# Table of Contents

1 Executive Summary

2 The Cloud Awaits

2 Past Cloud Adoption Challenges

3 DCIG’s Methodology for Data Collection

3 Nine Best Practices for Incorporating Cloud Services into a Comprehensive Backup and Recovery Strategy

3 Best Practice #1: Keep a Copy of Data Locally

4 Best Practice #2: Define the Scope of Cloud Services in One’s Backup and Recovery Strategy

5 We Walked Away from Recovery in the Cloud

5 Take a Tiered Approach to Assessing the Degree to Which to Leverage Cloud Services in One’s Backup and Recovery Strategy

6 Dell Data Protection | Rapid Recovery Keeps All Three Tiers of Cloud Services Options on the Table

6 Best Practice #3: Strive for Independence in Cloud Services

7 Cloud Independence Achievable with Forethought and Planning

8 Cloud Services Provider Market Share

8 Best Practice #4: Use a Single Solution to Manage All Data Stores in the Cloud

9 Best Practice #5: Seed a Large Backup and/or Large Recovery

9 Best Practice #6: Secure Any Data Sent Offsite

10 Three Factors that Influence When Portable Media is Needed for Seeding

11 Best Practice #7: Deduplicate Data Sent to and Stored in the Cloud

11 Three Tips to Shorten the List of Possible Cloud Services Provider/Backup Software Pairings

12 Best Practice #8: Keep a Replica of the Backup Software in the Cloud

12 Best Practice #9: Test, Test and Test Again

13 Two Technology Best Practices are Intangible

14 Incorporate Cloud Services into One’s Backup and Recovery Process with Confidence
Nearly every organization talks about using cloud services in some capacity. Whether they plan to store and share files in the cloud using a service such as Dropbox, run a specific application in the cloud such as Salesforce, or access and share video content using a provider such as YouTube, cloud services represents the future in terms of how organizations will access, share, and store applications and data going forward. This fixation that organizations have on cloud services makes it particularly puzzling that only a relatively small percentage of them incorporate cloud services into their existing backup and recovery strategy.

A recent survey found that only 35 percent of organizations with over 500 employees use cloud services in any form as part of their existing backup and recovery process. This relatively low level of adoption of cloud services among organizations comes as a particular surprise considering their traditional challenges in managing backup and recovery. One would think that they might be the first ones to jump on board with a new approach to backup and recovery that promises to solve these issues.

However, their past woes with backup and recovery might serve as the first indicator as to why organizations remain hesitant to adopt cloud services. Tired of backup and recovery initiatives that fail, only partially work, or do not meet their expectations, organizations want cloud services to come to them on their terms. They want some assurances that cloud services are sufficiently mature so they may easily and confidently incorporate them into their existing backup processes before they decide to adopt them.

Cloud services providers appear to have largely heard these organizational demands and frustrations. Cloud services offerings range from archiving and replication to recovery in the cloud are now sufficiently mature and ready for corporate use.

The key for organizations to successfully adopt these cloud services hinges on their adhering to best practices that incorporate cloud services as part of a backup and recovery process. Derived from interviews with storage providers, resellers, and other end-user organizations who already use the cloud in some way as part of their backup and recovery process, these following nine best practices emerged when backing up or recovering from the cloud:

1. Keep a copy of data locally
2. Define the scope of cloud services in one's backup and recovery strategy
3. Strive for independence in cloud services
4. Use a single solution to manage all data stored in the cloud
5. Seed a large backup and/or large recovery
6. Secure any data sent offsite
7. Deduplicate data sent to and stored in the cloud
8. Keep a replica of the backup software in the cloud
9. Test, test and test again

Implementing cloud services as part of an organization’s overall backup and recovery strategy is achievable. However, organizations should worry less about finding products that offer the cloud options they need. These products exist and are in production. Rather, the key to successfully implementing cloud services as part of one’s backup and recovery process largely hinges upon following these best practices to confidently move from considering cloud services to actually adopting them.

The Cloud Awaits

Any organization that still wonders if cloud services are ready for prime time only needs to look at the numbers. In the last few years cloud revenue has escalated as more organizations utilize cloud service providers as well as take advantage of the increased number of cloud-based applications.

To provide some perspective, as recently as 2008 cloud revenue only amounted to $5.6 billion. While a sizable number, by 2014 cloud revenue had climbed 10X to $56.6B. Now some analysts project it to more than double to $127 billion by 2018.\(^2\) Thanks to the availability of cloud services from dozens of cloud services providers including Amazon, Cloudera, Dropbox, Microsoft, and Rackspace among many others, it is reasonable to expect that more organizations will forge ahead with plans to store more of their data with these and other cloud services providers.

Yet organizations plan to do more than simply store data in the cloud. They will also utilize more applications offered by cloud services providers. Software-as-a-service (SaaS) applications are already growing at a rate six times (6x) faster than on-premise applications and represent 30 percent of all application spending. Customer relation management (CRM) software such as Salesforce represents how quickly SaaS is being adopted. Some estimate that by 2018 more than 60% of CRM applications will be cloud based.\(^3\)

This rapid adoption of cloud services among organizations, the increased number of cloud service providers available to store and host applications, and the number of applications available from cloud services providers lead to an obvious conclusion: cloud services are here to stay and will continue to grow for the foreseeable future.

Past Cloud Adoption Challenges

Organizations often view backup and recovery as, at best, a necessary evil. Historically it represents a cost to the business adding little value due to the infrequent use of recovery and, if and when it is called upon, does not work the way they anticipate.

