

## WHITE PAPER

---

# Backup as a Service: Approaches and Advantages for Service Providers and End Users

Sponsored by: EMC

---

Laura DuBois

March 2013

## EXECUTIVE SUMMARY

For over a decade, enterprise backup has been challenging. Troubleshooting failures, optimizing performance, managing media, and meeting backup windows result in operational overhead that offers no competitive advantage. While today's purpose-built backup appliances (PBBAs) are delivering enterprises significantly better service-level agreements (SLAs) over tape backup, sometimes the financial models just don't work. Cloud services, and the economics they embody, can fundamentally change the cost model around backup. Increasingly, firms are leveraging a range of new backup-as-a-service offerings as a means of outsourcing nonstrategic tasks as well as improving disaster recovery and service levels for certain workloads. Backup as a service also enables customers to reduce costs across datacenters while continuing to meet long-term retention requirements.

However, customers should not construe backup as a service as synonymous with online backup. Backup as a service provides a set of capabilities that pure software-as-a-service (SaaS) or public cloud-based online backup services do not. Specifically:

- ☒ Consultative approach to determine customer needs and an ideal solution
- ☒ Customized service contracts and products based on requirements and SLAs
- ☒ Managed services capabilities including configuration and implementation, monitoring, reporting, assistance with first backup and/or full recovery, and initiation of restores

EMC has a complete portfolio of backup products, including Avamar, NetWorker, Data Domain, and Data Protection Advisor, which service providers can use to optimize backup-as-a-service offerings for their customers. This paper outlines three fundamental approaches that customers can employ to effectively take advantage of backup as a service. It discusses EMC's robust backup and recovery solutions and how they are enabling hosted, replicated, and remote backup-as-a-service solutions today. To gain a better understanding of "real world" implementations, IDC spoke with several service providers using EMC products as essential elements of their backup-as-a-service deployments. Their insights are included in this document.

## SITUATION OVERVIEW

According to recent IDC cloud research, customers find that the benefits of cloud services stem from cost savings, faster time to market, user self-service, standardization of IT services, elastic scaling, and metering/chargeback. Some firms seek to use public clouds, leveraging third-party infrastructure and resources, often outsourcing their entire IT operation to a service provider that hosts and manages the infrastructure.

Over the past several years, and especially in certain industries, a barrier to public cloud adoption has often been the potential compromise to control and, in some cases, security. Other "pain points" can also be cited — for example, lack of interoperability standards, lack of data/system portability, and lack of migration choice/flexibility can be top concerns with the deployment of public clouds.

On the other hand, larger or more compliance-driven firms evaluate and deploy private clouds within their own datacenters or hybrid clouds with some workloads outsourced to an external cloud provider. Firms that are considering private clouds mention security, service reliability, and internal IT staff knowledge as concerns. Choosing both the right technology *and* the right service provider for hosted and/or managed private clouds can address these concerns to a great extent. According to IDC research, datacenters are considering and/or deploying private, hybrid, or hosted clouds for three leading reasons: to standardize IT processes, to lower costs by sharing resources, and to develop consumption-based pricing for chargeback.

The application of cloud services has been at the top of IT corporate strategy lists for the past 24 months. However, the concepts of remote vaulting and managed backup services have existed for over 10 years. Two factors have changed the market dynamics: One, existing service providers of these backup services are replacing their legacy backup infrastructure with new technologies. Two, as challenges with legacy customer datacenters and equipment refreshes have arisen, the interest in cloud approaches has increased. As a result, more service providers are trying to develop a cloud-based backup service approach. But for customers, this means that choosing the right approach has become a complex series of decisions. The following sections discuss the approaches and the advantages of each cloud methodology: public cloud, private cloud, and hybrid cloud.

---

### Public Cloud Backup

Public cloud backup means that a customer's on-premise data is copied and sent over a dedicated or shared (Internet) network to a service provider's datacenter. While a local agent resides on the system being protected, the backup infrastructure is hosted and managed by the service provider. Monthly fees for the "backup service" can be based on capacity, bandwidth, number of clients, etc. Public cloud backup services are ubiquitous and leverage shared (as opposed to dedicated) back-end infrastructure, also known as multitenancy. With public cloud backup, the economic savings can be attractive, although pricing and options vary. Customers should understand what a quoted price includes. High-volume services offered at the lowest \$/GB price will *not* result in high-touch customer service and monitoring against SLAs (and in many cases will not even include SLAs).

