Zertø

a Hewlett Packard
Enterprise company

# Data Protection 101 eBook

# Data Protection 101 eBook

# What is Data Protection?

Data drives business, but that data is at an increasing risk. As ransomware attacks become a new norm, organizations are experiencing more data loss and downtime from edge to cloud. An additional layer of security is needed in the form of data protection.

Data protection involves safeguarding important information from corruption, compromise, or loss. The importance of data protection increases in proportion to the amount of data that organizations create and store, which is growing at unprecedented rates. Organizations also have a decreasing tolerance for downtime that makes it impossible to access important information.

# What Requires Data Protection?

**All workloads**
Physical, virtualized, and containerized, including databases, cloud instances, DBaaS, and apps

**All tiers**
Mission-critical, business-critical, back office, and test/dev

**Need 24/7 Availability**

**No Data Loss or Downtime**

# Data Protection Trends

## The State of Data Protection

Businesses of all sizes are modernizing IT to increase productivity, reduce costs, and digitally transform. Unfortunately, the changing requirements for business continuity are often overlooked, and organizations have not made the same investments to modernize data protection as they have in other areas of IT. With the uptime and availability of critical applications and services becoming increasingly important, there is a widening gap between the level of protection that current solutions are delivering, and the level of protection organizations need.

These three factors are making data protection more challenging than ever:

1. The amounts and types of stored data continue to grow. And the diversity of places where data is stored is growing too.
2. The widespread adoption of hybrid cloud is changing data protection . For many enterprises, the cloud is becoming a preferred target for DR and backup. Backup as a service (BaaS) and DR as a service (DRaaS), with easy failover to the cloud, are attractive alternatives to traditional approaches.
3. Businesses are demanding a higher level of protection. Applications and services of all types have rising expectations for **service level agreements (SLAs)**. This translates to shrinking **recovery time objectives (RTOs)** and aggressive **recovery point objectives (RPOs)**.

# Data Protection Key Metrics

### Recovery point objective (RPO)

is the last point in time to which IT systems and applications can be recovered. It indicates the amount of data that will be lost, measured in elapsed time.

### Recovery time objective (RTO)

is the time that it takes to recover data and applications. It is measured as the length of time until business operations are back to normal after an outage or interruption.

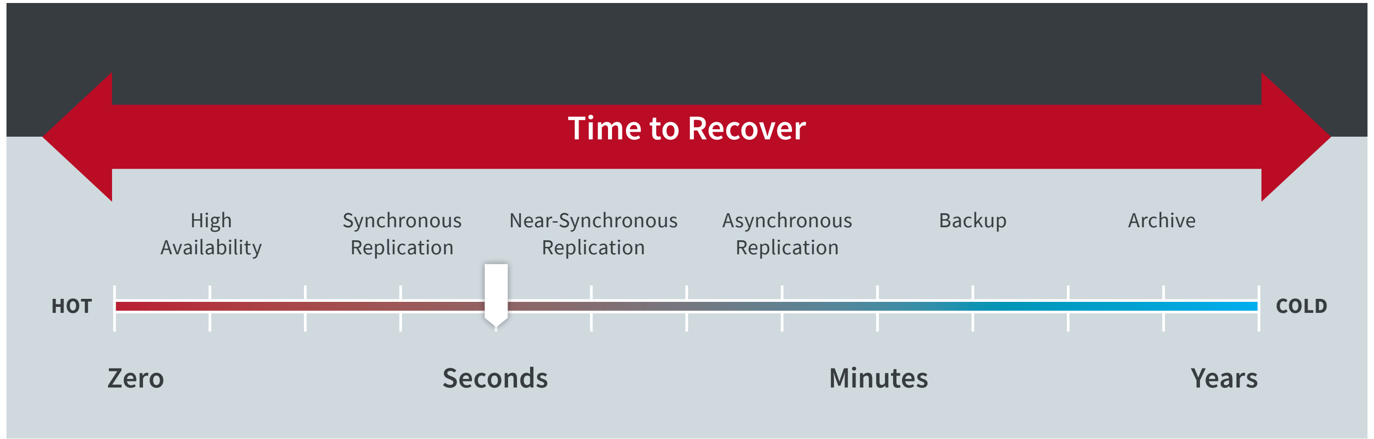
### Service level agreements (SLAs)

are requirements agreed upon by the business and the customer about the level of uptime and availability of their application.

