**Implement and manage storage in Azure**

Azure storage, azure account

Can be created using

* Portal
* Cli
* Powershell
* Management client libraries

However, it is usual done as a one-time operation. Portal is the most common choice

Basically, a container that groups a set of storage services together. These are:

* Azure blob
* Azure files
* Azure queues
* Azure tables

Storage account settings

Has polies that applies to all storage accounts.

* Subscription
* Location
* Performance, standard and premium offers different storage services
* Replication,
* Access tier, how quickly you’re able to access the storage data
* Secure transfer, determines protocols for access
* Virtual networks, allows inbound access requests from virtual networks that you specify

Data diversity

* Depending on some settings will determine how your storage account will be stored like 2 different regions might need 2 storage accounts
* Another might be data is proprietary to the public.
* Increased diversity meaning more storage

Cost diversity

* Multiple storage accounts can be used to reduce costs
* Settings for storage account can affect cost.
* Premium performance and hot access increases blob costs

Storage account settings

* Name; unique mix of lowercase and digits
* Deployment model, recommended use is ‘resource manager’
* Account kind, like a set of policies that determines which data services you can include and the pricing

Account kind

* **storageV2, supports all kinds pf storage with latest features**
* Storage V1(general purpose), supports all storage but not all features
* Blob, legacy kind

Recommended storage option is to use ‘Resource manager’ to create storage account and ‘Storage V2’ for option

**Choose the right disk storage**

Disk roles

* OS disk; usually in creation of VM. Storage of operating system max 4tb
* Data disk; can add one or more data disks. Used to store files such as database files, web, custom code
* Temporary disk, each vm has one and is used for swap files, page swaps. Usually lost in maintenance and not recommended to use for critical data

Ephemeral Disk

* saves data on a local storage
* Faster read/write
* Free at no cost
* Vm failure can cause all loss of data
* Faster to reset image to original boot state if using this disk

Managed Disk

* A virtual disk in which the physical infrastructure is managed by Azure
* Easy to manage and attach to VM
* Pros include
* Support for azure backup
* High scalability
* Simple scalability

Unmanaged Disk

* Created and managed manually
* Can be converted to managed disk using powershell

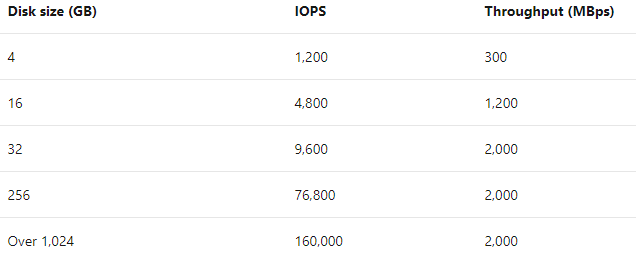


Disk performance

Measures

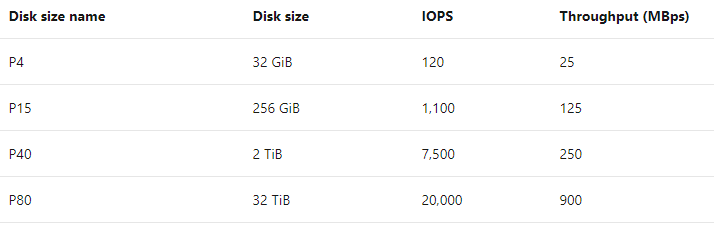
* Input/output; how fast read/write operations can be performed. IOPS
* Throughput; aka data transfer

