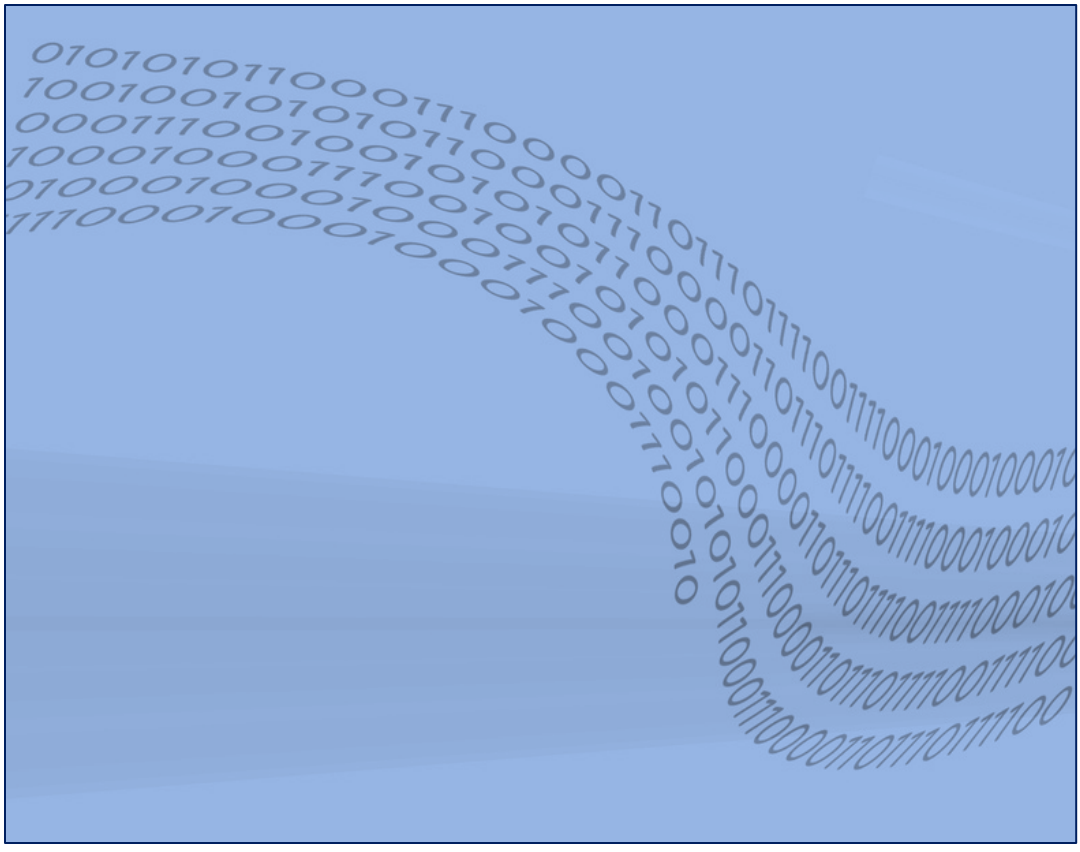
The Move to the Cloud

A Taneja Group Knowledge Brief



87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

Table of Contents

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

Authors

Arun Taneja, Founder & Consulting Analyst

Senior Analyst & Consultant

Jeff Byrne, Senior Analyst and Consultant

Senior Analyst & Consultant

Jeff Boles, Sr. Analyst / Dir. Validation Services

Christine Taylor, Analyst & Communications Manager

Introduction

Key Questions

Cloud storage can be a beast to wrangle. Deciding which applications to move into the cloud, understanding how to select and deal with a cloud storage provider, deciding on cloud storage solutions – none of these are easy. Is it worth it?

Cloud storage and application hosting can certainly be worth the investment but it's vital that you do into it with your eyes wide open. There are many, many questions you need to ask before entrusting your data to a cloud storage provider. Questions like:

#1. **What is your system uptime?** 99% is the minimum to look for, although that 1% down time can be a big hit on the business. Historical archives are one thing but when you host primary data such as web services, downtime annoys customers and can deeply impact the bottom line. Hardware uptime is also different from data accessibility: your cloud vendor's system can be technically up but your data may be unavailable due to a software or connection issue.

#2. **What data availability service levels do you support?** When you use a public cloud provider for backup and archives, then a 99% service availability and daily backup is usually acceptable. But when primary/active data enter the picture, such as web-based business data or analytics applications, the picture changes. Most cloud storage operates on very thin margins. Amazon and Google are not out there making money providing high data availability to cloud customers. This is why, for example, Amazon was happy to host giant pharma Eli Lilly's archives but not so happy to provide SLAs for active data analytics. It simply was not worth it to them to ramp up and provide the necessary level of SLA even for a large account.

#3. How easy is it to move my data to another provider? No cloud storage vendor makes it easy, they're not stupid and they want to retain customers. But you have the right to move your data to a provider that suits you better without getting locked in to an existing provider. We strongly suggest that vendors offer data migration at reasonable prices to facilitate mass data transfers in *and* out. Many customers do not want to go through a transfer even with tools available, so the anti-lock-in message works as a competitive differentiator.

#4. What data protection service levels do you provide? Since data protection and availability are the heart of useable storage, going without acceptable SLAs should be a deal breaker between IT and a cloud provider. Any reputable service provider will at least provide daily backups, but some critical data will require far more frequent backup, snapshots and/or replication. And if the public cloud provider is backing up their customers to removable tape (which they probably are), customers need to review their provider's tape retention, cycling and disposal policy.

#5. What is your level of performance? The answer to this will vary widely according to the type of data, usage, bandwidth, and storage performance. Keeping a copy of backup in the cloud is a very different performance matter than running analytics from the business on cloud-based data. The first usage case is very common and widely supported; the latter is much harder to achieve. The hybrid usage case is storing data in the cloud and letting the cloud provider run high performance applications on it from their own data center. This works well if the cloud provider is a subject matter expert in the usage case, such as eDiscovery services hosting.

#6. What applications can I best host in the cloud? Good application fits include test/dev workloads, personal productivity applications, collaborative and messaging applications, and virtualized applications. Intensive analytics and transactional applications are rarely suitable for running from your premises on your cloud-based data.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

The upshot is that corporate users should negotiate with their cloud vendors for workable service level agreements. Do not accept an oral agreement or a simple 99.9% uptime report. Insist on the service level agreements that you require. Do be prepared to pay more for higher level agreements, justifying your increased investment with the efficiency and scalability of cloud-based data and application hosting.

