datalink
An Insight company

White Paper

Considerations for Data Protection in the Hybrid Cloud

Table of Contents

Executive summary

This white paper describes some potential issues and pitfalls that can occur in regards to data protection of IT workloads hosted in the cloud. It also offers high-level guidance on ways to adequately protect your data and applications, especially when they reside on someone else's cloud.

Protecting your workloads and data in the cloud

Many companies are now somewhat familiar with the common data protection methods typically required to protect their own on-premises application workloads and data in their own private cloud environments. However, new cloud choices are bringing with them new questions about how best to protect an organization's application workloads and data – especially when the workloads are accessed from one or more public cloud providers, or from a hybrid cloud mix of public and private clouds.

More than ever before, companies are looking to public cloud services and external cloud providers to augment IT data center operations. Many have begun to use Infrastructure as a Service (IaaS) providers to host their virtual workloads, while others continue to use such services as well to promote faster, more agile development and testing. Still, organizations have seen the benefit of switching internal email operations to Software as a Service (SaaS) environments like Microsoft Office 365 via Azure.

> 38% of survey respondents have started to bring applications they pushed to the cloud back in-house due to high costs, security concerns, or different service levels than expected.
> – Datalink and IDG Research survey, August 2016.

These types of hybrid cloud use cases – and their related cloud savings – abound. Unfortunately, there are also just as many examples of unforeseen issues, unexpected data loss, and unwelcome cost-overrun surprises. This has caused many IT organizations to learn things the hard way about what it takes to adequately protect their data in the cloud.

Are there practices or good rules of thumb to prevent these types of issues from happening on your watch? Just as important, do they still allow your organization to embrace the best aspects and opportunities of hybrid cloud at the same time?

The answer to both questions is, "Yes." This white paper provides high-level guidance on ways to adequately protect your data and applications – especially when they reside on someone else's cloud. Also included are some of the potential issues and pitfalls that can occur in regards to data protection of IT assets in the cloud.

New venue, same old protection issues

In this time of transition to more cloud services, many environments are likely to support different types of application workloads in their hybrid cloud. Some workloads begin the hybrid cloud journey as legacy applications, which were originally on-premises in a corporate-owned data center infrastructure. These now find themselves partially or fully hosted in a public cloud environment. Other emerging workloads might be considered more "cloud-ready." These tend to be built from the ground up with more agility so that they can run more easily in either a public cloud, a private cloud, or both. Some might take advantage of technology like containers, open source software, and greater collaboration tools. A few common hybrid or public cloud uses cases appear as examples below.

Popular SaaS offerings include:

- HR functions, such as payroll or benefits management, are increasingly accessed from external SaaS-based cloud providers.
- Corporate email services, now offered via a managed SaaS model as well, have become a welcome OpEx alternative for companies tired of spending large CapEx dollars on their mail server infrastructures and on the cost of in-house email administration.
- CRM applications have become increasingly popular as cloud vendor SaaS offerings.

"Burstable" IT infrastructure that scales up or down quickly:

- Meeting peak workload needs that may be seasonally high demand, which only need this scale for short periods of time.
- Application development and testing functions in an organization often require variable amounts of IT resources. These environments tend to be well-suited for Infrastructure as a Service (IaaS) cloud providers. Sometimes an organization's programmers or testing teams may need a lot of compute power or storage for intensive phases. Then, their need may drop precipitously. Instead of taking a CapEx hit on infrastructure that may not be fully utilized all the time, companies are finding the pay-per-use model of IaaS a welcome change for their wallets.

Applications with new needs for scalability and access:

- Analytics applications are emerging as one use case that may do well in a public cloud, especially as many analytics data points become more accessible from the public cloud. Analytics associated with emerging technologies like IoT and newer data sources have seen some success in public or hybrid cloud environments.
- Collaboration-centric applications are finding a ready home in public and hybrid clouds. This is especially the case as more applications address the mobility of today's worker and workers' desires to access applications and data from multiple devices, whether from their office computers, laptops, tablets, or cell phones. With multiple points of potential access, such applications raise new questions about how to best protect and secure data from wherever it's accessed.

