



Azure Storage

Microsoft Services



Agenda

- Storage Overview
- Storage Services
- Storage Replication
- Storage Security

Azure Storage Overview

- Azure Storage is a scalable, durable, and highly available storage solution.
- Uses an auto-partitioning system that automatically load-balances your data based on traffic.
- Is accessible from anywhere in the world, from any type of application, whether it's running in the cloud, on a desktop, on an on-premises server, or on a mobile or tablet device.
- Supports clients using a diverse set of operating systems (including Windows and Linux) and a variety of programming languages i.e. .NET, Java, Node.js, Python, Ruby, PHP and C++.





Storage Services

Microsoft Services



Azure Storage Services

- Azure storage provides the following five services: Blob storage, Table storage, Queue storage, File storage and Disk storage.
- **Blob Storage** stores **unstructured object data or Blobs** and can be any type of text or binary data, such as a document, media file, or application installer.
- **Table Storage** stores structured datasets and is a NoSQL key-attribute data store, which allows for rapid development and fast access to large quantities of data.
- **Queue Storage** provides reliable messaging for workflow processing and for communication between components of cloud services.
- **File Storage** offers shared storage for legacy applications using the standard SMB protocol.
- **Disk Storage** provides persistent VHD disk storage without the overhead of managing a storage account.

Azure Storage Services

IaaS



Storage



Virtual
machines



Networking

PaaS



Existing
frameworks



Web
and mobile



Microservices



Serverless
Compute

Disks

Persistent disks for
Azure IaaS VMs

Premium Storage Disks
option: SSD based, high
IOPS, low latency

Files

Fully Managed File
Shares in the Cloud

SMB and REST access
“Lift and shift” legacy
apps

Blobs

Highly scalable, REST
based cloud object
store

Block Blobs: Sequential
file I/O
Cool Tier Available
Page Blobs: Random-
write pattern data
Append Blobs

Tables

Massive auto-scaling
NoSQL store

Dynamic scaling based
on load
Scale to PBs of table data
Fast key/value lookups

Queues

Reliable queues at
scale for cloud
services

Decouple and scale
components
Message visibility
timeout and update
message to protect
against unreliable
dequeueers

Built on a unified Distributed Storage System

Durability, Encryption at Rest, Strongly Consistent Replication, Fault Tolerance, Auto Load-Balancing

