

Cloud Computing

Caltech

**Center for Technology &
Management Education**

Designing Infrastructure Solutions on Azure



Design a Solution for Non-Relational Data

Learning Objectives

By the end of this lesson, you will be able to:

- 👁️ Analyze the Azure non-relational data
- 👁️ Illustrate Azure storage account
- 👁️ Analyze Azure files
- 👁️ Recommend a Solution for Encrypting Data



A Day in the Life of an Azure Architect

You are working as an architect in an organization, and you have been asked to choose the right type of Azure storage for the given use cases:

- For an application that will be used to store the e-documents as part of the digital transformation journey in your organization
- To ensure that the overall cost of data storage is minimized, and the documents are stored in such a way that, depending on the document age, they move to the right tier of the storage
- To take care of the storage access by keeping the data security in mind and providing some authentication method to access the data
- To build a solution for storing vast volumes of unstructured data

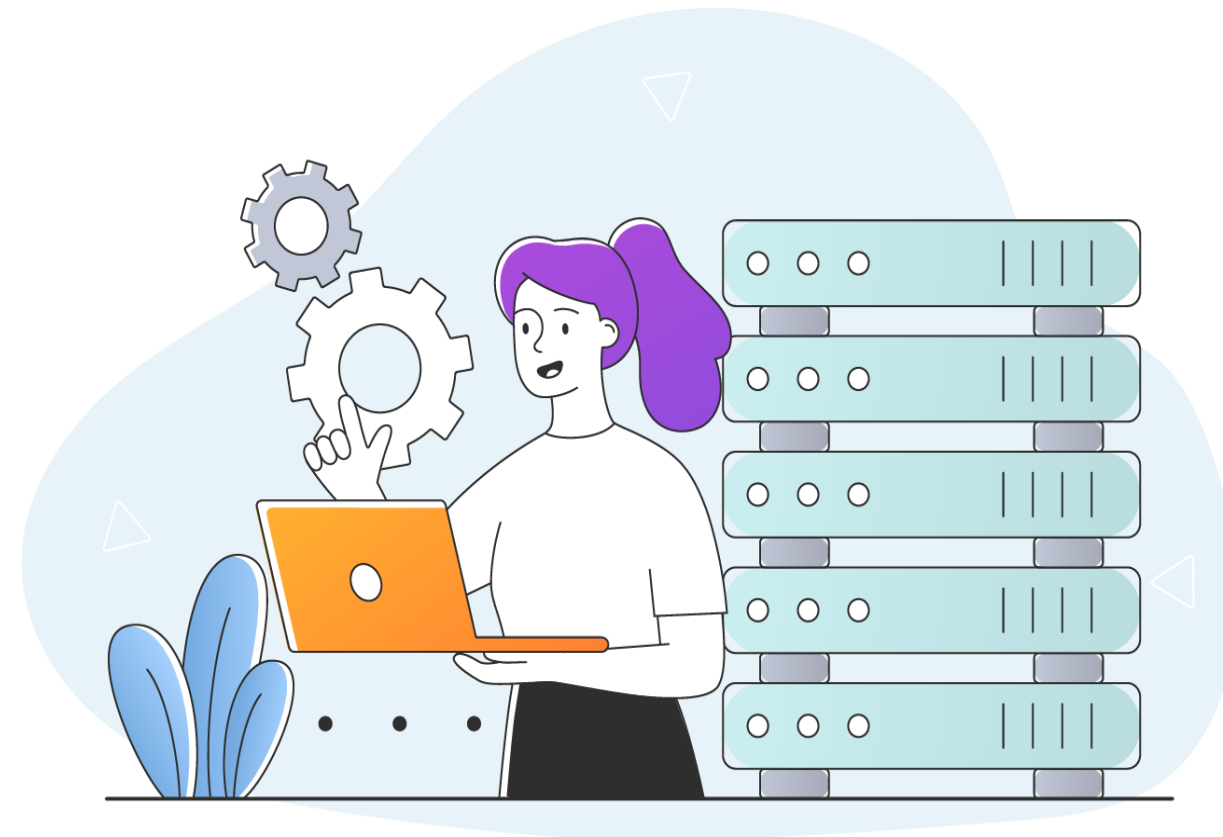
To achieve all of the above, along with some additional features, we would be learning a few concepts in this lesson that will help in finding a solution for the above scenario.



Overview of Azure Non-relational Data

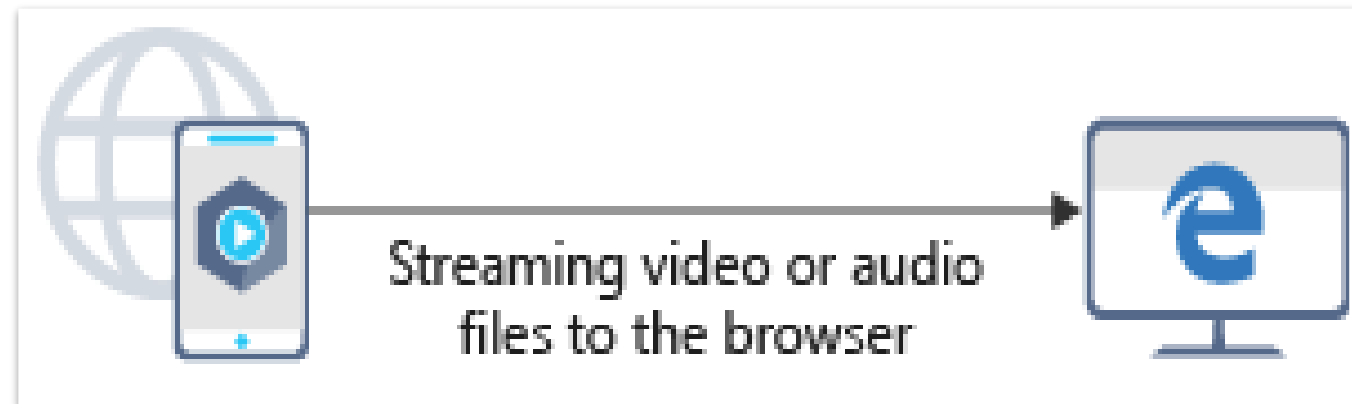
Azure Blob Storage

Azure Blob storage aids in the creation of data lakes for analytics and provides storage for the development of strong cloud-native and mobile apps.



Azure Blob Storage

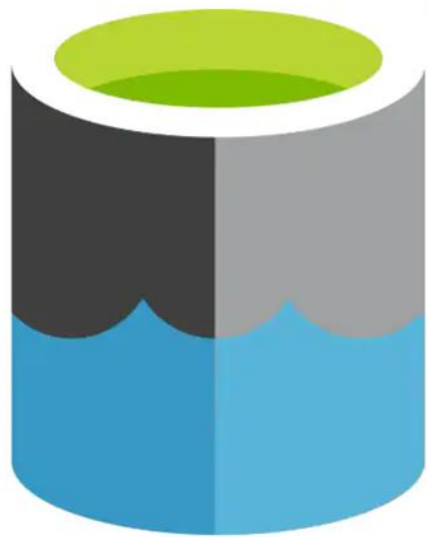
The features of Azure Blob storage are:



- Unstructured
- Highly scalable
- Composed of thousands of simultaneous uploads
- Accessible anywhere

Azure Data Lake Storage

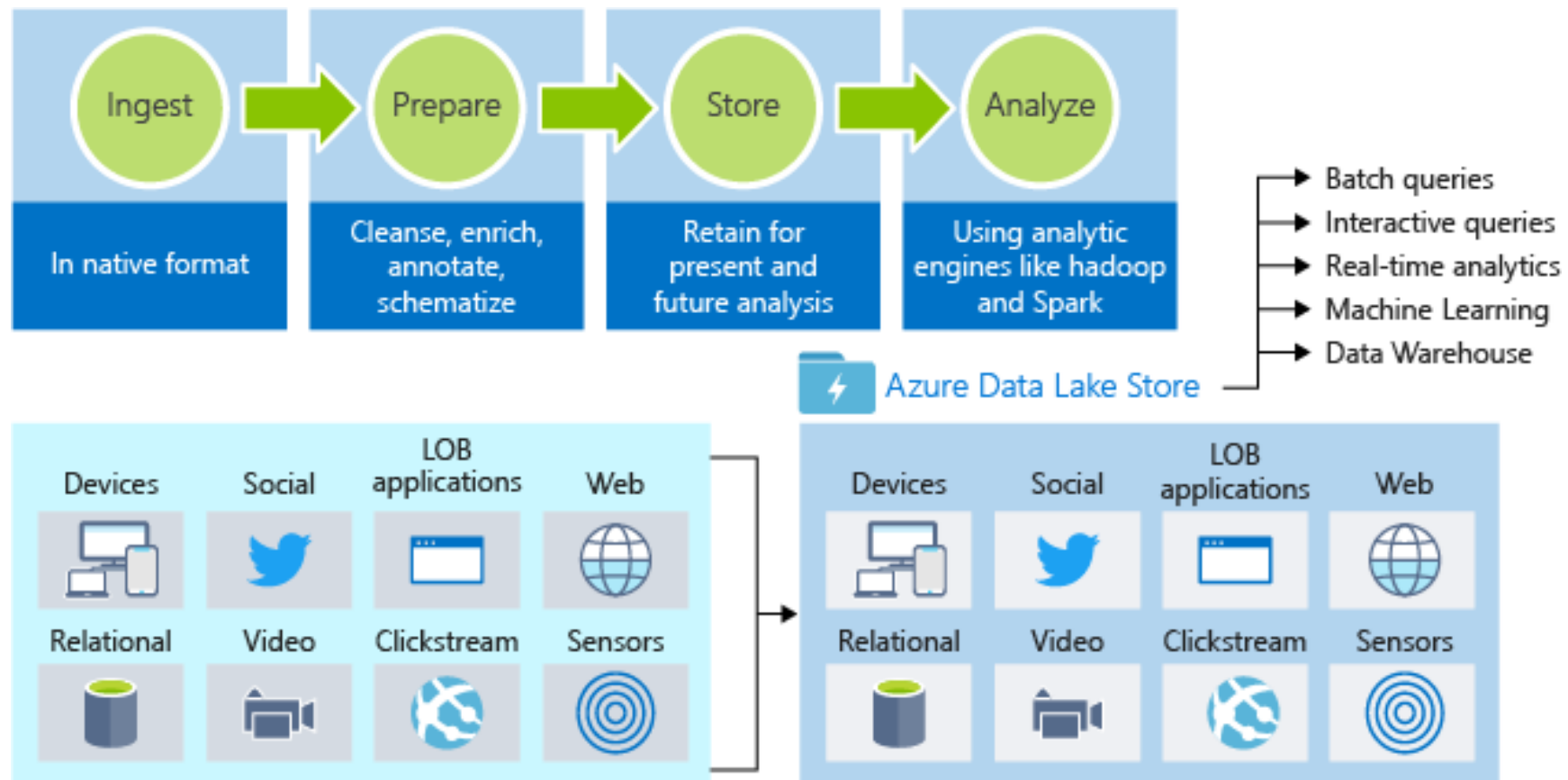
The Azure Data Lake platform is a cloud platform for big data analytics. It may store structured, semi-structured, or unstructured data indefinitely.