There is a big distinction between using public cloud backup for a few client PCs and/or a handful of servers for a small business and standardizing on the public cloud for thousands of PCs and many critical application servers. Firms of any material scale in terms of users, capacity, or server footprint must examine data volume, recovery time needs, and network capacity to properly determine which cloud model may work best. Companies considering public cloud backup must perform an additional analysis of WAN connectivity and bandwidth (GB versus MB).

Public cloud backup is often thought of as a panacea, but internal IT may still be accountable for many backup tasks, including monitoring backups and SLAs, managing capacity costs/thresholds, and initiating restores. Some customers employ "advanced" public cloud backup services offered by service providers as a way to further eliminate the company resource burden. Here the service provider takes on assistance with restores, monitoring SLAs and presenting reports, managing back-end capacity and notifying clients when they reach thresholds, and troubleshooting backup failures. Other distinguishing factors include virtual machine integration, inclusion of hybrid cloud options, broader application support, multiple datacenters, and more enterprise-grade service options.

---

## **Private Cloud Backup**

At a fundamental level, private cloud backup offers the same features as public cloud backup; the key difference is that private cloud backup is associated with only one firm, also known as single tenancy. A private cloud can be implemented and maintained either internally (the traditional way) by the customer or externally by a service provider, which can actually employ measures (e.g., firewalls) to emulate a fully internally owned and operated private cloud — the latter is a newer on-premise model that customers should consider as a viable option for cloud backup. A private cloud can be managed internally by the company's own IT staff or by a third party under a managed services agreement. Additionally, a private cloud can be run internally within the company's datacenter or hosted externally within a service provider's datacenter.

Having a dedicated private cloud, regardless of whether it is managed by a service provider or not, means that backup services can be made available to different organizations or divisions within the same company, leveraging a shared backup infrastructure. This approach is akin to traditional datacenter backup in that it still requires backup skills and infrastructure. It is different in that a normalized and standardized set of backup options are offered across all environments in the company rather than specialized backup solutions being deployed for each department, application, or compute family group. A private cloud backup service will offer economies of scale over traditional on-premise backup because of the shared and standardized infrastructure. However, unlike public cloud backup, private cloud backup allows the customer to retain control of its backups, its data, and its infrastructure.

With a wholly owned and operated internal private cloud, as opposed to a private cloud with some degree of service provider involvement, internal customer IT resources are still accountable for backups. The internal team still needs to do discovery and configuration, monitoring for success and failures, troubleshooting performance issues, measuring success against published SLAs, managing the

hardware, initiating restores and recoveries, and potentially issuing chargeback or show back reports to internal customers. While the internal operating and labor costs may be reduced because of the backup platform standardization, the capital costs and associated depreciation for the infrastructure remain on the firm's books.

An alternative that customers should consider to further reduce capital expenditures is utilizing a third-party service provider to manage and/or host their private cloud, whether virtually or physically. An external private cloud that is managed by a third-party service provider can completely offload internal IT from backup responsibilities and management. However, similar to public cloud offerings, private cloud offerings require firms to choose the right supplier based on service options (SLAs based on application tiers), SLAs (availability, reliability, performance, backup success rates, and recovery), internal application skills, contract terms, and more. When a hosted private cloud model is leveraged, backup costs can be shifted from a capex model to an opex model, while assets and resources remain onsite and under dedicated control.

The term "hybrid cloud services" (which cloud backup falls under) is used to describe the consolidated coordination/management of multiple cloud services and can be thought of as a merger of the public and private cloud backup solutions previously discussed. As such, some assets are privately held and controlled onsite by the customer, and some have a degree of service provider management or control. In this paper, IDC does not classify collaborations between cloud services and non-as-a-service IT as hybrid cloud services. The term hybrid cloud is used very loosely in the industry to include two use cases:

