With Zerto
DR STRATEGY	Backup to tape	Continuous replication
FILES AFFECTED	An entire file server at a manufacturing facility	A number of directories on a file server at a manufacturing facility
DATA LOSS	12 hours	10 seconds
RECOVERY TIME	2 weeks	Under 10 minutes

Complete Data Protection

It used to be the case that the mission-critical was protected, and the rest wasn't. It was typically a narrow band of protection when compared to the whole infrastructure. However, as businesses come to rely more heavily on their data and their applications for providing value to their customers, the scope of what must be protected is expanding.

The scope is changing, not only for *what* must be protected but also for *how long* it should be protected. As the value of data within organizations grows, it's justifiable to keep it for longer and longer periods. Sometimes it's not the business that wants to keep historical data around—in some cases, regulatory compliance demands it.

"Complete data protection" means being able to protect your data from 7 seconds to 7 years. Knowing that what you intend to protect is, in fact, protected is trickier than it might seem. Any backup admin will tell you that it's not the backup process completing successfully that's important; it's the *restore* completing successfully that's important!

The objective for an organization focused on IT resilience should be to:

- Protect all data worth protecting,
- at a useful level of granularity,
- with an acceptable RPO and RTO,
- at any scale,
- and regularly confirm that they can recover that data successfully.

A common practice in traditional backup strategy is what's known as the "3-2-1" strategy. It states that for data to be adequately protected, there needs to be at least three copies of it, on two different types of media, with at least one copy existing offsite.

EXECUTIVE CORNER

What Guarantees Do You Make To Your Business?

At the crux of data protection are RTO and RPO. You should understand these metrics inside and out if you're going to be helping shape your organization's push for IT resilience.

- **Recovery Point Objective.** The amount of data your business can tolerate losing, measured backward in time; for example, one day's worth of data, one hour, five minutes, 10 seconds from the incident. As a rule of thumb, the shorter the RPO, the higher the cost to implement that level of protection.
- **Recovery Time Objective.** The length of outage or service disruption that your business can tolerate before services are restored, measured forward in time from the incident.

Commonly, and traditionally, the way for IT departments to provide these levels of availability to their business is with backups. Backups provide a somewhat high (8 - 24 hour) RPO/RTO, and that may not be good enough for today's continuous availability needs. Backups are commonly run once per day; thus, you can't recover data any more recently than last night's backup. Many organizations stop here and consider this their data protection and recovery strategy.

Many organizations have yet to move beyond this solid but no-longer-sufficient data protection paradigm. But there's a whole new world out there awaiting organizations who are finding that so many backup use cases, like the ransomware examples noted earlier, aren't being adequately served by periodic backup technologies. Backup is not built to handle seconds of RPO at scale.

Figure 7: The 3-2-1 backup strategy.

Issues with Traditional Backup

A traditional backup strategy has served countless organizations well. And for some, it may still be an adequate way of thinking about protecting data. But with burgeoning data volume comes elongated backup job run time. For many organizations, the open data protection window is overnight or over a weekend when all employees have gone home. In many growing environments, the time needed to complete even an incremental backup regularly exceeds the available window in which to complete those backups. This causes a big problem when it comes to retention and compliance: backups in these environments often get skipped, fail to complete, or get canceled by an administrator when they begin to have a negative impact on production performance, since they're still running during working hours.

Many organizations today continue to deal with the backup technology of yore and its limitations:

- **It's time-consuming and expensive.** Virtual tape libraries and even physical tapes are expensive, and with physical tape comes significant risk of unrecoverability of data and an excessively high RTO. Tape libraries are a single point of failure.

- **Backup windows are too short.** Many companies have full backup jobs that run far longer than 24 hours, and for many companies, even incremental backup run times are encroaching on business hours. This is problematic both in terms of production workload performance and recoverability (if jobs fail to complete and have to be canceled).
- **Recovery takes too long.** Recovery time from traditional backups is long, complicated, and error-prone.
- **Many backup use cases are granular.** As the ransomware example shows, recovering from seconds before an attack or returning a deleted file from five minutes ago are typical backup requests that are difficult to do with snapshots or agents.

But all is not lost! We can port the ideas of traditional backup forward into a modern data protection strategy to lead us to IT resilience.

Modernizing the Idea of Backups

The concepts of RTO, RPO, and 3-2-1 are not outdated; in fact, they're more relevant than ever. Zerto replaces a traditional backup approach and protects vSphere and Hyper-V workloads by replicating them to alternate locations, whether to a second on-premises data center or to the cloud. Journaling delivers low RPOs, sometimes on the order of seconds, with a low RTO, measured in minutes or less. Compare that with traditional tape-based backups that are moved offsite for long-term retention, rarely tested to ensure recoverability, and that carry an RTO on the order of hours or days, and you can see why Zerto is a compelling alternative.

As always-on business becomes an expectation from your customers or your shareholders, periodic backups aren't going to cut it anymore; the RPO of periodic backups at scale is inherently limited. Zerto is uniquely positioned to change the game in this arena. The Continuous Data Protection engine that has always been the foundation of the Zerto solution is being enhanced to include journal-based backup

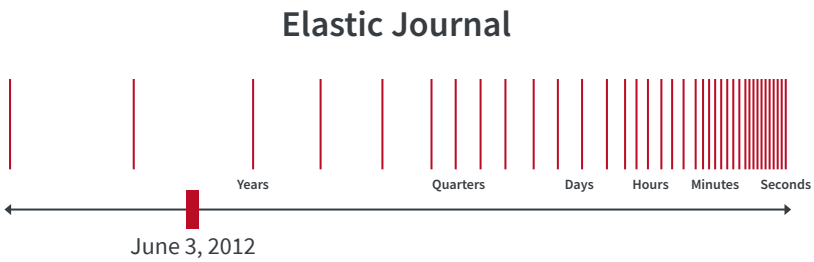


Figure 8: Journaling provides extremely low Recovery Point Objectives and Recovery Time Objectives, allowing businesses to get back on their feet faster.

and recovery. Now, instead of recovering from last night's backup, you can recover to the journal checkpoint from one minute before the incident you're responding to.