Azure Data Lake can store any data format, irrespective of company size.

Azure Data Lake Storage

Azure data lake storage features are:



- Structured and unstructured data
- Hyperscale repository for big data analytics workloads
- No limits on account sizes, file sizes, or the amount of data

Candidate Services

Category	Azure Data Lake Storage Gen1	Azure Blob Storage
Purpose	Optimized storage for big data analytics workloads	General purpose object store for a wide variety of storage scenarios
Use cases	Batch, interactive, streaming analytics, and machine learning data such as log files, IoT data, click streams, and large datasets	Any type of text or binary data, such as application back end, backup data, media storage for streaming, and general-purpose data
Structure	Hierarchical file system	Object store with flat namespace
Authentication	Based on Azure Active Directory identities	Based on shared secrets - Account access keys and shared access signature keys

Candidate Services

Category	Azure Data Lake Storage Gen1	Azure Blob Storage
Authentication Protocol	OAuth 2.0. calls must contain a valid JWT (JSON Web Token) issued by Azure Active Directory.	Hash-based Message Authentication Code (HMAC). Calls must contain a Base64-encoded SHA-256 hash over a part of the HTTP request.
Authorization	POSIX Access Control Lists (ACLs). ACLs based on Azure Active Directory Identities can be set at the file and folder level.	For account-level authorization – Use Account Access Keys
		For account, container, or blob authorization - Use Shared Access Signature Keys
Auditing	Available	Available
Encryption data at rest	Transparent and server-side	Transparent and server-side or client-side
Developer SDKs	.NET, Java, Python, and Node.js	.NET, Java, Python, Node.js, C++, Ruby, PHP, Go, Android, and iOS

Candidate Services

Category	Azure Data Lake Storage Gen1	Azure Blob Storage
Analytics Workload Performance	Optimized performance for parallel analytics workloads with high throughput and IOPS	Optimized performance for parallel analytics workloads
Size limits	No limits on account sizes, file sizes, or number of files	For specific limits
Geo-redundancy	Locally redundant (multiple copies of data in one Azure region)	Locally redundant (LRS), zone redundant (ZRS), globally redundant (GRS), read-access globally redundant (RA-GRS)

Azure Files

They are managed file shares in the cloud that are accessible via SMB. The files can be mounted concurrently by cloud or on-premises resources.

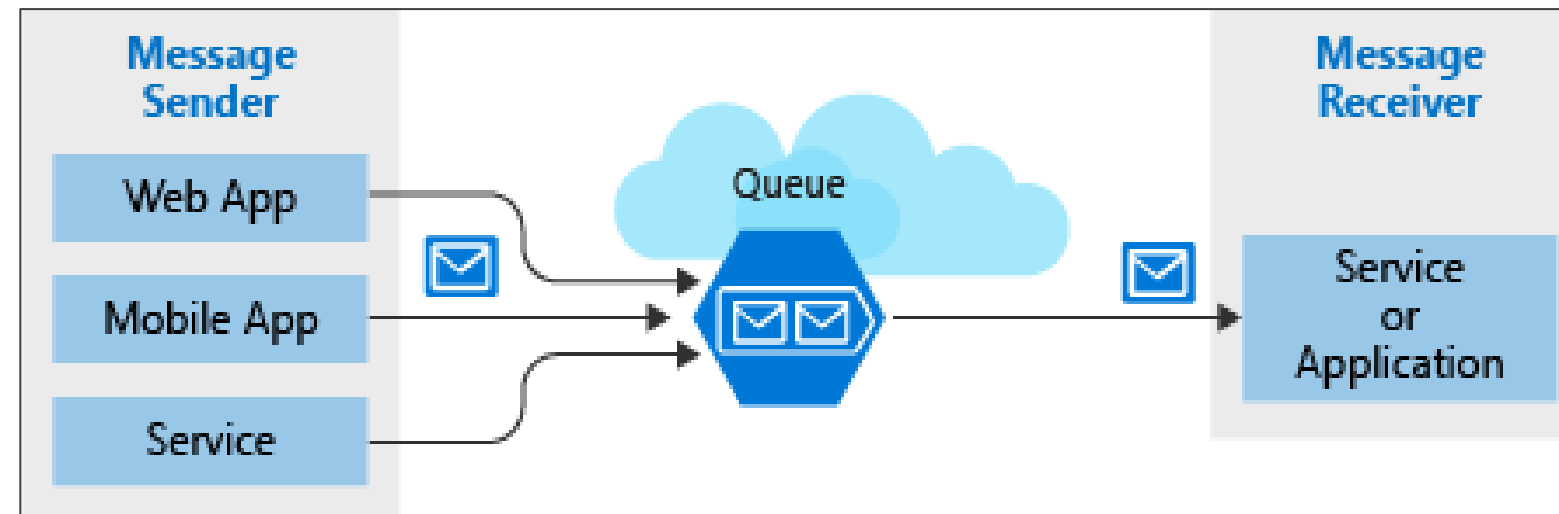


Common uses:

- Replace and supplement
- Lift and shift
- Azure File Sync
- Shared applications
- Diagnostic data
- Tools and utilities

Azure Queues

These are the features and uses of Azure queues:



Azure queues features:

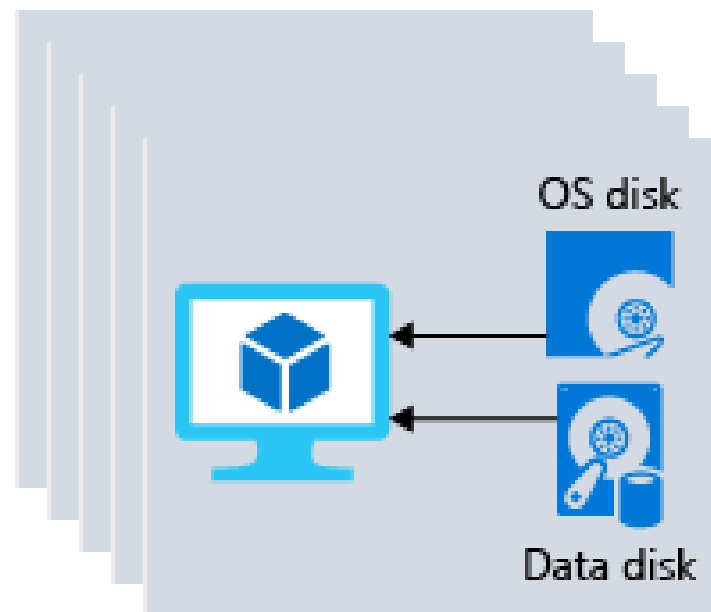
- Store large numbers of messages
- Can be accessed from anywhere in the world
- Provide asynchronous message queueing

Use queue storage to:

- Create a backlog of work
- Pass messages between services
- Distribute load
- Build resilience against component failures

Disk Storage

These are the disk storage features:

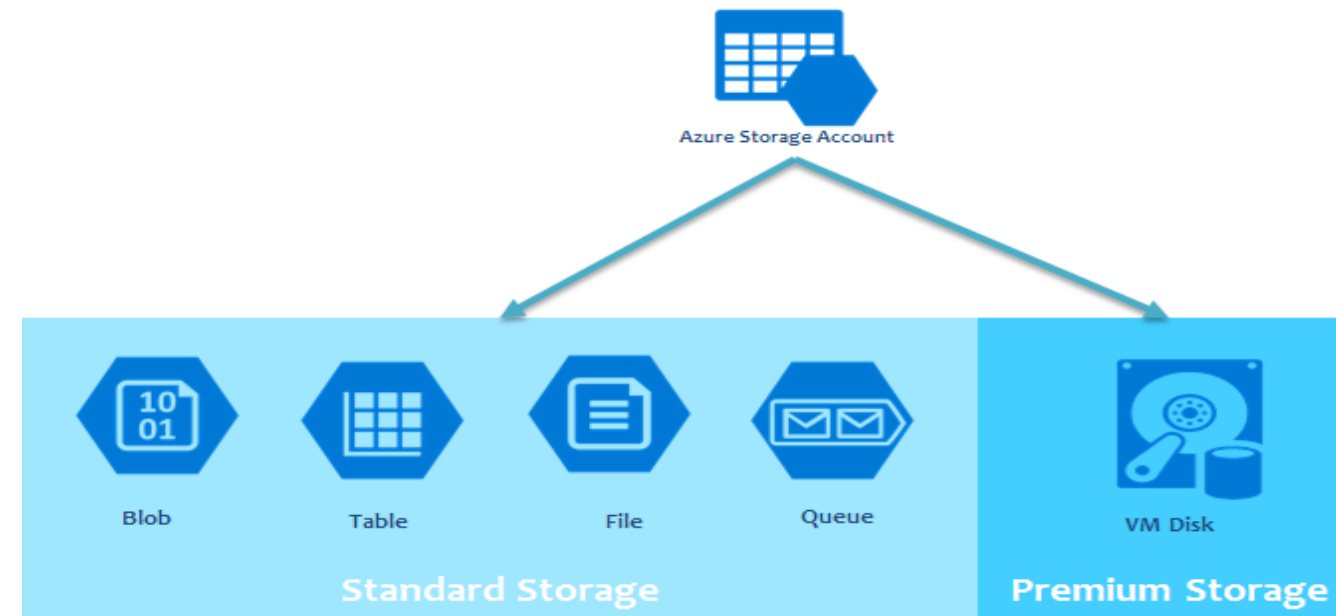


- It provides disks for virtual machines, applications, and other services.
- It is a persistent virtual hard disk storage.
- Disks can be managed or unmanaged.
- It has Solid-state drives (SSDs) and hard disk drives (HDDs)
- It includes premium SSDs for mission-critical applications.