Ultra-SSD; 300 – 2000mbps



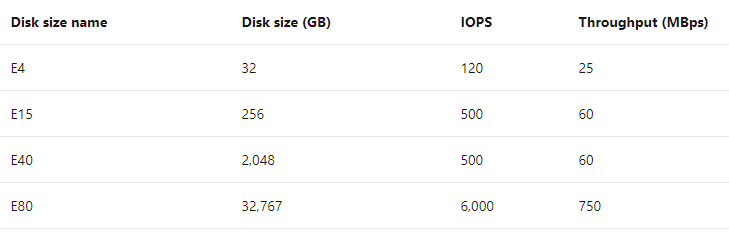
* High disk performance
* available is certain azure regions
* Can only be attached to vms available in those zones
* ES/DS v3 machines
* Only created as empty disks and used as **data disk**
* No disk support for snapshots, azure backup, site recovery

Premium SSD; 25 – 900mbps



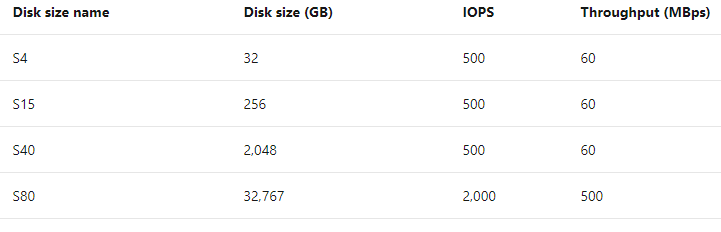
* Good for medium to large organizations
* Can migrate to premium, cannot adjust without detaching from vm
* Can be used with VMS outside of availability zones

Standard SSD



* Suitable for web servers and light enterprise apps can run on standard SSD

Standard HDD; 60 – 500mbps



* less critical workloads and to minimize costs and development or test environments

**Disaster Recovery for Data**

Azure storage allows for data replications across different regions for increased availability

Options include

|  |  |
| --- | --- |
| LRS – locally redundant storage | copies data across separate rack 3 times inside one region. If there is a datacentre failure, you could lose your data |
| Geographically redundant storage (GRS); | copies 3 times within region and 3 times in secondary region |
| Read-access geo-redundant storage (RGS), | secondary region is only available if main region is down. |
| Zone redundant storage (ZRS), | copies data in 3 clusters in a single region. If one datacentre is experiencing outage, then its available in another availability zone in the same region |
| Geo zone redundant (GZRS), | combines ZRS with GRZ. Copied across 3 availability zones in one region and 3 times to a secondary region |
| Read-access geo zone redundant storage (RAGZRS); | same as GZRZ but can read on secondary region |

* Paired regions; when you pair with another region to protect against regional outage

Switch replication

* Can switch replication depending on replication strategy

Convert account

* Can convert account using Portal, CLI, or Azure powershell

Live migration

* Done to avoid downtime or data loss but depends on how much data there is
* Can only be migrated to the same region
* If contains large files, then live migration to GZRS is not supported

Manual migration

* More flexible compared to live migration and can control the timing
* Can be done using ‘az copy’

**Failover Process**

How a system responds to a failure, by using the replicated data

1. Users use primary region when its up and running while data is asynchronously being copied to secondary
2. When primary region fails, users cannot access data in primary region and asynchronous replication stops
3. Failover then occurs and can be started by you or Microsoft. During process azure points users to secondary region instead of primary
4. When failover process is complete, your storage account is then updated to use LRS(locally redundant storage)

Data loss

* There isn’t a guarantee that all your data will be synced to the failover region. This is due to asynchronous replication having a delay in between
* When failover starts, primary region data is lost. Any data that wasn’t written to secondary before start of failover will be lost
* Use ‘last sync time’ to check how much data was lost after failover was started
* Check ‘last sync time’ when you select ‘prepare for failover’ option in portal