- ☒ **Capacity and/or processing overflow.** During seasonal, end-of-quarter, or other high spike periods, processing requirements may dictate the use of hybrid cloud services. Individual applications are deployed not only within a private cloud but also concurrently within a public cloud service for some period of time (aka cloud bursting). This class of hybrid cloud is explained here for completeness, but it is not discussed further because it is outside the scope of this paper.
- ☒ **Backup/recovery as a service.** Some public cloud backup services offer a hybrid option, which refers to the ability to create a local backup as well as a replicated, remote cloud-based backup. This local backup is created for faster recovery, when needed. Another form of hybrid cloud service could be considered recovery as a service, where applications are running on-premise in a private cloud, but in the event of a local failure, workloads effectively fail over to cloud infrastructure. (For more details on recovery as a service, refer to *Recovery as a Service: Leveraging the Cloud for Continued Operations*, IDC #236426.) EMC often sees end customers employing this model in conjunction with a qualified service provider.

In all three cloud backup models, customers can utilize service providers to fit their needs.

## **BACKUP CHALLENGES TODAY AND THE ROLE OF BACKUP AS A SERVICE**

The datacenter backup challenges that exist today are the same challenges that have existed for over 10 years. Issues have stemmed from data and infrastructure growth, an increase in distributed data the growing infeasibility of meeting backup windows, and the need to restore and recover data faster. While virtualization technology has helped solve this problem, the fundamental issue is the reality that firms sometimes view backup as inefficient to maintain internally. As a result, companies are increasingly looking to outsource their backups to a service provider when it makes sense, thus moving from a capex backup cost model to more of an opex backup cost model. One solution is not necessarily more expensive than the other; in the end, the decision comes down to a variety of other factors as well, as mentioned in the previous section.

An EMC service provider offering cloud backup put it this way: "One of the reasons why we've taken on backup as a service is it's purely a commodity-based service for organizations. It's mission critical in the fact that you need the data, but nobody wants to deal with the complexity."

In the datacenter today, both legacy and new applications and systems must be protected. Together, databases of multi-TB size and virtual machine proliferation have placed greater stress on the backup infrastructure and backup window. Critical applications cannot afford material downtime, so they often leverage snapshots and/or replication services for recovery, and at the same time, high-speed disk is increasingly used over tape for operational and disaster recovery of these applications. Although tape may still play a role in the datacenter, it's typically for lower-tier applications and/or longer-term archive. In effect, datacenters have a set of protection and recovery options available depending on the recovery time objective (RTO) and recovery point objective (RPO) requirements for a given workload. As such, service providers that once leveraged only tape are now also utilizing a new breed of data protection approaches to offer backup as a service for both physical and virtual infrastructure.

Recovery scenarios for centralized/distributed data, unstructured/structured data, and physical/virtual systems vary. Operational recovery for user error, application corruption, security events, and hardware failures must be satisfied, while disaster recovery is used in the event of catastrophic events. Traditional backup can struggle to meet a range of conditions around how data is recovered. Consider a series of virtual machines running on a physical host. This configuration will probably require item/object recovery, image recovery, application-consistent recovery, host-level recovery, and remote disaster recovery, and each recovery scenario may have different RTOs and RPOs. Depending on availability and RTOs for given workloads, customers can use service provider offerings that include a range of backup services.

In the same way that traditional backup approaches have been augmented with newer protection schemas such as continuous data protection, synchronous and asynchronous replication, and snapshots, tape as a backup target has been augmented — or in many cases, replaced — with disk. Many challenges are associated with tape, including risk from loss/compromise, media management requirements, capacity performance tuning and optimization, and slower restore/recovery times. Because of the risk of compromise, tape is being replaced with an electronic means of replicating backup data to secure remote sites. Additionally, the use of deduplication for backup and recovery has improved the economic viability of using disk for backup, enabling more firms to perform more operational recovery as well as disaster recovery from disk rather than tape.

## **BACKUP-AS-A-SERVICE APPROACHES AND EMC OFFERINGS TO ENABLE THEM**

EMC customers enjoy the benefits of the company's leading performance in the backup and recovery markets via Avamar, NetWorker, and Data Domain as well as in backup management via Data Protection Advisor. At the same time, investment in cloud services continues to be a top priority for CIOs. In IDC's *2013 CIO Sentiment Survey*, CIOs rated cloud computing as the most important IT initiative for 2013. Customers are often turning to service providers that offer cloud-based backup and recovery solutions that can continue to meet SLAs while potentially reducing cost. Fortunately, EMC backup and recovery offerings work effectively, whether they are employed directly or through a service provider.