Storage Account

Azure Storage

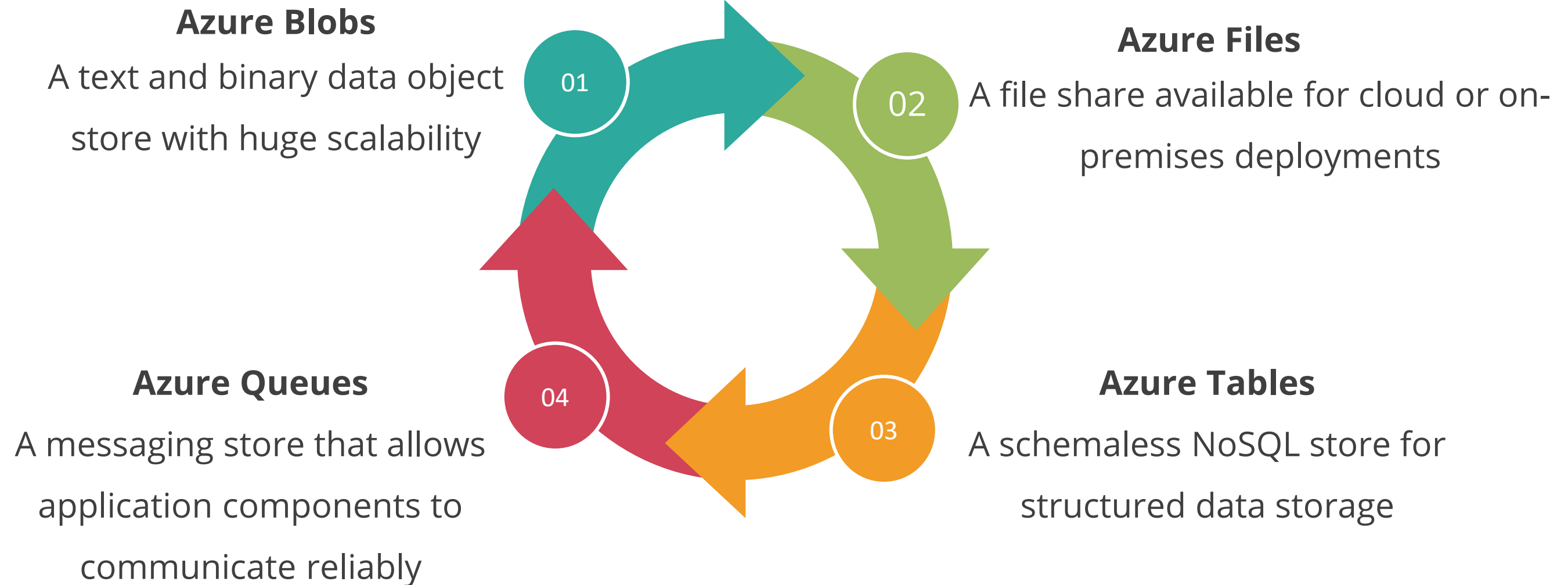
Azure Storage is a cloud storage platform designed for modern data storage scenarios. A massively scalable object store for data objects as well as disk storage for Azure virtual machines is available via core storage services (VMs).



The service encrypts all data written to an Azure storage account.

Azure Storage Services

These are the Azure Storage services:



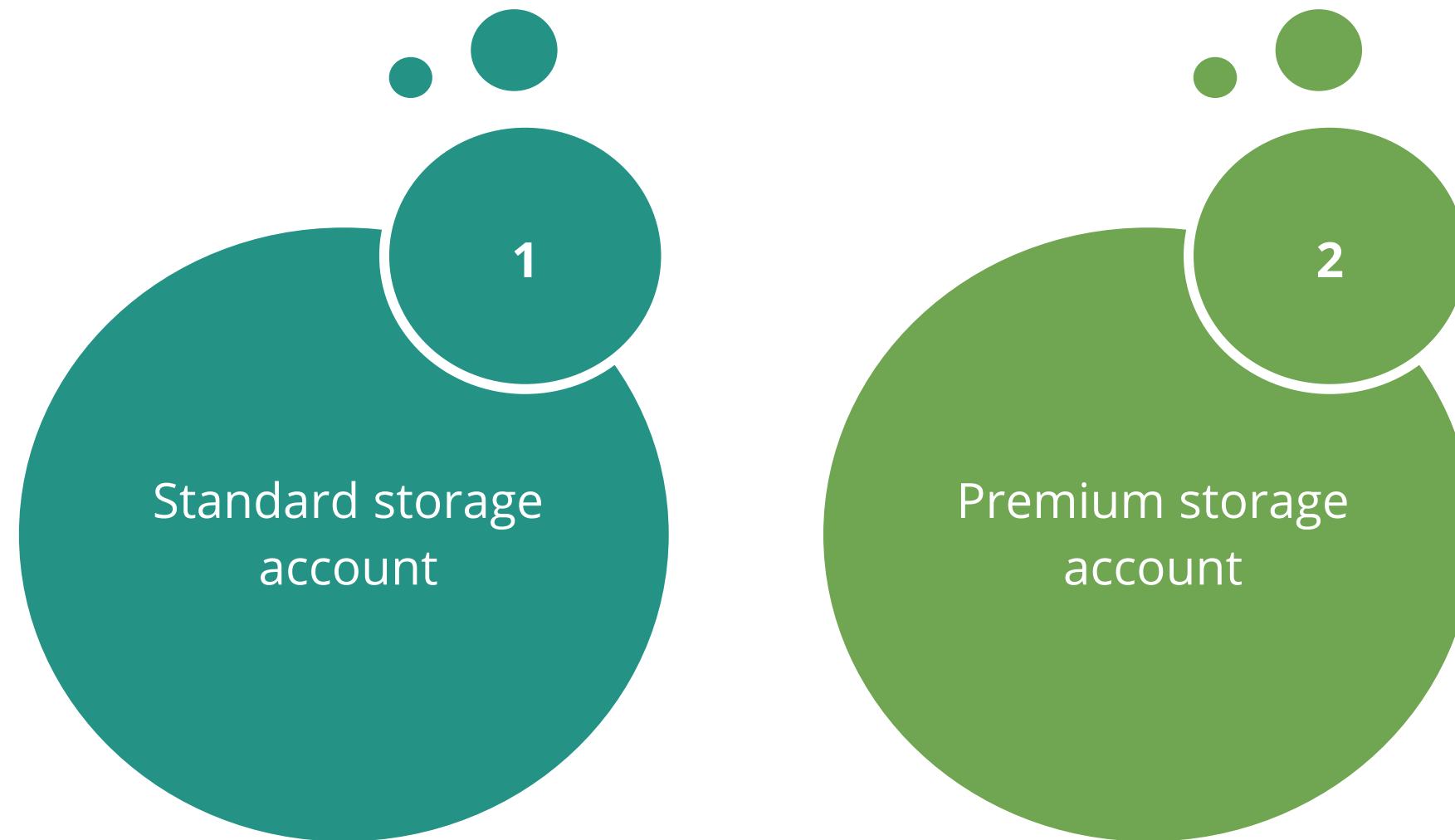
Azure Storage Services

The various Azure storage services are:

Storage Services	Description
Azure Blobs	Allow unstructured data to be stored and accessed as block blobs on a huge scale
Azure Files	Allow accessing fully managed cloud file shares from anywhere by using the industry standard Server Message Block (SMB) protocol
Azure Tables	Allow the user to store structured NoSQL data in the cloud and provide a schemaless key or attribute store
Azure Queues	Allow asynchronous message queuing between application components

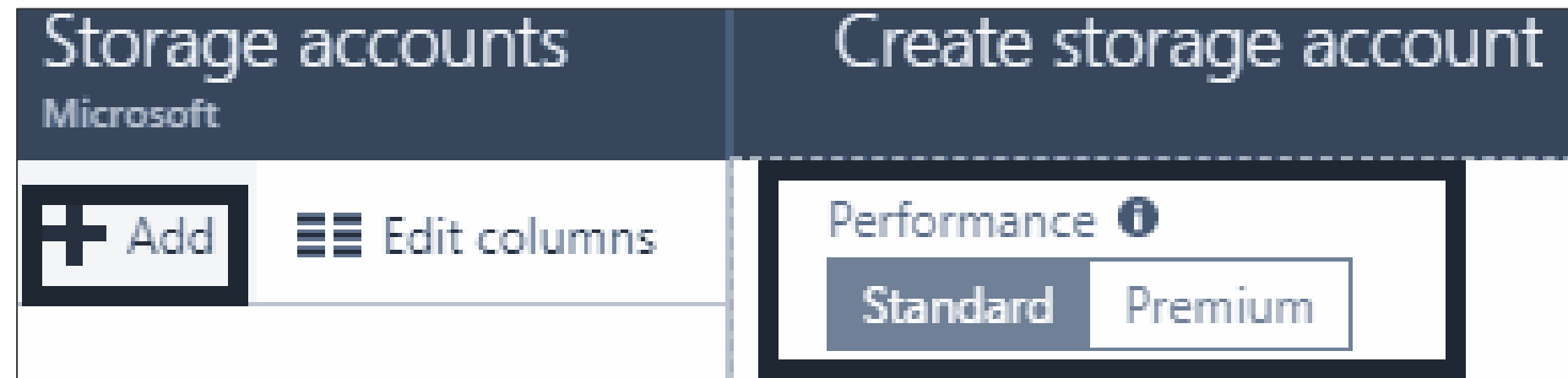
Standard and Premium Storage Accounts

Storage accounts are of two types:



Standard Storage Account

The standard storage performance tier can store Tables, Queues, Files, Blobs, and Azure virtual machine disks.



Most apps use standard storage, which is less expensive and slower.

Premium Storage Account

Create storage acc..

