AZURE FUNDAMENTALS

Azure Data Storage Options

-Microsoft Learning

Contents

Introduction

Suppose you manage an online sales learning portal for your organization. The majority of your sales team are often in different geographical areas, so the online learning portal is an essential requirement. It's even more important as your organization continues to increase the skills and knowledge enhancement training for the sales staff.

Your training data includes high-quality video, detailed sales simulations, and large repositories for maintaining student data and progress. Currently, all the training content is stored in your on-premises storage. You have an aggressive plan to add new courses and would like to avoid the need to continuously increase the local storage capacity. You're looking for a storage solution that is secure, durable, scalable, and easily accessible from across the globe.

Azure provides storage features that will meet all of your business needs.

Learning objectives

In this module, you will:

- Survey the data storage options in Azure
- Discover how Azure data storage can meet your business demands
- Compare Azure data storage with on-premises storage

Benefits of using Azure to store data

To address the storage problem for your online learning portal, you're considering storing your data in the cloud. But you're concerned about security, backup, and disaster recovery. On top of that, you're worried about how difficult it could be to manage cloud-hosted data. So, here's what you need to know.

The Azure data storage options are cloud-based, secure, and scalable. Its features address the key challenges of cloud storage and provide you with a reliable and durable storage solution.

Why store your data in the cloud?

Benefits of using Azure to store data

Here are some of the important benefits of Azure data storage:

- **Automated backup and recovery**: mitigates the risk of losing your data if there is any unforeseen failure or interruption.
- **Replication across the globe**: copies your data to protect it against any planned or unplanned events, such as scheduled maintenance or hardware failures. You can choose to replicate your data at multiple locations across the globe.
- **Support for data analytics**: supports performing analytics on your data consumption.
- **Encryption capabilities**: data is encrypted to make it highly secure; you also have tight control over who can access the data.
- **Multiple data types**: Azure can store almost any type of data you need. It can handle video files, text files, and even large binary files like virtual hard disks. It also has many options for your relational and NoSQL data.
- **Data storage in virtual disks**: Azure also has the capability of storing up to 8 TB of data in its virtual disks. This is a significant capability when you're storing heavy data such as videos and simulations.
- **Storage tiers**: storage tiers to prioritize access to data based on frequently used versus rarely used information.

Types of data

There are three primary types of data that Azure Storage is designed to hold.

1. **Structured data**. Structured data is data that adheres to a schema, so all of the data has the same fields or properties. Structured data can be stored in a database table with rows and columns. Structured data relies on keys to indicate how one row in a table relates to data in another row of another table. Structured data is also referred to as *relational data*, as the data's schema defines the table of data, the fields in the table, and the clear relationship between the two. Structured data is straightforward in that it's easy to enter, query, and analyze. All of the data follows the same format. Examples of structured data include sensor data or financial data.
2. **Semi-structured data**. Semi-structured data doesn't fit neatly into tables, rows, and columns. Instead, semi-structured data uses *tags* or *keys* that organize and provide a hierarchy for the data. Semi-structured data is also referred to as *non-relational* or *NoSQL* data.

3. **Unstructured data**. Unstructured data encompasses data that has no
   designated structure to it. This also means that there are no restrictions on the
   kinds of data it can hold. For example, a blob can hold a PDF document, a JPG
   image, a JSON file, video content, etc. As such, unstructured data is becoming
   more prominent as businesses try to tap into new data sources.

How Azure data storage can meet your business storage needs

Looking at the benefits of Azure data storage, you understand that it offers the best
options for storing your learning portal. Now let's explore the benefits and options in
detail to see how it fits your business needs.

How Azure data storage can meet your business storage needs

Azure provides several storage options that accommodate specific types of data
storage needs.

Azure SQL Database

Azure SQL Database is a relational
database as a service (DaaS) based
on the latest stable version of the
Microsoft SQL Server database
engine. SQL Database is a high-
performance, reliable, fully
managed and secure database.
You can use it to build data-driven
applications and websites in the
programming language of your
choice without needing to manage infrastructure.

You can migrate your existing SQL Server databases with minimal downtime using
the Azure Database Migration Service. The service uses the *Microsoft Data Migration
Assistant* to generate assessment reports that provide recommendations to help
guide you through required changes prior to performing a migration. Once you

assess and perform any remediation required, you're ready to begin the migration process. The Azure Database Migration Service performs all of the required steps. You just change the connection string in your apps.