Further, to allow you to craft a backup strategy that meets your unique business requirements, future iterations of the IT Resilience Platform will feature the Elastic Journal construct to allow you to meet both short-term retention requirements, such as for ransomware recovery, and long-term retention requirements, such as for regulatory compliance. An intelligent index and search service makes this entire elastic journal searchable, so you can find exactly what you're looking for in no time.

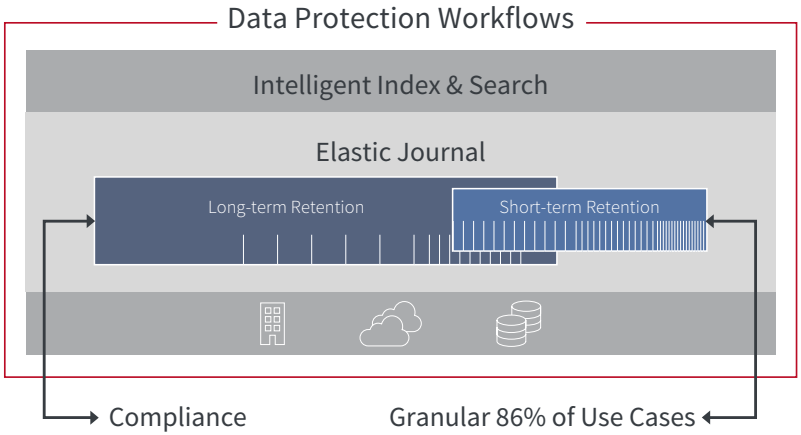


Figure 9: The Elastic Journal allows you to meet both short- and long-term retention requirements.

As an improvement on your old backup strategy, Zerto provides the following:

- Long-term retention and file-level recovery to meet data retention policies and compliance
- Application Consistency Grouping
- Automated testing to ensure recoverability with no impact on production
- Reports for full compliance, consistent with regulations like HIPAA, PCI, and GDPR

Now that you understand how to promote IT resilience by leveraging Zerto technology to protect yourself from outages (both planned and unplanned), ransomware attacks, and inadequate backups, it's time to shift the conversation toward mobility.

When everything is running fine, you may want to move workloads and data around as you upgrade, refresh, consolidate, and integrate data center infrastructure. There's an IT resilience story there, too, and you'll learn all about it in Chapter 3.

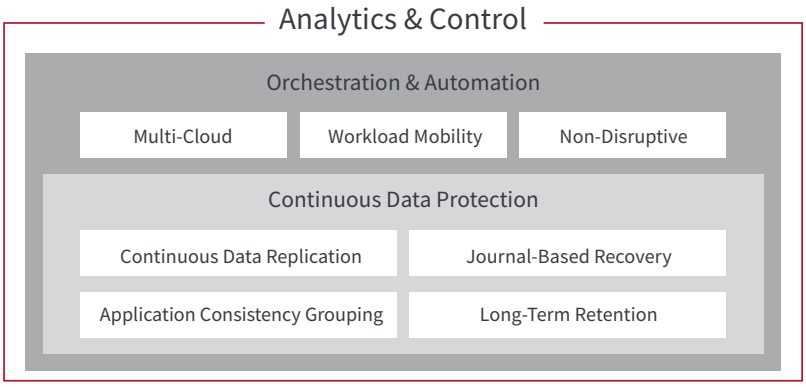


Figure 10: A complete data protection strategy encompasses both short- and long-term retention, and has orchestration and analytics wrappers.

CHAPTER 3

Enable Workload Mobility

The rate of change in data centers today is not only higher than ever, but it's also continually increasing. And with the hastening pace of technology improvements, organizations riding the digital transformation wave look to take advantage of that new technology as quickly as possible.

Moving out old technology in favor of new has historically been a painful ordeal. Most IT professionals have some battle scars from major migration projects, data center refreshes, and the like. As much as that was the accepted norm in the past, it's not going to be acceptable moving into the future. IT organizations will need the capability to shift data and workloads confidently, seamlessly, and with lightning speed.

That may mean moving workloads from one set of servers to another, from the Test environment to QA, from one data center to another, or from AWS to Azure to on-premises. But wherever workloads are moving, business leaders will expect it to be transparent and fast. And all the while, valuable business data should be protected from any risk. Easy, right?

Infrastructure Modernization

Traditionally, system upgrades were done via the "lift-and-shift" method, in which applications were logically lifted from the old servers and instantiated on newer hardware. This process was both labor- and time-intensive, and fraught with pitfalls. **Figure 11**

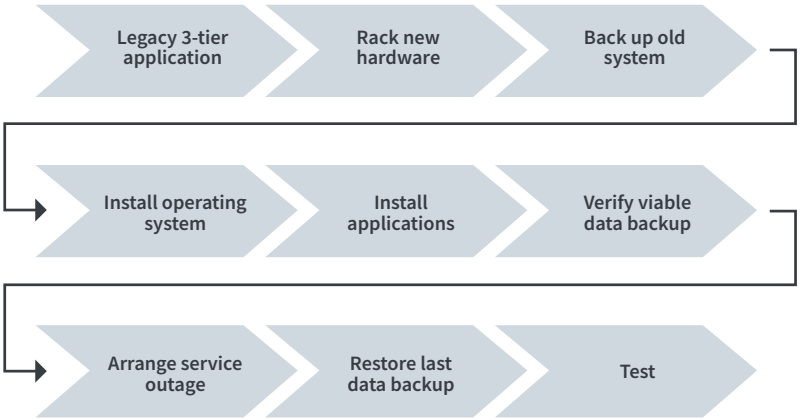


Figure 11: A typical "lift-and-shift" upgrade.

graphically shows a lift-and-shift process for a typical three-tier application in a data center undergoing a hardware upgrade.

The problem with approaching refreshes this way is that a significant amount of downtime is usually required, and there are so many unknowns that it's easy to fail at the first attempt of a forklift upgrade like this. With new workloads to support, like artificial intelligence processing, big data analytics tools, and ultra-critical virtual shopping carts at major online retailers, many businesses are moving from storage arrays based on spinning disks to more modern arrays that leverage solid-state media which outperform their old array(s) many times over. They need a way to swing these workloads over without interrupting their consumers.