# How to Keep Your Data Safe

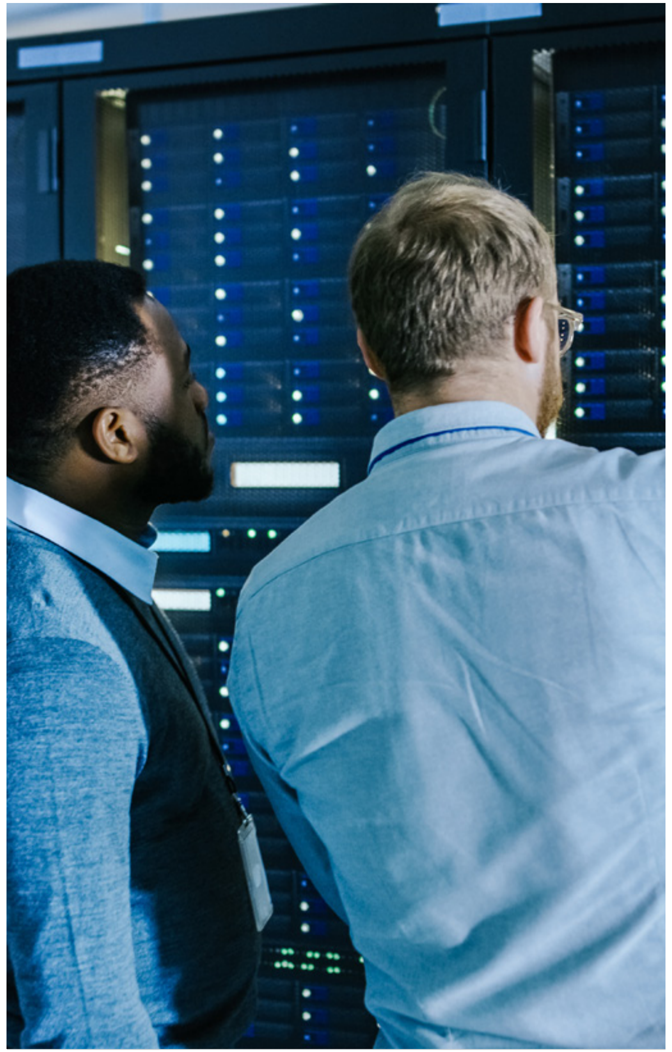
A good data protection strategy centers around your time to recover, which directly impacts RPOs, RTOs, and your overall SLAs.

## HERE ARE THE TYPES OF DATA PROTECTION THIS EBOOK WILL COVER:



# High Availability

High availability (HA) references an application's ability to continue serving clients who are requesting access to it. These services are hosted on nodes, which could be VMs, cloud instances, containers, or a combination of these types of compute resources.



## Two types of HA architectures:

### 1. Active-Passive

An application running in an active-passive cluster consists of two or more nodes where the first node is in an active state—hosting and servicing connection to the application for clients—and the others are in a passive or "standby" state. The standby servers act as a ready-to-go copy of the application environment prepared for failover in case the primary (active) server becomes disconnected or is unable to service client requests.

Should the primary server fail, processes running the services are moved to the standby cluster. This may take some time and runs the risk of service interruption to clients.

