Does Your Cloud Have a Silver Lining?

Getting the Most From Backup and Cloud Computing

UNITRENDS

7 Technology Circle
Suite 100
Columbia, SC 29203

Phone: 866.359.5411
E-Mail: sales@unitrends.com
URL: www.unitrends.com

Does Your Cloud Have a Silver Lining?

Clouds are all the rage right now. And the truth is that cloud-based computing can be a tremendous value. Cloud-based computing is not just an opportunity for a bunch of venture capitalists hanging out at Sand Hill Road, it's a wonderful chance for small and medium businesses to take a step toward aligning their information technology spend to increase their focus on revenue-generating opportunities.

At the same time, the level of hype surrounding cloud-based computing has never been higher. The purpose of this paper is to help you understand the advantages and the drawbacks of cloud-based computing - particularly as it applies to business continuity and disaster recovery.

What is a Cloud?

One popular definition, which we've modified slightly from the original, is that you know you're using cloud-based computing when the crash of a computer you've never heard of stops you from getting any work done.

Seriously, cloud-based computing is an information technology model by which resources, software, and data are provided as an on-demand service. The best metaphor is that cloud-based computing seeks to make information technology available in the same manner that your utility company makes electricity available - with a simple wire without your having to understand the details of how it's provided.

For those more technically inclined, a strict definition of a cloud is that they have five essential characteristics:

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

There are also different categories associated with the provisioning of cloud-based solutions. These are

- Private Clouds. Private clouds are single tenant - these are accessible only within some well-defined private domain. The most common use of private clouds is delivering a set of services within a single organizational entity.

- Public Clouds. Public clouds are multi tenant - these are accessible in a broad public domain to multiple organizations and individual users.

For those of you paying attention, you'll note that private clouds seem to violate the "broad network access" essential characteristic - for that reason there's some debate about whether a private cloud is really a "cloud." We'll leave it to others to argue how many angels fit on the head of a pin; for now we'll simply note that a private cloud is typically implemented to either have more control over resources, due to security concerns, or other factors.

We'll revisit these definitions later in the context of clouds associated with business continuity and disaster recovery; for now it suffices to state that the overall purpose of cloud-based computing is to lower your total cost of ownership and increase your return on investment with respect to information technology.

Categories of Cloud Solutions

There are several different categories of cloud-based solutions. These are

- Infrastructure as a Service. This category typically refers to the low-level provisioning of hardware (or virtual hardware) on which the user of the service generally has control over the operating system - and thus can execute any software that the user desires. Examples of this category are Amazon's EC2/S3, Mosso, and GoGrid.

- Platform as a Service. This category typically refers to a higher-level environment in which developers write custom applications. There are typically some restrictions on the type of software that the developer can write; however, there are higher-level services available for those developers to make writing those applications easier. Examples of this category are Google's AppEngine, Bungee Labs, or Heroku.

- Software as a Service. This category typically refers to an environment in which only special-purpose software is made available as a service via the Internet. Examples of this category are SalesForce, Microsoft BPOS, and NetSuite.

Advantages of Cloud Solutions

The fundamental advantages of cloud-based computing are

- Reducing costs. This advantage is primarily accomplished by eliminating capital expenditure and reducing operational expenditure. For multi-tenant public clouds, cost reduction is also possible by the pooling of information technology resources in order to be able to meet peak demand with effectively less information technology resources.

- Ability to quickly scale. This advantage refers to the fact that because cloud-based computing typically has excess capacity, that this excess capacity may be provisioned quickly and in a highly responsive and nimble fashion to those users requesting it. To put this more simply, there is no wait for new hardware and/or software to be ordered and installed.

- Well-defined service level agreements resulting from measured service. Typically cloud-based computing services are priced to some degree based on usage characteristics; thus the service level agreements associated with everything from availability to performance are more rigorously captured than alternative solutions.

Of course, there's no such thing as a free lunch - there are some disadvantages associated with cloud-based computing as well. These are discussed in the next section.