The following sections identify three possible backup-as-a-service deployment models and discuss how EMC backup and recovery solutions can help firms achieve their technical and business objectives for cloud-based backup. When customers evaluate backup services that are available today, they should pay careful attention to the service provider's level of support, which can translate into a range of capabilities such as predeployment planning, on-premise installation/configuration, initial backup policy setting, backup monitoring against SLAs, reporting and dashboard provisioning, and assistance with restores and recoveries.

---

### **Hosted Backup Services**

Customers often use service providers that host application infrastructure because they do not want to take on the responsibility for backups. Furthermore, a service provider can offer an economic benefit by leveraging a shared backup infrastructure and assist the movement from a capex model to an opex model, as previously mentioned. Customers can utilize a leveraged IT infrastructure that is shared across multiple service provider tenants or a dedicated (per tenant) backup infrastructure.

---

## **Replicated Backup Services**

Increasingly, firms are looking for more cost-effective, secure, and efficient ways to ensure that remote backups are created and viable rather than relying on manual tape collection and ensuring backups are sent offsite. In many firms, half the backup infrastructure supports local backups and the other half supports replicated backups from a secondary datacenter. In essence, each datacenter backs up the backups of the other datacenters. This can be a costly configuration that relies on the existence of a secondary datacenter. With datacenter consolidation and the increased evaluation of cloud services, customers can employ service providers for replicated backup services as a more effective means for disaster recovery.

With this type of service, customers can successfully perform backups at their physical/local site without owning, managing, or incurring the expense of a remote site for disaster recovery purposes. For such customers, a service provider offers its tenants replicated backup services on a leveraged (i.e., shared) storage grid or a backup platform. More recently, there has been interest in recovery-as-a-service options, which are also offered by service providers using EMC technology.

---

## **Remote Backup Services**

Firms looking for public cloud backup often employ service providers because of the economic benefits and advantages of offloading IT as well as consolidating underlying infrastructure. Often service providers offer dedicated network links to help customers perform direct backups over the WAN to their hosted backup IT environment. Initially, customers can target remote backup services for either PC or branch office data or offer a service for lower-tier applications or endpoint data and then expand from there. This use case typically encompasses lower-tier data and systems that reside outside the central datacenter. Additionally, backups are streamed from multiple tenants and tenant systems to a centralized backup cloud deployment in the service provider datacenter.

According to IDC research, the customer use cases for remote backup services today are PC backup, backup for smaller applications and collaboration systems, and backup for remote and branch office environments. Data volume and network bandwidth constraints are obstacles for firms moving more applications to a pure remote cloud backup service. Hybrid configurations, where a local backup device or cloud gateway serves as the initial backup target and restore/recovery location, can help. Here backups are sent to the service provider location using asynchronous replication, although the network bandwidth must be able to support the successful completion of the replication process before the next iteration starts. Additionally, efficiency technologies such as compression and deduplication that are used in EMC solutions can reduce the amount of data that must be transmitted over the WAN.

One service provider interviewed affirmed this integrated approach: "We have two backup services ... with a backup-to-disk destination on EMC Data Domain. We also offer auto-managed remote backup services through public networks using EMC Avamar."

## Choosing Which Service Is Right and How EMC Solutions Can Help

Choosing the right service requires addressing a complex series of decisions. Table 1 provides a comparison of the three backup-as-a-service approaches in terms of infrastructure requirements, networking and workload considerations, and recovery SLA expectations. Service providers can select which services and EMC products they should offer based on customer requirements.

**TABLE 1**

Comparison of Backup-as-a-Service Approaches

	Hosted Backup Services	Replicated Backup Services	Remote Backup Services
Connectivity — production servers/storage to initial backup	LAN	LAN for backup WAN for replication	WAN
Network communications; potential bandwidth and latency issues	Low	Low — backup Moderate — replication	Moderate
Secure VPN/networking requirements	No	Yes	Yes
Suitable workloads constraints	None	Capacity, change rate, and bandwidth	Capacity, change rate, and bandwidth
Supported backup applications	NetWorker, Avamar, other third party	Avamar, other third party	Avamar
EMC backup targets	Avamar, Data Domain	Avamar, Data Domain	Avamar
Recovery time expectations	Fastest option, all recovery is local	Operational recovery is local and fast, but disaster recovery is remote*	Recovery is remote — slowest option*

\*Recovery times will vary based on data volumes, network bandwidth, backup approach, use of recovery device, etc.

