Introduction

An exciting transformation in IT is the mainstreaming of “cloud” into almost every aspect of enterprise computing. Business value is the big reason behind such widespread adoption—the cloud greatly improves the economics of IT service delivery. In fact, IT executives tell ESG that increasing their use of cloud services as an alternative to in-house applications and infrastructure is a major cost-reduction driver for them this year, with 72% expecting their cloud investments to increase in 2016.\(^1\) Basically, they’re spending some money to save far more money.

One way to embrace the cloud is to leverage infrastructure-as-a-service (IaaS), which enables IT staff to perform common management/provisioning tasks using hosted services instead of traditional in-house hardware and software. ESG delved into how organizations leverage IaaS (see Figure 1) and found that data protection efforts (specifically backup, archiving, and disaster recovery) sit at the top of the IaaS use-case list right now.\(^2\)

---


\(^2\) ibid.
Considerations and Decisions Involved in Using Cloud Services for Data Protection

Data protection and the cloud intersect in many ways—not only as IaaS, but also as:

- **Backup-as-a-service**—BaaS is a third-party service that includes software to back up data into a cloud-based repository, which a subscriber typically pays for based on capacity needed/used. A BaaS architecture may incorporate an on-premises caching appliance or other onsite storage device for fast recovery, but its main purpose is to ensure backup data is stored safely in an internet-accessible facility.

- **Disaster recovery-as-a-service**—DRaaS enables virtualized servers and services to resume operating in a hosted cloud service center instead of from within the subscriber’s self-managed data center. Like BaaS, a DRaaS architecture might utilize onsite technologies, especially to support failover or network extension. An appliance provides the necessary orchestrated, cloud-based compute, storage, and networking.

- **Storage-as-a-service (for data protection)**—STaaS/dp leverages cloud-based storage as a tertiary repository supplementing onsite data protection technologies. It ensures that traditional backups and recoveries can occur onsite before data is replicated to the cloud for long-term retention and offsite protection.

With all the permutations, the presumption is that data eventually will reside in the cloud—with or without an intermediary on-premises capability to restore it. Such an arrangement, which is nascent but growing in popularity (see Table 1 for an example), is called “disk to cloud” (D2C) or “disk to disk to cloud” (D2D2C).

**TABLE 1. Using Cloud Services as Part of a Data Backup Process Is Increasing**

<table>
<thead>
<tr>
<th>Thinking about your organization’s environment today, which of the following best describes how the data backup process is generally managed?</th>
<th>2012 (N=330)</th>
<th>2015 (N=375)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is initially backed up to onsite disk storage, and a copy is then sent to a cloud storage service provider (D2D2C)</td>
<td>5%</td>
<td>16%</td>
</tr>
<tr>
<td>Data is backed up over the WAN to a cloud service provider (no onsite storage of backup data (D2C)</td>
<td>2%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Source: Enterprise Strategy Group, 2016

Both D2C and D2D2C have merit. The main deciding factor between them usually centers on the organization’s recovery requirements. Specifically, the service level agreements (SLAs) that IT departments maintain with internal business units and external customers sometimes demand a faster, more agile recovery capability than what a pure-cloud (i.e., D2C) arrangement can typically accommodate. The need to recover data locally tends to drive the “D2C versus D2D2C” decision.

Another decision is in play as well: This one involves choosing between BaaS and STaaS/dp. The choice boils down to whether an organization is ready to **abandon** its existing backup solution for a new BaaS offering, or whether it would rather **augment** its backup solution with cloud storage.

For many organizations, adding cloud-based storage to the existing data protection solution (i.e., choosing STaaS/dp) is preferable. Consider the following scenarios:

---


4 ESG believes equally compelling scenarios exist for BaaS and DRaaS, but they are outside the thematic scope of this D2D2C cloud-augmentation-centric piece.
• For organizations whose data protection software mechanisms perform satisfactorily, augmenting them with cloud storage is the least invasive option, requiring no new agent deployment, schedule definition, or training. Basically, if your existing approach works, don’t throw the baby out with the bathwater. Just add cloud storage to the tub.

• For organizations whose admins are accustomed to “doing their own thing” (e.g., DBAs use Oracle RMAN, vAdmins use Veeam to protect VMs, etc.), a centralized BaaS solution won’t hold as much appeal as a cloud gateway will. In this case, no matter which backup software methods are already in use, you just link them through a gateway to a common storage platform and avoid taking away people’s favorite backup tools. The cloud simply serves as “a behind-the-scenes detail.”

**Leveraging Cloud Storage for Data Protection—Pros and Cons**

Even narrowing in on the D2D2C best practice—in which an on-premises backup solution provides local recovery ability prior to replicating the data to the cloud—implementations vary (see Figure 2).

**FIGURE 2. Three Ways to Accomplish D2D2C**

Each configuration has positive and negative considerations.

**Considerations Regarding Cloud Connectors in Backup Software to Cloud Storage**

Just as onsite backup solutions know how to write to disk and tape, many backup software vendors have incorporated API-based cloud connectors into their offerings to support offsite protection.

• The pros—Cloud-connected software gives backup administrators precise control over which data is to be stored in the cloud. It’s the same flexibility they enjoy when they decide which data will go to disk versus tape for protection. For example, the backup software can be instructed to determine that all versions of 12 months of data are to be stored on deduplicated disk, whereas 10 years of monthly data should be stored on tape, and the last 90 days should be stored in the cloud.