**Secure Storage Account**

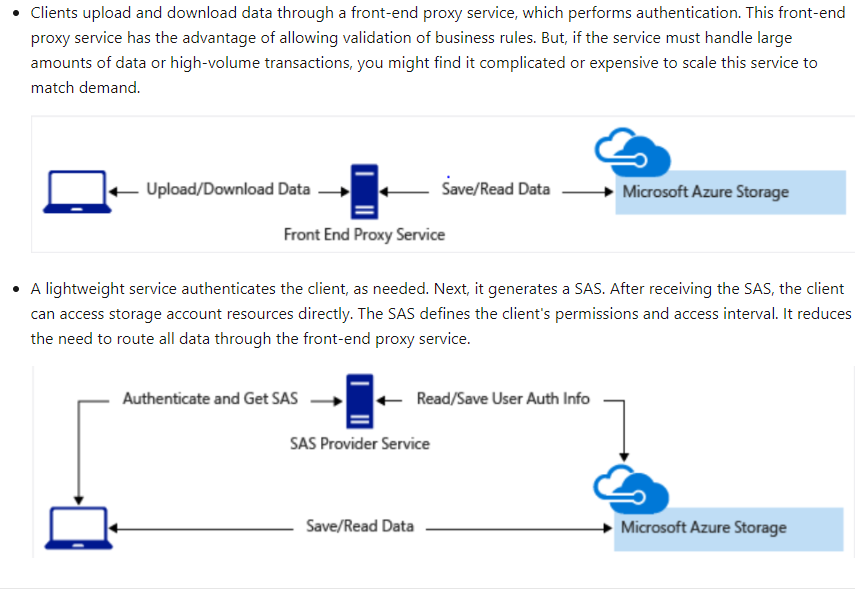
Benefits

|  |  |
| --- | --- |
| Protect data at rest | all data written to storage is automatically encrypted |
| Protects in transit | Encryption over data transfer. Always use HTTPS when communicating over public internet. |
| Supports browser cross-domain access | CORS with HTTPS allows cross domain access to services from another domain. This ensures that using CORS, web apps only load authorize content from authorized sources. |
| Control who can access data (RBAC) |  |
| Audit storage access |  |

**Storage account keys, best solution to control access to blog storage and queues**

* 2 keys primary and secondary
* Only regenerate periodically
* Change each trusted app to use secondary key
* Refresh primary key in the azure portal in case of keys being lost or stolen

**Shared access signatures, for untrusted clients**



Advanced threat protection

Azure defender

* Used as an extra layer of security to detect unusual activity and potentially harmful attempts at access to storage accounts
* Available for blob storage, azure files, and azure data lake storage.

When storage anomalies occur, you can receive an email notification with detailed information about it such as

* Storage acc name
* Nature of anomaly
* Event time
* Storage type
* Potential cause
* Steps for solution

Azure data lake storage security features

* Built on blog storage with all security features
* Has RBAC and provides ACL (Access control lists)
* Authenticates using azure active directory along with multifactor authentication

**Store and share file with Azure files**

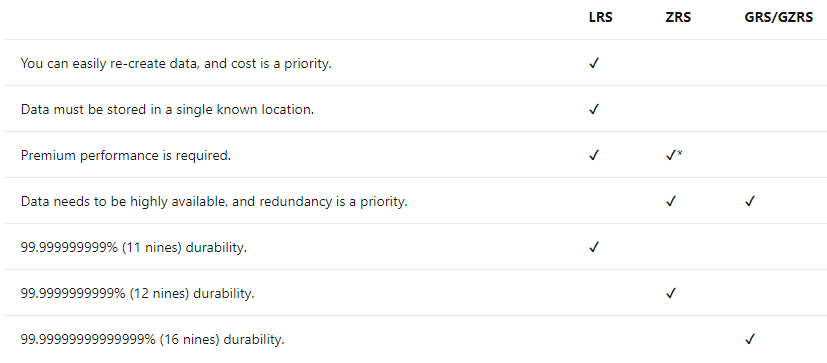
* Cloud based file share for storing and sharing files
* Think of it as a standard file share but hosted on azure
* Can be used to replace company NAS, network attached devices
* All data is encrypted in transit by using HTTPS
* Share files is cross platform
* Automate creation and editing of file shares using Powershell or CLI commands