Source: IDC, 2013

### **EMC Avamar**

Avamar is unique in that the product can serve as both a backup application and a disk-based backup target. It also provides remote replication for backup data. Avamar has tight integration with virtual machines, includes multitenancy features, and performs client-side deduplication. Avamar's sophisticated role-based access controls, encryption, deduplication, and low TCO software and system make the product ideal as the infrastructure for remote backup services. The concept of an Avamar domain logically separates user data within the system. Granular controls for domain administrators, operators, and users are available. Whole domain is a logical

demarcation that some customers require to physically separate their data, in which case a separate and dedicated Avamar system can be configured. If customers are already using Avamar for local on-premise backups and want a replicated backup service, Avamar can fit here too. For hosted backup services, depending on the data volume and types, Avamar can additionally fit into these environments.

### ***EMC Data Domain***

Data Domain deduplication storage systems reduce backup storage requirements by 10x to 30x with high-speed, inline deduplication. As purpose-built backup appliances, Data Domain systems offer a variety of features such as support for standard interfaces (e.g., NFS, CIFS, NDMP), Data Domain Boost for advanced application integration, encryption, network-efficient replication, and built-in data integrity protection for ensured recovery through the EMC Data Domain Data Invulnerability Architecture. Data Domain systems easily integrate into a variety of environments with support for leading backup and archive applications. In addition, these systems replicate only unique, compressed data across the network, which requires a fraction of the time, cost, and bandwidth compared with traditional replication methods. As a result, Data Domain systems are optimal for both hosted backup services and replicated backup services.

### ***EMC NetWorker***

NetWorker is EMC's enterprise data protection platform, which supports a broad range of clients and applications in both physical and virtual environments. With NetWorker, administrators can logically zone data, devices, and users in shared backup environments. Customer data can be securely separated or partitioned while being managed from a centralized server. Restricted data zones enforce a separation of tenants from all the other tenants in the system. Users within a data zone can be given limited credentials to be able to only ad hoc backups or request a recovery, which are functions that align nicely with tasks that tenants might want to perform. A single NetWorker backup server can be shared among tenants or dedicated to a specific tenant. NetWorker supports tight integration with both dedicated and shared Data Domain systems and Avamar grids as the back-end storage targets for backup data, along with the ability to leverage several other tape and disk backup systems. As a result of NetWorker's broad set of features, including support for tape, NetWorker is ideal for use in hosted backup services.

### ***EMC Data Protection Advisor***

Data Protection Advisor offers effective data protection management, enabling the visibility and insight to deliver data protection as a service, whether to other business units within a large company or external clientele. It is scalable to support functionality for organizations of all sizes that are looking for new economic and service models for their data protection needs. The automated collection, analysis, and monitoring of backup environment statistics enable more effective SLA management by service providers. Data Protection Advisor also has an analysis engine that can proactively identify potential problems on a critical client or for threshold exceptions, helping address events before they become larger problems. The product works across backup, replicated backup, and primary replicated environments, as well as across different backup applications.

## CONSIDERATIONS IN CHOOSING THE RIGHT SERVICE PROVIDER AND SERVICE

Customers need to evaluate the service provider and its contracts when considering a backup-as-a-service option. IDC recommends that customers consider the following factors, which are common issues:

