Whitepaper

10 Rules to Protect Workloads in a Hybrid or Public Cloud

Executive summary

A growing mix of organizations now look to external cloud providers to host either a portion, or all, of one or more of their application workloads. Knowing the best way to protect such workloads, however, remains a concern.

This whitepaper looks at some common mistakes and lessons learned about data protection in a public cloud or hybrid cloud environment. It also offers high-level guidance to better protect your data and applications, especially when they reside in someone else's cloud.

> Special note:
> This paper covers data protection of application workloads — hosted partly or in total — with an external, Infrastructure as a Service (IaaS) cloud provider. Applications hosted with Software as a Service (SaaS) or Platform as a Service (PaaS) providers are beyond the scope of this paper. Applications supplied by SaaS providers typically offer data protection services as part of their Service Level Agreement (SLA). Workloads hosted with a PaaS provider typically involve more upfront planning about data protection and Disaster Recovery (DR) within their early, application design phase.

Learning from others' cloud mistakes

Moves to cloud have garnered many success stories. But, they have also begun to highlight a few common missteps that can torpedo the savings companies hoped to gain from moving part, or all, of their workloads to the cloud. When it comes to protecting cloud-based workloads, most of these missteps often have one thing in common: They usually start with wrong assumptions or unrealistic expectations.

Organizations have come to expect that in-cloud data protection solutions provide granular restores and tiered storage, as well as the ability to deliver against SLAs and support newer technology deployments like containers.[1] Of course, this isn't always the case and may require specific expertise to enable such cloud-based capabilities in ways that empower the business and its data protection strategies.

These challenges raise questions about how well organizations plan for the potential risks of a workload move — in whole or in part — to the cloud. For many, unfortunately, the answer is "not too well."

To find success, organizations can begin by performing greater due diligence and planning surrounding workload migrations to cloud. This includes scoping the potential costs to move one or more workloads to the cloud. Such efforts also include careful planning regarding how workloads will be protected in a hybrid or public cloud setting.

10 rules to remember for data protection in the cloud

How can organizations avoid making costly mistakes regarding data protection of workloads in the public cloud? What should they know to help them be better prepared? Here are 10 general rules to keep in mind:

1. Don’t confuse data resilience with data protection

2. Remember the RTO

3. Shorter, stricter RTO/RPO goals often mean more $$$, even in the cloud

4. Data protection’s new frontier: Go beyond RTO/RPO

5. Interpreting “No SPOF” in the age of cloud

6. Backup copies: When, where, how many, what type?

7. Beware the cost of snapshots and data transport

8. Write it down, write it down: Process and paper trails

9. Become informed about emerging cloud backup trends

10. Be realistic about what you can do

Rule 1. Don't confuse data resilience with data protection

If you use an IaaS cloud provider to host some — or all — of your application workloads, don't assume that means the provider also offers adequate protection of your data or applications. IaaS providers often plan for "data resilience" not "data protection."

Data resilience means the cloud provider's data center environment has been designed to tolerate failure. For example, if a provider's server or storage system goes down, practices like data mirroring and redundant hardware might allow the same data and applications to become available on another system.

Data resilience is different from data protection, however. For certain use cases like above, data resilience might ensure the same primary data becomes available shortly after the event. What it doesn't do is protect primary data from being altered, corrupted, or accidentally deleted. This is where data protection tools come in. Instead of just having resilient copies of your data, data protection tools do more. They offer a safe, unchangeable secondary copy of your application data from a prior point in time. If your data becomes corrupted or is accidentally deleted, these tools let you restore the data from an earlier point before the corruption or deletion. Many data protection tools even help you protect and restore your data using the same compliance rules and processes that you follow in your own, in-house environment.

Talk to your cloud provider about the difference between data resilience and data protection. Ask to see what guarantees are in the SLA regarding each area. Unless you are working with a provider specializing in cloud-based backup and recovery services, you may find the IaaS provider doesn't go as far as you'd think to protect your data the way you would protect it in-house.

Rule 2. Remember the RTO

Not to be confused with the Alamo battle cry of 1836, "Remember the RTO" refers instead to important data protection fundamentals. These hold just as true for protecting workloads in the cloud as they do for protecting on-premises workloads.

Regardless of the underlying platform in use, data protection efforts for workloads should clearly define that workload's specific Recovery Time Objective (RTO) and Recovery Point Objective (RPO) goals.

In an outage or downtime scenario, this means asking:

- RTO: How quickly do I need this application workload to be back up and running?
- RPO: How much recent data can I afford to lose when I restore an application?

If the workload will be hosted in a public cloud, it also means asking the cloud provider:

- How can you support my RTO/RPO goals for this workload?
- What does your SLA promise regarding workload protection, recovery, and availability?
- How do your SLA end goals and requirements compare to or differ from our company's internal SLA requirements?

Rule 3. Shorter, stricter RTO/RPO goals often mean more $$$, even in the cloud

Those managing enterprise data protection and DR in the pre-cloud era may recall the cost discussions associated with manning second data centers and off-site DR. While cloud now offers an interesting way for some organizations to avoid the costly CapEx investment in remote DR, use of cloud to host workloads doesn't preclude the still-significant costs associated with proper backup and DR.