### 2. Active-Active

If an application's downtime is intolerable, then you can deploy an active-active configuration for higher levels of availability. An active-active cluster architecture runs the same service simultaneously on two or more nodes. In doing so, an application can serve multiple clients access by load balancing. Load balancing is the distribution of workloads across all nodes in a cluster to prevent application overload. As there are more nodes running the service for every client, there will be a drastic improvement in application response times. While this architecture can ensure that the application is always online, this comes at a cost.

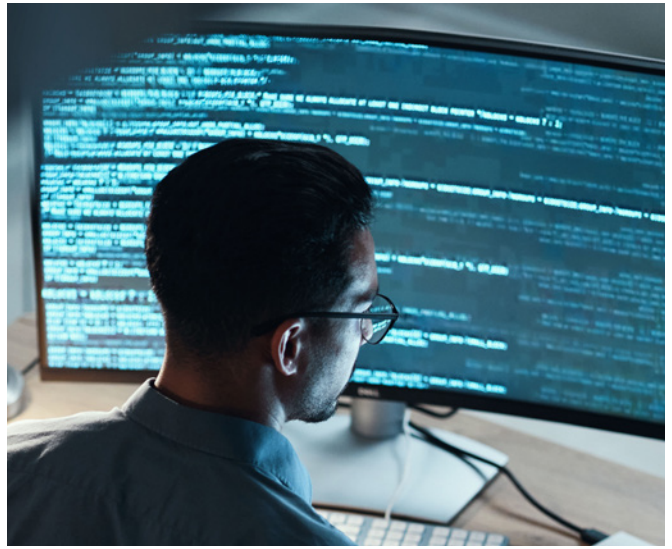
# HA Benefits vs Drawbacks

## Benefits

- **Eliminates** a single point of failure within the cluster
- **Balances** application workloads across compute resources and scales clusters to accommodate increasing demand

## Drawbacks

- **Time intensive**—multiple nodes are needed to maintain the promise of consistent application uptime
- **Performance issues**—shared storage does not scale well for applications that require dedicated resources
- **No orchestrated recovery**—an application could come back online (or, worse, scale to try to meet the demands of a DDoS attack) in a corrupt state



# Disaster Recovery

**Disaster recovery** involves preparing and mitigating potential data loss in the event of an unexpected outage

**Replication** is the underlying technology

Three types: synchronous, near-synchronous, and asynchronous

# Replication Is Key to Disaster Recovery

*Definition: The act of copying data and then moving data between a company's sites, whether those are data centers, co-location facilities, or public or private clouds.*

**RPO = 0** $$$

### Synchronous

Synchronous replication ensures all data is written to the source and target storage simultaneously, waiting for acknowledgement from both storage arrays before confirming the transaction. Synchronous replication relies on proximity between sites (to reduce latency) and enterprise-grade hardware. It runs the risk of quickly propagating ransomware, which dramatically extends recovery times.

**RPO > 1** $

### Asynchronous

Asynchronous replication uses storage snapshots to take a point-in-time copy of any changed data before sending it to the recovery site. The frequency is typically set on a schedule of hours, depending on the number and frequency of snapshots that the storage and application can withstand.

**RPO ≈ 0** $$

### Near-Synchronous

Near-synchronous replication maintains a very small time gap between the primary (source) and secondary (target) data copies. Unlike fully synchronous replication, which requires both copies to be updated simultaneously before confirming the transaction, near-synchronous replication allows a slight delay between updates.

This minimal delay, often just a few milliseconds, helps balance the trade-off between data consistency and performance. It ensures that the secondary copy remains up to date while still providing an improved responsiveness and reduced latency compared to fully synchronous methods.

# Backup

Backup is an essential part of data management. It's like insurance for your data: when you create a backup copy, you are insuring it against loss, corruption, and malware attacks. At its best, the backup copy is a perfect duplicate, readily accessible whenever it's needed. Backups rely on snapshot-based technology to ensure data is copied at a scheduled point in time.