Cloud-Enabled DR for SME: Ready for Prime Time?

By Jeff Boles and Jeff Byrne (original version published in InfoStor)

Have you experienced an unplanned outage lately?
Based on our conversations with IT administrators,
there's at least a 25% chance you have.

Those that have experienced an outage recently may no longer be in the IT trenches. But if you've survived a painful outage, there's a good chance you found out the hard way that a good number of the servers under your management have a critical role to play. Even a minor outage that renders one of those critical servers, or even a file, unavailable for a few hours can be extremely costly, if not devastating.

Small and mid-sized enterprises (SMEs) are particularly vulnerable to the consequences of such outages. These companies tend to be in a precarious position: they are large enough that an outage can do major damage in terms of compromised data or lost business, and yet many do not have adequate disaster recovery (DR) plans, tools and infrastructure in place to enable a rapid and effective response.

Until recently, effective DR practices required an IT practice unto itself – full blown duplicate IT infrastructure spread out to two or more different sites, along with all the hands-on systems and storage management that goes with it. The cost and complexity of these traditional approaches discouraged – and in many cases, prevented – SMEs from investing in DR planning and processes. But the advent of virtualization and cloud technologies has changed that picture dramatically.

With the recent emergence of virtualization and cloud storage technologies, it's now possible to do DR in the cloud. Such solutions run the gamut, from simply duplicating data to the cloud as a

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

form of off-site backup storage to creating a conduit between a customer's site and a remote virtualized compute and storage infrastructure in the cloud.

Cloud-based DR

In Taneja Group's view, cloud-based DR has a specific definition. It is more capable than just cloud backup, and it is more efficient than remote co-location of equipment. Cloud-based DR is the use of connectivity to compute and storage resources hosted on remote, elastic, multi-tenancy clouds to enable more cost-effective and flexible protection of data at a distance. In terms of DR, that's a recipe matched to SME needs.

The cloud can shrink the CAPEX required for traditional DR. There's no need to invest in a remote DR facility, and even on-going costs are minimal because the cloud is economically priced and can allocate capacity and performance on demand, enabling the customer to pay only for the resources consumed. Moving DR to the cloud can also increase the flexibility of DR configurations and practices, and since clouds are designed for remote management, it may speed recovery. Compared to cumbersome and expensive tape-based DR practices (both on-premise and off-site) such capabilities can make routine testing practical, and mean a solution really works when it is needed.

Cloud challenges for DR

But great technology aside, the promise of cloud-based DR has not yet been realized for many SMEs. Early adopters have run into a wide variety of complications, including challenges around access, security, ease of use, recovery time and effort, and provider lock-in.

Though one of the primary reasons users move to the cloud is to make their data universally accessible, native storage in the cloud often limits access. Cloud storage tends to be designed for programmatic access – an alphabet soup of HTTP REST and SOAP APIs and acronyms – which

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

means there might not be any out-of-the-box way to access it as regular storage. If you're thinking of building a cloud DR solution on your own, prepare for a lot of custom development.

The cloud also introduces latency and data movement challenges. For DR, you might not be able to ensure the right data is in the right spot at the right time, and you may have a train wreck of inconsistent data on your hands. Native, unassisted access to the cloud may leave data unencrypted in-flight or at rest. Moreover, if you try on your own to work with HTTP data stored as objects for DR – and need to invoke snapshots, backups, and other under-the-covers storage functions you take for granted today – you may never achieve a DR plan you can execute with confidence. Workarounds to these challenges may do little better than hamstring an approach to cloud DR, while a poor approach may require copying complete data sets, perhaps through a backup layer or, worse, through scripts. Aside from limiting breadth of support and/or adding to recovery times, administering such steps via command line or remote console can be daunting. And in the end, you will have achieved little more than cloud backup.

Not to be overlooked, the rush toward emerging opportunities also leaves many cloud-enabled DR solutions in support of only one, or at most two, backend cloud providers. If a solution has limited support, SMEs in effect realize a lock-in "double-whammy" – not only is the effort to move cloud data enormous, but now they need another device to boot, and they have to move data from two clouds across two devices. As a result, a "one-provider-fits-all" approach may wind up being a far from optimal fit.

- **Shaking up the classic DR approach.** Fortunately, innovators are tackling the challenges. They are better coupling workloads and data together, and providing simplified, easy-to-use, low-cost mechanisms to get both into cloud infrastructures, preserve data access, and manage them when there.

Since the challenge is data-centric, the leading innovators are storage vendors tackling the problem with technologies that push data into the remote cloud. By periodically synchronizing

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

data, vendors enable workloads to switch to another site – either another physical location or entirely within the cloud itself. And by taking on the challenges at the storage layer, vendors can leverage what are already significant disaster recovery capabilities unleashed by server virtualization.

- **Flexible integration with both sources and targets, including applications as well as cloud DR providers.** Solutions shouldn't force users to modify or customize their applications or file systems in order to take advantage of cloud-enabled DR. Nor should they entrap users into a single provider in what is today a rich cloud ecosystem of largely compatible compute and storage providers. While a number of vendors are coming at the challenge with backup technologies and agents, other vendors offer iSCSI appliances that store data locally while simultaneously connecting to public and private cloud storage services. Under the covers the latter solutions can act as a conduit for directly storing data in the cloud, even though the storage seems local. Some of these solutions enable simultaneous use of multiple storage providers, and some even allow users to move transparently to a new provider, while data is gradually migrated behind the scenes. Irrespective of architecture, though, the right solutions for connecting DR into the cloud will enable choice and cost competition across cloud providers, with the right tools to enable easy provider migration if that time comes.

- **Ready for the virtual infrastructure.** Key to making DR more affordable for SMEs is using a remote cloud where standby applications can be configured on virtual servers and can easily be turned up and down in a shared-cost, service provider cloud. There are various ways to move data to such virtual clouds so that virtual servers can access it just like physical and virtual servers do in the physical data center. Most approaches either include agents on physical and virtual servers - along with increased management complexity - or deploy virtual appliance versions of a particular storage technology in the virtual infrastructure. Vendors with physical and virtual appliance offerings may enable customers to satisfy their integration requirements in both local and cloud environments, and emulate the remote cloud

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

environment on their own virtual servers. Moreover, a virtual appliance is simple to implement and can improve the effectiveness of DR planning by significantly reducing testing costs and allowing DR scenarios to be validated without disrupting the business.