Azure Storage & Data Services

Unstructured Data

Blobs

Highly scalable, REST based cloud object store

Data Lake Store

HDFS as a service

mongoDB

Elastic scale
Cross platform

Files

Fully Managed File Shares in the Cloud

Queues

FIFO async messaging

Disks

Virtual Machine VHD files

Structured Data

DocumentDB

NoSQL document database service

Azure SQL DB

Fully managed database-as-a-service built on SQL

SQL Data Warehouse

Elastic data warehouse as a service

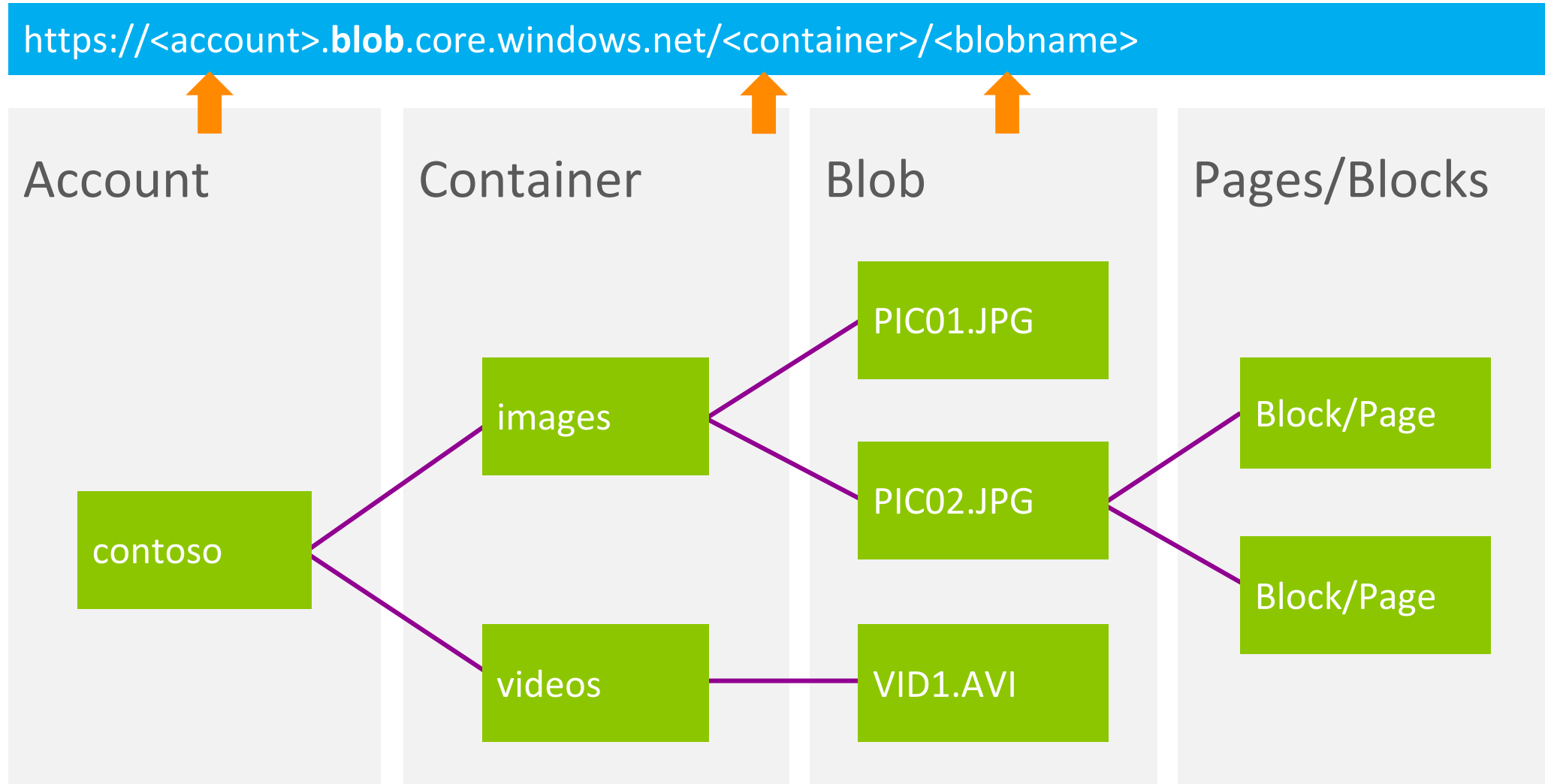
Tables

Key Value, high scale, auto-scaling NoSQL store

Blob Storage

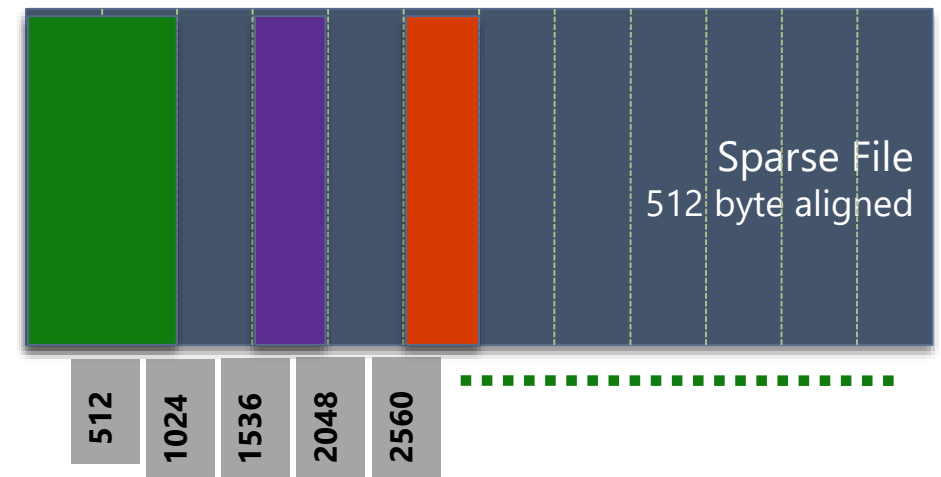
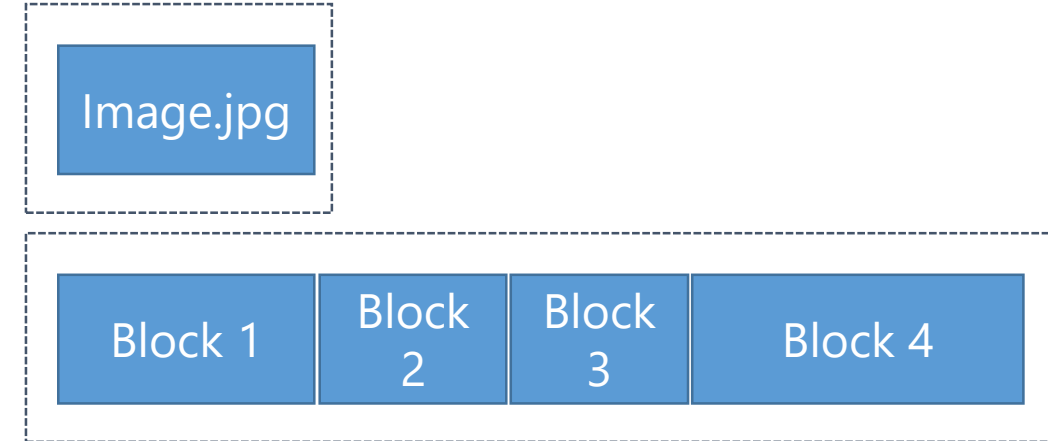
- Azure Blob storage is a service that stores unstructured data in the cloud as objects or blobs.
- Blob storage can store any type of text or binary data, such as a document, media file, or application installer.
- Blob storage is also referred to as object storage.
- Azure Storage offers three types of blobs: **block blobs**, **page blobs**, and **append blobs**.

Blob Storage Concepts



Blob Types

- Block blob
 - Targeted at streaming workloads or individual file uploads
 - Each blob consists of a sequence of blocks
 - Each block is identified by a Block ID
 - Size limit of **4.7 TB** per blob
 - Optimistic concurrency via Entity Tags (ETags)
 - Optimistic or pessimistic (locking) concurrency via leases
 - Manage leases from Azure portal
- Page blob
 - Targeted at random read/write workloads
 - Each blob consists of an array of pages
 - Each page is identified by its offset from the start of the blob
 - Size limit of **4 TB** per blob
 - Optimistic concurrency via Entity Tags (ETags)
 - Optimistic or pessimistic (locking) concurrency via leases
 - Manage leases from Azure portal



Blob Types

- Append Blob
 - An append blob is comprised of blocks and is optimized for append operations
 - When you modify an append blob blocks are added to the end of the blob by the Append Block operation
 - Updating or deleting of existing blocks is not supported
 - Does not expose its block IDs
 - Each block in an append blob can be a different size, up to a maximum of 4 MB and can include up to 50,000 blocks
 - The maximum size of an append blob is 4 MB X 50,000 blocks
 - Optimistic concurrency via Entity Tags (ETags)
 - Optimistic or pessimistic (locking) concurrency via leases
 - Manage leases from Azure portal

Blob Names

<https://contoso.blob.core.windows.net/vhds/OSDisk.vhd>

- Account Name: 3-24 characters, lower case only
- Container Name: 3-63 characters, lower case only
- Blob Name: 1-1024 characters, case sensitive
- Virtual directories within blob namespace

Hot & Cool Storage Tiers

- Azure Blob storage offers two storage tiers for object storage, **hot** and **cool** storage.
- Hot storage is optimized for storing data that is frequently accessed.
- Cool storage is optimized for storing data that is infrequently accessed and long-lived.
- Available for Blob storage accounts only.

Hot & Cool Storage Tiers

New tier for Blob storage

- For high volume infrequently accessed data
- Same API and durability; similar latency

Pricing to match workload

- Hot: Lower access costs
- Cool: Lower price per GB

Switch account tiers as needed

- No charge for Hot to Cool switch

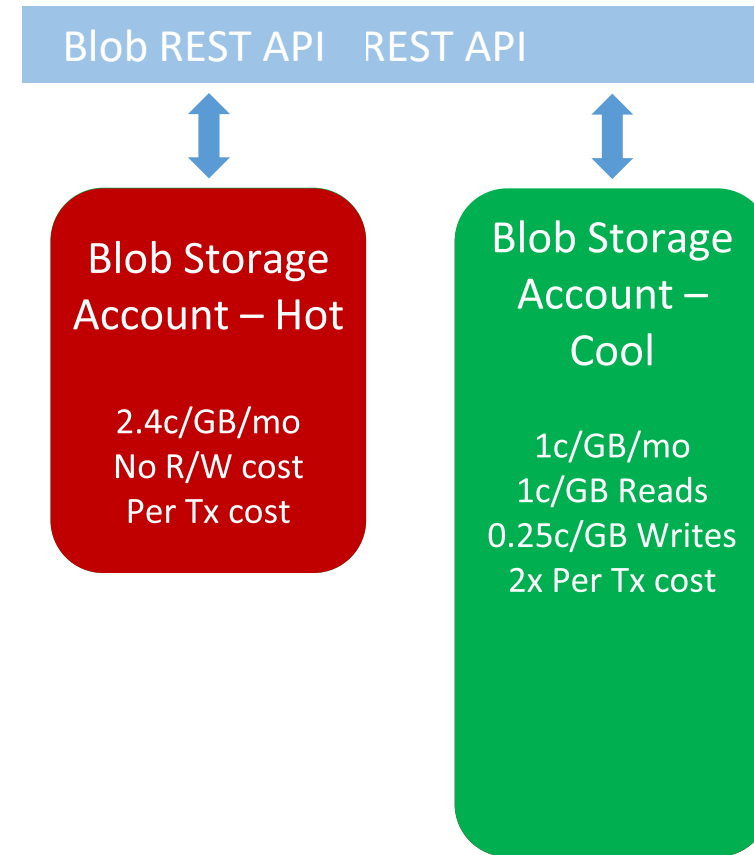
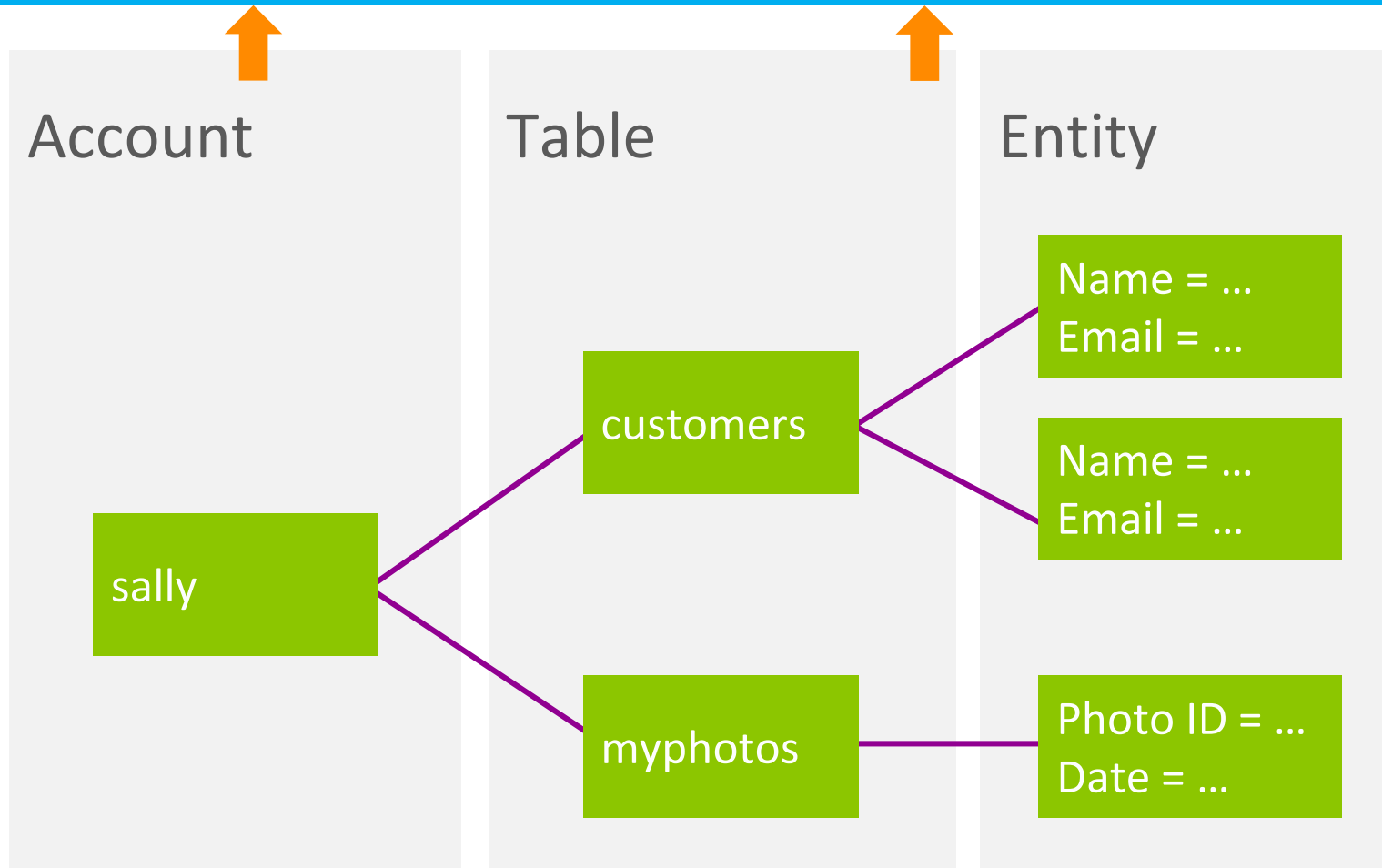


