Azure Storage

So far, we talked about “Compute” and “networking”. Now we talk about the third pillar of computing, which is “Storage”.

Azure Storage Services

Azure storage services consist of a few different storages.

- General purpose storage

- Disk Storage (virtual hard disks for virtual machines)

- File Storage

- Storage tiers

Note that storage consider as an infrastructure as a Service (IaaS).

Azure Storage (GPv2) / Standard Storage

It contains 4 types of data: Container, File, Queue, Table

Can hold up to 5 Petabytes (5PB) = 5 million GB

Pay for what you use (2 cents per GB per month) / Cheap

Not recommended for high-demand workloads

When creating a GPv2 storage account we have the option for a “**Data Lake**”. Data Lake is extremely large storage. Good for “big data” analytics.

When dealing with high-demand workloads, we may consider “**Premium Storage Options**”. This allows us to choose a couple different blobs (block blobs). This are for “Blob Storage” (Blob Storage is optimized for storing massive amounts of unstructured data) and not for queues and Tables. This uses premium SSD, and they triple the operations per seconds (OPS), lower latency (time to first byte), and more expensive.

There are multiple high-performance options: Premium SSD, Premium SSD v2, Ultra Disk

Container (Blob - Binary Large Object) Storage

These are files of any type (TXT, PDF, CSV, ...). Store loosely in a container (like putting everything in a box).

We can set privacy (private / public)

These are “unstructured data”

Container storage:

- Can create multiple containers
- Each container can contain blobs
- There is concept of folder (note: it is not like windows folders / no hierarchy)
- Only pay for what you use

In AWS, “storage account” is called “Storage Service (S3)”.

When creating a storage account we can choose “location”. We can create multiple storage accounts in any region of the world. It is important to keep the data close to the person / service consuming it (for access speed), and price varies by region.

Redundancy:

- Azure keeps 3 copies of your data by default
- Locally or Zone redundant
- Azure will almost never lose a file once it is successfully received it.

We can also choose “globally redundancy” for storage. In this case:

- Azure keeps 6 copies of your data (3 locally and 3 in another region)
- Honors data sovereignty laws

Access Tiers for storage accounts

There are 4 access tiers:

- Hot: The default / balanced
- Cool: Cheaper storage, expensive read/write operations
- Cold: Cheaper storage, expensive read/write operations
- Archive: Can not get immediate access to files, cheapest storage, most expensive operations

Let's assume a scenario where you want to transfer data (container) from one storage account to another.

One approach to do this is download the data and then upload them to the other account. This is not an efficient method (Azure also charge for "Egress")

Another option is "AzCopy". This allows you to copy of your data to, from, or between Azure storage accounts.

"Azure Storage Explore" is also an option and it manages Azure Storage resources from your desktop.

Azure File

Another type of cloud storage. This is a true hierarchical structure as you would see in operating system folders.

Why we need this instead of containers?

- This will replace on premises file storage with the cloud
- This is also called "List and shift migration". No need to make any changes how they operate on current architecture / cloud
- Additional benefits: redundancy, data recovery, failover benefits

There is also an option called "Azure File Sync". This is a hybrid option (on premises files with cloud options). This is known as "file tiering" (most popular files are stored locally and all files are stored in cloud). This enables distributed access (can access from anywhere) and backup options.

Azure Migrate

This is a tool set to assess and move on premises resources into Azure.

This includes multiple migration tools: Server migration, database migration, web app migrations (not only virtual machines, physical machines as well)

Before migrating, we check (research) which servers can be migrated as-is and which need upgrades (cost estimates, batch planning, streamline migration)

Key words: discover, asses, and migrate

- discover: The tool help to discover servers and databases in your environment

Data Box

Another issue to consider when moving cloud is amount of data (how to get these into the cloud). In some case it is impractical to upload some quantities of data into the cloud. To address this, Azure has “Data Box” (Data Box, Data Box Disk, Data Box Heavy). Azure will mail this to you, and we can mail it back with data. In these, data are encrypted and wiped-clean after use.