Whatever type of workload or application an organization uses in a hybrid or public cloud environment, one thing is clear: The workload's requirements for data protection do not change from those of your organization's original on-premises applications. In fact, you could argue that such data protection requirements can even become more stringent for applications hosted in the cloud.

What follows are several considerations to keep in mind when attempting to protect workloads in either a public or hybrid cloud.

1. **Remember and compare RTO and RPO goals**
2. **Don't keep all your eggs in one basket**
3. **Beware the cost of cloud services**
4. **Consider unique data protection requirements per application workload**
5. **Don't forget data protection processes and paper trails**
6. **Protect cloud-based workloads beyond backup**

Consideration #1: Remember and compare RTO and RPO goals

RTO and RPO goals should be carefully defined per workload, regardless of where the workload now resides: either on your organization's infrastructure or within a provider's cloud. It's important to keep in mind that a service provider's Service Level Agreement (SLA) end goals and requirements may differ substantially from your own organization's internal SLA requirements. Similarly, your organization's RTO and RPO definitions may differ from those of your provider. It is recommended to compare the two sets of requirements for important differences well before you need to recover from a disaster.

Consideration #2: Don't keep all your eggs in one basket

This consideration involves obtaining your own off-site copy of backup data from the cloud provider. According to common data protection wisdom, it makes sense to get a third copy of your data off-site. Just as organizations have long used tape vaulted off-site as a fallback position to their local, disk-based backups, it makes just as much sense to not put all of your data protection eggs (backup copies) in one basket (the same provider's cloud). Some areas to explore include:

- Insuring copies kept by your provider are geographically separated (just as you would geographically separate them from your private cloud environment)
- Considering setting up an offsite process that periodically transmits cloud-based backup data to your own data center (or even to another cloud provider)
- Using storage efficiency practices (compression, deduplication, etc.) to back up or take a snapshot of just the changed blocks in your workload since the last backup. How easy is it to transmit those changes outside of the cloud, let alone reassemble any changed blocks into a rehydrated working copy?
- Researching the cost involved in sending backup data sets outside of the provider's cloud. Based on tiered pricing, it can often cost substantially more than expected to send data out of the provider's cloud (compared to the cost of migrating data sets into the provider's cloud).

Time involved, bandwidth speeds and size of the "pipe" can also impact how long it takes to perform this type of process. In one case, a customer recognized they needed a 10GbE pipe to transmit a third copy of their cloud-based data back to their own data center. The size of the pipe wasn't a problem for the service provider. The problem occurred with the customer's originating virtual machine hosted in the service provider's cloud. Due to the compute resources required by the virtual machine to obtain and transmit the extra copy, the customer's unexpected compute costs more than doubled.

Consideration #3: Beware the cost of cloud services

This consideration shouldn't put fear in your heart. It's more a product of being forewarned and forearmed ahead of time in regards to a cloud provider's tiered pricing structures and how they can impact your data protection efforts.

For example, if an application workload requires frequent restores from prior backups, organizations should evaluate these needs based on the potential cost involved for the provider to allow this many restores. The amount of restores needed may push the customer's workload into a higher cost tier. This might then lead to unexpectedly high bills.

Consideration #4: Consider unique data protection requirements per application workload

Specific backup and recovery needs and options will vary in the public cloud. Application-based recovery can add another layer of requirements and potential complexity into data protection plans. By looking at each cloud-hosted application workload individually, you can better determine what type of protection is most needed to support that workload's day-to-day operations and to minimize the application's most common risks of data loss.

While a simple backup process might suffice for the majority of cloud-based workloads in your environment, there are some situations where application-level needs may dictate different data protection behaviors. Examples include:

1. **Mission-critical workloads** – Some workloads are so critical to the organization that their protection needs extend to also encompass business continuity and disaster recovery planning. Just as organizations often deploy different layers of protection within their own data centers, they should investigate what layers of protection may be warranted or available when the mission-critical workload is within the service provider's cloud. Some areas of inquiry might include:

   - ✔ The frequency and type of backups and recoveries available from the provider.
   - ✔ Where necessary, can you specify that replicated copies or snapshots of the workload be geographically separated from the primary copies, even specifying that they be saved to a different regional data center in another part of the country?