Data access method

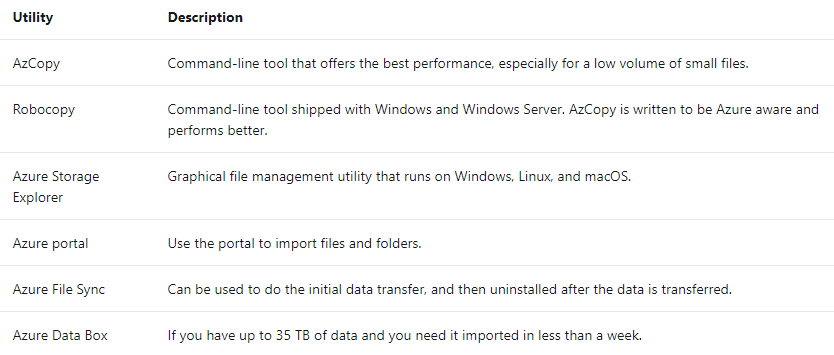
* **Direct access** via mounted drive. Use windows server and install on premises or Azure file to sync files between local shares and azure files.
* Azure file sync, or compute resources locally, slow internet connect it increases the need to run file sync
* **Direct cloud access** where apps are running on Azure and all locations that need access must have fast internet.

File redundancy option

* Standard, low cost low performance
* Premium, higher cost high performance but only available for ZRS and limited in some regions



Data migration Solution



Authentication for file share

* Azure active directory
* Shared key

**Secure access to files stored in Azure files**

* Only supports SMB 3.0, if versions are earlier, it needs to be disabled or uninstalled
* Needs communication through port 445
* Configure IP based firewalls

Enable and use Azure AD DS authentication

* Because the benefit is that it enables role-based access control
* Enable Azure AD DS for the tenant and have the VM in the same network

SMB file share permission levels

* Share reader
* Share contributor
* Share elevated contributor

**Upload/download using azure storage explorer**

* Is a GUI app developed to make management of data in storage account easier to use
* Available in most OS platforms
* Easy to connect and manage multiple storage accounts
* Can use interface to update and view entities in storage accounts
* Free to download and use
* Easy to manage across multiple subscriptions and storage accounts
* Can use to manage data stored in Azure Cosmos DB and Data lake

Local emulators

* Azure storage emulator, a local instance of MS sql server 2012. Emulates Azure Table, Queue, and blob storage
* Azurite, based on node.js, an open source emulator that supports most azure storage commands through API

Connecting to storage explorer

* Before connecting to azure storage explorer, you need 2 permissions
* Storage explorer with data-layer permission allows you to be granted at minimum a read data role. Data layer is used to access blobs, containers, and other data resources
* Management role; grants access to see lists of various storage accounts, containers, and service endpoints
* Connect using storage account key and name
* Connect using shared access signature URI

Add account using Azure AD

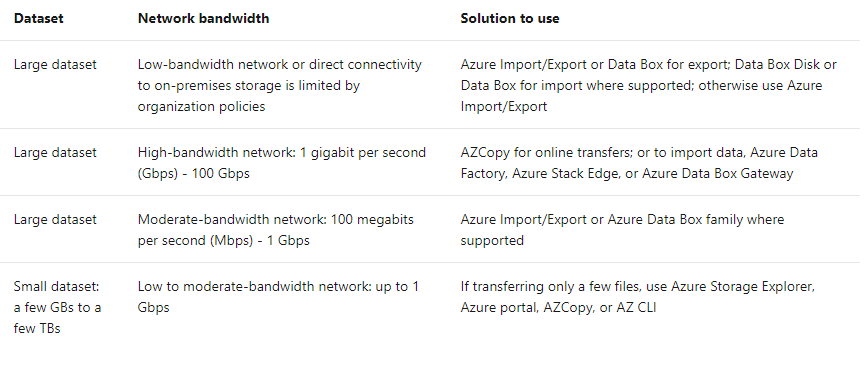
* Only uses this when the user can access the data layer
* Requires more config than other methods
* Account used must have correct permission and authorization to access resources