Aggravating the current situation, managing backup and recovery is not a core competency of the IT staff in many organizations. This results in these organizations possessing poorly implemented, inflexible solutions making them difficult to manage and challenging to scale. These factors suggest that backup and recovery is ripe for change.

Introducing cloud services into an organization’s backup and recovery strategy addresses many of the drawbacks present in legacy backup architectures. Incorporating cloud services as part of an organization’s broader backup and recovery plan gives organizations the new flexibility they need to more easily scale the amount of data they can protect. It also creates new options to archive and/or replicate data offsite, control and lower business costs, and makes it more cost-effective and practical for almost any sized business to implement disaster recovery offsite or in the cloud.

Hence the real question is not “Should organizations use cloud services?” as part of their backup and recovery plan. The answer to this question is, in short, a resounding “Yes!” The real question is, “Why are more organizations NOT already using cloud services as part of their existing backup and recovery strategy?”

This hesitation on the part of organizations to more aggressively adopt cloud services stems from two concerns. First, organizations still have a hangover from past bad experiences with backup and recovery software and, as such, remain cautious about trying anything new. Second, in terms of maturity, cloud services is still relatively young from the perspective of many organizations.

A recent survey produced by Spiceworks found that only 35 percent of those surveyed use the cloud as part of their backup and recovery strategy.\(^4\) A separate, older 2014 report found that only 17% of organizations had “embraced” the cloud at that time.\(^5\)

---

3. Ibid.
While organizations slowly adopted cloud services as part of their backup and recovery strategy from 2014 to 2016, the uptick was not stronger as many still have lingering concerns about:

- The steps they should take to adopt cloud services as part of their backup and recovery strategy
- The risks involved with adopting cloud services
- How to manage their current infrastructure and preserve existing investments

To address these concerns, organizations need a clear path forward. While the use of any cloud services warrants that organizations monitor and review them, the future of most organizational backup and recovery strategies will in some way include the cloud. As such, organizations need to identify and put best practices to place to successfully incorporate cloud services into their overall backup and recovery strategy.

**Nine Best Practices for Incorporating Cloud Services into a Comprehensive Backup and Recovery Strategy**

**Best Practice #1: Keep a Copy of Data Locally**

Successfully backing up and recovering data locally should proceed any organization’s discussions about using cloud services as part of their broader backup and recovery strategy. Cloud services are a two-edged sword in that they possess the capabilities to improve and even solve certain backup or recovery challenges. However, if organizations cannot yet successfully back up and recover data locally, implementing cloud services before they successfully back up and recover data locally only stands to aggravate existing problems and may even make them worse.

Keeping a copy of data locally facilitates:

- Completing backups and recoveries at LAN speeds
- Shorter backup windows
- Superior recovery time objectives (RTOs)
- Improved recovery point objectives (RPOs)
- Disaster recovery (DR) for mission critical applications

Using local disk (i.e. secondary storage not located in the cloud) as the primary backup target ensures the creation of local, reliable copies of data. Disk offers the best media as a backup target from cost and performance perspectives. Backing up and recovering data over a LAN provides the amount of bandwidth needed to deliver the highest throughput. Most organizations only need to keep their backup data on hand for a short period of time to satisfy the three most common reasons for restoring data:

1. Application and file restores
2. Virtual machine (VM) boots
3. Creating tape copies for archival and long term retention

To minimize the amount of local disk media needed while maximizing its available capacity, organizations should keep local backups at least one week on disk and no more than six months. Many organizations find storing a month’s worth of backups on local disk a good compromise. By retaining data for a month they can fulfill almost every recovery request as they may restore data from local disk.

In contrast, using the cloud as a primary backup target currently creates two specific challenges.

- **Increased backup and restore times.** Due to the inherent WAN bandwidth constraints associated with backing up or recovering data with a cloud services provider, backup and recovery almost always takes longer than using local disk. This results in backups not completing in designated backup windows, failed backups, and more time spent managing backups and recovering data.

- **WAN bandwidth restrictions.** Organizations only have limited amounts of available WAN bandwidth to send or retrieve data. Improving the speed of data movement to or from the cloud typically requires a significant increase in WAN bandwidth. This, in turn, increases the costs associated with using the cloud as a primary backup or recovery target.

---

DCIG's Methodology for Data Collection

DCIG spoke to various storage and solutions architects ranging from end-users to resellers and storage providers that had implemented cloud services, were involved in implementing cloud services, and/or were currently using cloud services as part of their broader backup and recovery strategy. From those conversations, DCIG gleaned specific best practices that these individuals followed or learned and subsequently adopted while incorporating cloud services into their backup environment.

All of the best practices included here were shared with these individuals prior to this paper’s publication. Each of these individuals agreed these are best practices that they would recommend to their peers and colleagues and others faced with the same decisions.
Currently the best use for cloud services in the context of backup and recovery is as a secondary target. By only storing a copy of backup data with a cloud storage provider for long term data retention and archive, organizations reap the primary benefits that cloud storage providers currently offer while minimizing their existing drawbacks. Consider:

- **Long term archive and/or data retention.** Many public cloud services providers offer very attractive rates for storing data on their cloud storage platform—as low as a penny per gigabyte (GB) per month. However, organizations may incur additional costs when they retrieve this data from the cloud services provider. By only storing data with the cloud services provider that is for archive and/or long term data retention, organizations reduce the likelihood that they will incur these extra charges.