These scheduled copies of an application and its associated data are used to recover an application from a copy generated at a pre-defined point in time. Backing up an application while leveraging a high availability architecture center will maintain copies of applications that are in a stable state, as well as maintain data for archival and compliance purposes.

**Drawback:** Backup operations are generally run during off-peak hours because copying large amounts of data takes time and has a performance impact on production environments. Whether the solution is using agents in the OS or snapshots on the virtual machines, the data is read directly from the production systems and sent across the network. This can degrade the performance of the application while it is still online, rendering it temporarily unusable.

### 3-2-1 Rule

**Generally, backup architectures are designed using the "3-2-1 rule." The 3-2-1 rule consists of three simple parts:**

- Retain three copies of data, including all production data and two backup copies.
- Store backup copies on two different types of storage, including any combination of on-premises, cloud, or offline options.
- Ensure one backup copy is stored at an off-site location.

By creating multiple copies of production data and storing them in disparate locations, organizations bolster their resilience against major disasters. Should something impact copy one, copy two will be safe in its own storage medium from which it can be safely recovered.

# How About Archiving?

Archives are backup copies of application data retained long term, generally for compliance and auditing purposes.

This data is stored on the cheapest available storage, which, depending on the data protection solution leveraged, can be anything from a file share to an object storage bucket. As days go by without this data being accessed, it "ages" within these repositories and is subsequently tiered into "colder" levels of storage. While tiering to these levels of storage is cheaper, there is a longer time and cost associated with retrieval.

Archiving solutions are best paired with data protection technology that offers a granular level of recovery.

|  | Backup | Replication |
|---|---|---|
| DEFINITION | Involves making a copy or copies of data at specific points in time. | Replication is the act of copying and then moving data between a company's sites. It is typically measured in Recover Time Objectives (RTO) and Recovery Point Objectives (RPO). |
| PURPOSE | Backup focuses on compliance and granular recovery, such as long-term archival of business records. | Replication and recovery focus on Disaster Recovery — quick and easy resumption of operations after an outage or corruption. Minimizing the Recovery Time Objective (RTO) is key. |
| USES | Backup is typically used for everything in the enterprise, from production servers to desktops. | Replication is often used for mission-critical applications that must always be up and running. |
| HOW IT WORKS | Backup typically relies on snapshots which are copies of the data set taken at a pre-determined point-in-time. | Replication can by synchronous, asynchronous or near-synchronous and may use continuous data protection to enable user to access historic data. |
| REQUIREMENTS | Backup requires a tape library (usually VTL doing disk-to-disk backup) and some place to store archived tapes. | Replication requires investment in another infrastructure in order to enable recovery and continued business operations. |
| BOTTOM LINE | Backup is a relatively inexpensive way to avoid complete data loss. Valuable for compliance. Does not ensure continuity or operations. | Replication is focused on ensuring that business applications and processes are always available, even after an outage. |

# Four Reasons Why Backup Is Not Replication

**Service Levels.** Backups typically occur once per day and at night, which means that the potential data loss could be days or more. When protecting the applications and data that matter to your business, this amount of data loss is unacceptable. Restoring from a backup, especially a tape backup, can take days; from disk it might be slightly faster—a few hours.

**Application Impact.** Backups rely on snapshot-based technology. The reason they are taken so infrequently is because this type of technology drains resources on the server. It is possible to take more frequent copies, but this comes at the expense of server resources and user productivity is significantly impacted.

**Reverse Protection.** Once applications and data have been made available at a target site, protection must be extended to include the new data that user are creating. A backup solution will not start taking backups and ship them back to the production site. Replication technologies will replicate back to the source site, ensuring the application is still protected both during and after an outage.

**Retention.** Backups are normally stored for a very long time for compliance and audit purposes. Depending on how often they occur, the recovery granularity can be hours, days, or more. Technologies that use continuous data protection (CDP) offer extremely granular recovery points, often separated by mere seconds. This gives your enterprise several points in time to recover to, just in case the last point in time is corrupted.