Similarly, many business leaders are finding that the budgeting and deployment models for new data center architectures, such as hyperconverged infrastructure, are very attractive and they'd like to move. Of course, getting workloads onto the platform poses the same sort of challenge that an array upgrade does.

THE 101

What is Hyperconvergence?

Hyperconvergence is a way of constructing private data centers that seeks to emulate public cloud consumption in terms of its operational simplicity, economic model, and scaling granularity. And it provides all of this without sacrificing the performance, reliability, and workload availability that businesses today expect.

The most exciting benefits for adopters of hyperconverged infrastructure include:

- **Focus on the workload:** For too long, infrastructure policy and management have focused on the wrong constructs. Managing LUNs and hosts and clusters is old school. In the post-cloud era, the workload should be the focus.
- **Data efficiency:** The nature of hyperconverged infrastructure lends itself well to a high degree of data reduction by way of global deduplication and compression, which leads to more approachable requirements for storage capacity, network bandwidth, and IOPS.
- **Elasticity:** The beauty of the cloud is that if you need to scale out or in, you just click a few times and it's done. Hyperconvergence focuses heavily on making scaling easy, and in bite-sized units; this model stands in stark contrast to the 3- or 5-year bulk purchasing model of traditional IT infrastructure.

Modernization Without Fear

Zerto technology is agnostic with regard to hypervisor, storage, and even cloud. It also has the built-in ability to fail over, fail back and regression test prior to cut over. The continuous protection features of the platform make IT refreshes easy, because you can have peace of mind about all kinds of workload movements.

Moving Workloads From Old Hardware to New Hardware

Whether it's from one storage system to another storage system or from an old server cluster to a new server cluster, this sort of migration happens regularly within organizations. Technologies like vMotion and Storage vMotion can simplify this transition, but it's not always as straightforward as it should be. For example, some migration scenarios aren't supported—such as where latency between source and target is too high.

With Zerto, this sort of migration can be done in more situations and with more assurance that everything will go smoothly. The continuous data protection that comes with the platform protects your workloads throughout the whole process, and the orchestrated failover and failback operations ensure that your workloads don't get stuck on the wrong side of a bad migration.

Moving From One Hypervisor to Another

As great as inter-node migration technologies are, they're suddenly useless when you're attempting to change platforms. Some organizations move from one hypervisor to another to decrease on-premises infrastructure spending as they move more of their business to the cloud. Other organizations adopt more robust hypervisor technology as their business grows. In either case, the movement of workloads from one to another (from VMware to Hyper-V, as an example) is daunting if you're going it alone.

There are various third-party tools to help, but the beauty of performing the migration with Zerto is once again that the applications are protected throughout the whole process, and you're assured the ability to cut over seamlessly and fail back just as quickly if anything should go wrong.

Moving Workloads From On-Premises to the Cloud

With each passing day, more businesses are recognizing the additional flexibility and availability that comes with leveraging a public cloud. The truth is that running infrastructure at scale is what a cloud provider does best, and they're *probably* better at it than most enterprises; after all, data center operations isn't the primary line of business for most companies. They sell groceries, manufacture canoes, broker insurance, weave baskets, cobble shoes, and design furniture; data center infrastructure is simply a means to deliver their goods and services to their customers. Wise business leaders recognize what they should and should not focus on, and in some cases, the choice to focus less on operating a data center and more on the core business is what leads an organization to begin a migration to the cloud.

Cloud migrations come with the same core challenges as moving from one hypervisor to another, and then they throw in a few more challenges as a bonus. For example, you're usually dealing with network throughput constraints as you perform the migration. Replicating workloads from on-premises environments to your cloud of choice with Zerto not only grants you some comfort with the process, but also makes the process simpler and faster. Replication can take place in the background, and once workloads are nearly synchronized, and RPO is very low, a cutover can take place. If all goes well, the workloads stay in their new home; if anything is amiss, a failback can be initiated, and all is well again within moments. Zerto supports the movement of workloads to Microsoft Azure, IBM Cloud, and Amazon Web Services, as well as a host of other Zerto-enabled cloud service providers.

Transitioning From SAN/NAS Storage to HCI Seamlessly

When hyperconverged infrastructure is first deployed into a data center, it often creates its own little infrastructure silo - a situation that corrects itself over time. As hyperconvergence becomes the norm in the data center, growing and changing is relatively easy. But getting started can involve heavy lifting.

This transition to a newer type of data center architecture is another use case in which Zerto excels. Just like the others, this job can be done safely and easily, since failover is automated and failback only takes a few clicks to initiate. Automated failover and failback is achieved with Virtual Protection Groups, which logically collect applications and their dependencies and control the failover/failback processes in an application-aware fashion. Application-consistent cutovers mean that application downtime for performing these types of infrastructure modernizations is near zero, or in some cases, *is* zero.

Refreshes aren't the only reason that workloads need to move, however. Next up are some non-upgrade reasons why you may be interested in Zerto.

Migrations and Consolidations

People have bought and sold companies for as long as commerce has been practiced. These transactions always lead to change, which frequently causes conflict. Two companies merging is one of the most challenging environments for change. Putting aside the human elements, which can be very delicate, the technical hurdles abound. How do you merge your user directories? How do you consolidate your data? How will you eliminate duplication of resources like HR systems, CRM databases, and service desk tools?

BRIGHT IDEA

Hardware Refresh, Minus the Impact to Production

Kroll Ontrack is the data recovery and data destruction business of KL-Discovery, a global provider of eDiscovery, information governance and data recovery solutions. During Ontrack's migration from a traditional data center architecture to one based on hyperconverged infrastructure from Nutanix, the migration team looked for ways to make the migration simpler and less prone to error.

Moving data from the old system onto the new system was a challenging part of the relocation project. Kroll Ontrack saw an opportunity to use Zerto to migrate all of the applications on the old kit to the new Nutanix kit. This approach allowed the VMs to run in production while data was replicated, with zero impact in the background.

Original projections for the migration project came in at 240 days, which meant not meeting target deadline. As a result, Kroll Ontrack architects and engineers from both Nutanix Services and Zerto put their heads together to devise a solution that resulted in lowering the total project time to just 40 days, and the migration was completed without any impact to production.