- **A high level of data security and resilience**. Leading solutions raise the bar on data security and integrity. Data can be highly secured with in-flight and at-rest encryption (AES is the standard today). Data integrity can be elevated by combing through and check-summing data more thoroughly as it is transmitted to the cloud. Furthermore, data stored by many of these cloud DR solutions can exploit the multi-site, automated replication capabilities of multiple cloud providers to protect offsite systems and data better than self-sourced solutions. Finally, the best storage technologies for building cloud DR can provide many flexible paths to securing data to meet a wide range of users' needs. If the public cloud is not an option, look for cloud solutions that support private clouds in any number of configurations – ranging from secure, completely partitioned hosted offerings to complete onsite implementations using low-cost storage such as EMC's Atmos. This combined set of capabilities will give SMEs peace of mind and enable them to more readily satisfy regulatory requirements for their key data and applications.

- **Closely integrated with existing and remote compute and applications.** A DR technology that approaches DR from the storage layer should reach well beyond just the conveyance of data from one location to another. Protecting consistent, known-good data is the primary task, and can be approached with a variety of technologies ranging from backup to snapshot technologies along with agents or providers that tie in at the application level to guarantee that applications are quiesced and data is consistent – Microsoft's VSS framework makes snapshots the de facto standard for integration with many applications. Moreover, since cloud DR will revolve around virtual infrastructures in the cloud, technologies should tie into frameworks for managing virtual compute, such as VMware's vCenter or Microsoft's SCVMM. More importantly, they should integrate with a system or toolset to orchestrate disaster recovery. Not all cloud DR solutions are built the same, and some may require steps

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

that span the better part of an hour, or multiple hours. By tying into virtual infrastructures, efficiently moving data, and employing sophisticated snapshot and synchronization techniques, solutions can deliver cloud DR in a fraction of that time.

- **Various paths to cloud-based DR.** The movement to cloud-based DR is being enabled by a broad range of vendors and approaches. Each vendor applies its own unique technology to move disaster recovery to the cloud instead of the traditional secondary data center. Some vendors will move backup data. Such solutions may require a recovery process in the cloud, but recovery can be "pre-staged" so that applications can be immediately restarted upon disaster.

Emerging cloud gateway vendors, when enabled by primary storage support and the ability to serve up storage from a virtual appliance, can in effect pre-stage data automatically and make it available to virtual servers in the cloud. Not all cloud gateway solutions can do this, but when they can they will further simplify the recovery process.

The many choices for cloud-based DR should allow SME users to more easily select a product that integrates with their existing technology and management practices, and should finally make DR practical for the SME.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

Cloud-Based Management

How to Avoid Vendor Lock-in

By Arun Taneja (original version published in SearchStorage)

Users are very concerned about putting too much of their data -- or more importantly, their most valuable or critical data -- into the cloud.

Put another way, how can users gain cloud storage benefits such as ease of accessibility and reduced costs, without the fear of vendor lock-in?

This issue is on the minds of many end users we speak with, and unfortunately, there's no simple or universal answer today. Fear of vendor lock-in is clearly a legitimate concern. There are no widely adopted standards in place today to ensure that customer data can be freely moved among different cloud storage service providers' sites. And it always seems to be more difficult and more costly to transfer data *out* of a cloud storage repository than it is to upload it in the first place. The more data a user has, the harder it is to move; and if the user wants to move the data to a new cloud provider, they'll probably have to pay for the bandwidth twice. This inhibits opportunistic migration and puts customers in a weak negotiating position relative to their cloud storage service provider.

So what steps can cloud storage customers take to reduce the likelihood of lock-in, and the cost and inconvenience that go along with it? Here are a few guidelines that we recommend users follow as they shop around for a cloud storage provider:

- **Read the fine print of each provider's policies.** If necessary, ask them directly how they facilitate moving customer data out of their cloud storage repository. Given the amount of data you'll be uploading and its expected growth over time, could the data be moved via the Internet back into your data center in a reasonable timeframe or to a different provider's site? As an alternative, if the volume of data is too large for digital transfer, can it be moved via a

portable storage device? What are the process, timeframe and cost required for each of these approaches? Unfortunately, today, pulling data out of most cloud storage solutions requires a brute-force approach that you, the customer, will be responsible for. But it's worth finding out the likely scenarios before you sign up with a provider. You may not like all the answers, but at least you'll have a feel for the magnitude of effort and cost you might incur down the road.

- **Ask the provider whether they offer data migration tools or services to facilitate the movement of large amounts of data.** For example, most providers require that customers moving data between clouds first download data to an intermediate location such as the customer's data center and then re-upload the data to a new cloud. However, a few public cloud vendors provide direct cloud-to-cloud data migration. It's just a start – the movement is inbound, not outbound – but it does save new customers inbound bandwidth charges. Some cloud gateway vendors are also making migration easier by integrating with different cloud storage APIs and then providing a standard file system interface to facilitate data migration between those clouds.

- **Choose providers that have pledged to support emerging industry standards.** We support standards such as the Cloud Data Management Interface (CDMI) standard created by the Storage Networking Industry Association (SNIA). The CDMI provides a standard functional interface that applications will use to create, retrieve, update and delete data elements in the cloud and, once adopted, will make it much easier to move data from one cloud to another. The OpenStack initiative is also attempting to create and enforce standards that will facilitate data movement across different platforms and providers.

As the market matures, we're hopeful that cloud storage service providers and technology vendors will rally around a handful of cloud standards that will put the data and vendor lock-in problem to rest. But in the meantime, *caveat emptor:* let the buyer beware.

Capacity Management In Virtualized Cloudy IT

By Mike Matchett

A few years ago, one of the big attractions of virtualization technologies was that they enabled highly responsive and even dynamic allocations of resources on demand. Many IT folks assumed this would alleviate the need for up front capacity planning. It didn't.

Now IT has at least three new capacity management challenges.

1. **Resource sizing.** The biggest one is sizing the resources needed for the entire resource pool. As we virtualize more and more of our mission-critical applications it's ever more important that the entire cluster be able to handle the aggregate demands of many kinds of applications co-hosted together. Despite increasingly popular modular scale-out virtual infrastructure solutions, this still requires capacity planning at the larger scale or you risk overspending on soon-to-be obsolete infrastructure or face severe performance bottlenecks at the worst possible times when critical applications peak together. Capacity planning has always been about right sizing the right infrastructure at the right time. Sure, hybrid cloud bursting is just around the corner for many as yet another reactive panacea to in-house resource constraints, yet it's still possible to overspend on cloud allocations, or under subscribe with poor resulting performance.