Mission-critical workloads with a very short RTO/RPO goal may still require significant investment in backup and DR, even though such workloads are now hosted partly or completely in a public cloud. While some workloads may manage with looser RTO/RPO requirements and less investment, others may still require more. Learn what else may be needed to ensure more mission-critical workloads remain available if they go down for some reason.

Rule 4. Data protection's new frontier: Go beyond RTO/RPO

Unfortunately, when it comes to protecting your workloads in someone else's data center, it's not just about data protection and recovery after some type of downtime event. Workload cloud migration plans should be broadened to encompass not just data protection but other risks to the workload as well. This includes potential security, governance, and compliance risks or requirements that should be addressed for that workload and its data in the cloud.

In our workshops and assessments surrounding workload platform alignment, we often encourage participants to look more closely at these types of risks as well as the wider business and technological goals for their workloads. This gives a great head start when planning an upcoming cloud migration of a specific workload.

Some questions to ask in these areas might include:

- Should my workloads or data live in only one cloud provider or multiple cloud providers?
- Should my data live in the cloud or external to the cloud? The latter "external" case means you might use the cloud for its compute features but choose to host your own data. This is not intuitive but it can sometimes be a compelling use case for certain environments.
- If my workload is in only one cloud provider, what happens if the provider has an outage? Will my business survive?
- What levels of encryption are involved or needed with data or copies in transit and at rest?
- Who has access to my primary or protected data?
- Are different workloads adequately separated and protected in different areas of the provider's cloud?
- Do different workloads have established role-based access procedures and processes with separate user logins and passwords, where needed, to prevent unwanted or accidental access to important areas?

Tales from the trenches: What not to do with logins and passwords
One client had three "shadow IT" workloads running concurrently at a large, hyperscale cloud provider. These included the client's production workloads, associated development and testing workloads, and backups of the primary workloads (stored on the provider's secondary storage tier). Users of all three workloads accessed them with the same login/password credentials. One day, an administrator's accidental error caused all three sets of workloads to be deleted.

Rule 5. Interpreting "No SPOF" in the age of cloud

Another data protection fundamental is to plan for no Single Point of Failure (SPOF). In an on-premises data center environment, that might mean backup copies stored locally on disk, others possibly stored in the cloud and, even, offline backup tape stored in a third location or in another data center.

With workloads in the cloud, to give further assurances against potential SPOF scenarios, the cloud provider may offer the option to host additional copies of your data elsewhere. This might include hosting a copy on a cheaper storage tier. Most also offer the option to house more copies in the same geographic "zone" or in one or more alternate geographic zones. One recommendation with considerable merit may even involve moving backups to a different cloud provider entirely. This could protect you if an event caused disruption across the primary cloud provider's environment. Not surprisingly, all such protection scenarios involve additional cost and should be carefully vetted to determine the best mix of cost vs. risk.

Moving workload data or backups from one provider to another is likely to cost more than moving between zones within the same cloud provider.

Word to the wise: Outages happen
In November 2020, AWS experienced a major outage in a U.S. data center due to a capacity addition. A little more than two weeks later, Google Cloud Platform (GCP) experienced a failure that impacted its authentication system.[2] Until significant changes are made to improve interoperability, the onus falls on organizations to manage this risk appropriately.

Rule 6. Backup copies: When, where, how many, what type?

The prior rule touches on the area of backup copies. How often workloads require backup and what type of backup should be carefully considered with workloads in the cloud. Database systems or other systems may need application-level backup and backup that allows them to restore from some type of sudden or rolling database corruption. For example, as an adjunct to other data protection efforts, you may opt to also perform a separate daily "snapshot" copy that does not get overwritten in case you need to "rollback" to a prior point in time.

Does the provider have these capabilities? If so, what do they cost?

Does the provider offer true, unchangeable backup copies of the original or only copies for resilience? (See Rule 1)

Special note about snapshots:

Another common data protection misstep is to assume snapshots are equal to backups. This is incorrect. With a snapshot, you create an efficient, immutable (unchangeable) copy of your data. Unlike backups, however, snapshot data usually remains on the same hard drive as the primary data. If the original hard drive(s) fail, your snapshots go with them. A true backup copy implies the copy is on a completely different system. Many organizations use snapshots to quickly recover their data, but they are still not to be confused with separate backup copies that are removed from the primary system. There are also different implementations of snapshots with different associated costs. See Rule 7 for an important cost metric associated with the use of some types of snapshots.

Rule 7. Beware the cost of snapshots and data transport

In the traditional world of backup and enterprise storage, disk-based snapshots have evolved to be a very rapid, space-efficient way to back up large blocks of application data. Many traditional snapshots consume very little space and can be considered essentially "free" beyond the infrastructure needed to support it.