# What is Cyber Recovery?

**Refers to the process and set of strategies** designed to restore and resume normal operations of an organization's digital systems and data from cyberthreats such as ransomware, data breaches, or other types of cyberattacks.

**Goal:** Recover and reconstruct compromised or lost data, systems, and services completely.

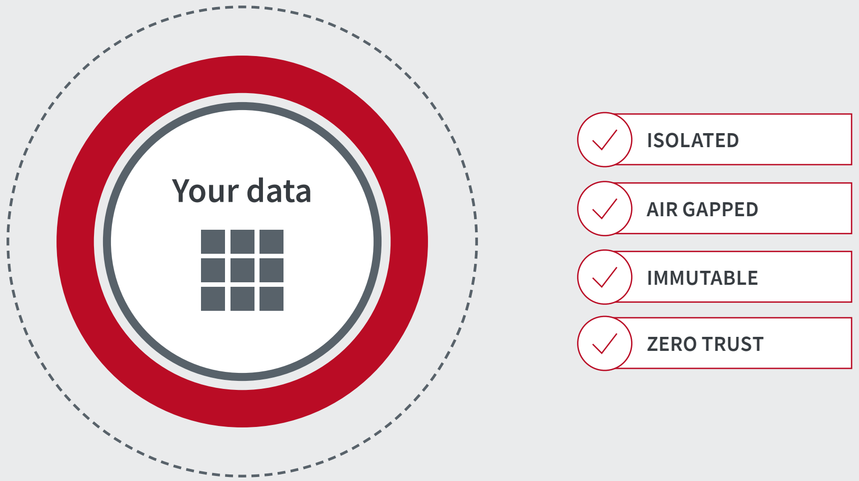
**Involves a combination of measures** usually provided by a vault, including:

- Immutable data copies
- Isolated and air-gapped networking functionalities
- Implementing zero trust technologies

# What is a Vault?

**Isolated, offline, and air gapped immutable data copies using zero trust architecture.**

Recovery option of last resort in worst case scenarios.



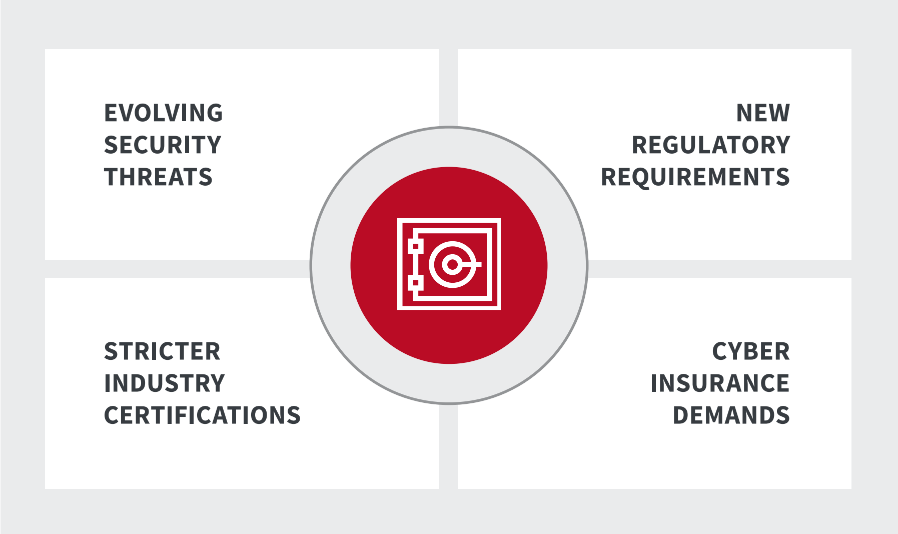
Your data

- ✓ ISOLATED
- ✓ AIR GAPPED
- ✓ IMMUTABLE
- ✓ ZERO TRUST

# Why a Vault for Cyber Recovery?

> "Isolated recovery environments (IREs) with immutable data vaults (IDVs) provide the highest level of security and recovery against insider threats, ransomware, and other forms of hacking."
>
> GARTNER