- **Manage data stored in the cloud with existing backup software.** A growing number of backup software solutions manage the placement, movement, and retention of backup data at both the organization’s site and the cloud services provider’s site. While backup software integration with cloud services providers varies and differs in capabilities, organizations may leverage this software to a single point of management for their data located both locally and in the cloud.

- **Minimizes or even eliminates the need to manage a growing storage infrastructure.** Using a cloud services provider to store data long term addresses concerns about overall data growth. Data typically grows 30 percent or more year over year in most organizations *(though estimates of annual data growth can vary widely.)* By storing that extra data with a cloud services provider, organizations eliminate the need to manage and store data in house while shifting the costs, management overhead, and infrastructure needed to store it long term with the cloud services provider.

- **Lays the foundation for offsite DR.** Many organizations envision doing DR at another site. But to achieve that objective, they first need to get their data offsite. By storing their data with a cloud services provider, they set the stage to eventually accomplish that objective. Further, a number of cloud services providers such as Amazon Web Services (AWS) and Microsoft Azure already permit organizations to setup replicas or standby systems at their site, perform DR, and even recover in different geographic regions by creating copies of data at other locations.

Successfully backing up data and recovering it locally to disk should currently be every organization’s primary objective. But once in place, leveraging cloud services makes a great deal of sense as it sets the stage for organizations to lower both their capital and operational costs while enabling them to achieve longer term objectives that previously were too costly and difficult to achieve.

**Best Practice #2: Define the Scope of Cloud Services in One’s Backup and Recovery Strategy**

Every organization’s data, environment, and business differ. As such, there is no “one size fits all” formula that organizations may follow to implement cloud services as part of their overall backup and recovery plan. This message came through loud and clear in the conversations that DCIG had with the individuals that it interviewed as each one leveraged cloud services differently and took different paths to arrive at its decision as to how to best implement them.

Where there is no single path that organizations may follow, cloud services providers offer four general ways for organizations to leverage their cloud services. Each offering may require progressively more capabilities and involvement on the part of the cloud services provider to deliver depending on the solution selected. These offerings include:

- Archiving to the cloud
- Replicating to the cloud
- Archiving and replicating to the cloud
- Archive, replicate and recover in the cloud

**Archiving to the Cloud**

When archiving to the cloud, the cloud services provider simply functions as a backup and recovery target that appears as a networked storage device to the backup software. Once in place, organizations only need to periodically verify that the cloud services provider offers sufficient bandwidth and storage capacity to archive their data to the cloud.

**Replicating to the Cloud**

Replicating data to the cloud services provider can also be very straightforward to implement, though the simplicity of the implementation depends heavily upon the solution and cloud services provider selected. To ease the implementation, an organization needs to select a replication solution that interacts with and is supported by the cloud services provider’s cloud.
We Walked Away from Recovery in the Cloud

Failure to grasp the full cost of recovery in the cloud surfaced at one large legal practice in the Northeast. When it first adopted cloud services, it implemented both archive and recovery services in the cloud. However, after using both of them for a period of time, it walked away from recovery in the cloud.

Initially, its IT manager was excited about the possibilities of bringing up any of the legal firm’s applications at any time with the cloud services provider. But once it implemented recovery in the cloud, reality set in. It simply could not justify the cost to run both its applications and supporting services in the cloud. It required the cloud services provider to operate a dedicated infrastructure which, over time, became cost prohibitive.

The legal practice ultimately decided to scale back its level of cloud services adoption to simply archiving data with the cloud services provider. This resulted in a configuration that aligned much more closely with its actual budget and needs. While it does not currently perform recovery in the cloud, the law firm may still reintroduce recovery in the cloud at some future point should it want or need to go down this path again.

Take a Tiered Approach to Assessing the Degree to Which to Leverage Cloud Services in One’s Backup and Recovery Strategy

Everyone wants access to all of their data for all of their applications immediately and who can blame them? It’s a great idea but unfortunately, “if you want to dance, you have got to pay the fiddler.” In other words, it is costly to offer 100% availability for every application.

To optimize cloud services as part of a broader backup and recovery strategy, one cloud solutions architect recommends that organizations break the application servers and VMs they want to recover in the cloud into three tiers:

1. Must Have
2. Need to Have
3. Nice to Have

Using a spreadsheet, create three tabs or, on a single tab, create column headings that have each of these three tiers listed at the top. Then go through the application servers in the environment and list each server under each tier as appropriate. The first step is the easiest. Organizations should assume that they will archive the data for all of their applications to the cloud. Even though many organizations may still use tape in whole or in part to archive their data, organizations should begin to view the cloud as the preferred repository for their archival data as a means to simplify the management of data in their environment.

Next, determine which applications, if any, belong in the “Must-Have” and “Need to Have” tiers. To do so, organizations must first either know or determine the RPOs and RTOs for each of these applications. By identifying what applications must be recovered in the cloud and by when they need to be recovered helps organizations:

- Determine what applications they actually need to recover in the cloud
- Establish where they need to recover these applications
- The time by which the applications need to be recovered

Organizations that make these assessments before they implement any tier of cloud services helps to ensure they make the most appropriate decision about the degree to which they want to use each of the different tiers of available cloud services.