2. **Virtual resource guessing games.** The second issue is that as we virtualize deeper into our mission critical applications portfolio, we simply can't continue to guess at what virtual resources might deliver satisfactory application performance and trust that the reactive system dynamics will smooth everything out. Virtualization is essentially sharing, and good

sharing schemes require a sound understanding of the resource demands required by each application within each VM in order to set the knobs and buttons to do the right thing at run-time. It's possible and maybe even desirable to oversubscribe the low-hanging fruit of servers in test and dev, but don't try that with your mission critical apps in production.

3. **Managing convergence.** Finally, much of what is happening in IT infrastructure these days is converging. It's no longer sufficient to examine performance or capacity plan silo by silo (if it ever really was). Today, it's critical that capacity management take a holistic view across servers, storage, networking, and any other critical resources. And with the advent of clouds, capacity management isn't limited to the data center anymore either. It's an enterprise function at the CIO visible level.

The bottom-line is that performance analysis and capacity planning disciplines aren't even close to dead, although there are fewer and fewer adherents who learned the formal discipline in big iron. What's needed for this new generation is a competitive approach to optimizing total IT spend for maximum business value that can be leveraged by the average virtual admin. *Old school* capacity planning might be dead, but long live the new virtual infrastructure capacity management!

Data Protection and Backup

Cloud Disaster Recovery Plan

By Arun Taneja (original version published in SearchStorage)

Organizations large and small should seriously consider incorporating cloud storage into their DR planning, testing and deployment.

Cloud storage offers greater versatility and data accessibility than most other DR options, and may also provide a significant cost advantage, particularly to small- and medium-sized businesses (SMBs).

Before we discuss how cloud storage can improve a DR plan, let's touch on the disaster recovery plan itself. All organizations should have a DR plan, even the smallest ones. In talking with end users, we're always surprised by just how many organizations don't have a DR plan, or have let their plans fall into abeyance. And if your organization has a DR plan, you must still invest the time to review and exercise it regularly -- to be sure the plan is functioning and to ensure you haven't overlooked opportunities to make your DR practices more rigorous and efficient.

Until recently, effective DR practices required an IT practice unto itself -- full-blown, duplicate IT infrastructure spread out to two or more different sites, along with all the hands-on systems and storage management that goes with that plan. The cost and complexity of these traditional approaches discouraged (and, in many cases, prevented) SMBs from investing in DR planning and processes. But the advent of virtualization and cloud technologies has changed that picture dramatically.

With the recent emergence of virtualization and cloud storage technologies, it's now possible to implement a cloud disaster recovery plan. So how are companies deploying cloud storage for DR purposes today? Users basically move or replicate data on a regular basis from the data center into a cloud storage repository, where that data can then be used for recovery in the event of a

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

disaster or other prolonged outage. For environments that require rapid recovery, this approach can be extended to cover application workloads as well. By periodically synchronizing data, vendors can enable workloads to switch or fail over to the cloud, where they can continue to run as long as the primary site is down.

Cloud-enabled DR delivers a number of advantages over traditional DR architectures, which generally involve data being moved or replicated to a physical, off-site facility. Because the cloud eliminates the need for customers to invest in a remote DR facility, the cloud significantly shrinks the CAPEX required for traditional DR. Ongoing operating expenses are also reduced as users no longer have to pay power and cooling costs for remote equipment. Because the cloud is economically priced and can allocate capacity and performance on demand, customers only have to pay for the resources consumed.

Moving DR to the cloud can also increase the flexibility of disaster recovery configurations and practices. And because clouds are designed for remote management, it may speed recovery. Compared to cumbersome and expensive tape-based DR practices (both on-premises and off-site), such capabilities can make routine testing practical and ensure a solution works when needed.

As you would expect with any emerging technology, cloud storage also introduces some potential issues that users must pay attention to. Before pursuing disaster recovery in the cloud, users should demand answers from cloud storage providers to the following questions:

How will my data be secured in the cloud? Does the provider offer encryption services, both for data in transit and data at rest? How are cloud storage users authenticated -- using passwords only or a two-factor authentication scheme?

- Does the cloud provider satisfy regulatory requirements for the regulations I care about? How is compliance measured and certified?

- What are the expected recovery times (RTOs) for the data I'll be storing in the cloud? Is the topic of RTOs covered in the service-level agreement (SLA)?
- Does the cloud storage provider have a demonstrated track record in meeting data availability and recovery requirements?
- Can the provider effectively match the virtualization infrastructure I have in-house to facilitate rapid recovery?
- Will the provider allow me to test and exercise DR processes on a regular basis to ensure my organization is prepared in the event of an outage?
- The answers to these questions will depend in part on what type of DR capability your provider offers. There are three types of cloud disaster recovery plan options to consider:
- ***Cloud backup solutions.*** For small organizations without large amounts of data, simple backup to the cloud may be enough protection for DR purposes. Backup tends to be simpler and less costly than more rigorous disaster recovery solutions, but recovery will likely take longer, since data will need to be restored from backup files or images. These solutions are available from cloud and virtualization startups, as well as traditional data protection vendors.
- ***Cloud gateway, on-ramp and integrated storage solutions.*** These offerings are considerably more functional than simple backup solutions, often including a local storage appliance and building in capabilities such as capacity optimization, automated replication and transparent scalability.
- ***Cloud-hosted DR solutions.*** This emerging class of solutions takes data and virtual server workloads off-site into a multi-tenant, hosted cloud and can enable full recovery of servers, data and applications.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

You will need to choose the solution that best fits your situation, considering everything from the value of your data to cost and recovery objectives. Whatever type of solution or provider you choose, there's no question cloud storage must now be taken seriously in your disaster recovery planning process and that creating a cloud disaster recovery plan can be a cost-effective and necessary choice for your company.

Cloud Backup Survey 2012: Stormy weather?