The cost of your storage account depends on the usage and the options [Learn more](#)

* Name ⓘ
anilpremiumaccount01 ✓
.core.windows.net

Deployment model ⓘ
Resource manager Classic

Account kind ⓘ
General purpose ▼

Performance ⓘ
Standard Premium

Replication ⓘ
Locally-redundant storage (LRS) ▼

* Subscription
Windows Azure MSDN - Visual Studio Ultr ▼

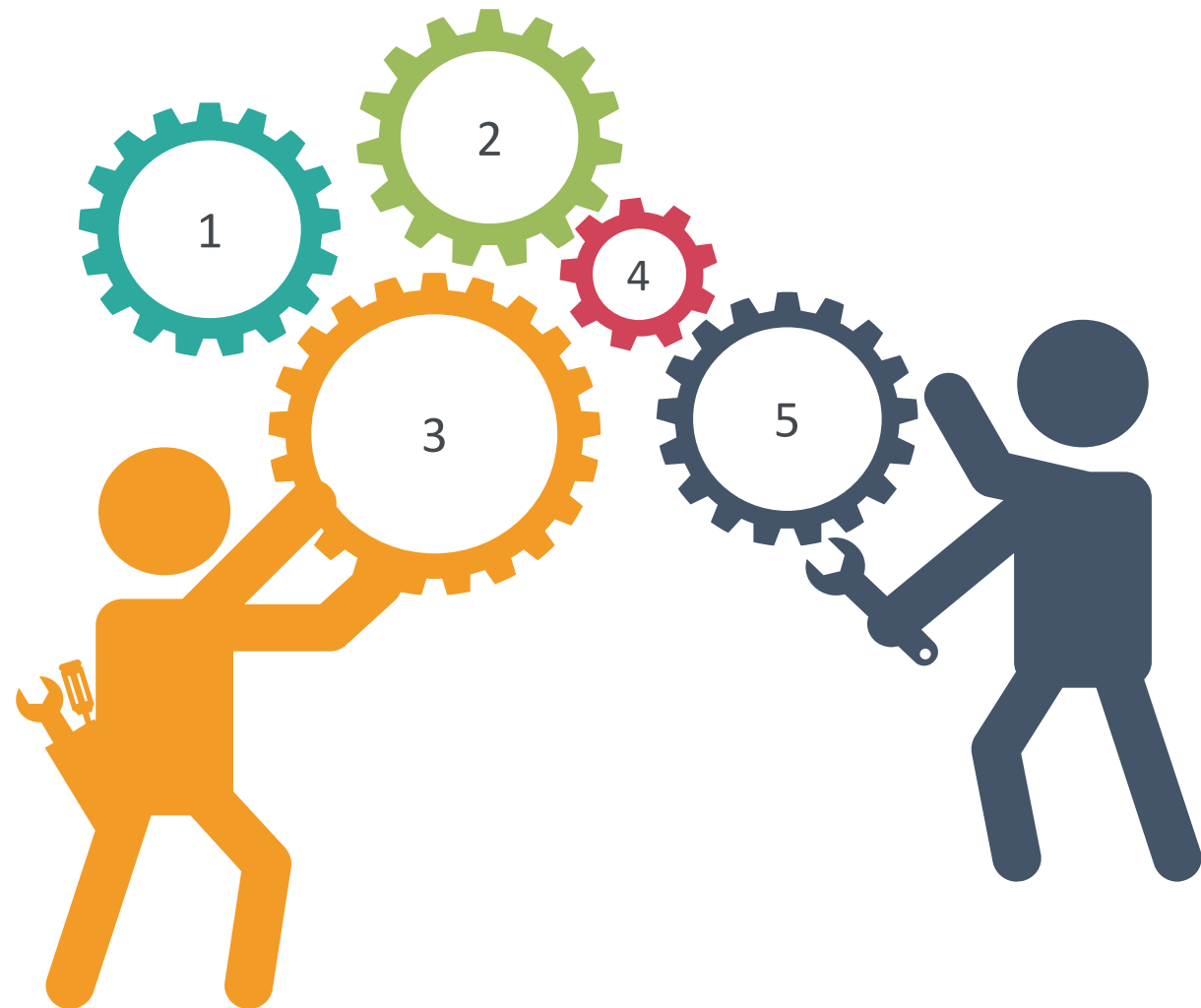
Premium disks aren't charged by transaction. It's more of a flat fee model.

Azure virtual machine disks are supported by the Premium storage performance tier only.

Transaction fees are not applied on premium disks.

Storage Types

There are five types of Storage:



Storage V2

Storage V1

Blob storage

Block Blob storage

File storage

Storage Types

Storage types	Supported services	Supported performance tiers	Replication options
Blob storage	Blob (block blobs and append blobs only)	Standard	LRS, GRS, RA-GRS
General-purpose V1	Blob, File, Queue, Table, and Disk	Standard, Premium	LRS, GRS, RA-GRS
General-purpose V2	Blob, File, Queue, Table, and Disk	Standard, Premium	LRS, GRS, RA-GRS, ZRS, ZGRS (preview), RA-ZGRS (preview)
Block blob storage	Blob (block blobs and append blobs only)	Premium	LRS, ZRS (limited regions)
File storage	Files only	Premium	LRS, ZRS (limited regions)

Design for Azure Blob storage

Blob Storage

Blob storage is an object storage solution for the cloud, which is optimized for storing unstructured data.



Blobs

REST-based object storage for unstructured data

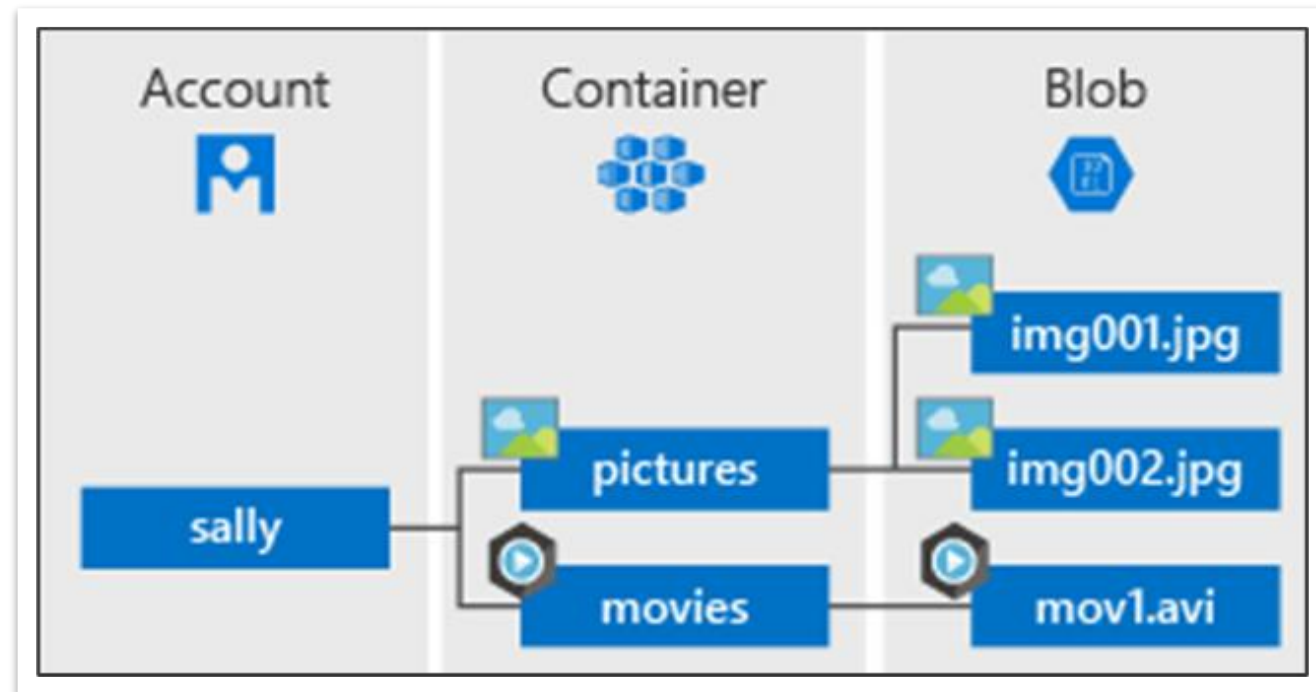
[Learn more](#)

[Explore data using Azure AD preview](#)

It can store any type of text or binary data. Also referred to as object storage.

Blob Storage

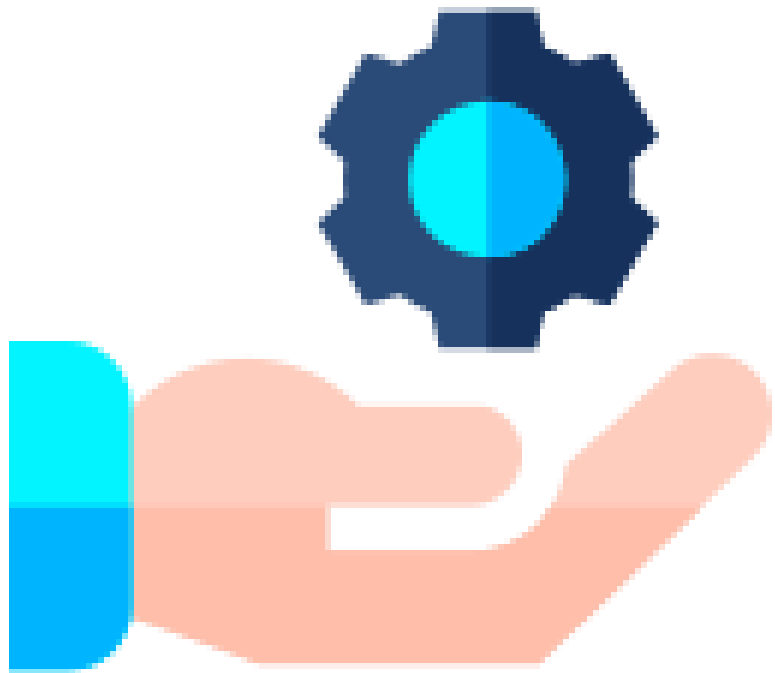
These are the common uses of blob storage:



- Serving images or documents directly to a browser
- Storing files for distributed access
- Streaming video and audio
- Storing data for backup and restore, disaster recovery, and archiving
- Storing data for analysis by an on-premises or Azure-hosted service

Blob Containers

These are the features of blob containers:



- A container organizes a set of blobs.
- A storage account can include an unlimited number of containers.
- A container can store an unlimited number of blobs.

Blob Performance Tiers

These are the blob performance tiers for storage:

Hot tier (inferred)

Optimized for frequent access of objects in the storage account

Cool tier

Optimized for storing large amounts of data that is infrequently accessed and stored for at least 30 days

Archive tier

Optimized for data that can tolerate several hours of retrieval latency and will remain in the Archive tier for at least 180 days

A user can switch between these access tiers at any time.

Design for Blob Performance Tiers

The following points apply to the access tiers:



- At the account level, the hot and cool access tiers can be configured but the Archive access tier cannot.
- During or after upload, the user can organize the hot, cool, and archive levels at the blob level.
- The data in the cool access tier can withstand a minor reduction in availability.
- The archive storage stores data offsite and offers the most cost-effective storage options.