Archiving, Replicating and Recovering in the Cloud

Introducing recovery in the cloud adds yet another wrinkle into the decision making process. As part of this configuration, organizations will expect to perform real-time or near real-time application recoveries locally and at the cloud service provider’s site.
To recover at the cloud services provider’s site, the cloud services provider will need to understand the application(s) replicating data to it. Further, it will need to become more deeply involved in an organization’s overall backup, recovery, and business continuity preparedness to ensure it understands the organization’s requirements and can successfully meet its expectations for recovery.

Of these four options, many organizations may initially only archive data in the cloud as archive typically requires the least amount of forethought and effort to implement regardless of what solution they have or choose. However, those organizations that invest a bit more time and effort in selecting a backup and recovery solution will find that archive, replication or both may be implemented with a cloud services provider with nominal amounts of effort and cost on their part.

Organizations that hope to either initially or eventually perform recoveries in the cloud need to prepare to invest some time and effort to make that a reality. They will want to identify a cloud services provider that offers all of these options (archive, replication, and recovery). They will also want to ensure they have staff on hand or a trusted partner to assist in the implementation and ongoing support. By choosing a cloud services provider that supports all three options, they can begin by archiving and/or replicating data to the cloud services provider knowing they have a roadmap to perform recoveries when they are ready at the cloud services provider’s site.

Best Practice #3: Strive for Independence in Cloud Services

Organizations should view cloud services in much the same context as they view any other datacenter resource: a replaceable component of their infrastructure. Subscribing to this view may be particularly difficult. The decision to use cloud services challenges organizations as it requires them to entrust the storage of their data with a cloud services provider. This decision can create both emotional and technical attachments to the cloud services provider.

Even though a cloud services provider will act as the custodian of an organization’s data, the organization must remember that it needs to remain in control of its data at all times and retain responsibility for it. While perhaps unpleasant to consider, organizations may need to change cloud services providers or introduce a new one into the mix at some future time for reasons such as:

1. The cloud services provider may go out of business
2. The cloud services provider may want to go a different direction and no longer offer the cloud services the organization wants or needs
3. The organization may need cloud services that the cloud services provider does not offer
4. Pricing structure may go up or change and the organization can no longer afford or justify using that cloud services provider for that service

IT professionals shared with DCIG that as they evaluated various cloud services providers and their respective service offerings, they first took the time to understand the impact of onboarding the cloud services provider of their choice. However they also investigated how much time, money and effort it would take to migrate to a different cloud services provider should business conditions warrant.
Best Practices for Incorporating Cloud Services into a Comprehensive Backup and Recovery Strategy

As part of an organization’s due diligence into incorporating cloud services into its backup environment, it needs to confirm that that the backup solution in which they are currently invested or in which they are looking to make an investment simultaneously connects to multiple cloud services providers. Ideally, the backup solution should view and treat cloud services from any provider from a completely agnostic viewpoint.

Achieving this ideal is currently very difficult, especially for those organizations that wish to replicate and/or recover data at the cloud provider’s site. These more advanced cloud services typically require the cloud provider to:

• Integrate with the organization’s backup software, appliance, or cloud gateway solution
• Offer specific options to recover applications at its site
• Facilitate recoveries back at the organization’s production or DR site (if a DR site exists)

Organizations may currently view and treat archiving to any cloud services provider as a commodity service since when organizations utilize cloud services in this manner, it primarily functions as a storage target. The only exceptions to that would be if the cloud services provider offers options to store archived data in a format that is more easily accessible, may be more quickly retrieved, or stored in specific geographic regions that satisfy some specific internal or external organizational compliance requirement.

Organizations may also potentially view replicating to a cloud services provider as a commodity service. While it does require the cloud services provider to natively support the specific backup and/or replication software used by the organization, if this support exists, replicating to the cloud services provider becomes a fairly straightforward implementation.

As soon as an organization moves up the stack to recovering applications with the cloud provider, tighter integration and longer term contracts with the cloud services provider may be needed. Absent these types of contracts, it varies by cloud services provider if it will provide the guarantees or levels of service that organizations will need to successfully recover their applications with the cloud services provider.

Cloud Independence Achievable with Forethought and Planning

Achieving cloud independence requires organizations to know how they want to incorporate cloud services into their backup and recovery process. However, as organizations evaluate cloud services, which ones they want to use, and under what circumstances, they should primarily view archive, the most basic of cloud services, as a ubiquitous offering.

Any organizations that wishes to move beyond archive to replicating data to the cloud and/or recovering applications in the cloud will need to select a solution that gives them the flexibility to use one or more cloud services providers.

To do so, organizations will want to adopt a backup solution that gives them as much flexibility as possible to use the cloud services provider of their choice. Dell Data Protection Rapid Recovery is one such solution that gives organizations the option to connect to multiple cloud services providers for replication, recovery, or both as it includes support for the following:

• Amazon Web Services (AWS)
• eFolder
• Microsoft Azure
• OpenStack compatible services such as Rackspace

If an organization’s goal is to archive and/or replicate to the cloud, Rapid Recovery supports AWS and Microsoft Azure for both of these services. It also supports OpenStack compatible clouds such as Rackspace for archiving and eFolder for replication. Yet whether archiving or replicating data to any of these cloud services providers, Rapid Recovery treats them as ubiquitous storage targets. Then, when an organization needs to locally recover data from any of these cloud services providers, Rapid Recovery simply retrieves the data.