Table Storage

- Azure Table storage is a service that stores large amounts of structured data in the cloud as entities within a table.
- Table storage contains the following components:
 - **URL format:** `https://<storage account>.table.core.windows.net/<table>` to access the tables.
 - **Storage Account:** All access to Azure Storage is done through a storage account.
 - **Table:** A table is a collection of entities.
 - **Entity:** An entity is a set of properties, similar to a database row and can be up to 1MB in size.
 - **Property:** A property is a name-value pair and each entity can include up to 252 properties to store data.

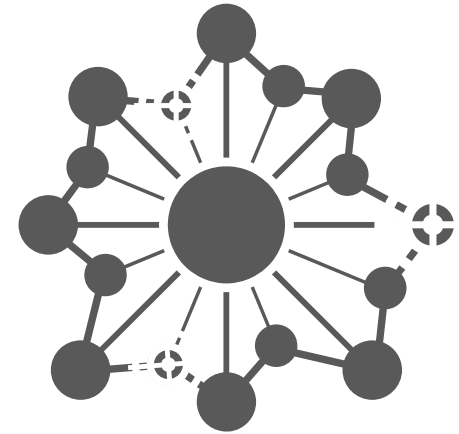
Table Storage Concepts

<https://<storage account>.table.core.windows.net/<table>>



Entity Properties

- Entity can have up to 252 properties
 - Up to **1 MB** per entity
- Mandatory Properties for every entity
 - PartitionKey and RowKey (only indexed properties)
 - Uniquely identifies an entity
 - Defines the sort order
 - Timestamp
 - Optimistic concurrency
 - Exposed as an HTTP eTag
- No fixed schema for other properties
 - Each property is stored as a <name, typed value> pair
 - No schema stored for a table
 - Properties can be the standard .NET types
 - String, binary, bool, DateTime, GUID, int, int64, and double2



No Fixed Schema



FIRST	LAST	BIRTHDATE	FAV SPORT
Wade	Wegner	2/2/1981	
Nathan	Totten	3/15/1965	
Nick	Harris	May 1, 1976	

Queue Storage

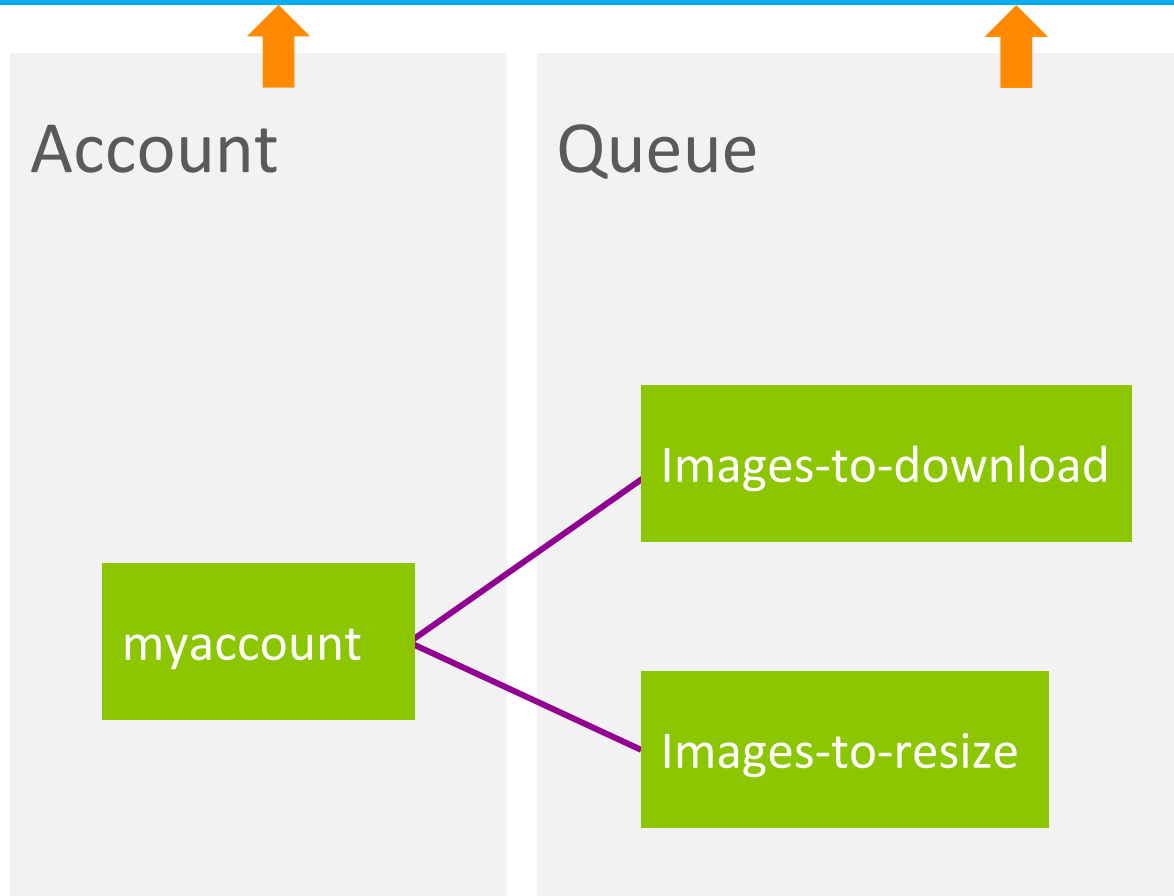
- Azure Queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS.
- A single queue message can be up to 64 KB in size, and a queue can contain millions of messages, up to the total capacity limit of a storage account.
- Asynchronous messaging for communication between application components, whether they are running in the cloud, on a desktop, on an on-premises server, or on a mobile device.