2. **Database corruption** – Database systems hosted in the cloud may need regular backup protection as well as protection from sudden or rolling database corruption. Common data protection tools used by the service provider, such as the use of disk-based snapshots, may not allow organizations to easily "roll back" to an earlier known, good state after they detect a corruption. Here, a separate daily snapshot or copy that is not overwritten may be warranted. Does the provider have these capabilities? If not, can you use an add-on or third-party tool to create another daily copy and host it in a different location?

3. **Accidental deletion of individual email** – Organizations now using cloud-based email services may find the service great for backing up current email. But, end users often make mistakes with emails and files. What happens if an email end user deletes an important email message by accident and doesn't discover it until a week later? Is there a way for them to get it back? If you expect these types of "user error" scenarios to happen often enough, it makes sense to see what recourse you have within the cloud provider's environment or by using third-party add-ons to fill any gaps.

Consideration #5: Don't forget data protection processes and paper trails

Regardless of whether many of your workloads are now hosted in the cloud, certain data protection practices should still be followed. The documentation of data protection processes, success/failure reports, DR testing, and the creation and maintenance of specific restoration procedures should still 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 to ensure everything is operating and in good order?
- ✔ 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 in this regard should be spelled out in detail within the vendor's Service Level Agreement (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?

Consideration #6: Protect cloud-based workloads beyond backup

Management, oversight, and protection of your cloud-based workloads cannot be neglected or delegated to the cloud provider just because the workloads are now housed in their cloud.

Beyond backup and routine data protection needs are also higher protections of your cloud-based data. This includes ensuring the proper levels of information governance, compliance, and security are in place for all of an organization's cloud-based workloads. This is especially true for highly regulated industries or for workloads involving proprietary or sensitive company information.

For example, email that relies on cloud services must be appropriately protected from loss. It also must be able to apply appropriate information governance policies regarding its use and access over time.

Some questions to ask in these areas might include:

- What levels of encryption are involved or needed with data or copies in transit and at rest?
- Who has access to your primary or protected data?
- Are different workloads adequately separated and protected in different areas of the provider's cloud, complete with different user login and password requirements to prevent unwanted or accidental access to key areas?

Trouble can occur when shadow IT teams in the same organization use the same cloud provider's services without the benefit of corporate oversight. A corporate policy with role-based access procedures and an established process to securely access cloud services (especially those accessed from a key cloud provider) can go a long way toward ensuring diverse workloads remain separated, secure, and relatively immune from unexpected surprises or cost overruns.

When role-based access to cloud workloads is not defined properly in advance, risks go up significantly. Server, storage, and network administration in the private cloud normally have separate credentials for access and modifications. Using a similar approach in the public cloud helps ensure proper security and governance.

The following is one example of what can go wrong:

> One customer had a large, shadow IT initiative running at a large hyperscale cloud provider. Three types of data were involved: The customer's production workloads, their development and testing workloads, as well as backups of the primary workloads stored on a secondary storage tier. Unfortunately, users of these three workloads accessed the services via the same sets of credentials (via the same login and the same password). Due to an error on the part of an administrator, however, both sets of workloads and their backups were accidentally deleted.

Learn about the key drivers for data center optimization in a recent IDG Research and Datalink study:

datalink.com/idg-survey-DC

Data protection in the hybrid cloud: it's time to take it seriously

The previous pages highlight a few of the data protection missteps that can occur when workloads are hosted in hybrid 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. For these reasons, many turn to Datalink experts and Datalink Managed Services to extend their own IT teams.

Working together with you, Datalink can help develop the right baseline plans to protect your growing workloads – wherever they reside.

Datalink 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

For more details on Datalink services, visit www.datalink.com/services

To receive the latest white papers and insight into data center technologies and practices, follow Datalink online:

twitter.com/datalinkcorp

blog.datalink.com

facebook.com/datalinkcorp

More about Datalink

Datalink is a complete IT services provider that helps companies transform their technology, operations, and service delivery to meet business challenges. Combining extensive experience, a full lifecycle of services and a comprehensive approach to producing IT innovations that empower positive business outcomes, Datalink delivers success across cloud IT transformation, next-generation technology, and security.

For more information, call 800.448.6314 or visit www.datalink.com.

An Insight company

datalink.com | insight.com

CDP-HC-WP-2.0.3.17