In cases where an organization wants to use Rapid Recovery to recover applications with the cloud services provider, a number of them already support Rapid Recovery for this role including AWS, eFolder, and Microsoft Azure. While organizations may not switch recovery from one cloud services provider to another with a mouse click, they can usually do so within a reasonable period of time (30 days or less.)

Organizations currently have a great deal of freedom to archive with multiple cloud services providers or migrate archives from one cloud provider to another. However, solutions such as Rapid Recovery give organizations more freedom to easily replicate to a number of cloud services providers. Further, Rapid Recovery makes it relatively straightforward to implement recoveries in the cloud and even makes it possible to change which cloud services provider performs the recovery should that become a requirement.
Best Pratices for Incorporating Cloud Services into a Comprehensive Backup and Recovery Strategy

Cloud Services Provider Market Share

Most organizations prefer to select and work with a cloud services provider with which they are familiar, have significant market share, and are financially stable. The two cloud services providers with the largest market share, Amazon Web Services with 29 percent and Microsoft Azure with 12 percent, are both well-known and financially stable. However, it is notable that cloud service providers, referred to as “All Others” in Figure 1, collectively have 41 percent market share, more than any individual cloud services provider.

Some IT managers to which DCIG spoke had already migrated from one cloud services provider to another. They found the vast majority of the tasks that they needed to perform could be fairly easily accomplished (relatively speaking.) They attributed the use of only one management console as contributing to their ease in managing these various cloud interfaces.

Solutions such as Rapid Recovery illustrate how organizations may use a single management console to manage multiple recovery types across both local storage and with multiple cloud storage providers. Using Rapid Recovery, organizations can perform any of the following types of recovery:

- Physical to Physical
- Physical to Virtual
- Virtual to Physical
- Virtual to Virtual

Organizations may perform these recoveries either locally or in the cloud with potentially multiple different cloud providers using only the Rapid Recovery management interface.

Best Practice #4: Use a Single Solution to Manage All Data Stores in the Cloud

Every organization to which DCIG spoke used a different solution to place and manage data stores in the cloud. These solutions included:

- Backup software
- Cloud gateway
- Deduplicating backup appliance
- Integrated backup appliance

Despite these differences in approaches in storing backup data to the cloud, these organizations notably only used one solution to manage backup data stores in the cloud.

Some IT managers to which DCIG spoke had already migrated from one cloud services provider to another. They found the vast majority of the tasks that they needed to perform could be fairly easily accomplished (relatively speaking.) They attributed the use of only one management console as contributing to their ease in managing these various cloud interfaces.

Solutions such as Rapid Recovery illustrate how organizations may use a single management console to manage multiple recovery types across both local storage and with multiple cloud storage providers. Using Rapid Recovery, organizations can perform any of the following types of recovery:

- Physical to Physical
- Physical to Virtual
- Virtual to Physical
- Virtual to Virtual

Organizations may perform these recoveries either locally or in the cloud with potentially multiple different cloud providers using only the Rapid Recovery management interface.


© 2016 DCIG, LLC. All rights reserved.
Best Practices for Incorporating Cloud Services into a Comprehensive Backup and Recovery Strategy

Using it, organizations may manage archiving and replicating backup copies to the cloud and even recover at the cloud services provider’s site.

**Best Practice #5: Seed a Large Backup and/or Large Recovery**

Every organization faces two potential challenges when using cloud services:

- Archiving and/or replicating a large amount of data to the cloud services provider
- A large local recovery using data restored from the cloud services provider

Performing either of these two tasks is akin to trying to drink a large, thick milkshake through a very narrow straw: one can do it but it takes a great deal of time and effort to complete. The first copy of data an organization archives or replicates to the cloud may be very large in size—potentially multiple terabytes. Transmitting this amount of data over an existing WAN link to a cloud services provider can take dozens of hours or even days to complete. In the same way, when an organization needs to recover a large amount of data from a cloud services provider, it can also take a long time to recover over the WAN link.

The prolonged periods of times needed to archive, replicate, or recover large amounts of data stem from how organizations size their WAN connections. The data they transmit daily only equates to about five percent of the total amount of data that organizations have under management. Further, organizations have about 24 hours to transmit these changes so they purposely limit their WAN bandwidth to control costs.

It is during those infrequent times when organizations need to move a large amount of data to or from the cloud, such as when creating the first archive, doing the first replication, or doing a large recovery, that this WAN bandwidth limitation surfaces. Moving all of this data in a period of time that organizations find acceptable dictates they mitigate or eliminate their dependency on the WAN link.

The definition as to what constitutes an acceptable period of time varies by organization. Many would say it should take no more than 24 hours to move data to or from the cloud if moving it during the week. If done over a weekend, it should take no more than 72 hours to complete.

To seed the cloud site or a recovery site, organizations need to create a one-time copy of the data on some type of portable media, typically disk or tape. The organization or the cloud services provider will then physically ship this media to the other location.

To accomplish this, organizations must evaluate both their backup software and the cloud services provider to verify the following:

1. The backup software must be able to create a copy of data on portable media. Once created, the backup software must be able to export this copy of data, rediscover it once it arrives at the cloud services provider’s site, import it so it can once again manage this data, and then synchronize this copy of data with the production site by transmitting any changes that have occurred since the copy of data was created.

2. The cloud services provider should offer the option to receive the physical media from the organization. It must also be able to create physical media to send back to the organization to seed recoveries at the organization’s site.