Since most organizations maintain some level of headroom in their infrastructure to allow for peak workloads and growth without worrying about hitting the ceiling, one of the first problems you may encounter during a merger or acquisition is that there's an excess of headroom—you've got data center resources coming out your ears! Not only that, but it's extremely likely that some data center systems are duplicated across environments.

And so begins the process of collapsing and decommissioning unneeded systems—and even entire data centers—into fewer numbers of higher-utilized ones. There are clear financial benefits to be gained here, even beyond that of increased utilization. By consolidating data center resources, you're also likely to see a reduction in operating costs, hardware maintenance, support contracts, and software licensing.

That said, consolidation—especially in the M&A case where two previously independent infrastructures are becoming one—is a tricky undertaking. Because there's often a very low degree of homogeneity between environments, extra care is required when marrying the two infrastructure domains.

Thoroughly Test Before Committing

Zerto is perfectly positioned to provide real benefits to any data center consolidation project. Zerto offers the ability to seed a new target with source data before enabling continual protection with the CDP engine and journaling processes. This makes the consolidation process easier, as you can stage the movement in the background rather than imposing a major outage on the business. There's no requirement to wait for full synchronization of data before testing can begin, as the target machine can be mounted at any point in time, isolated away from production systems. Further, the most fragile systems can be instantiated in a sandbox location and fully vetted before the production consolidation is performed, significantly reducing the risk of a production outage as a result of the migration.

Multi-Hypervisor Support

If an organization that is staffed, trained and tooled to support vSphere acquires a company which staffed, trained, and tooled on Hyper-V, there's a good chance that over time something will have to change. As consolidation efforts commence, part of the plan

will be migrating the current Hyper-V workloads over to a vSphere environment or vice-versa.

The broad hypervisor support of Zerto enables seamless cross-hypervisor migrations, which is handy when attempting to merge heterogeneous environments. And since there's generally some uncertainty around moving VMs from one hypervisor to another, you can sleep at night knowing that failing back a problem workload is just a click away.

IP Addressing Is No Problem

Resource consolidation projects aren't always clean affairs. When consolidating resources, you're likely to find disparate hardware and software; additionally, in most cases, the networking configurations won't neatly align with one another. As a result, your consolidation efforts will include applying new IP addresses to production workloads.

Application consistency with Zerto is provided via Virtual Protection Groups (VPG) that treat entire application stacks as a single entity for replication purposes, so you can ensure that you move a full application at once as you consolidate. Without such capability, you'd be left moving chunks of applications and manually synchronizing them. Moreover, VPGs allow administrators to specify characteristics such as boot order and to define the new IP address space for applications. So, during a migration for the purpose of consolidation, the VPG settings can ensure applications are brought up on the destination side with the appropriate network settings already configured.

Couple these operational advantages with near-zero RPO and RTO, and the benefits of Zerto for consolidation are obvious. Data center cutover is almost seamless. And if any issues crop up, failback is just as quick.

Testing and DevOps

The world has started to move on from legacy waterfall software development cycles, with big-budget release cycles and multiple new features per version. We've entered the world of agile development, with shorter release cycles and fewer new or improved features at one time. Modern software development methodologies have effectively merged application and operating system patch management with continual development practices in the sense that each new release feels more like a simple patch.

This has completely changed the rules for testing. The release cycles are now so short that the old style of QA, with its two-to-three-month testing cycle for each annual release, can't hope to keep up.

The new model has operations, testing, security, and developers knit together in tight teams; development cycles are contracted, and testing and deployment are built in at the source and fully automated. This revolution has been dubbed DevOps.

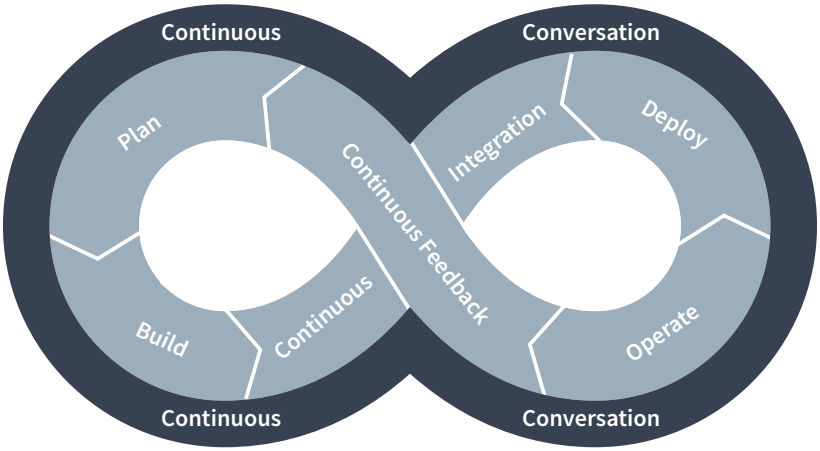


Figure 12: The DevOps circle of life.

Virtualization has helped this quickly-emerging development and operations paradigm by providing flexibility via snapshots for a point-in-time failback. But limitations in the technology keep it from being as flexible as needed. Further, significant work is still necessary to keep the environment up to date, as once a snapshot is deleted the data is gone.

New tools like Docker, Kubernetes, Ansible, and Puppet help the operations team quickly deploy the new releases. But development teams are expensive, and any amount of idle time waiting for a Development database to be refreshed from Production, or for various regression testing environments to be instantiated, is akin to throwing dollar bills in a fireplace.

Increase Code Quality

Development is really only as good as the copy of production data that developers can work with. If all their development uses synthetic test data meant to simulate production, results can be hit or miss and code defects can crop up more often than you'd like. The challenge has always been getting a current copy of production data in a reasonable amount of time.

Zerto makes development and testing easy. Firing up a clone of your production environment is simple, as Zerto can allow multiple copies of data in different locations in a one-to-many configuration. Your production data can be replication to DR facilities for protection, but also to Development, Test, and QA environments to ensure that developers and testers are working with the latest version of production data.

These production replicas coupled with the use of the Zerto API can allow the orchestration of test builds and deployments without any manual intervention. Being able to reliably test on good data can significantly reduce the number of code defects deployed to production.