- Seeding.** How will initial backup data migrate to the service provider infrastructure? Is the migration included in the agreement, or is it an extra cost?
- Full recovery.** In the case of a disaster recovery event, can the service provider ship the device to the recovery site and provide full recovery from a local site versus trying to perform recovery over a network?
- Local data protection laws or cross-border compliance.** Does the service provider assure its customers that local data protection laws are satisfied? Are local laws mixed with international laws, and if so, how does the service provider address this? Are data privacy requirements in leading geographic locations ensured, and does a second site exist in a compliant region or locale?
- Metered billing.** Can the service provider support fully automated metering features integrated into the accounting system to produce an invoice on usage?
- Backup reporting.** What level of backup reporting is available and with what level of detail? Do reports include data on backups that never started, backups that failed, and completed backups; time of day; start/stop; and how much data was transferred? Does the service provider notify customers about missed and failed backups?
- Segregation.** Can the service provider segregate data and systems based on requirements such as physical systems, different privileges, and roles?
- Dedicated.** Some companies are not comfortable with their data being comingled on the same infrastructure as that of another customer. They may have backup window or retention requirements that necessitate a separate infrastructure. Can the service provider support this?
- Bandwidth.** Does the service provider have dedicated bandwidth out to a customer at LAN-like speeds while backing up offsite?
- SLAs.** What are the service provider's standard and published SLAs? Are these SLAs embedded in the contracts or found elsewhere and thus able to be changed? What are the service provider penalties for SLA violations? How are SLAs monitored and measured? What are the financial consequences in the event that the service provider does not meet the following SLAs?
  - Backup success rates
  - Backup service and infrastructure availability/uptime
  - Backup job completion

- Recoverability of the backup data
  - Recovery times
  - Retention times/periods
  - Response times based on severity
- Restore accountability.** Who is responsible for restores? Is it a self-service restore model, or will the service provider do the restore? If self-service, is there a 24 x 7 NOC so that a customer can always call to pull that data back?
  - Data protection guarantees.** Can the service provider offer a data protection guarantee/SLA? How are events, such as a backup failing twice in a row without a success in between those failures, handled?
  - Staff qualifications.** Does the service provider have staff trained on backup troubleshooting, performance optimization, and SLA monitoring? What applications and systems are they trained on? Are they certified on a given technology?
  - SLA tiering.** Does the service provider have different offerings based on different SLA requirements? For example, a tier 1 application may have one set of RTO and RPO requirements that is more expensive, but this does not mean that the same service/service fees should exist across all workloads.
  - Pricing.** How does the service provider's pricing compare with the cost of alternatives and on-premise strategies? A cloud service may be the least expensive, but it may not meet a customer's SLAs or other requirements. An enterprise-grade cloud service may be more than an on-premise investment, but it also could eliminate onsite IT costs to allow redeployment of personnel to more strategic projects.

## CHALLENGES

Clearly, cloud computing and IT as a service are changing the landscape for how firms procure and manage IT. However, the reality is that firms are still analyzing how best to leverage cloud services in a private, public, or hybrid fashion and considering what workloads are best suited to each of these deployment models. Firms that have reached an extensive level of virtualization are looking to take the next step in deploying a private cloud. However, this requires integrating different infrastructure components and streamlining, through integrated workflows, the provisioning of new applications and infrastructure. Today, many of these workflows remain standalone and separate. Other firms are considering hosted application services or infrastructure-as-a-service (IaaS) offerings for specific workloads and continue to look to public cloud for collaboration, messaging, and information-sharing use cases.

However, firms are embracing a range of cloud offerings for backup and disaster recovery. Firms see backup and disaster recovery as workloads that can and should be outsourced. They offer little competitive advantage and consume disproportionate amounts of internal IT staff time to manage on an ongoing basis and too much of the

capital budget. For these reasons, coupled with compressing recovery time objectives and tighter SLAs, these secondary storage workloads are optimal candidates for backup as a service.

## **FUTURE OUTLOOK**

Once the customer has taken that first step to leverage the cloud, there's plenty of opportunity to build upon the virtual infrastructure to improve IT agility and performance. Service provider cloud-based infrastructure allows customers to easily leverage online storage, long-term archiving, and data mining. Customers should be looking ahead to see how backup data can be mined and leveraged for business analytics. Many service providers are already looking into how analytics can be offered to customers. Once backup data is secured, the cloud can be used as a location not only for backup but also for recovery.

Increasingly, today's service providers and end users are looking to extend the range of services to include recovery as a service, longer-term archive, and data analytics. These present strategic opportunities for service providers to help businesses transition from using cloud services for an operational advantage to using cloud services as a revenue-generating mechanism. Firms and service providers alike can leverage EMC's range of competencies across analytics, backup and recovery, storage, cloud computing, and big data technologies as they build clouds of the future.

---

### **Copyright Notice**

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2013 IDC. Reproduction without written permission is completely forbidden.