By Christine Taylor with Ashar Baig (original version published in InfoStor)

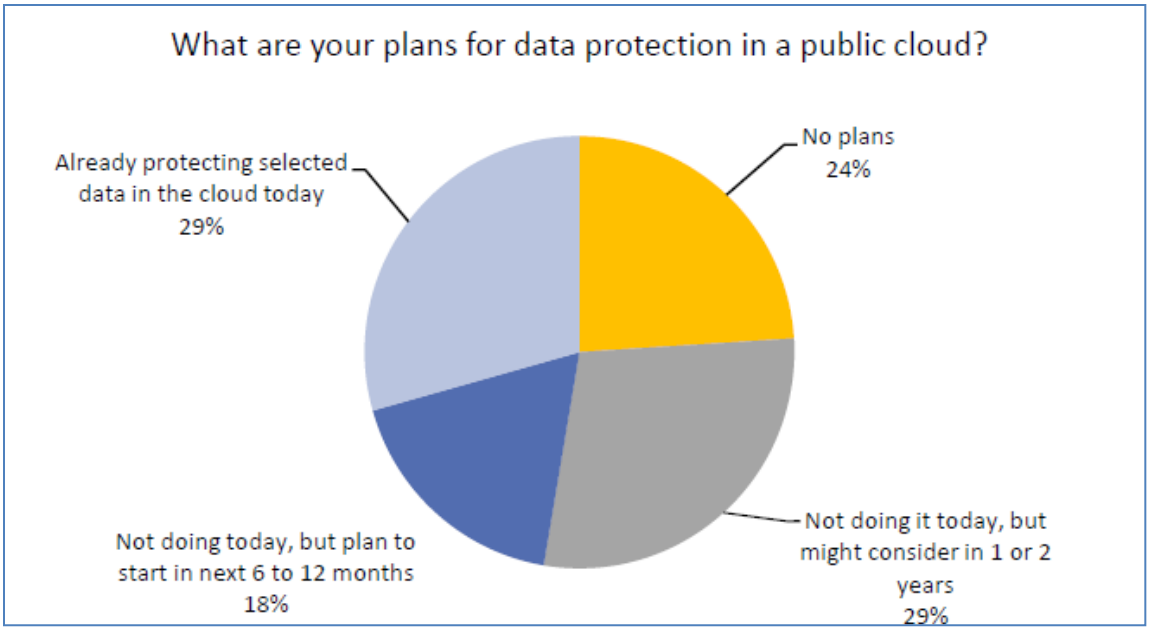
Taneja Group and *InfoStor* jointly ran a survey asking IT managers about their experiences and plans in using public clouds to protect data and host applications.

We specified public clouds because of their huge capacity and wide availability at low cost. Taneja Group surveyed 150 IT respondents representing company revenues from a low of less than $1 million to greater than $1 billion. The $50-$100 million segment yielded the most respondents but all revenue segments were fairly evenly divided. A slight majority of respondents were from the private sector with the rest coming in from government, education, and non-profits.

We concluded that there is a lot of interest *and* a lot of uncertainly around public clouds, and public cloud vendors have their work cut out for them to capitalize on these big opportunities. We suggest concentrating on customer education and corporate service level agreements that benefit both corporation and vendor.

Plans for Data Protection in the Cloud

Only 29% of the respondents were already protecting data in the cloud, leaving 71% who were not – a distressing picture for cloud vendors. 58% were however planning on moving some data protection to the cloud between 6 months and 2 years, yet 24% had no plans to do so.



Those who have or were planning to have data in the cloud varied in their reasons. 59% of respondents reported that potential CAPEX savings were important to them, and they would move to the cloud if it would help them eliminate secondary DR sites and redundant DR backup storage systems. 59% also found cloud storage scalability to be attractive. This is no surprise: on-premise storage scalability is a real issue involving budgetary costs, data migration, provisioning, downtime, data center space considerations, and energy costs. IT respondents felt that gaining scalability with cloud-based backup was an attractive proposition for both CAPEX and OPEX savings over in-house storage growth. Simplifying backup management gained a 52% response.

Concerns

These responses all point to the attractiveness of a cloud-based storage solution for cost and IT resource savings. However, if cloud-based storage is so attractive why aren't more IT administrators storing to the cloud?

The answer to that question was startling with its solid list of objections and concerns. Cloud vendors are going to have to do a better job of answering these objections in the coming months and years in order to spur more mid-sized and enterprise adoptions.

Value	Percent
Security concerns	55%
Data availability concerns	45%
Data accessibility concerns	45%
Internal IT policies	56%
Regulatory compliance issues (PCI, HIPAA, SOX, etc.)	47%
Inability to meet RTOs (Recovery Time Objectives)	49%
Cost concerns (e.g. data transfer or storage costs)	43%
Fear of service provider or vendor lock-in	47%

The startling variety of these objections should give cloud vendors pause. Granted they represent an opportunity, but so many of the marketing pieces we see and hear simply do not address any of these concerns. They are the sizzle and not the steak. But IT professionals are not starry-eyed personal consumers; they are charged with valuable data protection and they are wary. They should be. How should the public cloud vendors answer an IT administrator's very real concerns about data availability and accessibility? Or figure out a way to meet compliance or eDiscovery needs, or avoid cloud vendor lock-in? This is where the rubber hits the road. Most sales teams worth their salt know how to answer these concerns with individual clients, but we are not seeing nearly the level of customer education that we would like to see. And widespread reluctance to move data to the cloud is the result.

For example, the vendor lock-in objection is a real concern. Moving data between cloud storage service providers is not easily done, and it is in the vendors' interest to keep it that way.

Unfortunately, the difficulty has a long-term effect on making customers leery of moving data into the cloud at all when it is terrifically hard to move out of it again. They must face the expense of mass data migration twice – once when moving to a cloud provider to begin with, and then to move to another one. The vendors know that inertia and a reluctance to pay for the second migration is a strong motivator for customers to keep their data with one service provider, but there are many customers who are reluctant to move their data to the cloud in the first place because of these considerations. We don't blame cloud vendors for hanging on to their customers, but they should at least offer data migration tools or services to facilitate mass data transfers as necessary.

Additional objections ran neck and neck with vendor lock-in. Data accessibility and availability need to be defined. That data will generally be accessible and available is a given with large public cloud providers like Amazon and Google, who aren't going anywhere. (Note that data availability may be a big concern with smaller service provider firms. Some of these smaller firms offer better service level agreements than the big public clouds do, but the customer must be certain that the service provider is in it for the long, long haul.) The question for entrenched cloud providers is how *soon* data will be accessible and available given an urgent requirement, and how long it will take to retrieve that data. Answers involve the amount and type of data the customer has stored on the public cloud, and what (if any) migration tools exist to quickly retrieve the data.