Drawbacks of Cloud Solutions

Of course there are drawbacks with respect to cloud-based computing as well; these include

- Performance and latency. WAN performance has not and will not keep pace with microprocessor, memory, disk, and LAN performance. Any cloud solution must take into account not only the gap between on- and off-premise performance, but must also address the fact that this gap will only widen over time. This issue affects both private and public clouds equally.

- Security and privacy. Security and privacy, for both at-rest and in-flight data, must be taken into account. Different industries have different requirements so it's important that security and privacy be addresses in a flexible manner. This issue affects primarily public clouds.

- Reliability and availability. The information infrastructure that constitutes the cloud as well as the connection to the cloud must be both reliable and highly available. This issue affects both private and public clouds equally.

- Manageability. Because clouds can be somewhat amorphous in nature to the end user and because users perceive a loss of control when using clouds, the manageability of the cloud is very important. In private clouds, this is typically less of an issue since the entire infrastructure can be managed by the one tenant of the cloud. In public clouds, manageability has to be a key component of the architecture since otherwise a user can end up with a lowest common denominator solution.

- Monitoring. Trust is the single biggest issue raised by consumers of cloud-based services. Monitoring is incredibly important as one way of helping establish trust. This is a major issue for public clouds - and has also been called out as an important issue in private clouds as well.

Cloud Storage and Cloud Backup: The Problem

While there are a lot of different cloud-based computing solutions out there, what we'll focus on in the remainder of this document is cloud storage and cloud backup. Cloud storage simply refers to the use of the cloud as a replacement for some NAS (Network Attached Storage) centralized storage - some vendors advertise their cloud storage as "your file server in the cloud." Cloud backup refers to the use of the cloud as a mechanism to protect data.

"Pure" cloud storage and cloud backup (where "pure" denotes using only the cloud instead of some type of on-premise mechanism coupled with the cloud) alone suffer from one drawback first and foremost: the aforementioned gap between on-premise and off-premise (cloud) performance due to the WAN used to connect the two. Practically, the problem is that while the price per gigabyte of storage has been dropping at a tremendous rate with 7200RPM 1TB drives priced below $90 at the time this paper was being written - the rate at which the price per megabit per second of WAN bandwidth has dropped has been relatively sluggish.

For that reason some type of caching mechanism for both is needed. For pure cloud storage, the typical mechanism is some type of local per-client replication - such that the data is replicated directly on the client where the data is being used as well as in the cloud. This works as long as the change rate of the data is relatively low.

Pure online backup is offered by an ever-growing number of vendors. And most of these online backup vendors do a credible job of backing up a few tens of gigabytes of data. The real problem with online backup is recovery. You might not care that it takes a month or more to ship your first terabyte up through the Internet to the online backup vendor. However, most people don't have a month to wait for that terabyte to be downloaded back from their online backup vendor when a hard drive or a complete system is lost.

Cloud Backup: Enabling Rapid Recovery Through On-Premise Backup Appliances

So we've established that pure online backup of more than a few tens of gigabytes to some type of cloud has a distinct problem with respect to recovery performance. And yet the advantages with respect to cloud-based computing as it pertains to backup remain:

- Enabling disaster recovery by moving the data away from the premises.
- Reducing costs via eliminated capital expense and lowered operational expense.
- Ability to quickly scale.
- Well-defined service level agreements resulting from measured service.

The answer to rapid recovery while taking advantage of cloud-based computing is an on-premise backup appliance that can quickly serve to backup your information technology infrastructure and serve as a gateway point to a cloud-based disaster recovery service.

An on-premise appliance offers not only local backup but also offers dedicated in-flight deduplication. Once the cloud is loaded with your initial set of data this in-flight deduplication not only compresses your data but also deduplicates that data before it is sent upstream to the cloud. This is done with no impact to the clients (i.e., servers, PCs, workstations, and notebooks) that are being protected by the on-premise backup appliance. It means that once your initial set of data has been uploaded, much more data can be protected than can physically move over the WAN in any given time. Thus it is possible for hundreds of gigabytes and more to be protected and kept synchronized with the cloud each day with a relatively small amount of bandwidth.