The following illustration shows the types of data from the online learning portal scenario that would be stored in an Azure SQL database.
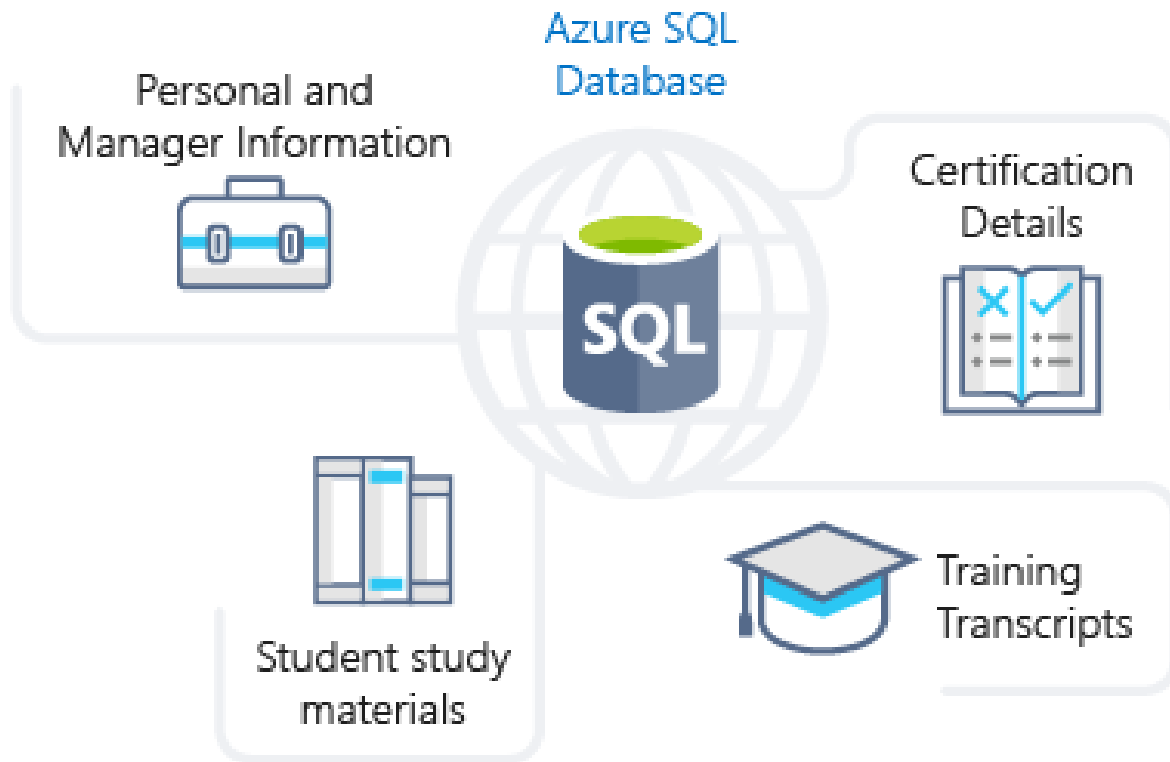
Azure Cosmos DB

Azure Cosmos DB is a globally distributed database service. It supports schema-less data that lets you build highly responsive and **Always On** applications to support constantly changing data. You can use this feature to store data that is updated and maintained by users around the world. The following illustration shows a sample Azure Cosmos DB database that's used to store data that's accessed by people located across the globe.
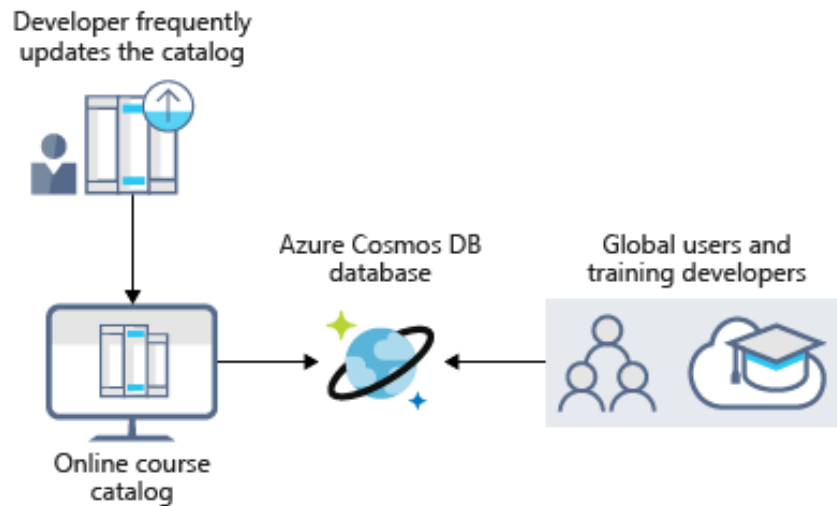
Azure Blob storage