**Increasing pressure on organizations**



EVOLVING SECURITY THREATS

NEW REGULATORY REQUIREMENTS

STRICTER INDUSTRY CERTIFICATIONS

CYBER INSURANCE DEMANDS

# Three Layers of Data Protection

**Building a resilient data protection strategy means having layers of recoverability.**

| | High Availability | Replication | Backup |
|---|---|---|---|
| **RPO** | 0 | 5–15 seconds | 12–24+ hours |
| **RTO** | 0 | Minutes | Hours/Days |
| **Technology** | Synchronous | CDP | Snapshot |
| **Solution Type** | Hardware | Software | Both |
| **Main Use Cases** | Natural disaster, hardware failure | Natural disaster, hardware failure, human error, ransomware | Regulatory and compliance, long-term retention, last-resort recovery |
| **Bandwidth** | High | Low | High during backup window |
| **Workload Tier** | Mission-critical | Mission critical and business critical | All tiers |

# DR and Backup as a Service

## What is BaaS?

**Backup as a service (BaaS)** is a cloud-based solution that offers automated data backup, storage, and restoration processes for businesses. It ensures data protection, reduces the risk of data loss, and provides efficient recovery options in case of system failures, data corruption, or disasters. BaaS simplifies the management of backup tasks, enhances business continuity, and minimizes downtime, all through a subscription-based model.

## What is DRaaS?

**Disaster recovery as a service (DRaaS)** is a solution offering data and system recovery after a disruption. It involves backup, replication, and rapid recovery, reducing downtime and data loss. This service is cost-effective and aids business continuity, eliminating the need for on-site infrastructure.

## BENEFITS OF HAVING BOTH

$ $

### Control Costs

Gain greater predictability of storage costs and choose the DR strategy that is right for you.

### Diversify Data Protection

Gain confidence with target site diversification. Take advantage of the extended global network of DR sites afforded by managed service providers.

### Take DR to the Cloud

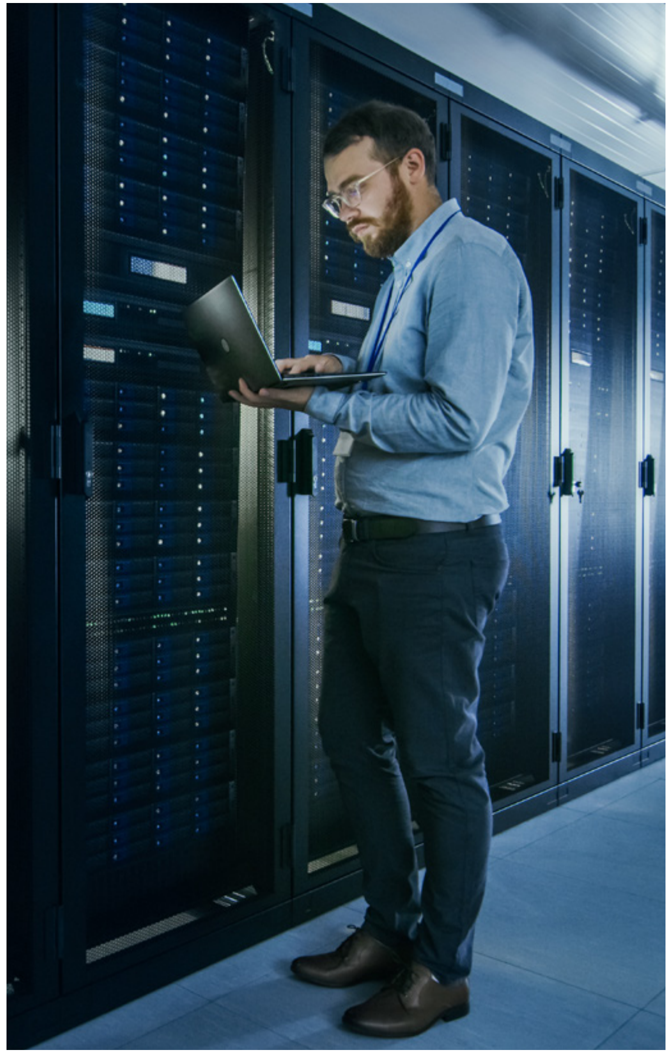
Leverage a DRaaS provider to be your guide to the cloud.