Make Changes with Confidence

Because it's so easy to spin up real-time replicas of production workloads, developers don't have to guess at what might happen when they push a particular code update into the production instance. Instead, they can clone a copy of production that has the latest data to a sandbox area where it can't impact production users. They can then apply their changes and observe the results with zero risk. The level of confidence this practice brings is a perfect example of increasing IT resilience.

CHAPTER 4

Multi-Cloud Agility

For all the tales that analysts and bloggers regale you with about the pay-as-you-go OpEx nature of the public cloud, cloud adoption is about much more than just saving a few bucks. Sure, cloud computing advantages include spot instance pricing and consumption-based billing, which are wonderful tools, but there's much more to the story. In reality, cloud adoption is about *accelerating business growth.* It does this chiefly by:

1. Reducing time to value of new initiatives
2. Re-focusing IT on providing business value, instead of struggling to "keep the lights on"

To realize the full potential of cloud computing and leverage it in this way, some level of agility is required. The benefits of cloud begin to disappear if you become stuck, locked in, and unable to operate in a cloud-y manner.

Migrating applications can be hard, even when you're in the safe, comfortable, familiar confines on your own data center environment. Migrating applications to, from, and between clouds takes workload migration challenges to a whole different level. Moving into the next decade, many businesses will choose to adopt not just one, but many, clouds, both public and private, and for a variety of reasons:

- To increase flexibility
- To optimize for performance and cost
- To hedge against lock-in

Note, however, that the more clouds you utilize, the trickier things get. Each cloud you adopt features a unique administrative interface, unique service types, and unique performance and availability characteristics. This makes managing multiple clouds particularly challenging, and it can quickly become overwhelming. In this chapter, you're going to learn how to enhance your multi-cloud agility to extract the full potential of cloud computing as a practice. You'll also learn how Zerto can help you accomplish your cloud goals, however nebulous they may be, without losing your mind!

Cloud Integration and Migration

As an IT service delivery method, cloud computing offers flexibility beyond anything previously available in IT. The ability to dynamically expand and contract resources, with always-on access, has revolutionized the industry: not only applications and data, but business practices, as it allows a utility-like consumption-based cost model for resources from public could providers such as AWS and Microsoft Azure.

But there's a catch: *getting there*.

It's hard to move applications from a traditional, on-premises environment into a public cloud. The challenges with this sort of migration abound, but two major challenges are almost universally faced by companies who undertake this transformation:

1. Moving applications with a high rate of change is taxing to the network, and in some cases, it's impossible to replicate data fast enough to get the destination in lockstep with the source.
2. Moving multi-part applications with complex dependencies is hard to understand and even harder to execute.

Moving Workloads

When planning a data migration, the need to determine how much data is being moved and how long the transfer will take over the existing Internet connection is one of the first things you need to come to grips with.

And now, it's time for a bit of math, so please hang in there!

To calculate the time required to move a given set of data, this formula can be used[9]:

$$\text{Number of Days To Complete Transfer} = \frac{\text{[Total Bytes to Transfer]}}{\text{[Megabits per second of connection]} * 125 * 1000 * \text{[Network Utilization]} * 60 \text{ seconds} * 60 \text{ minutes} * 24 \text{ hours})}$$

Remember that any bandwidth used for data migration won't be available for an organization's typical Internet traffic, potentially putting at risk other key services if proper QoS policies and traffic shaping rules aren't applied. In addition, your organization may be concerned with moving sensitive business information from the internal network to a third-party cloud provider using the very public Internet as the transport mechanism.

Given physical access to the destination site, both of these problems would often be mitigated by a seeding process where the data is physically transported to the destination. There is a funny saying that "Despite incredible advancements in networking technology, the interstate highway system still has the highest bandwidth and throughput around—if you can stomach the latency."

Some cloud providers do provide a way to transfer data in bulk. AWS Import/Export Snowball is one example. However, in some cases this isn't an option; unlike a data center migration (including a co-location move) in which you own the hardware and have access to the rack, physical access to the cloud provider's environment to seed

[9] https://d1.awsstatic.com/whitepapers/Storage/An_Overview_of_AWS_Cloud_Data_Migration_Services.pdf

your environment may be curtailed. This will obviously increase the time to value, as all data will need to traverse the network.

Using the formula as an example for a 750TB dataset over a legacy European E1 line, the transfer time is significant:

$$\frac{750 * 1024\text{^}4}{2.048\ (\textit{line speed in Mbps}) * 125 * 1000 * .80(80\%\ \textit{utilization}) * 60 * 60 * 24}$$

This dataset, utilizing this particular link, will take more than two years to copy. You can see why some organizations have no choice but to move data with a Snowball-like device. Let's improve the bandwidth situation and try again. Now assume that the business has a symmetrical 1Gbps link. The math looks like this:

$$\frac{750 * 1024\text{^}4}{1{,}024\ (\textit{line speed in Mbps}) * 125 * 1000 * .80(80\%\ \textit{utilization}) * 60 * 60 * 24}$$

Now the migration takes approximately 93 days. Given that a wholesale migration of 750 TB is not a likely starting point for most organizations, it's generally a little easier to get started if you have some decent connectivity and <500 TB of data to move at the outset.

Cloud vendors do provide many tools to aid in this process. It's in their best interest to make it as easy as possible to move data to their platform. However, when dealing with multiple clouds and a diverse data set, it's highly desirable to have a tool that unifies the experience. That's where Zerto comes in.

Because migrating data takes time, and migrating datasets which have a high rate of change is even harder, Zerto is uniquely positioned to help you through this process. The near-zero RPO means that, given the proper bandwidth, even datasets which have a huge daily change rate can have replicas synced to within minutes of the primary workload. When it comes time to swing over to the destination, you instruct Zerto to fail over the VPG in question with grace, and you

BRIGHT IDEA

Cloud Migration Made Easy

Maritz is a sales and marketing services company that designs and operates employee recognition and reward programs, sales channel incentive programs, and customer loyalty programs. Maritz helps businesses achieve their full potential through inspiring and motivating employees, channel partners and customers.