However, snapshots in the world of cloud can mean something very different. Instead of traditional, space-efficient snapshots, native snapshot capabilities offered by some cloud providers may consume the same amount of storage space as the primary data. That means if your primary data is a full terabyte, a snapshot of that data will be another full terabyte. Since many providers charge for allocated space, whether you use the data or not, this can end up costing quite a lot.

Typically, efforts to move copies of your data from place to place — or from provider to provider — may also cost more than you anticipated. Often, moving data out of the cloud provider's environment costs more than moving the data into it.

Tales from the trenches: Reining in snapshot cost overruns in the cloud

An international company had moved many workloads to a hyperscale cloud provider. To help with backup, the company had also begun to use the cloud provider's native snapshot capabilities. It wasn't long before the cost of storage for these snapshots came to the company's notice. We recommended the company replace the provider's inefficient snapshot process with a more traditional snapshot method available from a third-party tool. Just this change alone resulted in an 8X savings on the company's monthly cost of cloud storage.

Rule 8. Write it down, write it down: Process and paper trails

Even though workloads are in the cloud, data recovery is still your organization's responsibility. As such, it's critical to document and periodically update the process and responsibilities surrounding data protection and DR of workloads hosted partially or fully in a provider's cloud. Such documentation includes:


Data protection processes


Success/failure reports


Results of DR testing


Specific restoration procedures — which should be performed and audited on a periodic basis

Some questions to consider here might include:

- Who manages day-to-day data protection of cloud-based workloads?
- How easily can you restore your data from the cloud? What specific restoration steps are required? When should restoration be performed? Who is authorized to perform it? If you are relying on the vendor to perform some of these steps, their obligations should be spelled out in detail in the SLA.
- In an emergency, what assurance is there that important workloads will be restored or rehosted in a reasonable amount of time? Do you have to pay more for premium or faster restoration services? (Again, check the SLA to ensure these types of details are being addressed.)
- What is the provider's track record when problems arise?

Rule 9. Become informed about emerging cloud backup trends

There are many rapidly maturing approaches and technologies in the world of backup and the cloud. Some vendors now offer capabilities to protect on-premises data by sending data protection copies to public clouds. Many vendors are also developing cloud-native data protection solutions specifically designed for how data is stored and managed in a public cloud. Still other vendors are touting the potential reuse of public cloud backup copies for development and testing efforts.

If you are considering such options, look closely at the vendor's costing models and technical capabilities. Also learn how they would work with your organization and its corporate governance rules.

> As you conduct your search, be sure to look at the capabilities of not just emerging vendors but also your traditional data protection vendor and its competitors.

These features are evolving rapidly and can sometimes be hard to follow. Many of our clients find they can avoid the time and expense of doing this research on their own by seeking expert guidance. The Insight team spends considerable time with all vendors, new and emerging, to validate and even advise the manufacturers themselves.

Rule 10. Be realistic about what you can do

The previous pages highlight a few of the data protection challenges to be aware of when workloads are hosted in public cloud environments.

Unfortunately, many organizations don't always have the staff resources to perform the levels of due diligence required for adequate protection of their cloud-based workloads. Evaluate your team's capabilities and knowledge. You may find you need some help.

If so, why not look into some of Insight's services for training, assessment, architecture, or managed services to help extend your own IT teams?

Working together with you, Insight can help you identify data protection requirements, define strategy, and architect the right data protection processes for your growing workloads — wherever they reside.

Insight services span a wide range of areas to assist customers with various levels of data protection need, including:

Monitoring and reporting services, such as cloud monitoring and managed monitoring

Data center transformation services to aid your organization's own journey to the cloud

Various managed services, such as managed backup, managed archive, managed private cloud DR, and managed private cloud IaaS

Getting help and more information

We have helped mid- and large-size organizations navigate to public and hybrid clouds and optimize resulting business value. From developing cloud strategies, assessing workloads, and choosing best-fit cloud platforms, to creating cloud-ready IT governance models, security practices, and service catalogs, we help organizations accelerate smart IT transformations.

Leverage our additional resources for further details:

- Whitepaper: "Managing the Public Cloud: Who Owns What?"
- Whitepaper: "Workload Migration: Public or Private — 9 Core Principles"
- Whitepaper: "Moving Workloads to the Public Cloud? Don't Forget About Security."
- eBook: "Key Considerations When Migrating Workloads to Public Cloud"
- eBook: "4 Best Practices for Ransomware Readiness"
- Webcast: Prevent and Recover: Mitigating the Threat of Ransomware
- Webcast: Data Loss Prevention: The Role of Data Governance, Discovery, and Classification

---

1 Bertrand, C. (April 2021). The Evolution of Data Protection Cloud Strategies. ESG.

2 Zuo, T. (2021, March 17). Commercial Cloud Outages Are a Wake-Up Call. Nextgov.com.

Meaningful solutions driving business outcomes

We help our clients modernize and secure critical platforms to transform IT. We believe data is a key driver, hybrid models are accelerators, and secure networks are well integrated. Our end-to-end services empower companies to effectively leverage technology solutions to overcome challenges, support growth and innovation, reduce risk, and transform the business.

Learn more at:
insightCDCT.com | insight.com