Using a WAN link to send and receive data from the cloud is the best and most reliable way for organizations to archive, replicate, and/or recover data to meet their needs on a day-to-day basis. However, organizations cannot beat the bandwidth of a truck or a plane during those infrequent times when they have to move large amounts of data. While these modes of transporting data are not as elegant as moving data over a wire, in these circumstances they are faster and more efficient and cost effective.

**Best Practice #6: Secure Any Data Sent Offsite**

Organizations should always secure any data sent offsite—whether sending it to the cloud or another location owned or managed by them. Securing data takes two forms:

1. Encrypting data as it is transmitted, stored or both
2. Authentication to ensure only the appropriate individuals access it

**Encryption**

Organizations should view the encryption of any data that leaves their premises as a prerequisite and especially when sending it to the cloud. To do so, organizations must verify they have options available to them to encrypt their data at a sufficiently high level of encryption (minimally AES-256) as it is transmitted or “in-flight” to another location as well as when the data resides or is “at-rest” in the cloud services provider’s site.
Organizations may encrypt their data using different options prior to sending it to the cloud services provider. They may use the application that owns and manages the data, the backup software, or a replication appliance. However, most organizations will find that backup software provides the simplest and easiest means to encrypt the data prior to storing or transmitting it offsite as well as managing it after it is encrypted.

Ideally, organizations should encrypt data both as it is in-flight and at-rest with the cloud services provider. However, both of these options may not be possible to implement, especially in circumstances where an organization replicates data and/or recovers applications with the cloud services provider.

As such, organizations should always minimally encrypt data when transmitting it to the cloud or another site regardless if they move it using portable media or over a WAN link. They should also perform at-rest encryption of any archives stored in the cloud. Performing at-rest encryption of application data that is part of a replication pair is likely not an option in many circumstances.

**Authentication**

Most organizations only access and recover their archives through their backup software. This puts the onus on organizations to ensure that the individuals who need access to data in the cloud have the proper permissions to first log into the backup software so they may retrieve and restore the data they need. While an acceptable and widely used approach to securely access to backup data, organizations should take the following steps to ensure it works as they intend.

- **Confirm the organization still wants to follow this strategy and has the right tool to execute upon it.** Using one tool to store data to and retrieve it from the cloud does simplify administration. However, the tool can become problematic if it does not possess the right features to execute upon this strategy, such as offering the needed levels of encryption or supporting multiple user logins with different administrative permissions.

- **Verify the individuals who need access to the backup software have it.** The number of individuals who can access and log into an organization’s backup software to back up and recover data is usually quite small and may be only one person in small organizations. Organizations should minimally confirm that at least two people have logins to the backup software and ideally a few more. Further, verify individuals such as application

### Three Factors that Influence When Portable Media is Needed for Seeding

No exact formula exists to determine exactly when an organization should seed the cloud using portable media for an initial archive or replica or when it should use portable media to seed a recovery from the cloud. However organizations may look at three factors to determine whether or not they will need to use portable media to seed the cloud or seed a recovery from the cloud.

1. **The amount of time available to move all of their data.** Organizations need to internally determine how quickly they need to send or retrieve all of their data from the cloud. If they need to move all of their data offsite or recover it within 72 hours, they will want to have portable media available as a seeding option. If they need to move it all or recover it all within 24 hours, they should have a portable media option available to them.

2. **The amount of data to send or receive from the cloud.**
   - **a.** If an organization has 10GB or less of data to move or recover, it can almost always move this amount of data in less than 72 hours and potentially under 24 hours over a T-1 WAN link. In this scenario, no portable media option may be needed.
   - **b.** If the amount of data to move or recover is between 10GB and 30GB, organizations should have a portable media option available to them to move the data in under 24 hours. However, if they have a T-1 WAN link, they can potentially move this data in about 72 hours.
   - **c.** Organizations with over 30GB of data to move or recover should definitely plan on having portable media available to them as an option to move data, especially if they need to recover it in 24 hours or less.

3. **The amount of WAN bandwidth to the cloud services provider.** A high bandwidth WAN link such as a T3 line to the cloud services provider can alleviate an organization’s need to rely as heavily upon using portable media to seed the cloud or a recovery. Even in this situation, however, organizations with large amounts of data to move (500GB or more) should still have a portable media option available in the event their high bandwidth link goes down, they need to recover in less than 24 hours, or the amount of data that they need to copy or recover takes more than 24 hours to move.
owners and/or database administrators also have some form of access to the backup software or a means to contact the people that do.

- **If using Active Directory or LDAP, verify your backup software integrates with it.** Many organizations use Active Directory or LDAP to authenticate users on their corporate network. If so, verify the backup software integrates with these directory services and ideally uses them as a means to authenticate other individuals logging into the backup software as a means to control who accesses what data.

- **Maintain secure log of user logins and passwords used over time and during what period of time they were used to retrieve older backups.** User logins and passwords change over time. However to recover older data, organizations often need the login and password combination used at the time the data was protected.

Finally, organizations should be aware that many cloud services providers offer multi-site replication which can result in them storing copies of an organization’s backup data in multiple geographically dispersed sites. Maintaining copies of your backups in other locations provides an added layer of protection against any disasters that affects an organization’s datacenter as well as the primary cloud provider’s datacenter where the organization’s data is stored. Where possible, and if an organization can afford it, they should look to take advantage of cloud services providers who have at least two (2) locations where an organization can store multiple copies of its data offsite.