Maritz looked to enhance their DR capabilities while reducing long-term capital expenses by leveraging the public cloud. Maritz leveraged AWS' Elastic Compute Cloud (EC2) to significantly reduce the time required to obtain and boot new server instances and only pay for capacity they actually use. By leveraging the Zerto IT Resilience Platform™, they were able to seamlessly migrate complete workloads to AWS in an automated manner with minimal configurations to protect their critical environment.

"In addition to the cost savings, the resilience Zerto and AWS provide for DR gave us the confidence to migrate production workloads to AWS."

- Andy Wolfe, Technical Architect at Martiz

smile and watch without fear as it happens. Within moments, your high-change-rate workload is running in a new cloud, and you're headed home for the day with a pat on the back from your team.

Hybrid Cloud, Multi-Cloud

In today's enterprise, some corporations are unwilling to go all-in on a single cloud provider, no matter how functional their offerings. The risks are too great, and the rewards for your loyalty too few. Rather,

most businesses serious about IT resilience adopt either a hybrid cloud or a multi-cloud strategy.

Leveraging multiple clouds at once *sounds* great. But in practice, managing multi-cloud environments can be troublesome due to issues like VM format differences.

The hybrid cloud model uses public cloud for all of its good parts, but doesn't rely solely on it. The model marries both public and private clouds and uses whichever destination makes the most sense for the application workload and the consumers.

Hybrid cloud environments are often slightly easier to manage from the perspective of VM deployment, as corporations tend to select a provider consistent with their current virtualization model—for example, OVH or IBM Cloud if the company uses vSphere, and Azure if their environment is based on Hyper-V.

Multi-cloud deployments are trickier, though, because workload formats are invariably misaligned. For example, because OVH and IBM Cloud are VMware-based clouds they utilize VMDK and VMX as the disk and machine formats of choice. AWS, Oracle Cloud, and Azure use other formats. There are methods of deploying identical VMs in different cloud providers and ways to transport VMs (OVF and OVA) so that when they're imported, they'll be deployed in the correct format for the chosen cloud. The problem with these tools is that they're time-consuming; and more importantly, they're not dynamic.

Multi-Cloud Drivers

Regardless of these challenges, organizations continue to press on into this increasingly multi-cloud world. There are a number of reasons, some of which are described here, that make the headache of figuring out how to do multi-cloud necessary.

COMPLIANCE REQUIREMENTS

Compliance is one possible shortcoming of a single cloud platform. For data sovereignty reasons, it could be the case that there aren't enough viable availability zones or regions to guarantee compliance with the 3-2-1 practice for data protection. The requirement that data not leave a nation is further complicated by the fact that even if a cloud provider has multiple regions in a locale, there's no guarantee that those locations will support all the required services.

Organizations outside the U.S. face special challenges, as very few providers have the level of investment in regions per country or geo-location to provide true resilience. For example, the U.K., with the world's sixth-largest economy, upon leaving the European Union will not have a single public cloud provider that can provide three separate regions for their data within the country. At the time of this writing:

- AWS has a single point of presence
- Azure has two
- Oracle Cloud has two

AVAILABILITY CONCERNS

Another reason for casting a wider cloud net may be a lack of trust in a single provider to guarantee the safety of the corporate crown jewels. Several high-profile outages of major cloud providers have caused real financial problems for their customers. As great as public cloud providers are at their craft, and as much as they design their infrastructure to eliminate downtime, it's bound to happen. Having a second cloud to depend on when another fails can take your business to the next level of IT resilience.

In a related vein, it's becoming more and more common for companies to leverage Disaster Recovery-as-a-Service (DRaaS) offerings. Zerto has DRaaS partnerships with 350-plus cloud providers; it makes sense to place your disaster recovery data in a different cloud than the one your primary workloads run in.

DEEP DIVE

What Is Disaster-Recovery-as-a-Service?

You've likely heard of the "as-a-Service" movement, which is a different way of delivering products. The most well-known type is Software-as-a-Service, and the idea is that rather than deploy software on your own infrastructure, you consume software via the Internet that is managed, upgraded, and secured by the software provider. Here are some of the more popular examples of software-as-a-service:

- Office 365
- Salesforce
- Workday

Disaster-Recovery-as-a-Service (DRaaS) applies this model to protecting your business from outages. A DRaaS provider hosts infrastructure resources and disaster recovery software to which your business will replicate data and workloads.

In the event of a disaster at your primary site, you can fail over to your DR site with the help of your DRaaS provider. The beauty of DRaaS vs. building your own DR site in a data center you manage is that the DRaaS provider does all the heavy lifting: they operate the data center facility, keep the hardware up to date, and troubleshoot on the DR site.

When disaster strikes, it's nice to have someone on your team adept at recovering quickly. DRaaS providers can also help you proactively test recovery on a regular cadence to ensure you're always prepared.

AVOIDING LOCK-IN

Finally, fear of vendor lock-in is real and has a solid basis. One public cloud vendor may have added a feature you've needed for awhile, while another doesn't even have it on their development roadmap

yet. Changing cost models can be a concern over time, too; that's why, as your needs and workloads change, a different cloud may be a better fit in the future than the one(s) you started with.

Given these and other considerations, maximizing your IT resilience likely requires having the ability to shift your data to multiple cloud locations that are independent of one another.

Protection and Migration on a Hybrid/Multi-Cloud Platform

Because so many hybrid cloud deployments use a like-for-like hypervisor base, the task of moving workloads is much simpler. There's no requirement to do a data transformation of the underlying disks and virtual hardware during transfer. There can still be some challenges around data transfer rate, as was discussed earlier in this chapter, but it's generally an approachable task.

Multi-cloud migrations, on the other hand, aren't quite so simple. Many cloud providers claim that migration is as simple as a drag-and-drop between the two differing cloud environments. On its face, this may be the case, but if you're moving between a VMware-based cloud and a Xen- or Microsoft-based cloud environment, the reality will be far, far different; this isn't a standard vMotion or Storage vMotion event, after all. There's a data transformation process that needs to take place on the VM being migrated. This conversion process on its own poses significant risk.