Support Tiering for Storage Accounts

These are the usage scenarios for different tiers:

Hot tier (inferred)

Actively used data or data that is expected to be read from and written to regularly

Cool tier

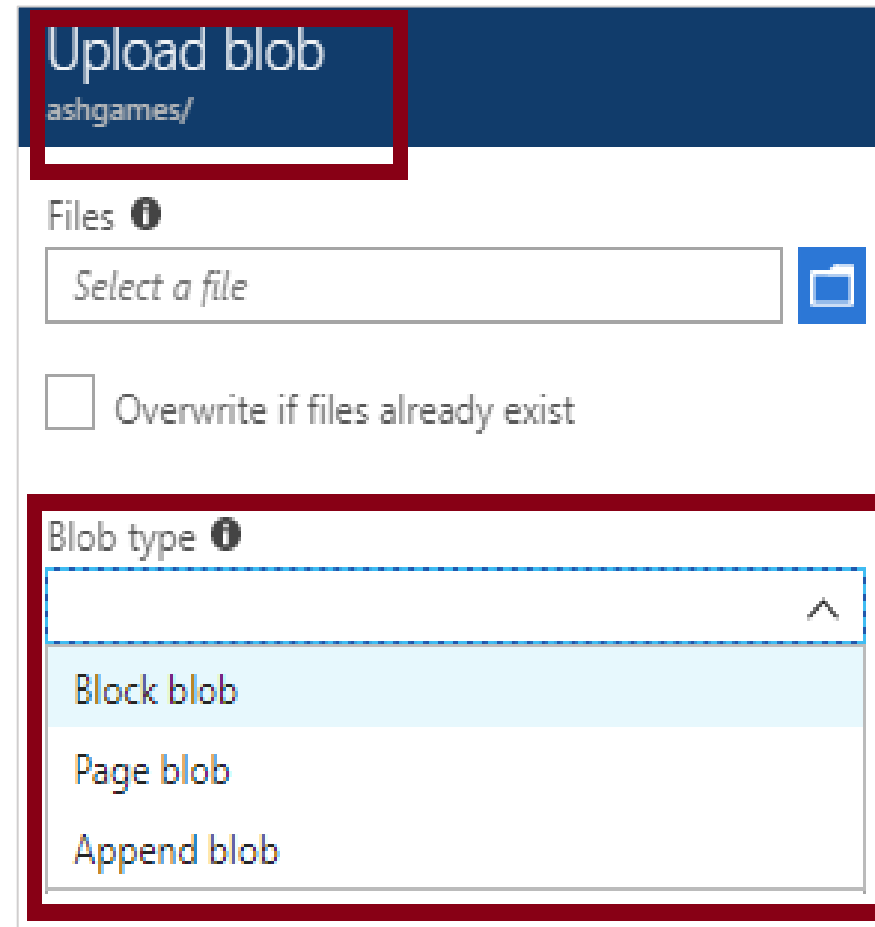
Backup and catastrophe recovery on a short-term basis

Archive tier

Datasets for long-term backup, secondary backup, and archiving

Uploading Blobs

The blob types that the user can upload are:



The screenshot shows the 'Upload blob' interface. At the top, there is a header bar with the text 'Upload blob' and 'ashgames/'. Below this, there is a 'Files' section with a text input field labeled 'Select a file' and a folder icon. Below the input field, there is a checkbox labeled 'Overwrite if files already exist'. Below the checkbox, there is a 'Blob type' dropdown menu. The dropdown menu is open, showing three options: 'Block blob', 'Page blob', and 'Append blob'. The 'Block blob' option is currently selected and highlighted in light blue.

- **Block blobs (default):** Useful for storing text or binary files
- **Page blobs:** More efficient for frequent read/write operations
- **Append blobs:** Useful for logging scenarios

A user cannot change a blob type once it has been created.

Azure Files

Azure Files

Azure Files provides actively managed cloud file shares that can be accessed using the industry-standard Server Message Block (SMB) or Network File System (NFS) protocols.



Files

File shares that use the standard SMB 3.0 protocol

[Learn more](#)

Azure Files is used to:

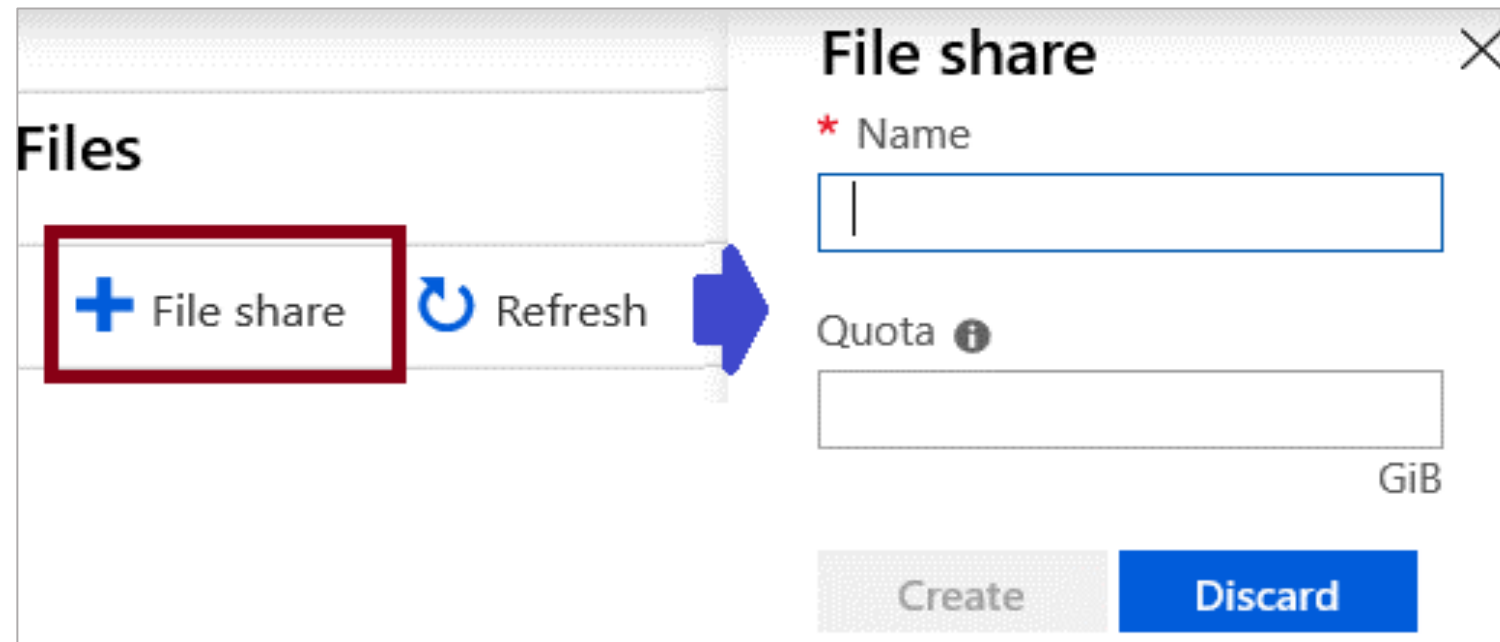
- Replace and supplement on-premise file servers
- Lift and shift applications
- Sync Azure Files
- Share applications
- Diagnose data
- Avail tools and utilities

Azure Files vs. Azure Blobs

Feature	Description	When to use
Azure Files	SMB interface, client libraries, and a REST interface that allows access from anywhere to stored files	<ul style="list-style-type: none">• Lift and shift an application to the cloud• Store shared data across multiple virtual machines• Store development and debugging tools that need to be accessed from many virtual machines
Azure Blobs	Client libraries and a REST interface that allows unstructured data (flat namespace) to be stored and accessed at a massive scale in block blobs	<ul style="list-style-type: none">• Support streaming and random-access scenarios• Access application data from anywhere

Creating a File Share

► In Portal:



The screenshot shows the Azure Portal interface for creating a file share. On the left, under the 'Files' section, there is a '+ File share' button highlighted with a red rectangular box. To its right is a 'Refresh' button with a circular arrow icon. A large blue arrow points from the '+ File share' button to a 'File share' dialog box on the right. The dialog box has a title bar with a close button (X). Inside, there is a 'Name' field with a red asterisk, a 'Quota' field with an information icon, and 'Create' and 'Discard' buttons at the bottom. The 'Quota' field is labeled 'GiB'.

► In PowerShell:

```
# Retrieve storage account and storage account key
$storageContext = New-AzStorageContext <storage-account-name>
<storage-account-key>
```

```
# Create the file share, in this case "logs"
$share = New-AzStorageShare logs -Context $storageContext
```


Mapping File Shares (Windows)

The methods to use an Azure file share with Windows are as follows:

Connect

test

Windows Linux MacOS

Drive letter

Z

To connect to this file share from a Windows computer, run these PowerShell commands:

Alternatively, run this command if the key doesn't begin with a forward slash:

```
net use Z: \\rgtest11.file.core.windows.net\test /u:AZURE\rgtest11 k3b JEEy0YX41g HtX J==
```

When connecting from a computer from outside Azure, remember to open outbound TCP port 445 in your local network. Some Internet service providers may block port 445. Check with your service provider for details.

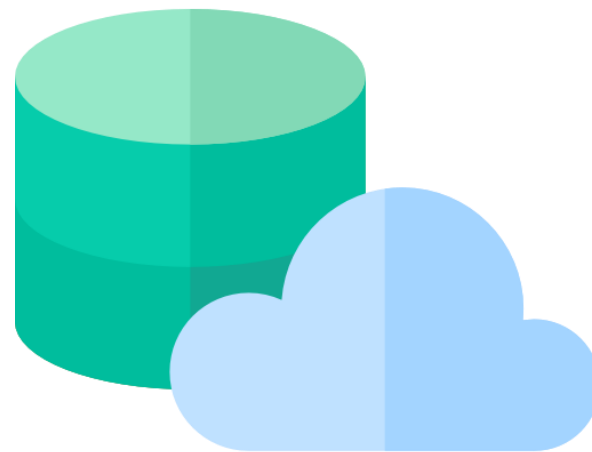
- Mapping drive letter
- UNC path
- Account user
- Storage account key

✓ Ensure port 445 is open

Design for Azure Disk Solutions

Data Disk

A data disk is a managed disk that is attached to a virtual machine and used to store application data and other important information.



The maximum capacity of each data disk is 32,767 gibibytes (GiB).

Data disks are labeled with a letter of user's choice and registered as SCSI drives.

