Cloud Backups - Purpose Built vs. Public Cloud

by George Crump, Lead Analyst

Leveraging the cloud as part of the data protection process is increasing in popularity, and with that popularity come a lot of options for IT planners to sift through. One area of consideration is the cloud destination. There is a tendency to assume that all clouds are created equal and that as long as the destination is some type of cloud, then all is fine. In reality, all clouds are not created equal. Many IT professionals would be surprised to learn that the type of cloud chosen can specifically impact backup and recovery performance.

Cloud Options

The typical business or enterprise cloud backup solution often leverages a hybrid appliance on the customer premises. The software sends data from the servers, desktops and laptops being protected to the appliance. Then, as that appliance collects data, it is replicated to the cloud provider. The key differentiators between solutions are how well the software does the job of collecting, securing and sending data, how well the software recovers data and applications and the type of cloud data center used. Again, the type of data center can directly impact the backup and recovery capabilities of the solution.

Many cloud backup solutions start as a software only package that leverages a generic cloud like Amazon, Rackspace, Google or some other provider. But these clouds also support thousands of other non-backup related applications. For example Amazon hosts both Netflix and Dropbox. Google of course supports the entire Google ecosystem. Vendors that choose this option are typically software development focused and don't want to or can't invest in the backend infrastructure.

A purpose built cloud is one designed solely for a specific purpose. In this case, cloud backup. These vendors have typically created the software to perform the backups as well as built out a data center to store this data. They more than likely have built a second data center to store a redundant copy of data in case the primary data center fails. This obviously is a much more significant investment. A purpose built backup cloud means all the resources of the cloud data center are focused on receiving inbound backup jobs, initiating recovery requests and, increasingly, actually running the application if Disaster Recovery as a Service (DRaaS) is offered by the provider.

Generic Support

The number one consideration when selecting any data protection solution is quality of support. At some point almost any data protection and disaster recovery solution will need support. Getting support should be the number one concern for potential cloud backup customers. But, it is not just their access to support they should be concerned with, especially in the generic cloud use case. Certainly the software developer should be able to support a problem in their code but what if there is a problem in the cloud or the way their software interacts with the cloud? What types of support will that developer receive and how quickly will they be able to resolve a problem that very well may be out of their hands?

The purpose built cloud backup provider, since they own the data center, the equipment in that data center, the on-premises appliance and the software they should be able to more fully troubleshoot any problem and find resolution more quickly. The generic provider only controls one of these variables, where the purpose built backup provider controls them all.

The Uniqueness of Cloud Backup Bandwidth

Beyond support though there are infrastructure issues to consider. Infrastructure being tops among them. Cloud backup is unique in that most of the data transfer is over the span of about half dozen hours. For example a US based provider would likely receive the bulk of their jobs between 11pm eastern time and finish receiving them by 5am pacific time.

Compare this to other non-backup applications where much of the traffic is spread throughout the course of the business day. The exception is of course streaming services like video and music, which also both reach their peak in the evening and late evening hours.

Each of these inbound data streams require bandwidth at the provider and compute processing power to receive those jobs, index them, and store them on disk. That bandwidth utilization may have to be divided up among a wide variety of other processes in the generic cloud. Again, streaming services are often at their peak at the same time backup jobs are trying to complete. A purpose built cloud has none of these concerns; it can focus all of its bandwidth on receiving backup jobs and sending recovery data.

Restores also use network bandwidth as they transfer data to the originating business. During the day these restores may have to compete with other business applications. Daytime is also when file sync and share processes are at their busiest; they synchronize data as it changes throughout the business day. In the generic cloud use case, restores will probably occur when all the other applications are at their busiest.

Again the same spin-up questions need to be asked. In the purpose built cloud no other unrelated tasks are running so managing recovery during the daytime should not be an issue.

Disaster Recovery As a Service

Increasingly, backup providers are offering DRaaS. If one of the cloud backup solution customers has a server or data center outage, the provider can host the specific application in their cloud, saving time to recover the data across the internet. This obviously takes compute and storage resources. DRaaS may be especially challenging for a software developer using a generic cloud provider, as they do not have direct control over the network and compute resources.

In the worst case situation of a regional disaster, where dozens or hundreds of businesses may declare a disaster at once; the provider will have to provide compute and storage resources to all the businesses at the same time. In this situation the compute and storage requirements could be massive and expensive.

Costs

Most generic cloud providers charge for compute time used, quality of compute, storage capacity used and data transmitted out of the cloud. They are variable costs that tend to increase as the backup software developers' offering becomes more popular. The more compute power (server instances) that develop and needs to receive inbound backup jobs and to transmit recovery requests the more expensive their costs becomes. Obviously the amount of capacity required directly impacts the backup software provider's monthly bill from the generic cloud provider.

While the generic cloud has a clear advantage on startup costs, it may be surprising to learn that long term costs can actually be more expensive. If the cloud backup software solution is successful, they could be storing petabytes of data in the cloud. The recurring charge for storage can be substantial. Some cloud backup software developers that leverage the generic have indicated that their monthly cloud storage cost alone is over a $1,000,000 each month.

Cloud backup software developers that leverage the generic cloud shield their customers from these variable costs as best they can. But the nature of the cloud pricing model will tend to drive pricing up over time and may influence the quality of the solution. In an effort to save money they may choose to process certain requests more slowly. And still these developers are at the mercy of all the other functions happening in the cloud that may impact the responsiveness of their application.

The purpose built cloud has a much higher start up cost. Companies choosing this approach have to be well funded and prepared for a long term payoff on their

investment. The purpose built cloud is typically paid for as capacity is bought. But it is a one-time charge, not a recurring charge. This should lead to a less expensive solution over time.

Potential customers of these solutions should take these issues seriously, even if they don't directly impact them upfront. Backup is a long-term commitment. The cloud backup solution chosen will likely store your data for the next five to ten years, if not longer. Understanding the long term viability of not only the cloud backup solution but also its business model is critical to a long lasting relationship.

Purpose Built Cloud Backup vs. Generic Clouds

As mentioned, the primary advantage to using a generic cloud for backup is it allows the backup software developer to rapidly offer a solution. The downside for that software developer is that their data is at the mercy of the cloud provider as to what resources they have available to speed operations as well as cost variability. Also the generic cloud provider may give priority to other applications that represent higher revenue.

To the customer, the use of a generic cloud could also lead to unpredictability. While this impact can be mitigated through the use of multiple clouds or precise cloud management within the software, many developers are locked into a single solution or don't have the ability to manage their cloud instance well. They view the cloud primarily as a dumping ground. It is important to ask this type of cloud backup provider how they are providing data security and redundancy. For example, if they are using Amazon as their destination, are they leveraging recoverability zones?

The purpose built cloud, while harder and more expensive to build, is singularly focused on providing backup and recovery. It is less impacted by external influences like sharing bandwidth amongst other applications.

Conclusion

For the IT planner or business owner the primary interaction will be with the software that is installed on their premises. Certainly the quality and capabilities of that component is a critical point of evaluation, but it is one that we believe is going to be done. Evaluation of the capabilities of the cloud destination is also critical since that is where the organizations data will "live" long term and it is a part of the evaluation process that may be overlooked. It is critical that the organization understand and test each cloud type as best as possible for the above conditions to confirm that their cloud destination will be up to the task at hand.