As your cloud presence becomes more expansive, protecting data also becomes trickier. Which data copy is considered authoritative, for example? Where are backups stored? Can you even use the same backup technology in all your clouds? Technologies like distributed DNS and bi-directional replication are only a part of the solution.

As multi-cloud usage becomes more viable as an IT delivery method, it's obvious that inter-cloud workload mobility is an absolute necessity. The ability to move seamlessly is a must, whether it's

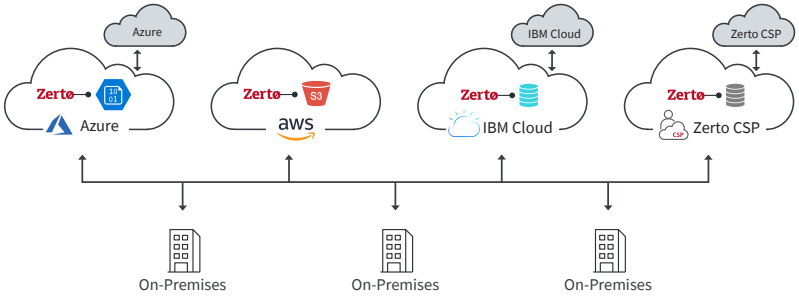


Figure 13: Hybrid cloud/multi-cloud with Zerto.

from Cloud A to Cloud B, between local and remote data centers, or a remote data center and hybrid cloud (and back again).

Zerto's replication and journaling technology greases the skids for these situations. With the ability to deliver near real-time replication using the native APIs and functions of whatever cloud platforms are involved, Zerto provides any-to-any replication between clouds. You can even configure bi-directional replication so that multiple clouds protect each other and can take over for the other at a moment's notice. Leveraging multiple clouds like this allows for the selection of the best cloud for the job at any particular point time. This has significant benefits:

- Cloud provider isn't providing the services you need, but another provider does? Move.
- Cloud provider has a major outage, costing your business dump-trucks full of cash? Fail over to a different cloud.
- Cloud provider fails to meet their SLAs? Leave.

That refreshing feeling you're having right now as you think about easily moving back and forth between clouds? That's the feeling of freedom.

Analytics Across Clouds

Growing your cloud footprint across multiple cloud providers can drastically increase your level of IT resilience and provide availability and flexibility like you've never known. But if it seems too good to be true, it probably is! There are downsides that come with a strategy that spans multiple cloud environments.

One of the biggest is that it introduces complexity. And adding a second cloud doesn't just double the complexity: the clouds influence each other, and there are intermediate factors like the network connections at a cloud exchange or the direct connection between your data center and a cloud provider. The complexity of a multi-cloud strategy is greater than the sum of its clouds.

Some of the areas most prone to additional complexity in a multi-cloud model include:

- Monitoring
- Reporting
- Troubleshooting
- Capacity planning
- Compliance

Cloud providers often provide tools to help with these sorts of activities, but of course, the tool only works with their cloud. It takes a special umbrella management tool to take a look at all your clouds and make sense out of what you're seeing. It's not a silver bullet to solve all of your multi-cloud challenges, but a number of these multi-cloud challenges are addressed by a SaaS tool offered by Zerto that integrates with any Zerto-enabled deployment.

Zerto Analytics

One of the immediate benefits you'll notice about Zerto Analytics is that because it's a SaaS tool, getting started means there's nothing to deploy; simply log in to the myZerto portal and begin configuring. Zerto Analytics provides visibility for protected environments across clouds and hypervisors, giving you the insight *across* clouds that no single provider can give you.

MONITORING

The tool provides easy-to-use, intelligent dashboards which show things like cloud-to-cloud replication status and metrics. This makes it simple to see which of your workloads are actively protected, which (if any) are lagging, and which may need some extra attention. It can also alert you if there's a problem.

Since you may be using a one-to-many or an any-to-any replication model, Zerto Analytics provides myriad views to show how any set of sites are interacting. And importantly, you'll be able to see how the

Figure 14: Get analytics from any device, simply by logging into the myZerto portal.

connections between the sites are performing and track down any issues in between them.

Finally, since Zerto Analytics hooks into Zerto Virtual Managers (which already have insight into your most important workloads), you'll be presented with a real-time visualization of data center health across all your clouds to proactively move and recover data.

REPORTING

Gathering and compiling useful data – especially for compliance auditing and SLA purposes – can be especially tricky in a multi-cloud deployment where each UI is unique and the metrics each provides are slightly different. Zerto Analytics has you covered here, too.

Acting as an aggregator for all your protected multi-cloud workloads, the SaaS tool is capable of generating reports about all the important aspects of your environment, such as network performance. The real-time dashboards can help when troubleshooting, but the reporting is pure gold for capacity planning exercises, post-mortem analysis of problems, and ongoing assurance of performance and availability.

Dynamic RPO and Journal reports can help to optimize IT efficiency and drive your businesses toward continual service improvement and, ultimately, a greater level of IT resilience.

CHAPTER 5

Level-Up Your IT Resilience

IT resilience today is more than backup/DR; those things, in fact, are just a starting point. Instead, it's a holistic approach to ensuring the systems that are becoming more and more critical to the business are always available and protected from outages, both planned and unplanned. It's the evolution beyond just redundancy and into flexibility and control.

Because today's business environments are much more reliant on IT systems to deliver value than ever before, companies are very risk-averse regarding change because of the potential loss of data or outages caused by failed updates and upgrades, which can lead to financial and reputational damage. Companies are often more afraid of a planned outage turning into a service disruption or a ransomware attack than they are about natural disasters like fires, floods or equipment failures.

Zerto is in the middle of this sea change to IT resilience. With their focus on the three pillars of protection—continuous availability, workload mobility and multi-cloud agility—Zerto is a bulwark against failure. Bi-directional replication and native cross-cloud replication promise to provide multiple, consistent data islands which are unlikely to fail simultaneously.

In addition, the ability to migrate large workloads and complex application stacks as part of a protection group (VPG) is important in aiding companies on their journey to the cloud. Zerto's ability to reverse the replication flow at the flick of a switch makes workload mobility a reality, and something more than just a one-way street.

Zerto's approach to journaling is similarly revolutionary, compressing RPO and RTO down to seconds and minutes.