Customers need to know this going in. They must write service level agreements for prompt data transfer from their cloud vendors, and include both digital data transfer objectives and physical media. It's pretty ridiculous that cloud vendors have to ship data on tape or disk to a customer, but the reality of bandwidth is that the old-fashioned truck can be a whole lot faster than the information highway. We urge customers to write the data transfer objectives into their SLA requests. Public clouds that cater to the corporate market will be open and ready for meaningful SLAs, but if the cloud vendor balks, understand why and what you will be giving up if you accede to their demands.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

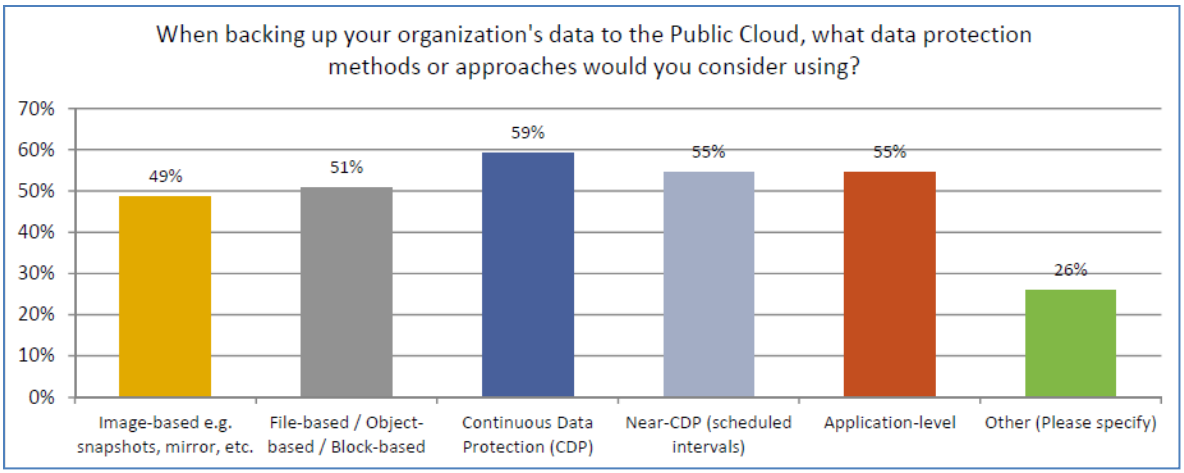
Applications

Data protection as backup is one thing; data protection with active applications running in the cloud is another. We also asked about application types that IT might be willing to run from the cloud if they were confident about their data's safety. The applications ran the gamut of functionality.

Value	Percent
Non-critical apps and data e.g. file and home directories	49%
Microsoft Exchange, GroupWise, Lotus Notes, etc.	41%
Microsoft SharePoint	36%
Databases (SQL, Oracle, DB2)	43%
Enterprise Resource Planning (ERP)	41%
Customer Relationship Management (CRM) applications (Oracle/Siebel, SAS, SAP, Salesforce, etc.)	34%
Analytical Tools	39%

Application in the cloud choices are frequently driven by concerns about performance, particularly network and storage IO. Not every application is suitable for running the cloud in the first place. Large data sets suffer from data transfer considerations and applications sensitive to latency are usually a loss on the cloud front. Applications that are a good fit include test/dev workloads, personal productivity applications, collaborative and messaging applications, and virtualized applications.

Of the applications listed above, most of them are provisionally comfortable in the cloud with the exception of data analytics and ERP. Analytics on small datasets can work just fine, but large-scale analytics sucks up cloud resources. There was a reason that Eli Lilly dumped its plans to launch large-scale analytics on Amazon's public cloud. ERP is also a hard choice to make for cloud, since these applications require high throughput and low latency, both requirements that cloud-based infrastructure is hard put to fill. They also contain sensitive data, large datasets, high availability requirements and compliance concerns: all requirements difficult to fulfill in the cloud.



Priorities

Our remaining question centered on data protection priorities in the cloud.

The specific responses ranged close to each other from 49% to 59%. The top number was continuous data protection (CDP), an interesting choice since CDP backup is usually only practiced with Tier 1 application data, not the type of data most often found in the cloud. The second most popular choice was related: Near-CDP or frequent backup at scheduled intervals (often processed as scheduled snapshots). Additional choices are driven by data types such as file, object or block-based; pointing to a desire for flexible storage platforms in the public cloud.

Last year Taneja Group analysts Jeff Boles and Jeff Byrne asked similar questions in a report for *InfoStor*. ("Is cloud-enabled DR ready for prime time?") Cloud-based DR is not the same animal as cloud-based data protection (DR is a more complex cloud infrastructure) but several of the report's conclusions hold true for this recent survey as well. Would-be corporate cloud users are facing worrisome challenges around access, security, ease of use, recovery time and effort, vendor lock-in, eDiscovery/compliance and application availability.

Some public cloud vendors are tackling the challenges and are providing simplified, easy-to-use, and cost-effective ways to get data safely in and out of the cloud infrastructure. These tend to be the vendors who specialize in supporting corporate cloud clients. However, some public cloud computing providers are unwilling to reach the level of SLAs that corporate users need. Frankly we think they are leaving vast amounts of business on the table. Public cloud customers should always do due diligence using SLAs before they place their data and applications in the cloud, not afterwards. The public cloud vendors who are better able to serve their corporate customers will reap the benefits, and will go a long way towards spurring reluctant business consumers towards the cloud.

Performance in the Cloud

Flash storage technology and cloud service providers' needs

By Jeff Byrne (original version published in SearchStorageChannel)

Data storage is a crucial part of a solution provider's offering.

Regardless of whether you're a large-scale cloud infrastructure as a service (IaaS) provider or a small regional managed service provider (MSP), a strong storage platform can push you ahead of the competition and help you quickly expand your business. On the other hand, a weak storage platform could become the Achilles' heel that sets a provider back, whether it is from a missed business opportunity or service-level agreement (SLA) penalties.

Businesses that have sizable storage budgets and employees with specialized skills are well-suited to traditional storage systems. In contrast, service provider operations designed for scale with margin-optimizing efficiency require considerably more storage with considerably less complexity. Many traditional vendors, in trying to meet these requirements, have compelled providers to make sacrifices in the key areas of performance, cost and availability.