Queue Storage

- Queue storage contains the following components:
 - **URL format:** `https://<storage account>.queue.core.windows.net/<queue>` to access the tables.
 - **Storage Account:** All access to Azure Storage is done through a storage account.
 - **Queue:** A queue contains a set of messages, all messages must be in a queue and queue names must be all lowercase.
 - **Message:** A message, in any format, of up to 64 KB, the maximum time that a message can remain in the queue is 7 days.

Queue Storage Concepts

<https://<storage account>.queue.core.windows.net/<queue>>



<https://myaccount.queue.core.windows.net/images-to-download>

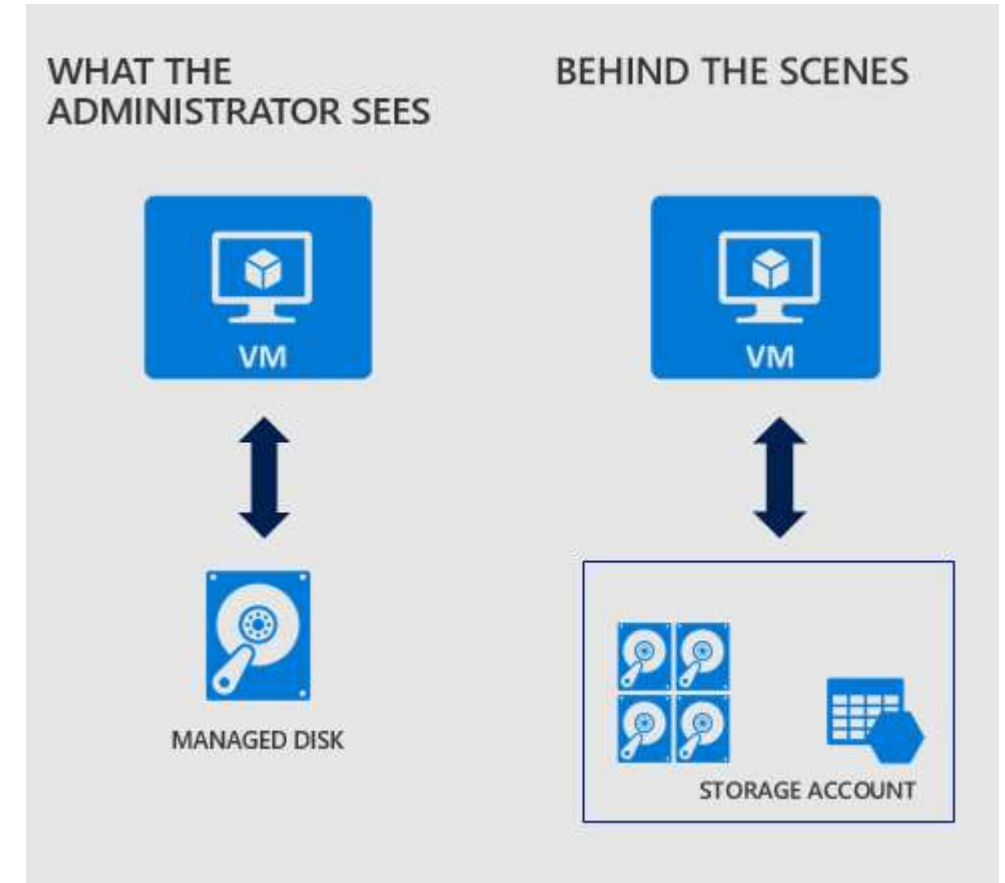
Disk Storage

- **Unmanaged Disks:** This is the initial storage model where you manage the storage accounts that are used to store the VHD files that correspond to your VM disks.
- **Managed Disks:** This is the new storage model where Microsoft manages the storage accounts that are used to store the VHD files that correspond to your VM disks.



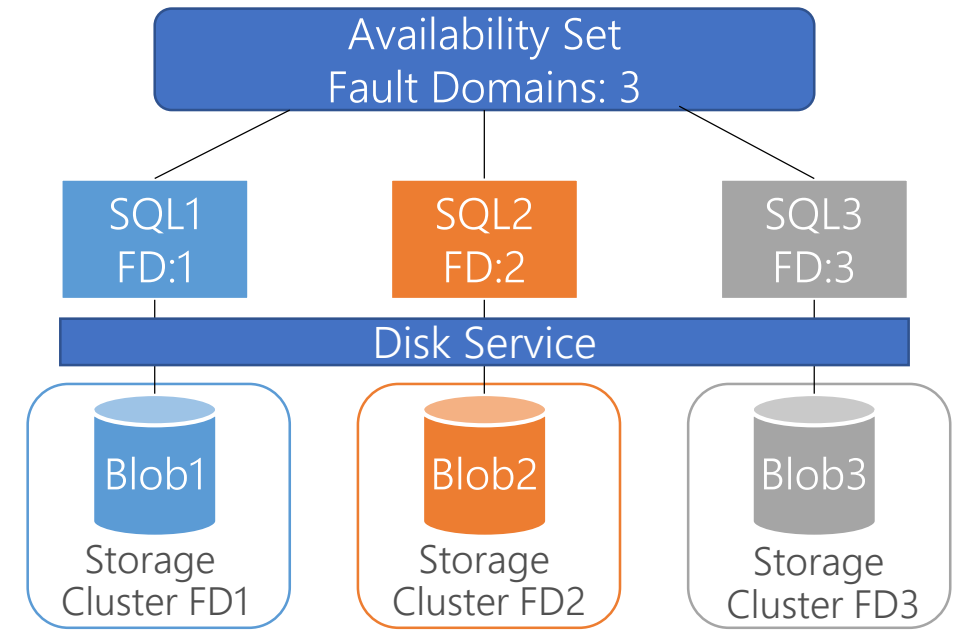
Managed Disks

- What are Azure Managed Disks?
 - Azure Managed Disks are VHD's that are stored in a Microsoft managed storage account.
- Administrators do not have access to the managed disk storage account.
- **Note:** Managed disks are not a replacement for other storage account services i.e. Blobs, Tables & Queues.



Managed Disks Benefits

- Simple and scalable VM deployment
 - No 40 disk or 20 000 IOPs per storage account limit.
 - Up to 10 000 disks per subscription.
 - Up to 1000 VM Scaleset images.
 - No storage account management.
- Better reliability for Availability Sets
 - Ensures that the disks of VMs in an Availability Set are sufficiently isolated from each other to avoid single points of failure.
- Better security
 - Use Azure Role-Based Access Control (RBAC) to assign specific permissions for a managed disk to one or more users.
 - Supports granular permissions.
- Supports a read-only shared access signature (SAS)



Managed Disks Performance Tiers

- Managed Disks offers 2 performance tiers: Premium (SSD-based) and Standard (HDD-based).