Ultimately, of course, Zerto can't stop any organization from being attacked. The bad guys are out there, and they're not going away. They also can't stop natural disasters from happening. But ZERTO *can* protect an organization from the outages caused by these types of problems.

Moreover, as businesses embark on a digital transformation journey, there is assuredly planned downtime in their future. Zerto can help minimize the risk and reduce the duration of that downtime.

For all these reasons, Zerto represents a new opportunity for companies: the ability to upgrade your infrastructure and protect your bottom line without leaving yourself exposed and vulnerable to the brand damage, financial loss, and morale depletion that major outages can cause. Businesses don't need to suffer through those things any longer!

About Zerto

Zerto helps customers accelerate IT transformation by eliminating the risk and complexity of modernization and cloud adoption. By replacing multiple legacy solutions with a single IT Resilience Platform, Zerto is changing the way disaster recovery, data protection and cloud are managed. With enterprise scale, Zerto's software platform delivers continuous availability for an always-on customer experience, and simultaneously simplifying workload mobility to protect, recover and move applications freely across hybrid and multi-clouds.

Zerto is trusted by more than 6,000 customers globally, and is powering resilience offerings for Microsoft Azure, IBM Cloud, AWS, Sungard AS and more than 350 cloud service providers. Zerto's award-winning solution provides enterprises with continuous

data replication and recovery designed specifically for virtualized infrastructure and the cloud. Zerto Virtual Replication is the industry's first hypervisor-based replication solution for tier-1 applications, replacing traditional array-based BC/DR solutions that weren't built to deal with the virtual paradigm.

The Zerto IT Resilience platform can be installed, configured and replicating applications in less than one hour, providing simple VM-based replication and enabling RPOs of seconds and RTOs of minutes. Why not go to www.zerto.com/trial and click to download a free trial today?

IT Resilience Checklist

The following checklist provides an organized way for you to take an objective look at your organization's resilience and see where the gaps are. The considerations map directly to the chapters in this book, so if you find yourself to be particularly weak in a given area, you should give that chapter special attention.

CONSIDERATION	SUB-CONSIDERATION	DESCRIPTION
Continuous Availability	Outages and Disruptions	In organizations undergoing or that have undergone a Digital Transformation, downtime and loss of data are simply unacceptable. Moreover, even performance degradation is a form of outage. ☐ Can you provide an RTO that's measured in minutes instead of hours, if required? ☐ Do you have push-button failover capability so that you can restore service while you fix a problem?
	Ransomware Protection	The question is not whether ransomware will infect your organization, but when. Are you prepared? ☐ Do you have a short (minutes or seconds) RPO so that your workforce won't lose all the work they've done since yesterday? ☐ Can you recover an unaffected copy of only the infected subset of your data, as opposed to recovering an entire storage volume?
	Complete Data Protection	The scope is changing, not only for what must be protected but also for how long it should be protected. ☐ Are you practicing the 3-2-1 strategy for sound data protection? ☐ Is it simple for you to test failovers and restores and confirm the integrity of your data? ☐ Can you fine tune your retention policies from seconds to years and guarantee compliance with regulations that apply to your industry?

CONSIDERATION	SUB-CONSIDERATION	DESCRIPTION
Workload Mobility	Infrastructure Modernization	Technology is always changing, and performing upgrades and refreshes is a never-ending chore. But it can become much less daunting with the proper tools in place. ☐ Can you move virtual machines – seamlessly and without interruption of service – from old hardware to new? ☐ From one infrastructure architecture to another? ☐ From one hypervisor type to another? ☐ From on-premises to the cloud?
	Migrations and Consolidations	As businesses change and grow, the IT infrastructure needs to change with it. If your company were to change significantly due to a merger or acquisition, or if your business makes a significant pivot that requires a substantially different infrastructure, are you prepared to make the change? ☐ Can you easily test your migration/consolidation to confirm functionality by performing failover and failback operations without disruption? ☐ Do you have the flexibility to choose which hypervisor, which cloud, and which management tools you use because you have sufficient workload mobility to relocate anywhere? ☐ As you swing over application stacks to a new home, will the boot dependencies be automatically considered, and will the network interfaces have correct new IP addressing, or will the reconfiguration at the destination be a highly manual effort?
	Testing and DevOps	Since agility and speed are everything to modern business, the ability to produce high-quality code (as free of defects as possible) and to deploy to production frequently is paramount to success into the future. ☐ Can your development team work with a fresh copy of masked production data at any time they want? ☐ Is automated testing and user testing streamlined so that defects are caught before they impact production? ☐ Is your development team able to push frequent but small updates to production, decreasing the mean time to resolution for failed deployments?

CONSIDERATION	SUB-CONSIDERATION	DESCRIPTION
Multi-Cloud Agility	Cloud Integration	It's hard to move applications from a traditional, on-premises environment into a public cloud. But being able to do so can unlock huge opportunities for your business. ☐ Do you have the capability to migrate applications with a high rate of change to the cloud without major downtime? ☐ Are you confident in your ability to deconstruct and move applications with complex interdependencies without breaking them?
	Hybrid- and Multi-Cloud	Leveraging multiple clouds at once sounds great. But in practice, managing multi-cloud environments can be troublesome due to issues like VM format differences. ☐ If you so choose, do you have the ability to replicate between multiple clouds without having to refactor applications and manually convert and re-IP virtual machines? ☐ Are you locked in to a particular cloud provider or are you free to leave whenever you please (because you have the tools to do so easily)?
	Analytics Across Clouds	One of the biggest downsides to a multi-cloud infrastructure model is that it introduces complexity. It takes a special umbrella management tool to take a look at all of your clouds and make sense out of what you're doing globally. ☐ Can you monitor across clouds to make sense of what's going on across your organization, rather than just within a single cloud provider? ☐ Can you provide reporting across clouds to ensure you're meeting compliance and SLA metrics universally?

CONSIDERATION	SUB-CONSIDERATION	DESCRIPTION
Operational Efficiency	Tooling	It's imperative that a fast-paced IT organization has a streamlined toolset. Using a different point solution for every problem makes administration cumbersome and decreases visibility. ☐ Do you have a single platform for backup, replication, automation and migration?