With this in mind, service providers looking for a better approach have increasingly invested in flash storage technology. Flash technology has promised -- and largely delivered -- better IOPS and increased throughput for critical, I/O-intensive applications, though often at a much higher cost per GB than traditional hard disk drives (HDDs). But performance is just one attribute cloud service providers are looking for in a storage solution. Also high on their list are requirements such as:

- **High availability and reliability.** Cloud service providers can't afford storage infrastructure downtime, lest they risk breaking SLA-driven availability commitments and irreparably damaging their reputations.

- **Cost-effective scalability.** Flash storage must scale easily and non-disruptively to accommodate rapidly growing capacity requirements, and at a cost that enables service providers to maintain affordable and competitive prices.

- **Built-in operational efficiencies.** In a similar vein, cloud service providers are looking for storage systems that are frugal in terms of space, power and cooling requirements, since every dollar of cost means one less dollar flowing to the bottom line.

Unfortunately, most flash storage systems today don't meet these requirements. We believe this gap between solution requirements and actual capabilities, along with some confusion about how and where flash technology can best be deployed, has inhibited the adoption of flash storage among cloud service providers. Let's take a closer look at these requirements. (Though flash storage technology can be deployed in various ways to meet service provider challenges, we'll focus on all-flash array solutions.) It's important to note that many enterprises have similar needs. So while emerging solid-state products satisfy service provider requirements, they also meet those of enterprises looking for high-performance, reliable and scalable storage.

- **Requirement: Availability and reliability.** For service providers looking for a storage platform, high availability is clearly one of the top requirements. Most cloud IaaS providers' SLAs are based on minimum availability guarantees, meaning storage must operate without disruption in the event of outages or planned maintenance activities. Unfortunately, many early flash storage offerings have been built around single controller systems, and have lacked the availability features needed to meet these requirements.

- **Requirement: Scalability.** While flash storage performance generally lived up to service providers' expectations, scalability of flash systems didn't.

- **Requirement: Operational efficiencies.** In the highly competitive cloud service provider space, efficiency advantages can make or break profitability. Service providers are constantly looking for ways to pack more storage wallop into a smaller footprint, and to wring every last

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

watt of power and BTU of cooling efficiency out of their arrays. Unfortunately, most flash storage offerings to date have fallen short of meeting those efficiency goals but there are a few who are meeting – and raising – the bar.

Flash memory has the potential to take cloud storage availability, scalability and efficiency to a whole new level. Innovative players are bringing flash storage products to market that help realize the potential of solid-state storage.

A Guide to Moving Applications to the Cloud

By Arun Taneja (original version published in SearchStorage)

We're often asked "Which of my applications should I run in the cloud?" When we dig deeper, we find this question is primarily motivated by concerns about performance. In particular, network and storage I/O performance.

It's a very good question, and one that application owners and IT infrastructure managers alike should take very seriously.

Before we tackle the concern about moving applications to the cloud and application performance, let's provide a slightly broader context because there are a number of other issues users should think about as they consider whether to move specific applications to the cloud.

- **Security**: Users should ask their cloud provider what type of security they offer, both in the form of data encryption and access controls, to ensure application data will be secure and that the provider authenticates users before they're allowed to log in and access sensitive applications.

- **Regulatory compliance:** Some business-critical applications, particularly in financial services and health care, will have strict requirements that may not be easily satisfied in the cloud

- **Application availability:** This will likely be a concern for many business-critical applications, and in such cases, users should demand that their availability needs can comfortably be met based on the terms of the provider's service-level agreement (SLA).

Unfortunately, public cloud computing providers aren't yet willing to commit to the kinds of availability levels that most corporate users demand for their business-critical applications.

Given this backdrop, let's address the question of performance. What types of applications should users avoid running in a public cloud, purely from a performance standpoint? First, latency-sensitive applications aren't a good fit. Latency across the Internet can vary, but increases with distance, which is typically beyond a user's control. Second, applications with large datasets are problematic because uploading large amounts of data is time consuming and costly in terms of bandwidth. Third, applications that require special hardware, such as a graphics processor for rendering, aren't candidates for the public cloud.

On the flip side, compute-intensive applications tend to be a good fit, particularly those with small datasets. Applications with spiky or fluctuating workloads, as long as they're not latency sensitive, can also be good candidates to run in the public cloud.

So what do these rules of thumb mean for the suitability of running particular applications in the cloud? Let's explore that further and consider the fitness of some common applications in two categories: those that are generally a good fit for the cloud, and those that are a questionable or poor fit.

Applications that are a good fit for moving to the cloud

- **Dev/test apps.** These tend to be quite suitable for the public cloud. The largest percentage of compute instances on major cloud sites (like Amazon Web Services) are dev/test workloads. The build and test process tends to be compute-intensive, and therefore a natural fit for cloud computing.

- **Personal productivity apps**. Word processing, spreadsheet and presentation design software tend to be a good fit. These applications are based on unstructured data and generally don't require low latencies or sub-second response times. Vendors such as Microsoft have developed SaaS bundles of productivity applications that are hosted in the cloud.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

- **Collaborative apps.** Social networking, web conferencing and other collaborative apps are good for the cloud, especially since many of these solutions were written for the cloud in the first place. Legacy apps such as SharePoint have also been adapted to run in the cloud.

- **High-performance computing (HPC) apps.** Based on their compute-intensive nature, HPC apps are usually a good fit for cloud compute farms, as long as their data needs can be managed.

- **Virtualized apps.** Given their compact footprint and built-in efficiency, most virtualized apps are highly suitable for the cloud.

- **Disaster recovery (DR) apps.** As discussed in our related tip on cloud storage gateways, DR apps can be an excellent fit for the cloud because the cloud provides a cost-effective, universally accessible recovery platform. Vendors are now busy building or reengineering DR apps to make them cloud-savvy.

- **"Big data" apps.** The data mining and analytics of big data applications, such as those running in Hadoop clusters, make them good candidates for cloud-based processing. One caveat: It may be costly and time-consuming to move large amounts of data into the cloud, so if this is a requirement, you'll need to decide whether it's worthwhile.

Applications that are a poor fit for moving to the cloud

- **Mission-critical apps**. Mission-critical apps such as ERP suites tend to have all the wrong characteristics to be hosted in the cloud. They tend to be transaction-intensive, with high throughput and low latency requirements. They contain sensitive data and often large datasets, and have high availability requirements. Some such apps also have regulatory compliance needs that may be difficult to meet in the cloud. So all in all, these aren't good candidates for the cloud.