**Azure Import/Export Service**

* Transfers large amounts of data between on-premises storage and azure storage account without tying up network bandwidth
* Import; writes data to disks then sends to azure datacentre where Microsoft copies the data then sends back to you
* Export; prepares blank disks and then same process as import
* Only works in azure blog storage, not azure files

for the import/export process to work you must

* Active azure sub and an account holding azure blog storage
* Supported system running windows
* Bitlocker enabled
* WAImport version 1 downloaded and installed from Microsoft data center
* Same process required for import/export too



**Moving large amounts of data using Azure Data Box**

Offline

* Data box disk; 35B connect and copy over USB
* Data box; 80tb per order. Connect and copy using network device like SMB and NFS
* Data box heavy; 800TB, use high throughput network interfaces to connect and copy data

Online – like network share

* Data box gateway; based on virtual machine

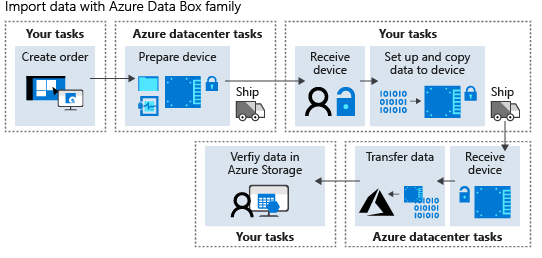
Security protections

Data box disk

* Tamper resistant and supports update
* AES 128 bit security
* Unlocked with storage key
* Protected by azure security features
* Once uploaded to azure it is then wiped clean

Data box heavy and data box

* Secured with AES 256 bit encryption
* Unlocked only with password from azure portal
* Protected by azure security features



Data transfer method

Az storage explorer

* Transfer files occasionally

Physical device

* Network bandwidth is low and expensive to use and has large volumes of data

Azure data box

* Convenience of using ms hardware

Network transfer

* If you have plenty of network bandwidth
* Regular transfer

Azure data factory

* If you want data to be transformed, changed without using code

Programmatic method

* If you want custom code and more control over transfer
* Powershell, azure CLI, az copy

**Azure File Sync**

Terminology

|  |  |
| --- | --- |
| Storage sync service | High level resource |
| Sync group | Topology for set of files and folder |
| Registered server | The trust relationship between on-premises server and storage sync service |
| Azure file sync agent | Downloadable package that enables windows server to be synced with az file share |
| Server endpoint | Specified location on registered server like a file or local disk |
| Cloud endpoint | File share that’s part of a sync group |
| Cloud tiering | Optional feature that allows frequently accessed files to be cached locally |

How it works

* Azure file sync agent is installed on the on-premises location and communicates with the azure sync service
* Uses USN journaling on the windows server pc to auto start the sync session when there is a file change on-premises
* When changes are made and detected it’s then replicated to the azure file share
* Azure file changes are not visible for 24 hours

Deployment process

1. Check if on-premises server OS and server can support azure file sync
2. Create storage account and must contain file share, storage sync service, and a sync group
3. Install file sync agent
4. Register windows server pc with storage sync. After installing sync agent you’re prompted to register the sync agent
5. Created endpoint on server pc

Evaluation

* Azure file sync has system requirements before you start

Possible problems

* Antivirus
* Backup
* Encryption

System requirements

* Windows server 2012, 2016, 2019 in either datacentre or standard edition
* 2gm or more
* Latest windows patches
* NTFS file storage

Common Problems

* Cant mount file share; firewall issue like port 445 is disabled
* Files or directories won’t sync; invalid file names
* Use azure monitor and view event logs to check status of files and to troubleshoot possible problems
* Performance monitor; to monitor the windows server

RDP – remote desktop connection

* Connection to the UI of windows desktop computers
* Sign in to remote physical or virtual machine
* Requires an RDP client