Disk Types Available in Azure Managed Disks



Disk types are:

Ultra-disk

Premium SSD

Standard SSD

Standard HDD

Data Disk

	Ultra disk	Premium SSD	Standard SSD	Standard HDD
Disk type	SSD	SSD	SSD	HDD
Scenario	IO-intensive workloads such as SAP HANA , top tier databases (for example, SQL, Oracle), and other transaction-heavy workloads.	Production and performance sensitive workloads	Web servers, lightly used enterprise applications and dev/test	Backup, non-critical, infrequent access
Max disk size	65,536 gibibyte (GiB)	32,767 GiB	32,767 GiB	32,767 GiB
Max throughput	4,000 MB/s	900 MB/s	750 MB/s	500 MB/s
Max IOPS	160,000	20,000	6,000	2,000

Disk caching

A disk cache is the method of reducing the time it takes to read or write data from or to a hard drive.



A disk cache is also a part of random-access memory (RAM) that has been set aside. It stores previously read data as well as data from nearby data regions that are expected to be accessed next.

Azure Disk Encryption

It assists the users in protecting and safeguarding their data in order to meet the organization's security and compliance obligations.



ADE encrypts the OS and data disks of Azure virtual machines (VMs) inside a user's VMs by utilizing the CPU of their VMs through the use of Linux's DM-Crypt feature or Windows' BitLocker feature.

Storage Security

Storage Security

High-level security capabilities for Azure storage are:



- Storage encryption services
- Authentication with Azure AD and RBAC
- Client-side encryption, HTTPS, and SMB 3.0 for data in transit
- Azure disk encryption
- Shared access signatures – delegated access

Storage Security



Authorization options:

Azure Active Directory (Azure AD)

Shared key

Shared access signatures

Anonymous access to containers and blobs

Storage Account Keys

- Azure creates two keys- primary and secondary for each storage account
- Either of the keys provides full access to the account
- Keys should be regenerated regularly or if they are compromised

The screenshot shows the Microsoft Azure portal interface for a storage account named 'AccountName'. The 'Access keys' page is displayed, showing the storage account name and two keys (key1 and key2) with their corresponding connection strings. Red boxes highlight the storage account name, key1, and key2 sections.

Storage account name
AccountName

key1
Key
[Redacted]Qg==

key2
Key
[Redacted]

Connection string
DefaultEndp

Private Endpoint and Service Endpoint

These are the features of Azure private endpoint and service endpoint:

Private Endpoint

- Private endpoints can be used to access Azure PaaS services within the VNet using a private IP address.
- On a user's VNet, it receives a new private IP address. When the user transmits traffic to a PaaS resource, the traffic is always kept within their VNet.

Service Endpoint

- Service endpoints allow direct and safe access to Azure PaaS services via an efficient path via the Azure backbone network.
- Traffic continues to leave user's VNet and reach the PaaS service's public endpoint.

Shared Access Signature (SAS)

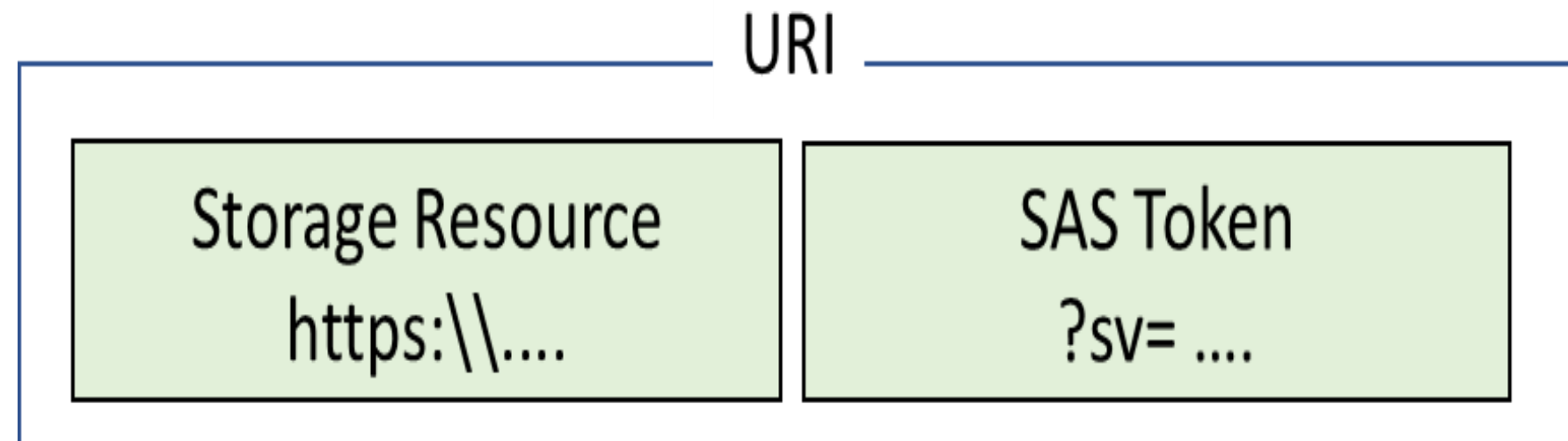
SAS provides delegated access to resources and also grants access to clients without sharing storage account keys.



The account SAS delegates access to resources in one of the storage services that are Blob, Queue, Table, or File service.

URI and SAS Parameters



A SAS is a signed URI that points to one or more storage resource and it consists of a storage resource URI and the SAS token.



Includes parameters for resource URI, storage services version, services, resource types, start time, expiry time, resource, permissions, IP range, protocol, and signature.

Storage Service Encryption

Encryption

 Save  Discard

Storage service encryption protects your data at rest. Azure Storage encrypts your data as it's written in our datacenters, and automatically decrypts it for you as you access it.

By default, data in the storage account is encrypted using Microsoft Managed Keys. You may choose to bring your own key.

Please note that after enabling Storage Service Encryption, only new data will be encrypted, and any existing files in this storage account will retroactively get encrypted by a background encryption process.

[Learn More about Azure Storage Encryption](#)

Encryption type

☒ Microsoft Managed Keys

☐ Customer Managed Keys

- Protects data for security and compliance
- Encrypts and decrypts data
- Encrypts through 256-bit AES encryption
- It is enabled for all new and existing storage accounts and cannot be disabled
- It is transparent to users

✓ You can use your own key

Customer Managed Keys

Encryption type

- ☐ Microsoft Managed Keys
- ☒ Customer Managed Keys

i The storage account named 'storage987123' will be granted access to the selected key vault. Both soft delete and purge protection will be enabled on the key vault and cannot be disabled. [Learn more about customer managed keys](#) 

Encryption key

- ☐ Enter key URI
- ☒ Select from Key vault

Key vault and key *

Key vault: keyvault987123

Key: storagekey

[Select a key vault and key](#)

- Use the Azure Key Vault to manage the encryption keys
- Distinguished personal encryption keys can be created and stored in a key vault
- Azure Key Vault APIs are used to generate encryption keys
- Custom keys give more flexibility and control

Best Practices



Always use HTTPS to create or distribute a SAS

Reference stored access policies, where possible

Use near-term expiration times on an ad-hoc SAS

Have clients automatically renew the SAS, if necessary

Be careful with SAS start time

Best Practices



Be specific with the resource to be accessed

Understand that your account will be billed for any usage

Validate data written using SAS

Don't assume SAS is always the correct choice

Use storage analytics to monitor your application

Assisted Practice

Manage Azure Storage

Duration: 10 Min.

Problem Statement:

You are given a project to evaluate the use of Azure storage for storing files residing currently in on-premises data stores. While majority of these files are not accessed frequently, there are some exceptions. You would like to minimize cost of storage by placing less frequently accessed files in lower-priced storage tiers. You also plan to explore different protection mechanisms that Azure Storage offers, including network access, authentication, authorization, and replication. Finally, you want to determine to what extent Azure Files service might be suitable for hosting your on-premises file shares.

Assisted Practice: Guidelines

Steps to manage Azure storage are:

1. Log into the Azure portal at <https://portal.azure.com>
2. Create a Storage account
3. Manage Blob storage
4. Manage authentication and authorization for Azure Storage
5. Create and configure an Azure File share
6. Manage network access for Azure Storage



Assisted Practice

Manage Storage Access Keys

Duration: 10 Min.

Problem Statement:

You are given a project to manage storage access keys to authorize access to data in your storage account via Shared Key authorization.

Assisted Practice: Guidelines

Steps to manage storage access keys:

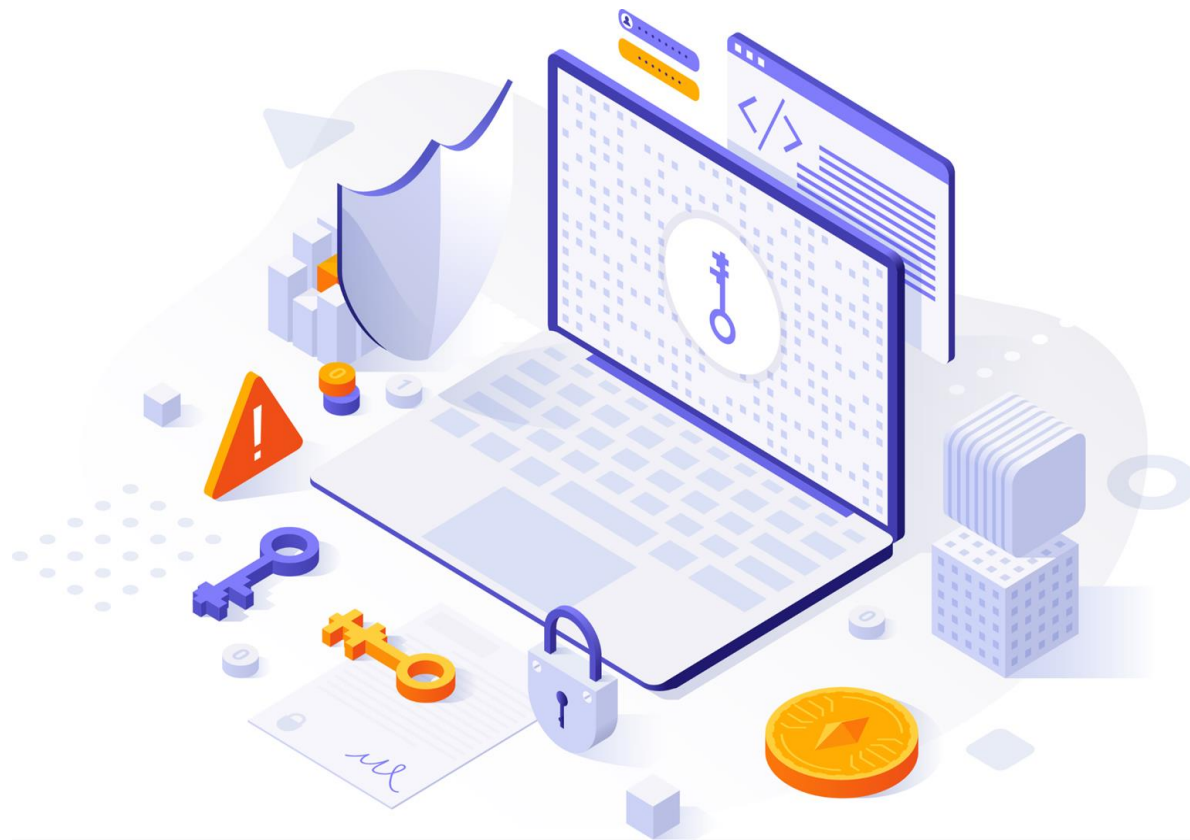
1. Managing storage access keys
2. Installing storage explorer
3. Accessing storage with storage explorer using storage access keys



Recommend a Solution for Encrypting Data

Data Encryption

Encryption is the process of making data unreadable and unusable.