- **Network-intensive apps.** Unless you have access to fast, high-quality network resources, applications that continually transmit and receive large amounts of data won't be a good fit. Such applications will often require access to, or integration with, other applications to share data. A whole host of considerations enter in here, but suffice it to say, caveat emptor.

Keep in mind that these are general guidelines, and your decision about moving applications to the cloud should be based on your own situation, including application performance needs, budgetary constraints and a whole lot more.

Cloud Gateway Options and Use Cases

By Arun Taneja (original version published in SearchStorage)

Since cloud storage gateways first burst on the scene a few years ago, they've rapidly evolved and taken on different roles in connecting in-house users and applications to the cloud. But what is a cloud gateway?

In simple terms, a gateway is an on-premises device, usually taking the form of a hardware or software appliance that connects local applications to cloud-based storage. Legacy applications tend not to speak the same language as the public cloud, so a gateway must also translate between the traditional storage-area network (SAN) or network-attached storage (NAS) protocols employed in the data center, and the REST API-over-HTTP protocols used in the cloud. This translation happens in the background, enabling these incompatible technologies to communicate transparently.

Given its built-in connectivity and protocol translation capabilities, a gateway makes remote cloud storage look just like the iSCSI or NAS storage that sits in your data center today. This enables in-house users who might be skeptical about the cloud the opportunity to try out cloud storage using a “toe in the water” approach -- the gateway provides transparent access while masking cloud-specific technologies. In this respect, cloud gateways have no doubt accelerated the adoption of cloud storage, even if only for a small subset of a customer's overall data.

A majority of gateways also provide at least one additional feature that makes cloud storage more palatable: they contain some amount of local storage, which is generally used as a cache to improve performance. A growing number of gateway solutions also provide one or more storage capacity optimization technologies, such as compression and/or data deduplication, which reduce

the amount of data that flows between the customer's data center and the cloud. This benefits customers in two ways: reduced bandwidth charges and increased performance.

Most cloud gateway solutions now support several major public (or virtual private) clouds on the back end. This is an important factor to look for because it will enable you to avoid getting locked into a single cloud provider.

While many assume that all cloud storage gateways are alike, this is not the case. Vendors have designed their gateways for specific functions, some relatively broad and others quite specific. Let's look at some of the major use cases for a cloud storage gateway, along with examples of solutions for each:

- **Backup**: Cloud-enabled backup provides customers with a highly scalable and elastic repository for their backup data. Gateways providing backup capabilities can take the form of appliances that serve as dedicated, on-premises backup targets that connect to cloud storage or general-purpose gateways that are used for backup (delivered by a whole host of traditional data protection vendors. Gateway solutions targeted for backup generally provide snapshot capabilities, and most offerings treat the cloud as a tier to which data is migrated during a data protection lifecycle. Look for solutions that provide local caching and optimize the cloud connection, and also optimize the cloud-stored data to reduce the amount of data transfer.

- **Archiving**: In this use case, the cloud is employed as an archive tier, to store relatively inactive unstructured data that consumes lots of space. Fast access times are usually not a priority, but cheap storage is essential. For example, deep file archives might be moved to the cloud to free up local storage capacity for other purposes. As in the case of the backup use case, the on-premises gateway serves as the access point to data in the cloud. Gateways deployed for this use case tend to have fully versioned file systems and strong data encryption features, and more advanced solutions offer quality of service network settings and data retention policies. The Nasuni Filer is an example of an on-premises appliance that can be used for this purpose.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

- **Disaster recovery (DR):** This is similar to the backup use case, though selected vendors are marketing their cloud gateway solutions specifically for disaster recovery. Gateways targeted at Learn More DR will often replicate data to the cloud on a continuous basis. These solutions are typically optimized for network efficiency, only sending data that has changed to the cloud. For a more complete DR solution that includes compute and storage, look for solutions that provide data mirroring from on-premises applications to applications running in an associated compute cloud. Most gateways addressing DR aren't yet optimized for the recovery phase, but look for a new wave of solutions over the coming year that accelerate and automate the recovery of applications and data in the event of an on-premises outage.

- **Collaboration**: Some gateway solutions provide the ability for distributed teams to collaborate around specific workflows or projects, such as artists editing video files across multiple locations. To deliver such capabilities, gateway-enabled collaboration solutions will ideally provide a global name space built on an application-aware global file system, with strong data encryption, access controls, local performance and snapshot protection. Panzura is an example of a cloud-tiered NAS gateway well suited to collaboration use cases.

- **Cloud-integrated enterprise storage:** The ultimate category of optimized cloud access solutions goes beyond mere gateway functionality. Solutions in this newly emerging category are engineered for the demands of primary storage, and enable customers to extend their primary storage transparently into the cloud, without compromising availability or performance.

Given this broad range of cloud gateway offerings and functionality, customers should decide upfront what use cases their solution will be deployed for, and shop carefully with that in mind.

Learn More: Taneja Group

More Resources from Taneja Group

More Cloud Resources

Taneja Group Announces a New Managed Service Provider (MSP)-focused Consulting Practice for Vendors http://us2.campaign-archive2.com/?u=99aeeaaecf98d38d4e1404801&id=f78742e383&e=2d5e5b0a15

Do you want your products to be used to create a cloud or be used with it? We can help storage vendors understand the MSP landscape and develop MSP strategy, plans, programs, messaging, positioning, value propositions, lead generation and thought leadership campaigns as well as on-boarding, product stickiness strategies and programs to grow revenues from existing MSP customers. We will introduce you to key stakeholders within large and small MSPs.

Taneja Group Newsletter on Cloud

Taneja Group sends out newsletters devoted to the hottest topics in today's storage and server markets. We have published a newsletter on the cloud; contact christine.taylor@tanejagroup.com for a copy.

Taneja Group advises clients on developing, differentiating and marketing their products and solutions in a crowded market. We provide complete marketing guidance including product management, messaging, positioning, validation reports, social media, and launch support. Our technology clients include cloud application delivery and storage, physical storage, virtualization, server, and info governance companies. These technology vendors are looking for unbiased assessments conducted by the Taneja Group professionals who have years of end user and vendor experience. http://tanejagroup.com/contact/.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557