Premium Managed Disk Type	P4	P6	P10	P20	P30	P40	P50
Disk Size	32 GB	64 GB	128 GB	512 GB	1024 GB (1 TB)	2048 GB (2 TB)	4095 GB (4 TB)
Standard Managed Disk Type	S4	S6	S10	S20	S30	S40	S50
Disk Size	32 GB	64 GB	128 GB	512 GB	1024 GB (1 TB)	2048 GB (2 TB)	4095 GB (4 TB)

Managed Disks Images & Snapshots

- **Images** is a feature that allows you to capture, in a single image, all managed disks associated with a running VM.
 - You can create an image from your custom VHD in a storage account or directly from a running VM.
- A **Managed Snapshot** is a read-only copy of a managed disk which is stored as a standard managed disk.
 - With snapshots, you can back up your managed disks at any point in time.
 - These snapshots exist independent of the source disk and can be used to create new Managed Disks or attach to a new VM.
- **Azure Backup service** can also be used with Managed Disks to create a backup job with time-based backups, easy VM restoration and backup retention policies.

NAME ▾	SOURCE DISK
 MDDC1-NTDS	Data1
 MDDC1-OS	MDDC1

Managed Disks Pricing

- When using Managed Disks, the following billing considerations apply:
- **Storage Type** – Billing of a managed disk depends on which type of storage you have selected for the disk
- **Disk Size** - Azure maps the provisioned size rounded to the nearest Managed Disks option
- **Number of transactions** - Billed for the number of transactions performed on a standard managed disk
- **Outbound data transfers** - Data going out of Azure data centers incur billing for bandwidth usage
- **Managed Disk Snapshots** (full disk copy) - The cost of a managed snapshot is the same as that for standard managed disk

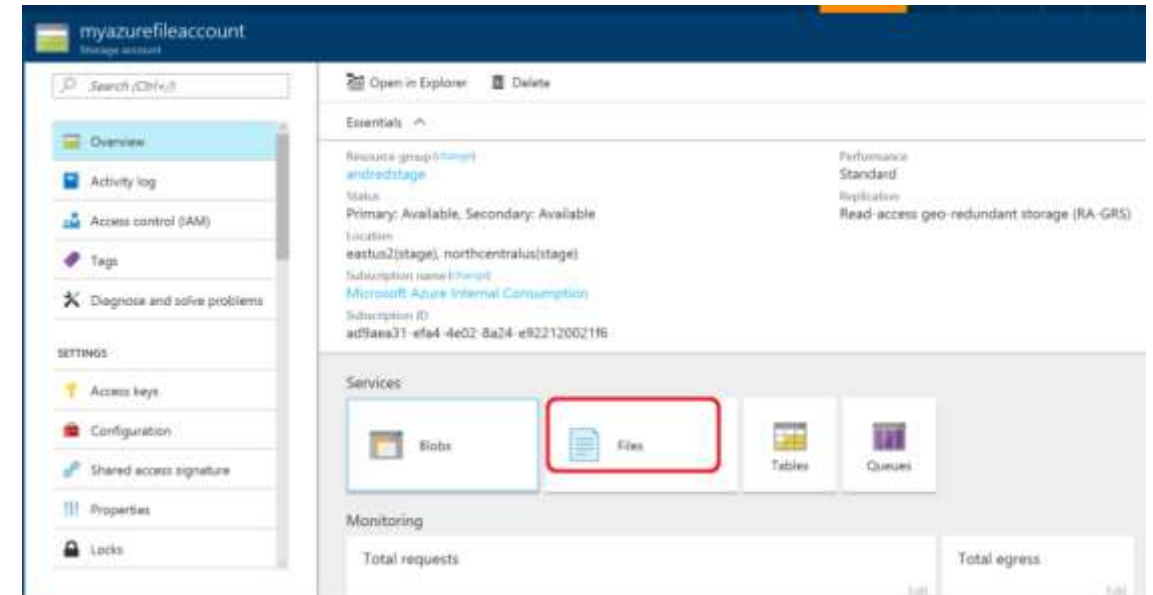
Managed Disks Migration

- You can migrate to Managed Disks in following scenarios:

Migrate...	Documentation link
Convert stand alone VMs and VMs in an availability set to managed disks	Convert VMs to use managed disks
A single VM from classic to Resource Manager on managed disks	Migrate a single VM
All the VMs in a vNet from classic to Resource Manager on managed disks	Migrate IaaS resources from classic to Resource Manager and then Convert a VM from unmanaged disks to managed disks

Azure Files

- Azure Files offers fully managed file shares in the cloud that are accessible via SMB
- Can be mounted concurrently by cloud or on-premises deployments of Windows, Linux, and macOS via the Internet
- Can be cached on Windows Servers with Azure File Sync (preview) for fast access near where the data is being used
- Support for file share snapshots (Incremental)
- Maximum of 200 snapshots