# Recap on Data Protection

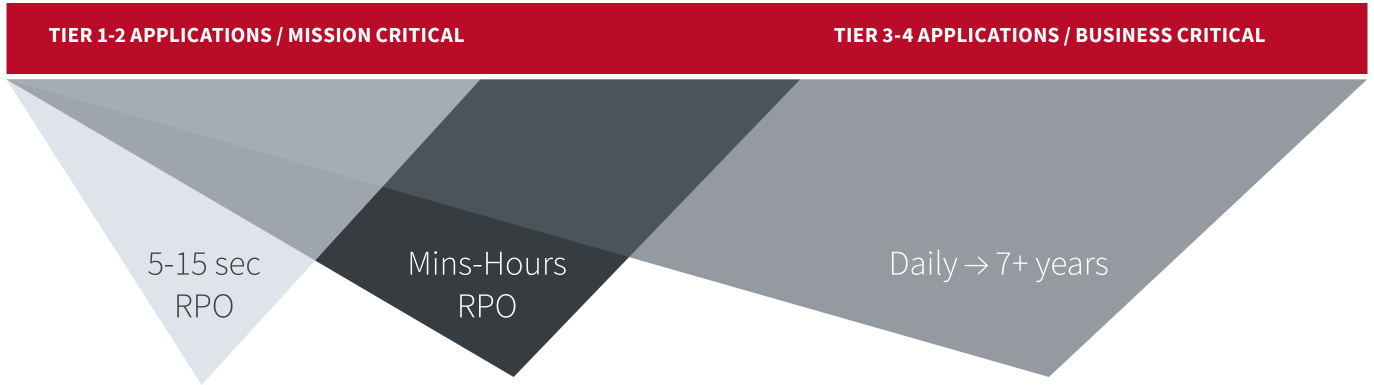
**Data Protection involves** safeguarding important information from corruption, compromise, or loss.

**A good data protection strategy** is focused on the time to recover.

**A strong data protection strategy** deploys multiple data protection technologies to enforce layers of recoverability.



# Data Protection with HPE and Zerto



TIER 1-2 APPLICATIONS / MISSION CRITICAL
TIER 3-4 APPLICATIONS / BUSINESS CRITICAL
5-15 sec RPO
Mins-Hours RPO
Daily → 7+ years