• The cons—Not all data protection software has a native cloud connector optimized to overcome the security and bandwidth constraints common to modern enterprises.
Considerations Regarding Storage Replication Between On-premises and Cloud-based Storage Appliances

A number of protection-storage deduplication appliances now offer virtualized versions, and ESG expects these virtualized dedupe appliances to become increasingly common in cloud-hosted environments. By taking advantage of the same reliable appliance-to-appliance replication technology that physical solutions use, IT teams can replicate from on-premises dedupe appliances to cloud-hosted dedupe appliances with ease.

- **The pros**—One or all of an organization’s data protection tools can just write data to the local appliance while remaining blissfully unaware of what is happening behind the scenes in regard to deduplication and cloud vaulting. Another pro is that deploying this architecture is basically as simple as changing or preferably adding a disk target system to the existing data protection software.

- **The cons**—Although some MSPs offer massive-scale dedupe appliances with multi-tenancy, not all storage solutions have adequate cloud-service-provider-hostable or virtualizable offerings.

For the service providers that do offer robust virtualized and/or multi-tenant appliances, their subscribers who have already installed disk solutions onsite will find replication to be an unobtrusive way to consume cloud storage.

Considerations Regarding Gateways with Local Protection Storage Transparently Utilizing Cloud Storage

Perhaps the most exciting but still-emerging cloud storage option is the cloud gateway.

In many ways, it evokes the appealing functionality and value of yesteryear’s deduplication appliances. Where those appliances provided an all-onsite solution that could be replicated to another appliance elsewhere, a cloud gateway boasts on-premises deduplication abilities while being able to send data directly to cloud storage without needing a hosted or multi-tenant remote target.

Cloud gateways provide the best of both worlds. Specifically, cloud storage can be natively accessed via APIs through a centralized storage platform, and no data protection software products will need to be updated or cast aside.

- **The pros**—One could argue that the most efficient way to “add cloud” to an existing backup architecture is via a cloud gateway. It will appear as a simple NFS or CIFS repository that transparently and efficiently leverages cloud storage. After deploying the solution, one simply points the backup software to the new disk system.

- **The cons**—Cloud gateways represent an additional capital cost, and appreciable differences exist in deduplication efficiency, WAN optimization capability, and the diversity of public clouds supported. An argument can be made for choosing a best-of-breed deduplication solution that extends to the cloud versus settling for a mediocre gateway.

Considering Google Cloud Storage as Part of a Data Protection Strategy

One name that is synonymous with “technology disruption” is Google. After entrenching itself as the undisputed leader of what had once been a crowded search-engine marketplace, Google broadly diversified, including launching cloud-based applications in 2006. By 2015, Google was offering Nearline (see Figure 3) as a featured element of its Google Cloud Storage platform.
It should come as no surprise that when Google announced it could deliver economical and durable cloud storage, backup software and appliance vendors Veritas, Commvault, EMC, NetApp, Actifio, and Unitrends were eager to offer support for the new tertiary storage platform.

**This Cloud Is Far More than ‘Another Container’**

Although many organizations continue to assess the value of cloud storage by first investigating its cost in dollars per gigabyte, the advantage of cloud-based storage in a data protection architecture isn’t equated only to bottom-line dollars. Organizations should also look at the agility that they may gain by moving beyond an all-onsite disk-based solution toward having on-demand data in a durable, hyperscale offsite repository.

So, besides protecting it, what agile tasks can one do with data in the cloud? Here are three examples. Consider how:

- The “warm” offsite copy of the data might be used as part of a disaster recovery preparedness plan.
- The extra or alternative copy of the data might aid test/dev efforts or analytics-related activities by combining cloud storage with cloud compute. In other words, it may be possible to unlock additional business value from that otherwise dormant data.
- The data may be used to accelerate end-users’ productivity by opening up their access to it through interesting complementary technologies such as the [Google Cloud Vision API](https://cloud.google.com/vision).

**Additional Differentiator**

Organizations assessing cloud storage solutions for data protection should note that Google boasts an interesting differentiator: Whenever possible, Google keeps data on its network, which is purportedly one of the most secure and sophisticated in the world. The setup improves network performance reliability and reduces latency, which supports shorter recovery point objectives (RPOs) and thus ultimately enhances IT agility levels.
The Bigger Truth

Recognizing that it is essentially no longer possible to have a serious IT transformation conversation that doesn’t include “the cloud” in some manner, the real questions become:

- How can I use the cloud in an evolutionary, if not revolutionary, way?
- To which cloud should I entrust my data?

In regard to the first question, many organizations are discovering that cloud-based tertiary storage, especially in support of a holistic data protection or DR preparedness initiative, is one of the most effective ways one can advance a data protection strategy and overall IT vision. It should come as no surprise that cloud-based storage is frequently being sought out these days as a tertiary repository beyond on-prem disk-based recovery, often at the expense of tape.

And in light of the second question, the most important considerations in play are cost efficiency, data durability, and the assurance that the chosen provider will have your data when you need it. One of the most notable IT disruptors of our generation—from its variety of services to its worldwide pervasiveness—is Google, and its Nearline offering aligns with many of the coveted trends in cloud-based data protection storage.