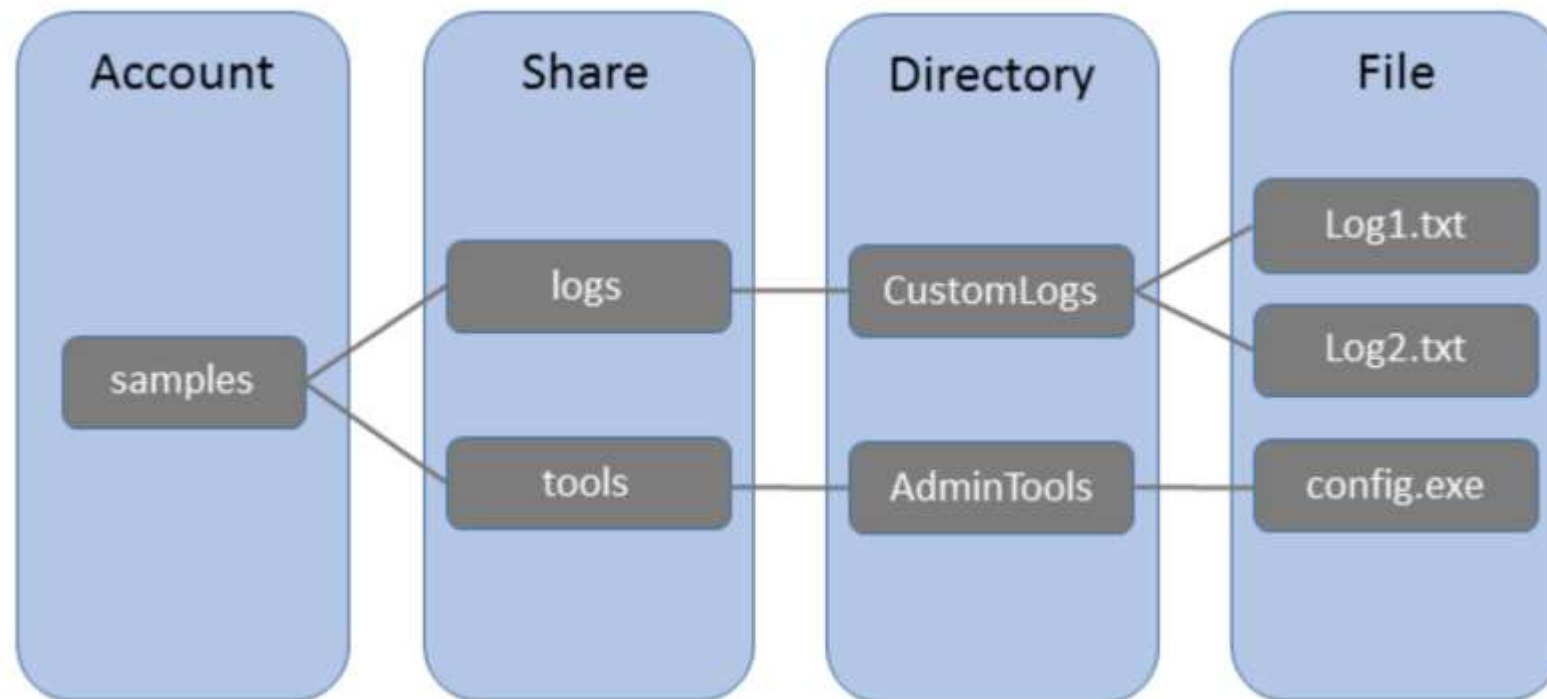


Azure Files Benefits

- **Shared access** - Using the SMB protocol you can seamlessly replace your on-premises file shares with Azure File shares without worrying about application compatibility.
- **Fully managed** - Azure File shares can be created without the need to manage hardware or an OS.
- **Scripting and tooling** - PowerShell cmdlets and Azure CLI can be used to create, mount, and manage Azure File shares as part of the administration of Azure applications.
- **Resiliency** - Azure Files has been built from the ground up to be always available.
- **Familiar programmability** - Applications running in Azure can access data in the share via file system I/O APIs.

Azure Files Concepts

- A share can have multiple directories
- All directories and files must be created in a parent share
- An account can contain an unlimited number of shares, and a share can store an unlimited number of files, up to the capacity limits of 5TB



Azure Files vs Blobs

Description	Azure Blobs	Azure Files
Durability Options	LRS, ZRS, GRS (and RA-GRS for higher availability)	LRS, GRS
Accessibility	REST APIs	SMB 2.1/3.0 (standard file system APIs) REST APIs
Connectivity	REST – Worldwide	SMB 2.1 - Within region REST – Worldwide
Endpoints	https://myaccount.blob.core.windows.net/mycontainer/myblob	\\myaccount.file.core.windows.net\myshare\myfile.txt https://myaccount.file.core.windows.net/myshare/myfile.txt
Directories	Flat namespace however prefix listing can simulate virtual directories	True directory objects
Case Sensitivity of Names	Case sensitive	Case insensitive, but case preserving
Capacity	Up to 500TB containers	Up to 5TB of files
Throughput	Up to 60 MB/s per blob	Up to 60 MB/s per share
Object size	Up to 4 TB/blob	Up to 1 TB/file
Billed capacity	Based on bytes written	Based on file size

Azure Files vs Disk

Description	Disk	Azure Files
Relationship with Azure VMs	Required for booting (OS Disk)	
Scope	Exclusive/Isolated to a single VM	Shared access across multiple VMs and also on-premises
Snapshots and Copy	Yes	No
Configuration	Configured via portal/Management APIs and available at boot time	Connect after boot (via net use on windows)
Built-in authentication	Built-in authentication	Set up authentication on net use
Cleanup	Resources can be cleaned up with VM if needed	Manually via standard file APIs or REST APIs
Access via REST	Can only access as fixed formatted VHD (single blob) via REST. Files stored in VHD cannot be accessed via REST.	Individual files stored in share are accessible via REST
Max Size	4TB Disk	5TB File Share 1TB file within share
Max 8KB IOps	500 IOps (Basic Storage)	1000 IOps
Throughput	Up to 60 MB/s per Disk	Up to 60 MB/s per File Share

Azure Files – Windows Client OS Support

When a client accesses Azure File Storage, the actual SMB version used will depend on the client OS being used.

Windows Version	SMB Version	Mountable in Azure VM	Mountable On-Premises
Windows Server semi-annual channel ¹	SMB 3.0	Yes	Yes
Windows 10 ²	SMB 3.0	Yes	Yes
Windows Server 2016	SMB 3.0	Yes	Yes
Windows 8.1	SMB 3.0	Yes	Yes
Windows Server 2012 R2	SMB 3.0	Yes	Yes
Windows Server 2012	SMB 3.0	Yes	Yes
Windows 7	SMB 2.1	Yes	No
Windows Server 2008 R2	SMB 2.1	Yes	No

¹Windows Server version 1709.

²Windows 10 versions 1507, 1607, 1703, and 1709.

Azure Files – Linux Client OS Support

- Linux SMB client does not support encryption
- Mounting from Linux in a different region to the Azure File share requires SMB 3.0

Linux distributions	SMB Version Supported
Ubuntu Server 14.04	SMB 2.1 and 3.0
Ubuntu Server 15.04	SMB 2.1 and 3.0
CentOS 7.1	SMB 2.1 and 3.0
Open SUSE 13.2	SMB 2.1 and 3.0
SUSE Linux Enterprise Server 12	SMB 2.1 and 3.0
SUSE Linux Enterprise Server 12 (Premium Image)	SMB 2.1 and 3.0

Azure Storage Accounts

- Azure storage provides three types of storage accounts, **General Purpose v1**, **v2** and **Blob**.
- **General purpose v1** storage accounts give you access to Blobs, Tables, Queues, Files and Azure virtual machine disks under a single account and has two performance tiers, **Standard** and **Premium**:
 - **Standard** storage performance tier uses HDD disks and allows you to store Blobs, Tables, Queues, Files and Azure virtual machine disks.
 - **Premium** storage performance tier uses SSD disks which currently only supports Azure virtual machine disks.
- **General purpose v2** storage accounts give you all the features of **v1** plus Hot and Cool storage tiers
- **Blob** Storage Accounts are specialized storage accounts for storing unstructured data as blobs in Azure Storage, optimised for block or append blob storage, not page blobs.

Standard Storage Account

- 500 TB limit per storage account
- 200 storage accounts per region
- Up to 20,000 IOPS Per storage account and Up to 500 IOPS per VHD
- Encryption at Rest by default

Premium Storage Account

- Only supports Locally Redundant Storage (LRS)
- Must use B-series, DS-series, DSv2-series, DSv3-series, GS-series, Ls-series, M-series, and Fs-series VMs
- Cannot be mapped to a custom domain
- Storage analytics not currently supported
- No support for Block blobs, Append blobs, Azure Files, Azure Tables or Azure Queues only Page Blobs for Virtual Machines (aka VHD's)

Azure Premium Storage Scalability

Premium Disks Type	P4	P6	P10	P20	P30	P40	P50
Disk size	128 GB	512 GB	128 GB	512 GB	1024 GB (1 TB)	2048 GB (2 TB)	4095 GB (4 TB)
IOPS per disk	120	240	500	2300	5000	7500	7500
Throughput per disk	25 MB per second	50 MB per second	100 MB per second	150 MB per second	200 MB per second	250 MB per second	250 MB per second