**Best Practice #7: Deduplicate Data Sent to and Stored in the Cloud**

Backup in and of itself is merely a copy of an organization’s data. As such, once the backup software creates the initial copy of data, the differences between that first copy and subsequent backups vary little—usually no more than five percent.

Introducing deduplication into the backup process minimizes the amount of backup data transmitted to and from the cloud services provider as well as reduces the amount of data stored at the cloud services provider’s site. Using deduplication prior to sending the data to the cloud services provider can reduce backup data stores by up to 95 percent assuming that organizations can achieve a 20:1 ratio using deduplication.

**Three Tips to Shorten the List of Possible Cloud Services Provider/Backup Software Pairings**

The proper choice of either a cloud services provider or a backup software solution represents a difficult decision for any organization. However, the difficulty of making the best decision for each of these solutions compounds when one considers that each pairing of these technologies creates its own set of trade-offs.

For example, each backup software product provides different options to archive, replicate, and/or recover data in the cloud. Further, each product does not work equally well and in the same way with all cloud services providers. Similarly, cloud services providers offer different features and do not work equally well (*or may not work at all*) with certain backup software solutions. It is when one pairs each of these technologies and then factors in the capital and operating expenses associated with each one that the viable cloud strategies available to an organization can again become rather murky.

To help shorten the list of possible technology pairings, organizations should verify each one can minimally perform the following:

1. **Stores data both locally and in the cloud.** Any technology pairing should minimally store backup data locally no less than one day and ideally up to one week or more while tiering all older copies of data to the cloud. A local copy ensures fast local recoveries. Tiering all older copies of data off to the cloud mitigates local storage costs and the overhead associated with longer term data management.

2. **Offers an option to quickly copy data to the cloud.** Once data is backed up locally, organizations should have an option to quickly (within 24 hours) copy replicate data along with application recovery points offsite to the cloud. While all viable technology pairings should support tiering data to the cloud, the speed at which this may be accomplished may vary due to bandwidth limitations at the organization’s site, the cloud services provider’s site, or both.

3. **Backup to the cloud.** In small offices, remote sites or locations with minimal amounts of data or very small data change rates (*less than one percent*), organizations may want the option to directly backup to the cloud without any local backup. Verify the technology pairing supports this configuration if this is a requirement.

Organizations need to keep in mind that no backup software/cloud services provider pairing yet addresses every possible backup and recovery scenario. This makes it imperative that organizations first define what they need to accomplish with both their backup software and the cloud services provider before settling on a technology pairing. Once an organization defines the conditions, it may then select a technology pairing that matches their requirements or most closely aligns with them.

© 2016 DCIG, LLC. All rights reserved.
Deduplication serves two important purposes in this process:

1. **Reduces WAN bandwidth requirements when moving data both to and from the cloud.** The less data that organizations send, the smaller they need to size their WAN link to the cloud services provider. Data deduplication serves to reduce the amount they spend on the WAN link and optimizes available bandwidth.

2. **Reduces data storage capacity requirements with the cloud provider.** Data storage costs with public cloud services providers have admittedly dropped drastically in recent years with some now charging as little as one cent per month per gigabyte (GB) for storage capacity. By storing data in a deduplicated format, organizations can maximize their available capacity used in the cloud while controlling their monthly operational costs.

Despite deduplication's advantages, organizations must still utilize it thoughtfully. Backup software, backup appliances, and cloud gateways each implement deduplication differently so they do not all achieve the same deduplication ratios.

A solution such as Dell’s Rapid Recovery deduplicates data that it archives or replicates to maximize available WAN bandwidth and storage capacity. However, for mission critical applications that organizations need to restore quickly, Rapid Recovery offers a Virtual Standby feature. Using this feature, organizations may perform a fast recovery of the application (15 minutes or less) either locally or in the cloud without having to wait for Rapid Recovery to rehydrate the data.

**Best Practice #8: Keep a Replica of the Backup Software in the Cloud**

A vulnerability that almost every organization must manage is its dependency upon its backup software to recover or restore data. If an organization loses its backup server, it cannot initiate either a recovery or a restore until the backup software itself is up and running as it has the specific metadata about each backup needed to recover the data. This metadata contains:

- The instructions about how frequently backup copies were made
- When the backup copies were created
- How long the copies were held for
- Where the copies were stored
- Who has the authority and permissions to recover the data
- The encryption keys

Backup software contains metadata that provides the structure of backup policies. Should an organization lose this metadata or be unable to access it, they put their ability to recover any or all of their data at risk.

Fortunately, organizations have a number of methods at their disposal to protect and recover this critical cog in their backup process. An organization should minimally protect its backup software’s metadata and server configuration by backing it up and storing a copy of it both locally and with a cloud services provider. Once protected, an organization may recover to a physical or virtual machine either locally or residing in the cloud.

Organizations should ideally have a replica of their backup software server residing with the cloud services provider. In this way they may immediately begin to perform recoveries of other applications at the cloud services provider’s site should that become a necessity. By having both a local copy of the metadata as well as one in the cloud, organizations create another layer of availability and recoverability to address those circumstances where either the local backup server or the physical site where it resides become unavailable.

**Best Practice #9: Test, Test and Test Again**

An organization may follow every best practice and put every piece of technology in place needed for its backup environment and then still have to face this harsh reality: There is no guarantee that everything will work either initially or long term.

Until an organization tests each of the components both individually and then together as part of their overall backup process, it should assume these component will not seamlessly work together when someone flips the switch to turn them on.