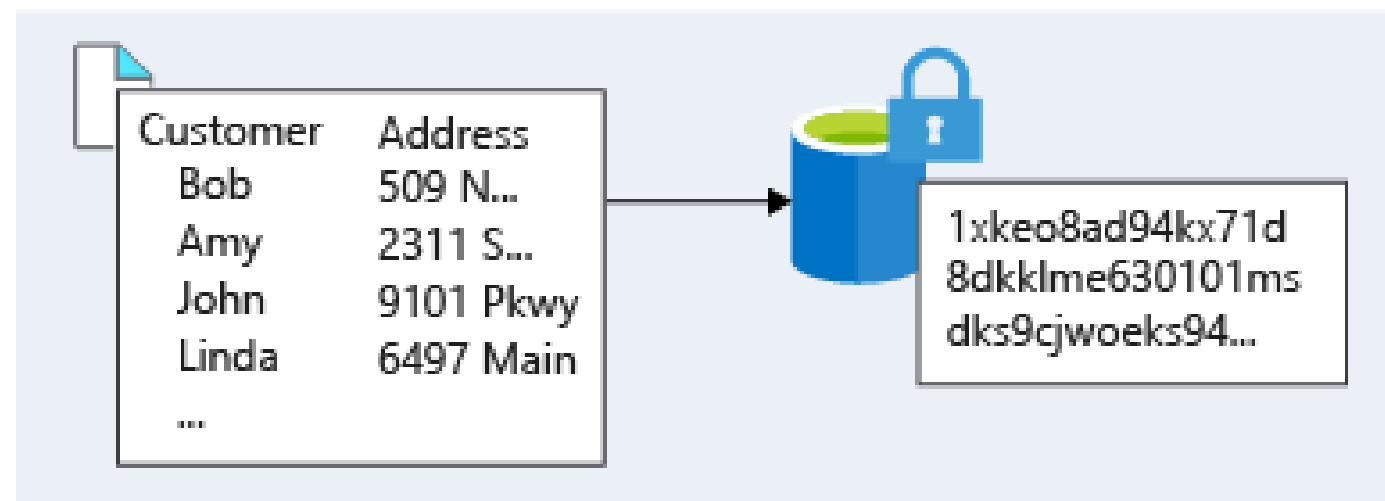


There are two top-level types of encryption:

- **Symmetric encryption:** Uses the same key to encrypt and decrypt the data
- **Asymmetric encryption:** Uses a public key and private key pair

Encryption at Rest

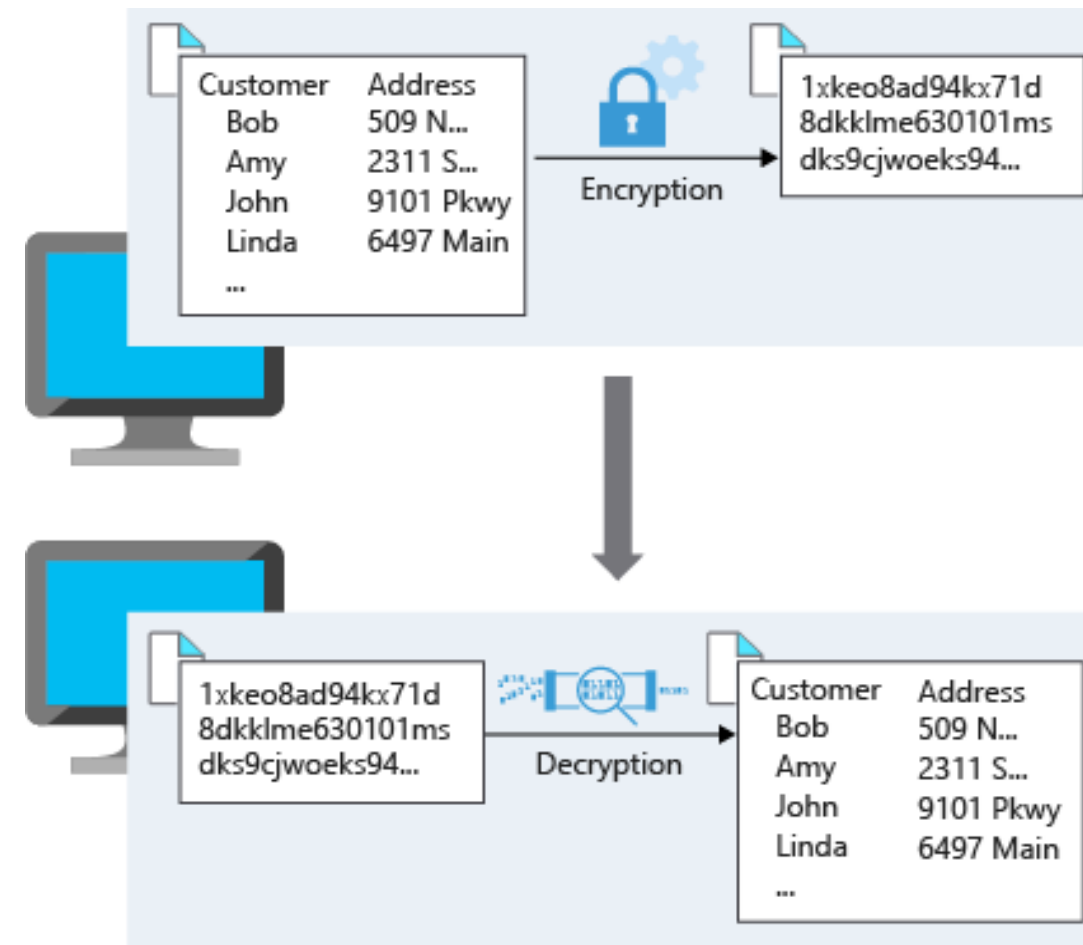
Data at rest is the data that has been stored on a physical medium.



- It is unreadable without the keys and secrets.
- Azure Disk Encryption uses Windows BitLocker, Linux dm-crypt.
- Azure Storage and Azure SQL Database encrypt data at rest by default.
- Users must use Azure Key Vault to maintain control of keys.
- They must encrypt drives before writing sensitive data.

Encryption in Transit

Encrypting data in transit protects the data from outside observers.



Microsoft uses Transport Layer Security (TLS) to protect data when it is traveling between the cloud services and customers.

Identify and Classify Data

These are the built-in roles that identify and classify data:

Data Classification	Explanation	Example
Restricted	Data classified as restricted poses a significant risk if it's exposed, altered, or deleted. Strong levels of protection are required for this data.	Data containing SS numbers, CC numbers, and personal health records
Private	Data classified as private poses a moderate risk if it's exposed, altered, or deleted. Reasonable levels of protection are required for this data. Data that is not classified as restricted or public will be classified as private.	Personal records containing information such as an address, phone numbers, and personal health records
Public	Data classified as public poses no risk if exposed, altered, or deleted. No protection is required for this data.	Public financial reports, public policies, and product documentation for customers

Encrypting Raw Storage

Azure Storage Service Encryption (SSE) for data at rest protects data to meet organizational security and compliance commitments.



Encrypting Raw Storage

The Azure storage platform automatically encrypts data with 256-bit Advanced Encryption Standard (AES) in:

- All Azure Storage services including Azure Managed Disks, Azure Blob storage, Azure Files, Azure Queue storage, and Azure Table storage
- Both performance tiers (Standard and premium)
- Both deployment models (Resource Manager and classic)

Encrypting Virtual Machines

These points guide encrypting virtual machines:

Azure Disk Encryption (ADE) encrypts Windows and Linux IaaS virtual machine disks.

ADE uses BitLocker on Windows and the DM-Crypt feature of Linux.

ADE is integrated with Azure Key Vault.

