# An Overview of the Microsoft hybrid cloud storage solution

The cloud is front and center in a lot of technology discussions these days. And everybody wants to know how they can take advantage of it. Here is an idea for people managing data storage.

Use it to solve some of your biggest storage problems like data growth and capacity management, backup, archiving and disaster recovery protection. How, you may ask, by integrating cloud storage services with local storage. Let me explain.

First, we need to think about what cloud storage is good for and what it’s not. It grows better than any other kind of storage. You can add a little or a lot of storage without having to stop working or moving data between different storage systems. Of course, it has the cloud pay as you grow scalability model, which means you don’t have to pay a lot for storage upgrade when all you need is a little bit more. It’s also good at being resilient, because there is redundancy built in the cloud storage services. A single upload of data is automatically replicated. So you don’t have to upload it numerous times for different backup jobs or for archiving or for capacity expansion. Finally, you can keep data in cloud storage for as long as you want. And you don’t have to worry about copying it from system to system over the years, *as they wear out it our replaced*.

Okay, so what’s cloud storage not good at. Cloud storage has much higher latency and much less throughout compared to traditional storage. So it’s good for light to moderate workloads like you have a file servers but it’s not so good at doing a lot of work in a hurry like database servers do. In other words cloud storage is much more suited to Tier 2 applications such as office files, SharePoint, document management, archival storage and many VMware environments than it is for Tier 1 applications like transaction, processing and corporate email. But here is a saying, most companies have lot more Tier 2 data with moderate performance requirements than they have Tier 1 data with heavy duty performance requirements.

So, let’s take a look at Tier 2 data. While the first things you notice is that a large percentage of Tier 2 data is consuming capacity but isn’t being used. In other words, it’s inactive or cold. When we think about data growth and its associated capacity problems, we tend to focus on the incoming data that’s been created. But the real problems is all the old cold data that keeps accumulating. Incoming data and the data that’s active in the system is called the working set. The working set is constantly evolving as new data that is added to the system and as all data becomes inactive. If you break the problem down to storing these two types of data: working set and cold data, you can imagine a solution that integrates the SAN storage with cloud storage putting the working set on the SAN part and the cold dormant data in the cloud part. Microsoft refers to this type of solution as hybrid cloud storage and it’s comprised of a cloud integrated storage system that on promises and cloud storage.

Cloud integrated storage was a technology developed by a company called StorSimple, which was acquired by Microsoft and it’s now a part of the hybrid cloud storage solution along with windows Azure Storage. The cloud integrated storage system has all the same technologies as any other leading enterprise SAN storage system. It has SSDs, SASS drives, automated tiering, deduplication, compression, thin provisioning, application consists snapshots and certified support from VMware. It’s a dual controller Architecture, which means all components can be hot swaped and it features non-disruptive software upgrades. It’s a fully-featured enterprise ISCSI SAN storage system with the big difference being its ability to use Windows Azure Storage as a Tier.

When you think about it, using cloud storage as a tier for cold data takes care of another problem at the same time: archiving. Due to the needs archiving can easily be archived at cloud storage where can be quickly accessed by information workers who might need it again. Busy storage administrators don’t even need to get involved. But as long as we’re talking about archiving, what about backup? Cloud integrated storage systems are easily set up to run daily cloud snapshots that backup and archive data to Windows Azure Storage. You don’t have to worry about mixing a backup tapes with archive tapes, because there aren’t any tapes. And in case you are wondering data uploaded with cloud snapshots is encrypted with AES-256 technology. So what about DR(disaster recovery)? Cloud storage is made for it. Disaster Recovery from Windows Azure Storage can be done for many location. It’s location independent. System administrators can access dr data in the cloud from any other site with an internet connection and a cloud integrated storage system. Customer don’t have to spend money on the recovery facility because cloud storage is the recovery facility. And there is no need to provision special networks for replication because a cloud connection is all you need. One of potential problems with a full recovery from the cloud is the limited throughput to download dr data. For example, if you use cloud storage as a virtual tape device, you will probably take a longer than you though it would to download a virtual tape. If you think you could take days, he might need to check dead. It could take weeks. But storsimple develop a cloud restore technology. The idea is simple: leave the cold data in the cloud and only restore the data that applications need, in other words: the working set. This is a huge time saver when you really need to save time. With the Microsoft hybrid cloud storage solution, Windows Azure Storage is always available as a tier. Which means data is not download at first and then restored. It is simply accessed and added to the working set for the cloud integrated storage system. RTOs or the recovery time objectives are amazingly short. So it turns out that cloud storage can solve a lot of big problems in enterprise storage if it’s integrated with local SAN storage and used as a TIER and a repository for backup and archive and DR. Just like everything else the cloud touches, enterprise storage is never going to be the same again.

Thanks for watching.

## Microsoft Azure Fundamental

Data storage options (home grown technologies)

* + Relational data (Azure SQL)
  + Serialized Object Graphs in JSON (DocumentDb)
  + Flat semi-structured data (Azure Storage Tables)
  + Files (Azure Storage Blobs)
  + Serialized Object Graphs in Binary/ Base64 (Azure Storage Blobs)
  + Messages between systems or processes (Azure Storage Queues)
  + Cache (Redis)
  + Quick, simple back-end web service w/a data store for mobile device front ends (Azure Mobile Services)
  + Application data for PHP apps (MySQL managed by ClearBd)
  + NoSQL data (MongoDb managed by MongoLabs)
  + Big Data (Microsoft HDInsight)