In the remainder of this paper, we're going to explore both the drawbacks associated with general cloud-based computing presented earlier in this paper and the unique issues associated with cloud backup with on-premise appliances. In each case we'll also present how those issues may be solved with Unitrends' all-in-one on-premise appliance with either private or public cloud-based vaulting solution.

Quickly Uploading Protected Data to the Cloud

If it takes 30-60 days to recover (download) 1TB of data using a 1.5Mbps T1, then it takes at least that long to upload 1TB of data to the cloud. And believe it or not, that's actually the best case. It assumes not only 100% of the bandwidth can be dedicated to the upload, but it assumes a "symmetric" WAN. WANs can be characterized into two types: symmetric and asymmetric. Symmetric WANs upload and download data at the same rate; asymmetric WANs typically upload data at a slower rate than they download data.

Many cloud-based data protection vendors, even those with on-premise backup appliances, force you to wait weeks and months to upload your data over the WAN. A few offer a faster way to protect your data. This faster way involves the shipment of media from the on-premise backup appliance for fast initial seeding of the vaulted data. After that, you can immediately begin protecting your data on a daily basis.

Quickly Recovering from Real Disasters

It's one thing to recover a file, a group of files, a directory, or even a volume. It's another to recover not only a whole system (a BareMetal image) but an entire information technology infrastructure consisting of multiple protected computers and to do so quickly. And yet in the case of a natural or man-made disaster, this is precisely what the information technology department has to do to keep a business up and running.

For that reason, it's important that your cloud-based data protection vendor be able to allow you to quickly recover from a disaster. Some offer the ability to ship you your data on various media. Make sure that you understand how this works - procuring and installing a backup server, storage, networking, an operating system, backup software, and various other software and then restoring that shipped media is pretty complex. Look for vendors who offer strict SLAs concerning shipment of one of our backup appliances for truly rapid recovery from disasters.

Security and Privacy

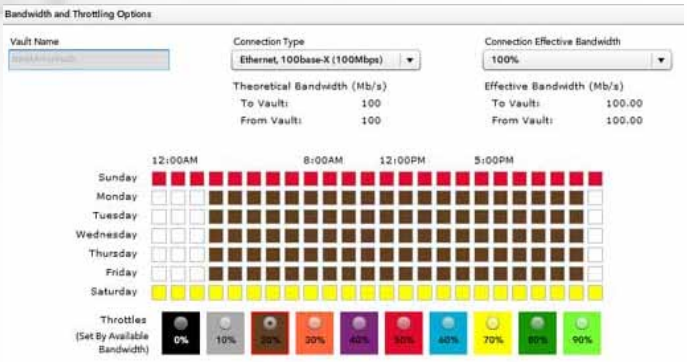
As noted previously, security and privacy, for both at-rest and in-flight data, must be taken into account. Different industries have different requirements so it's important that security and privacy be addresses in a flexible manner.

Ensure that your vendor supports in-flight and at-rest data security via encryption technology. Also make sure that this encryption technology is optional - encryption can have a tremendous impact on data protection so you want to make sure you have the flexibility to employ it where you need it and not employ it where you don't.

Reliability and Availability

The information infrastructure that constitutes the cloud as well as the connection to the cloud must be both reliable and highly available. Many cloud-based vendors tout their cloud infrastructure - of course, that's important. But the most common types of problems are between the customer's premise and the cloud.

Require UDP-level VPN (Virtual Private Network) technology from the on-premise site to the cloud in order to increase the quality (and security) of the WAN. This is particularly important if your WAN connection to the Internet may have occasional packet drops.