Storage Replication

Microsoft Services



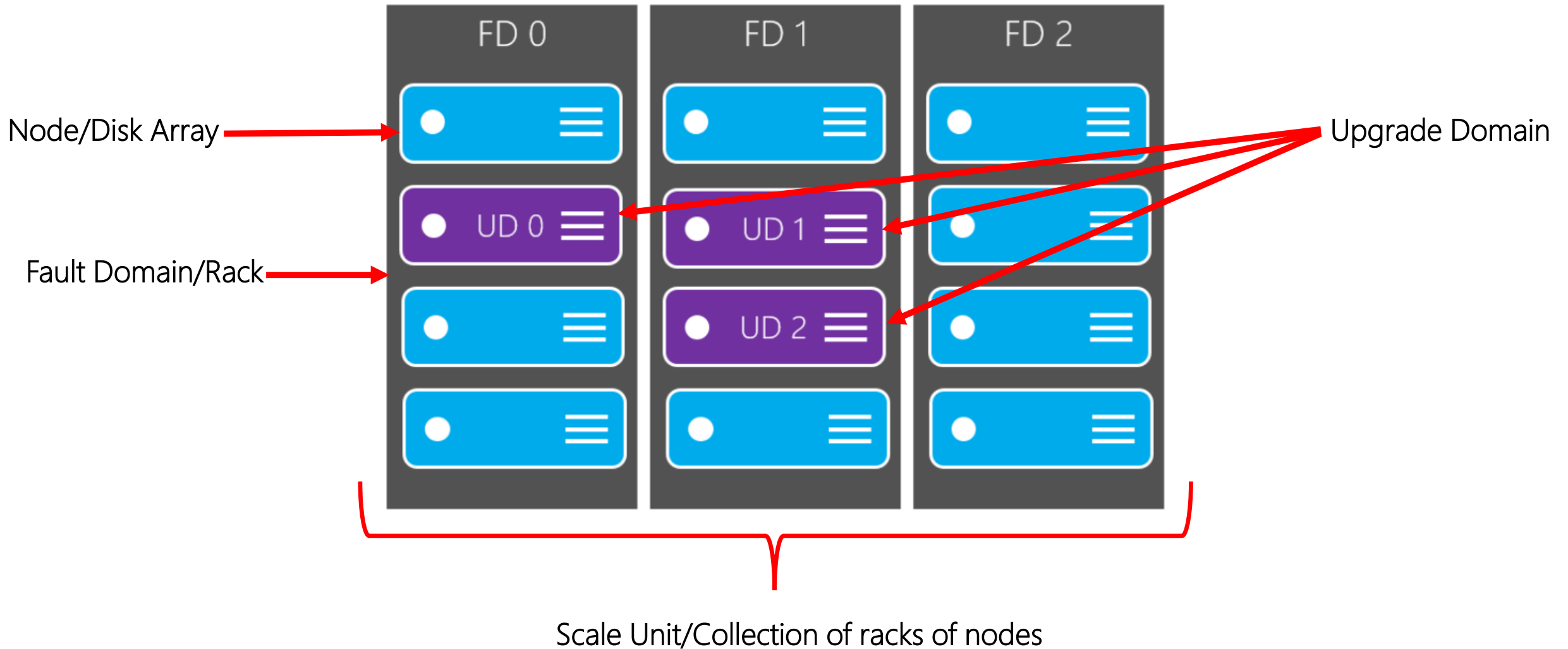
Storage Replication

- Data in a Microsoft Azure storage account is always replicated to ensure durability and high availability.
- Replication copies your data, either within the same data center, or to a second data center, depending on which replication option you choose.
- Replication protects your data and preserves your application up-time in the event of transient hardware failures.
- If your data is replicated to a second data center, that also protects your data against a catastrophic failure in the primary location.

Storage Replication Terminology

- **Storage Node** – A storage node is a disk array.
- **Scale Unit** - A storage scale unit is a collection of racks of storage nodes.
- A **Fault domain** (FD) is a group of nodes that represent a physical unit of failure and can be considered as nodes belonging to the same physical rack.
- An **Upgrade domain** (UD) is a group of nodes that are upgraded together during the process of a service upgrade or rollout.

Storage Replication Terminology



Storage Replication Types

- Locally redundant storage (LRS)
- Zone-redundant storage (ZRS)
- Geo-redundant storage (GRS)
- Read-access geo-redundant storage (RA-GRS)

Locally Redundant Storage (LRS)

- Locally redundant storage (LRS) replicates your data three times within a storage scale unit which is hosted in a datacenter in the region in which you created your storage account.
- A write request returns successfully only once it has been written to all three replicas.
- These three replicas each reside in separate fault domains and upgrade domains within one storage scale unit to ensure that data is available even if hardware failure impacts a single rack or when nodes are upgraded during a rollout.

Zone-Redundant Storage (ZRS)

- Zone-redundant storage (ZRS) replicates your data synchronously across datacenters within a region in addition to storing three replicas, providing higher durability than LRS.
- Data stored in ZRS is durable even if the primary datacenter is unavailable or unrecoverable.
- Available for block blobs, non-disk page blobs, files, tables, and queues in general purpose v2 storage accounts.

Geo-Redundant Storage (GRS)

- Geo-redundant storage (GRS) replicates your data to a secondary region that is hundreds of miles away from the primary region.
- If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region is not recoverable.
- An update is first committed to the primary region, where it is replicated three times, then the update is replicated asynchronously to the secondary region, where it is also replicated three times.
- With GRS, both the primary and secondary regions manage replicas across separate fault domains and upgrade domains within a storage scale unit.

Read-Access Geo-Redundant Storage (RA-GRS)

- Read-access Geo-Redundant storage (RA-GRS) provides read-only access to the data in the secondary location, in addition to the replication across two regions provided by GRS.
- Secondary endpoint is similar to the primary endpoint, but appends the suffix –secondary to the account name e.g. if your primary endpoint is **myaccount.blob.core.windows.net**, then your secondary endpoint is **myaccount-secondary.blob.core.windows.net**.
- The access keys for your storage account are the same for both the primary and secondary endpoints.



Storage Security

Microsoft Services

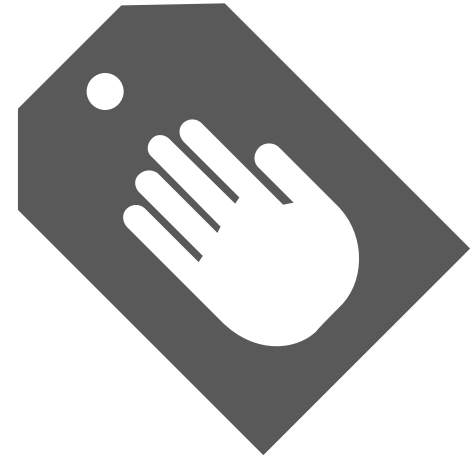


Storage Security

- Microsoft Azure Storage provides simple security for calls to storage service
 - HTTPS endpoint
 - Digitally sign requests for privileged operations
- Two 512-bit symmetric keys per storage account
 - Can be regenerated independently
- More granular security via Shared Access Signatures (SAS)
- Azure Storage does not authenticate users

Shared Access Signatures

- Fine grain access rights to storage entities (blobs/tables etc)
- Sign URL with storage key—permit elevated rights
- Revocation:
 - Use short time periods and re-issue
 - Use container-level policy that can be deleted
- Two broad approaches:
 - Ad hoc
 - Policy-based



Ad Hoc Signatures

- Create short-dated SAS
 - Signedresource Blob or Container
 - AccessPolicy Start, Expiry, and Permissions
 - Signature HMAC-SHA256 of above fields
- Use case
 - Single use URLs
 - For example, provide URL for the client to upload to container

```
https://...blob.../pics/image.jpg?  
sr=c&st=2009-02-09T08:20Z&se=2009-02-10T08:30Z&sp=w  
&sig= dD80ihBh5jfnpym05Hg1IdiJIEvHcJpCMiCMnN%2fRnbI%3d
```

Policy-Based Signatures

- Create container-level policy
 - Specify StartTime, ExpiryTime, and Permissions
 - Also created in the Azure Portal
- Create SAS URL
 - Signedresource Blob or Container
 - Signedidentifier optional pointer to container policy
 - Signature HMAC-SHA256 of above fields

```
https://...blob.../pics/image.jpg?  
sr=c&si=MyUploadPolicyForUserID12345  
&sig=dD80ihBh5jfnpym05Hg1IdiJIEvHcJpCMiCMnN%2fRnbI%3d
```