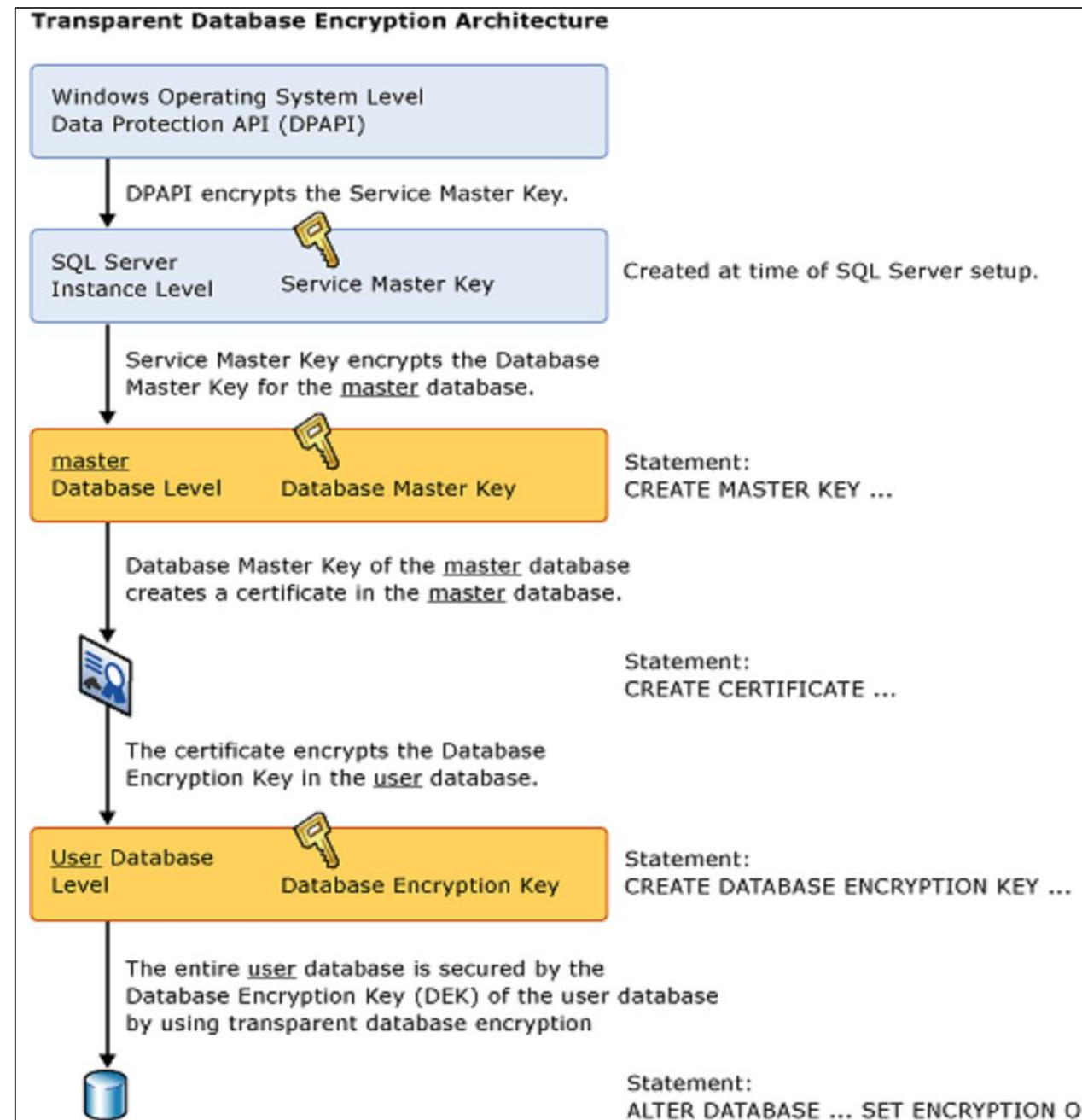
IaaS VMs are secured at rest by using industry-standard encryption.

IaaS VMs boot under customer-controlled keys and policies.

IaaS VMs can be audited in Azure Key Vault.

Encrypting Databases

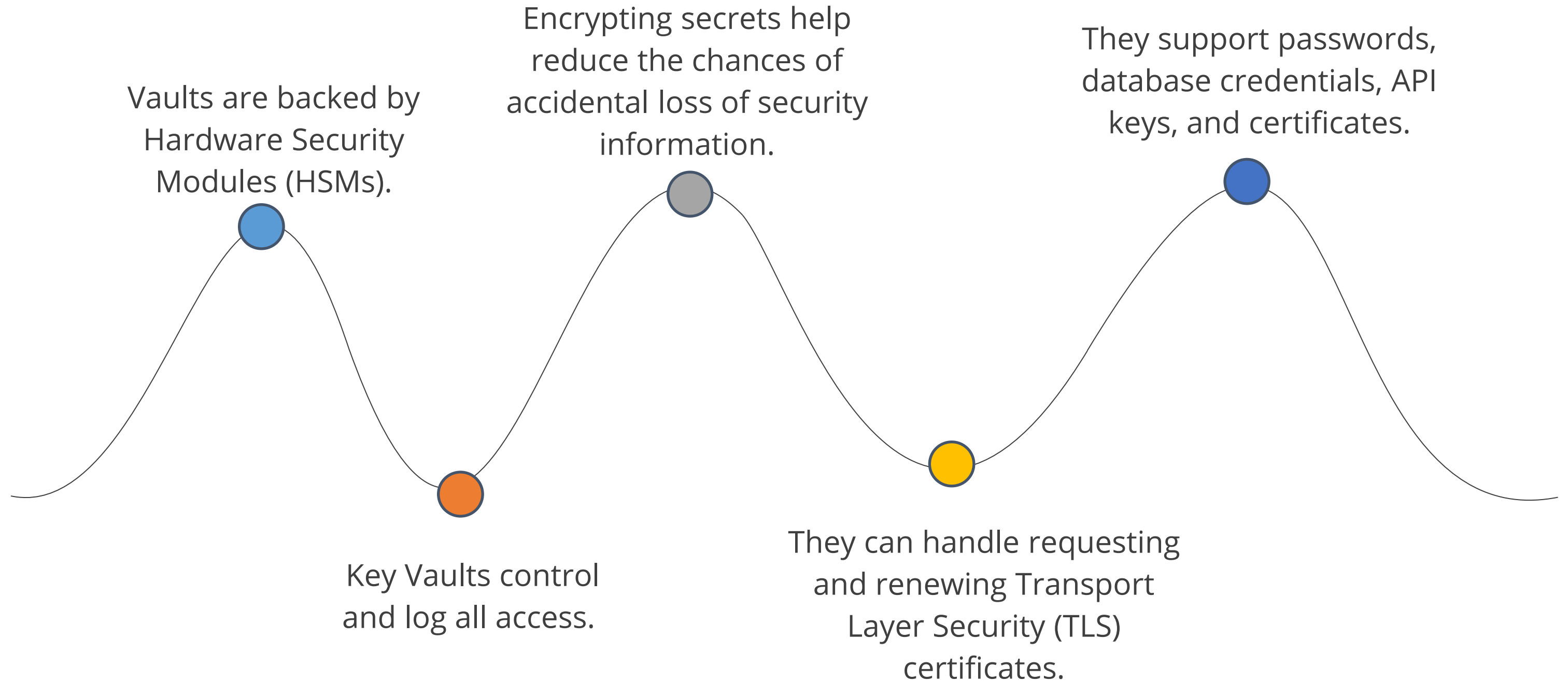
Transparent data encryption (TDE) helps protect the Azure SQL Database and Azure Data Warehouse.



Features:

- Real-time encryption and decryption
- Enabled by default
- Symmetric key called the database encryption key
- Unique encryption key per logical SQL Server
- Bring-your-own-key, supported with keys stored in Azure Key Vault

Encrypting Secrets



Key Takeaways

- Azure Blob storage aids in the creation of data lakes for analytics and provides storage for the development of strong cloud-native and mobile apps.
- They are managed file shares in the cloud that are accessible via SMB. These can be mounted concurrently by cloud or on-premises resources.
- Encrypting data in transit protects the data from outside observers and provides a mechanism for transmitting data while limiting risk of exposure
- Azure Storage Service Encryption (SSE) for data at rest protects data to meet organizational security and compliance commitments.



Implement Storage Account

Duration: 10 min.

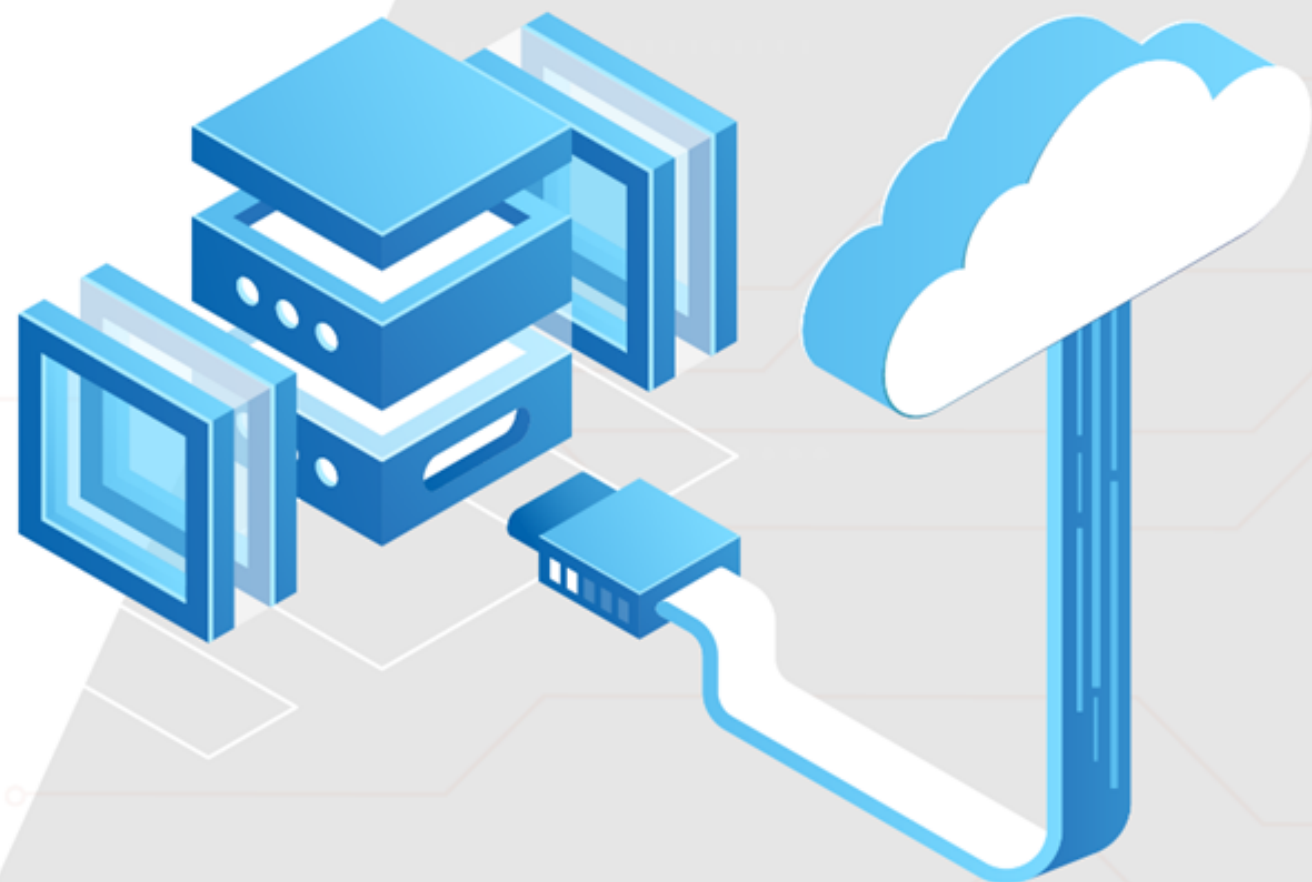


Project Agenda: To create a storage account and a container for storing images.

Description: You have been given a project to create a data store which could store the images being used by an e commerce application. As a part of this you need to create storage account and create a container with public access enabled which would store the images used by e commerce application.

Perform the following:

1. Create a storage account
2. Create a container
3. Adding images to the container



Thank you