Azure Blob Storage is *unstructured*, meaning that there are no restrictions on the kinds of data it can hold. Blobs are highly scalable and apps work with blobs in much the same way as they would work with files on a disk, such as reading and writing data. Blob Storage can manage thousands of simultaneous uploads, massive amounts of video data, constantly growing log files, and can be reached from anywhere with an internet connection.

Blobs aren't limited to common file formats. A blob could contain gigabytes of binary data streamed from a scientific instrument, an encrypted message for another application, or data in a custom format for an app you're developing.

Azure Blob storage lets you stream large video or audio files directly to the user's browser from anywhere in the world. Blob storage is also used to store data for backup, disaster recovery, and archiving. It has the ability to store up to 8 TB of data for virtual machines. The following illustration shows an example usage of Azure blob storage.

Streaming video or audio
files to the browser

Azure Data Lake Storage

The Data Lake feature allows you to perform analytics on your data usage and prepare reports. Data Lake is a large repository that stores both structured and unstructured data.

Azure Data Lake Storage combines the scalability and cost benefits of object storage with the reliability and performance of the Big Data file system capabilities. The following illustration shows how Azure Data Lake stores all your business data and makes it available for analysis.
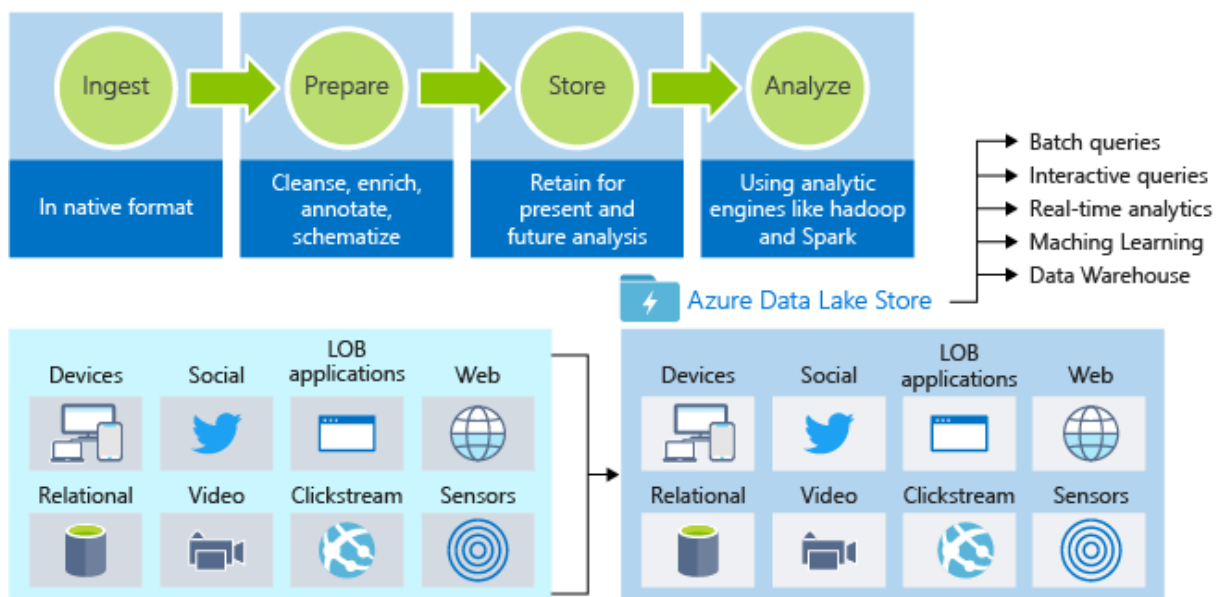
Azure Files