Since no provider yet offers a turnkey, end-to-end integrated backup solution that guarantees recoverability, it falls to the organization to verify the operability and interoperability of all of the different components within its backup infrastructure. To do so, an organization should take a two-pronged approach to testing.

**Test Local Backups and Recoveries**

An organization should ideally test the impact of backups and recoveries on its local network to verify they work. Once successful, it should then look to try run the same backups again, only this time throttling the amount of network bandwidth that the backup application can access to measure the impact on how long it takes the backup or recovery to finish or determine whether or not it can complete. Ideally, these types of tests should be done for applications, files, and entire systems.
As an organization performs these tests, it should record the steps and document the results to create standard operating procedures that become part of its backup playbook that it can follow going forward. Only once an organization has confidence that its local backup and recovery processes work should it turn its attention to testing cloud services as part of its overall backup process.

**Best Practices**

Testing cloud services as part of the overall backup process takes on a slightly different slant. At this point an organization should know its local backups and recoveries work and that they have recoverable copies of data.

Now an organization may focus its testing on archiving, replicating, and/or recovering copies of data with the cloud services provider. The scope of the testing will somewhat hinge on the cloud services that the organization uses. If only archiving and/or replicating data, the testing may be straightforward as confirming that the data successfully arrives at the cloud services provider’s site and that it may be retrieved and recovered locally in the time frame needed.

If recovering applications or data in the cloud, the time to test those features will be, by necessity, more extensive. An organization will need to verify that application recoveries work as intended, to possibly include testing of failover and failback. Only once an organization completes these tests should it deem the project of introducing cloud services into the infrastructure as complete.

Finally, an organization should establish a schedule to regularly retest all of the components in its backup infrastructure. Every organization routinely adds and removes components from its infrastructure. Organizations can be assured that within 3-6 months of completing the testing some part of its environment will have changed. As such, an organization should ideally retest all components no less than once a year.

### Two Technology Best Practices are Intangible

Organizations may pick all of the right technologies and follow the technical best practices to the letter and still find that they do not achieve their desired objectives for reasons that have nothing to do with the technology. Two best practices that organizations should follow do not depend upon technology at all. Rather, certain intangibles exist that organizations should have in place and as part of their culture to ensure the success of any technology initiative.

**Intangible #1: Know Your Organization’s Appetite for Change**

Knowing and understanding an organization’s technical make-up is always a key to successfully making any changes or determining if they are even necessary. Storing data in the cloud and/or doing recoveries in the cloud is likely a big change for many organizations after years of storing and managing data locally.

The cloud clearly beckons but organizations need to have a good grip on their internal appetite for this level of change. Try to ease into the adoption of cloud services by gradually introducing them to targeted individuals and departments before jumping in with both feet.

As the organization becomes more familiar with cloud services in general and the specific features they use, the organization’s probability of long term success increases with each passing day. Moving too quickly may create issues that negatively impacts the entire organization and may only serve to slow down or even halt the adoption of cloud services as a whole.

**Intangible #2: Communicate with Your Organization**

Incorporating cloud services into an organization’s overall backup and recovery process impacts everyone. This put the onus on the individual or team leading the cloud services initiative to make an effort to understand their needs and align the capabilities of the cloud services offerings with them if possible.

The individual or team should seek to communicate with the rest of the organization about the new possibilities for recovery that using the cloud creates. Educate them about the different cloud services offerings currently available as well as seek to understand their specific requirements for recovery times, appetite for data loss and data retention requirements.

Document this feedback and rank their needs in order of importance to the business. Keep in mind the organization’s ability to meet their expectations within the constraints of the proposed infrastructure. Once the assessment is complete, share the results with them as well as with the management team. By involving these various individuals and teams in the process, it increases the likelihood they will take a more active role in the introduction of cloud services into the environment.

Gathering this information and then sharing it should help the organization feel more confident about using cloud services as well as the individuals managing its implementation. It also provides the criteria for organizations to create a feature matrix that they can use to evaluate cloud vendors and the services they offer.

This will codify the buying process and build the foundation for an RFP or RFI that is well documented to match the specific needs of the organization. It also sets the stage for vendors to compete against specifications based on input from internal stakeholders to help ensure that the cloud services offerings procured align with what the individuals in the organization indicated that they wanted.
Incorporate Cloud Services into One's Backup and Recovery Process with Confidence

It is no secret that organizations of all sizes want to incorporate the cloud as part of their backup and recovery process. Confidently and successfully implementing it is another story altogether. Like almost other strategic initiative, introducing cloud services into the backup process requires significant planning and foresight as well as insight from others who have already gone down this road to avoid making their mistakes.

These best practices should help organizations successfully navigate the introduction of cloud services into their backup processes. By talking to individuals, consultants, and practitioners who have already paved a highway to the cloud and combining that with DCIG’s own high level of visibility into the multiple products on the market that support cloud services, organizations can have a high degree of confidence that they can implement these best practices to reap the best of what cloud services offer now without experiencing some of its undesirable drawbacks.

Implementing cloud services as part of one’s overall backup and recovery strategy is an achievable objective for almost any size organization. The key to success is less about finding products that support cloud initiatives as these products already exist and are in production. Rather, the key to successfully implementing cloud services as part of one’s backup process largely hinges upon organizations following these existing best practices so they may confidently move from viewing cloud services as a strategic initiative to actually adopting and implementing them.