### Disaster Recovery

Zerto for VMs and Cloud

HPE GreenLake for Disaster Recovery

### Cyber Recovery

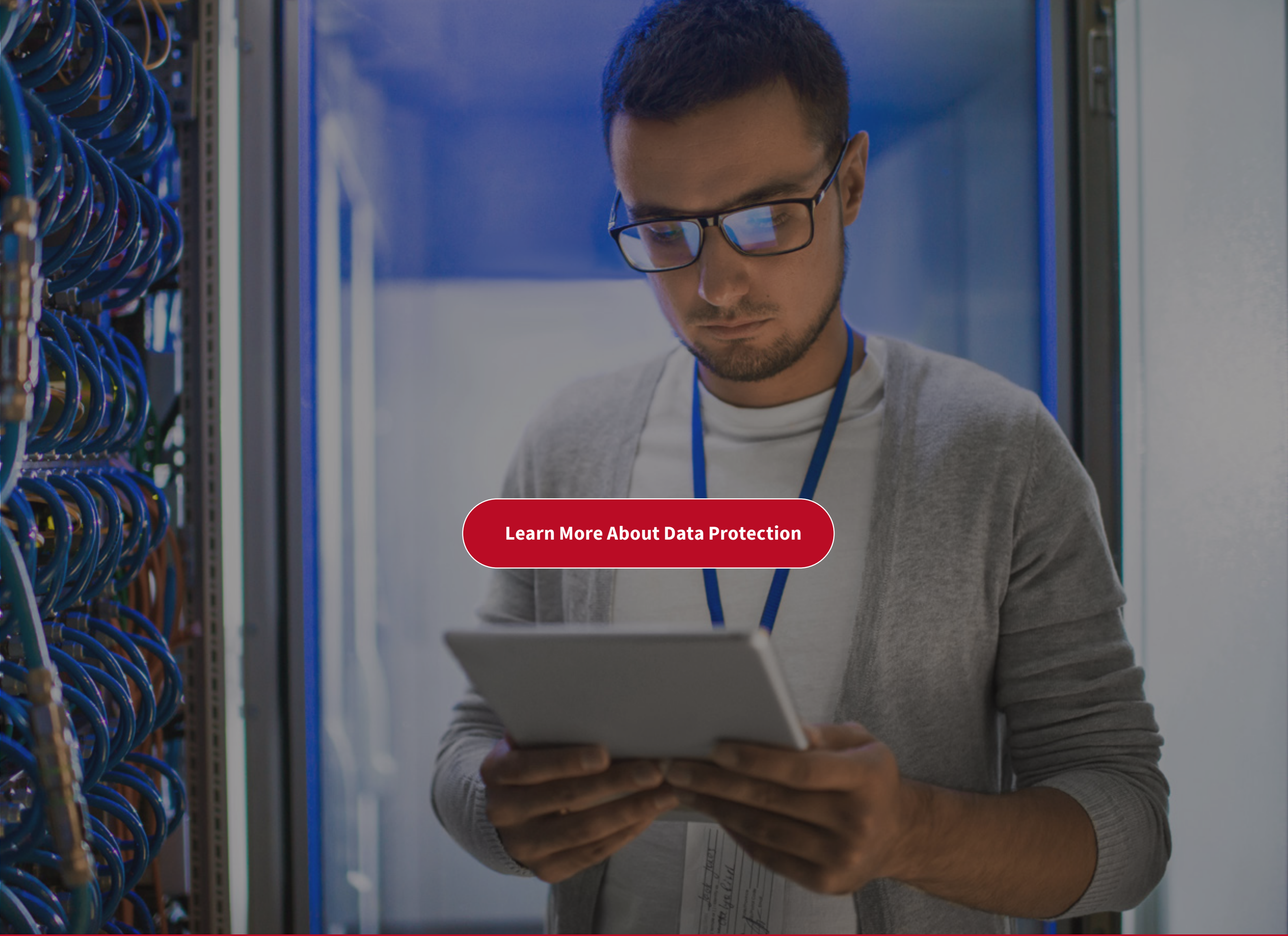
Zerto Cyber Resilience Vault

### Backup & long-term retention

Zerto Backup for SaaS, powered by Keepit

HPE GreenLake for Backup and Recovery

Unified portfolio running on best-in-class HPE storage, compute, networking



Learn More About Data Protection

## About Zerto

Zerto, a Hewlett Packard Enterprise company, empowers customers to run an always-on business by simplifying the protection, recovery, and mobility of on-premises and cloud applications. Zerto eliminates the risk and complexity of modernization and cloud adoption across private, public, and hybrid deployments. The simple, software-only solution uses continuous data protection at scale to solve for ransomware resilience, disaster recovery, and multi-cloud mobility. Zerto is trusted by over 9,500 customers globally and is powering offerings for Amazon, Google, IBM, Microsoft, and Oracle and more than 350 managed service providers. **www.zerto.com**