Azure Files offers fully managed file shares in the cloud that are accessible via the industry standard Server Message Block (SMB) protocol. Azure file shares can be mounted concurrently by cloud or on-premises deployments of Windows, Linux, and macOS. Applications running in Azure virtual machines or cloud services can mount a file storage share to access file data, just as a desktop application would mount a typical SMB share. Any number of Azure virtual machines or roles can mount and access the file storage share simultaneously. Typical usage scenarios would be to share files anywhere in the world, diagnostic data, or application data sharing.
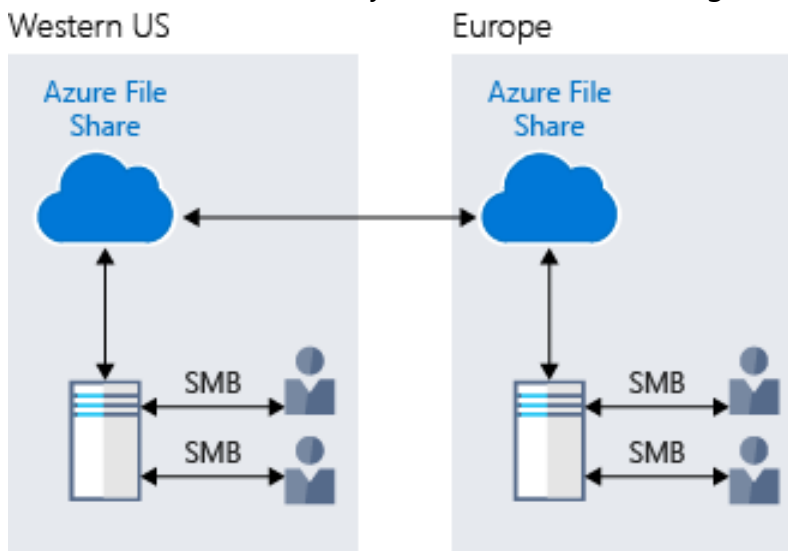
The following illustration shows Azure Files being used to share data between two geographical locations. Azure Files uses the Server Message Block (SMB) protocol that ensures the data is encrypted at rest and in transit.
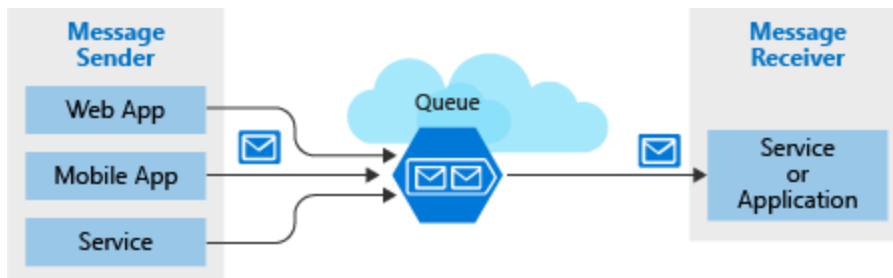
Azure Queue

Azure Queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world.

Azure Queue Storage can be used to help build flexible applications and separate functions for better durability across large workloads. When application components are decoupled, they can scale independently. Queue storage provides asynchronous message queueing for communication between application components, whether they are running in the cloud, on the desktop, on-premises, or on mobile devices.

Typically, there are one or more sender components and one or more receiver components. Sender components add messages to the queue, while receiver components retrieve messages from the front of the queue for processing. The following illustration shows multiple sender applications adding messages to the Azure Queue and one receiver application retrieving the messages.

You can use queue storage to:

- Create a backlog of work and to pass messages between different Azure web servers.
- Distribute load among different web servers/infrastructure and to manage bursts of traffic.
- Build resilience against component failure when multiple users access your data at the same time.

Disk Storage

Disk storage provides disks for virtual machines, applications, and other services to access and use as they need, similar to how they would in on-premises scenarios. Disk storage allows data to be persistently stored and accessed from an attached virtual hard disk. The disks can be managed or unmanaged by Azure, and therefore managed and configured by the user. Typical scenarios for using disk storage are if you want to lift and shift applications that read and write data to persistent disks, or if you are storing data that is not required to be accessed from outside the virtual machine to which the disk is attached.

Disks come in many different sizes and performance levels, from solid-state drives (SSDs) to traditional spinning hard disk drives (HDDs), with varying performance abilities.