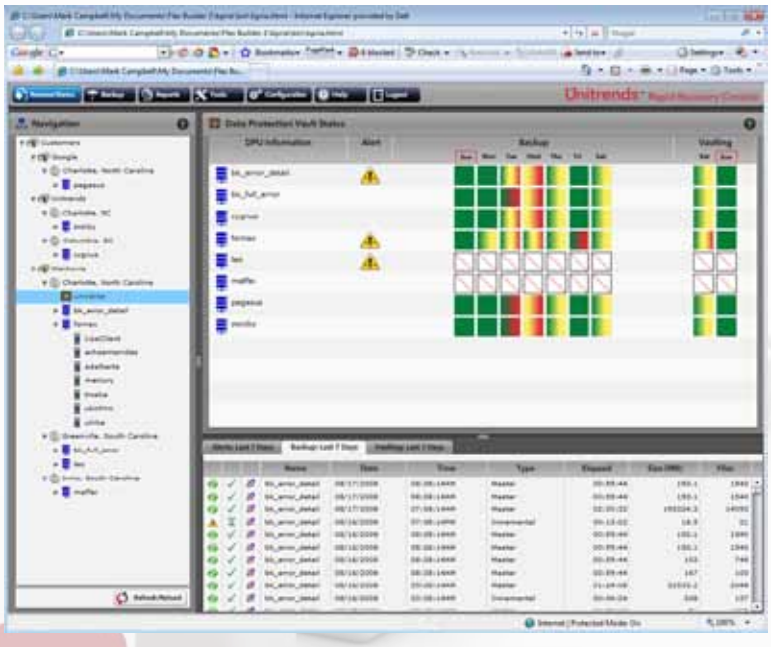


Manageability

Because clouds can be somewhat amorphous in nature to the end user and because users perceive a loss of control when using clouds, the manageability of the cloud is very important. Fortunately, with an on-premise appliance, this is made much easier. The ability to manage, and in particular the flexibility around throttling vaulting bandwidth (see screen shot to the right for how this is performed in Unitrends' cloud-based solution) to the cloud, is of paramount importance. The ability to select which clients are vaulting and which aren't (so that you can use alternative disaster recovery methods such as rotational archiving - which will be discussed in more detail later in this paper) is also very important.

Monitoring

Trust is the single biggest issue raised by consumers of cloud-based services. Monitoring is incredibly important as one way of helping establish trust. The tools and mechanisms that a vendor provides for monitoring are the backbone for establishing this trust.



It's important that vendors provide clear monitoring tools that can be accessed locally and remotely that enable the historical and real-time monitoring of the cloud-based disaster recovery service. You want your cloud-based disaster recovery service and your on-premise appliance to offer a single integrated view.

As an example, the figure to the right of this section (above) depicts the initial screen of the Unitrends user interface. A simple "green", "yellow', "red" scheme is used to show the success of both on-premise backups and off-premise cloud-based vaulting.

Support of Alternative Disaster Recovery Methods

Many vendors have a "monetization strategy" that centers around recurring revenue based on the number of gigabytes that are protected in the cloud. What you as a customer want is the ability to use any disaster recovery method, including but not limited to electronic vaulting of data to private and public clouds, to create the most efficient and cost-effective disaster recovery plan possible for your unique circumstances.

The alternative to electronic vaulting to clouds is rotational archiving. This can be done with either disk or tape (although recovery from tape can be problematic.) Vendors such as Unitrends offer D2D2x (Disk-to-Disk-to-Any) strategies, where there is the capability of the customer to use disk, tape, and the cloud as tertiary storage for the purposes of disaster recovery.

About Unitrends

Unitrends offers a family of affordable, all-in-one on-premise backup appliances that support virtual and physical system backup and disaster recovery via disk-based archiving as well as electronic vaulting to private- and public-clouds. Unitrends is customer-obsessed, not technology-obsessed, and is focused on enabling its customers to focus on their business rather than on backup.

For more information, please visit www.unitrends.com or email us at sales@unitrends.com.

7 Technology Circle, Suite 100
Columbia, SC 29203

Phone: 866.359.5411
E-Mail: sales@unitrends.com
URL: www.unitrends.com

UNITRENDS