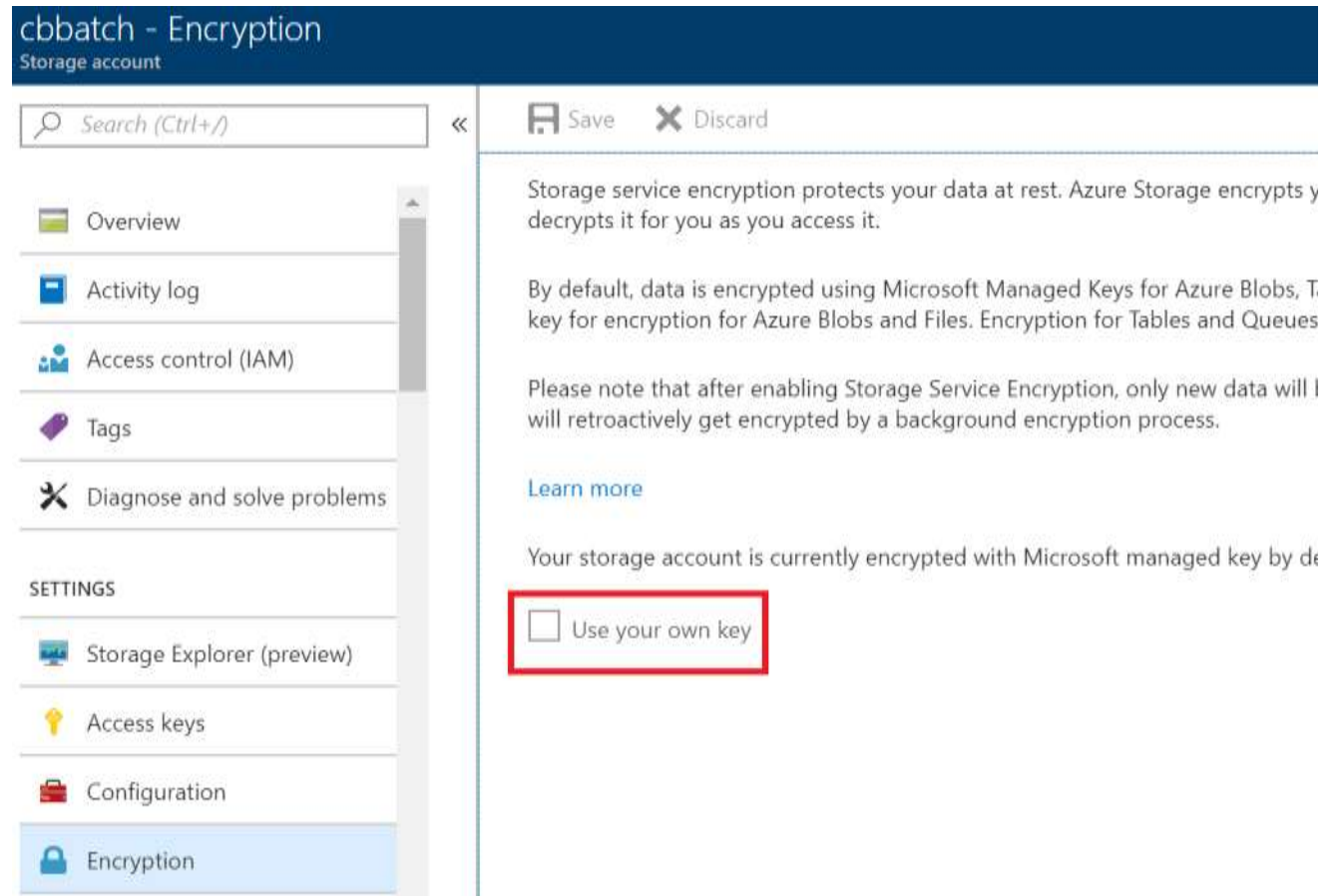
- Use case
 - Providing revocable permissions to certain users/groups
 - To revoke: Delete or update container policy

Storage Service Encryption (Encryption at Rest)

- Automatically encrypts your data before persisting it to Azure Storage, and decrypts the data before retrieval
- Enabled for all new and existing storage accounts and cannot be disabled
- Encrypted using 256-bit AES encryption, Microsoft managed keys
- Automatically encrypts data in:
 - Both performance tiers (Standard and Premium)
 - Both deployment models (Azure Resource Manager and classic)
 - All of the Azure Storage services (Blob storage, Queue storage, Table storage, and Azure Files)
- No additional cost

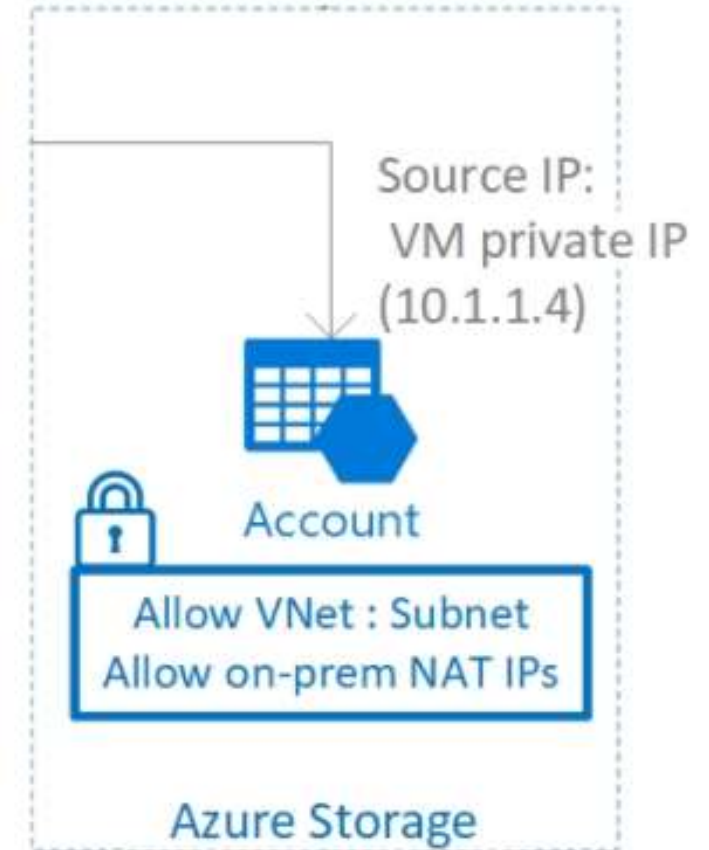
Storage Service Encryption with Customer Managed Keys

- Allows you to specify your own encryption keys
- Create your own encryption keys and store them in a key vault, or you can use Azure Key Vault's APIs to generate encryption keys
- Custom keys give you more flexibility, so that you can create, rotate, disable, and define access controls.
- Custom keys also enable you to audit the encryption keys used to protect your data.



Storage Account Firewall

- Azure Storage provides a layered security model allowing you to secure your storage accounts to a specific set of networks by means of firewall rules
- When firewall rules are configured, only applications from allowed networks can access a storage account
- When calling from an allowed network, applications continue to require authorization e.g. a valid access key or SAS token to access the storage account
- Must be configured in addition to virtual network service endpoints to allow traffic from a specific virtual network



Storage Account Firewall Benefits

- Improved security by restricting access to your storage account to select networks
- More control by granting access to traffic from specific Azure Virtual networks, allowing you to build a secure network boundary on a per application basis
- Better flexibility by granting access to public internet IP address ranges, enabling connections from specific internet or on-premises clients
- Can be applied to new or existing storage accounts

Demo: Create & Explore a storage account & Enable a storage account Firewall





Lab: Implementing Azure Storage

Microsoft Services