When working with VMs, you can use standard SSD and HDD disks for less critical workloads, and premium SSD disks for mission-critical production applications. Azure Disks have consistently delivered enterprise-grade durability, with an industry-leading ZERO% annualized failure rate. The following illustration shows an Azure virtual machine using separate disks to store different data.
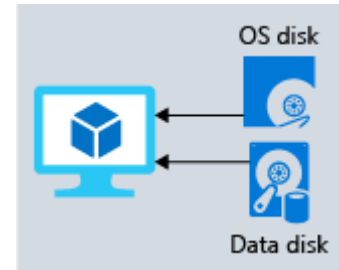
Storage tiers

Azure offers three storage tiers for blob object storage:

1. **Hot storage tier**: optimized for storing data that is accessed frequently.
2. **Cool storage tier**: optimized for data that is infrequently accessed and stored for at least 30 days.
3. **Archive storage tier**: for data that is rarely accessed and stored for at least 180 days with flexible latency requirements.

Encryption and replication

Azure provides security and high availability to your data through encryption and replication features.

Encryption for storage services

The following encryption types are available for your resources:

1. **Azure Storage Service Encryption (SSE)** for data at rest helps you secure your data to meet the organization's security and regulatory compliance. It encrypts the data before storing it and decrypts the data before retrieving it. The encryption and decryption are transparent to the user.
2. **Client-side encryption** is where the data is already encrypted by the client libraries. Azure stores the data in the encrypted state at rest, which is then decrypted during retrieval.

Replication for storage availability

A replication type is set up when you create a storage account. The replication feature ensures that your data is durable and always available. Azure provides regional and geographic replications to protect your data against natural disasters and other local disasters like fire or flooding.

Comparison between Azure data storage and on-premises storage

Now that you know about the benefits and features of Azure data storage, let's see how it differs from on-premises storage.

The term "on-premises" refers to the storage and maintenance of data on local hardware and servers. There are several factors to consider when comparing on-premises to Azure data storage.

Cost effectiveness

An on-premises storage solution requires dedicated hardware that needs to be purchased, installed, configured, and maintained. This can be a significant up-front expense (or capital cost). Change in requirements can require investment in new hardware. Your hardware needs to be capable of handling peak demand which means it may sit idle or be under-utilized in off-peak times.

Azure data storage provides a pay-as-you-go pricing model which is often appealing to businesses as an operating expense instead of an upfront capital cost. It's also scalable, allowing you to scale up or scale out as demand dictates and scale back when demand is low. You are charged for data services only as you need them.

Reliability

On-premises storage requires data backup, load balancing, and disaster recovery strategies. These can be challenging and expensive as they often each need dedicated servers requiring a significant investment in both hardware and IT resources.

Azure data storage provides data backup, load balancing, disaster recovery, and data replication as services to ensure data safety and high availability.

Storage types

Sometimes multiple different storage types are required for a solution, such as file and database storage. An on-premises approach often requires numerous servers and administrative tools for each storage type.

Azure data storage provides a variety of different storage options including distributed access and tiered storage. This makes it possible to integrate a combination of storage technologies providing the best storage choice for each part of your solution.

Agility

Requirements and technologies change. For an on-premises deployment this may mean provisioning and deploying new servers and infrastructure pieces, which is a time consuming and expensive activity.

Azure data storage gives you the flexibility to create new services in minutes. This flexibility allows you to change storage back-ends quickly without needing a significant hardware investment.

The following illustration shows differences between on-premises storage and Azure data storage.

Needs		On-premise	Azure Data Storage
Compliance and Security	1	Dedicated servers required for privacy and security	Client side encryption and encryption at rest
Store structured and unstructured data	2	Additional IT resources with dedicated servers required	Azure Data Lake and portal analyzes and manages all types of data
Replication and High Availability	3	More resources, licensing, and servers required	Built-in replication and redundancy features available
Application sharing and access to shared resources	4	File sharing requires additional administration resources	File sharing options available without additional license
Relational Data storage	5	Needs a database server with database admin role	Offers Database-as-a-Service option
Distributed storage and data access	6	Expensive storage, networking, and compute resources needed	Azure Cosmos DB provides price-winning distributed access
Messaging and load balancing	7	Hardware redundancy impacts budget and resources	Azure Queue provides effective load balancing
Tiered storage	8	Management of tiered storage needs technology and labor skillset	Azure offers automated tiered storage of data

Summary

In this module, you explored the benefits of using Azure to store your data. Azure provides the following features:

- Storage of both structured and unstructured data
- High security that supports global compliance standards
- Load balancing, high availability, and redundancy capabilities
- The ability to send large volumes of data directly to the browser using features such as Azure Blob storage

Ultimately, the capabilities of Azure data storage make it an ideal platform for hosting any large global application or portal.