Azure storage services

Data Types

- Structured - Data that can be represented using tables with very strict schema. Each row must follow defined schema. Some tables have defined relationships between them. Typically used in relational databases.
- Semi-structured - Data that can be represented using tables but without strict defined schema. Rows must only have unique key identifier.
- Unstructured - Any files in any format. Like binary files, application files, images, movies, etc.

Storage Account

- Group of services which include
    - blob storage,
    - queue storage,
    - table storage, and
    - file storage
- Used to store
    - files,
    - messages, and
    - semi-structured data
- Highly scalable (up to petabytes of data)
- Highly durable (99.999999999% - 11 nines, up to 16 nines)
- Cheapest per GB storage

Blob Storage

- BLOB – binary large object – file
- Designed for storage of files of any kind
- Three storage tiers
    - Hot – frequently accessed data
    - Cool – infrequently accessed data (lower availability, high durability)
    - Archive – rarely (if-ever) accessed data

Queue Storage

- Storage for small pieces of data (messages)
- Designed for scalable asynchronous processing

Table Storage

- Storage for semi-structured data (NoSQL)
    - No need for foreign joins, foreign keys, relationships, or strict schema
    - Designed for fast access
- Many programming interfaces and SDKs

File Storage

- Storage for files accessed via shared drive protocols
- Designed to extend on-premises file shares or implement lift-and-shift scenarios

Disk Storage

- Disk emulation in the cloud
- Persistent storage for Virtual Machines
- Different
    - sizes,
    - types (SSD, HDD)
    - performance tiers
- Disk can be unmanaged or managed

Database services

Cosmos DB

- Globally distributed NoSQL (semi-structured data) Database service
- Schema-less
- Multiple APIs (SQL, MongoDB, Cassandra, Gremlin, Table Storage)
- Designed for
    - Highly responsive (real time) applications with super low latency responses <10ms
    - Multi-regional applications

SQL Database

- **Relational database** service in the cloud (PaaS) (DBaaS - Database as a Service)
- **Structured data service** defined using schema and relationships
- **Rich Query Capabilitie**s (SQL)
- **High-performance**, reliable, fully managed and secure database for building - applications

Azure SQL product family

- Azure **SQL Database** – Reliable relational database based on SQL Server
- Azure **Database for MySQL** – Azure SQL version for MySQL database engine
- Azure **Database for PostgreSQL** – Azure SQL version for PostgreSQL database engine
- Azure **SQL Managed Instance** – Fully fledged SQL Server managed by cloud provider
- Azure **SQL on VM** – Fully fledged SQL Server on IaaS
- Azure **SQL DW (Synapse)** – Massively Parallel Processing (MPP) version of SQL Server

Azure Marketplace

- Think of it like an “Azure Shop” where you purchase services and solutions for the Azure platform
- Each product is a template which contains one or multiple services
- Products are delivered by first and third-party vendors
- Solutions can leverage all service categories like IaaS, PaaS and SaaS

Describe some of the solutions available on Azure

What is Internet of Things?

Internet of Things (**IoT**) is a network of internet connected devices (**IoT Devices**) embedded in everyday objects enabling sending and receiving data such as **settings** and **telemetry**.

Azure Iot Hub

- Managed service for bi-directional communication
- Platform as a Service (PaaS)
- Highly secure, scalable and reliable
- Integrates with a lot of Azure Services
- Programmable SDKs for popular languages (C, C#, Java, Python, Node.js)
- Multiple protocols (HTTPS, AMQP, MQTT)

Azure IoT Central

- IoT App Platform - Software as a Service (SaaS)
- Industry specific app templates
- No deep technical knowledge required
- Service for connecting, management and monitoring IoT devices
- Highly secure, scalable and reliable
- Built on top of the IoT Hub service and 30+ other services

Azure Sphere

- Secure end-2-end IoT Solutions
    - Azure Sphere certified chips (microcontroller units - MCUs)
    - Azure Sphere OS based on Linux
    - Azure Security Service trusted device-to-cloud communication

What is Big Data?

Big Data is a field of technology that helps with the **extraction**, **processing** and **analysis** of information that is **too large or complex** to be dealt with by traditional software.

The three V's rule

Big data typically has one of the following characteristics

- **Velocity** - how fast the data is coming in or how fast we are processing it
    - Batch
    - Periodic
    - Near Real Time
    - Real Time
- **Volume** - how much data we are processing
    - Megabytes
    - Gigabyte
    - Terabytes
    - Petabytes
- **Variety** - how structured/complex the data is
    - Tables
    - Databases
    - Photo, Audio
    - Video, Social Media

Azure Synapse Analytics

- Big data analytics platform (PaaS)
- Multiple components
    - Spark
    - Synapse SQL

        - SQL pools (dedicated – pay for provisioned performance)
        - SQL on-demand (ad-hoc – pay for TB processed)
    - Synapse Pipelines (Data Factory – ETL)
    - Studio (unified experience)

Azure HDInsight

- Flexible multi-purpose big data platform (PaaS)
- Multiple technologies supported (Hadoop, Spark, Kafka, HBase, Hive, Storm, Machine Learning)

Azure Databricks

- Big data collaboration platform (PaaS)
- Unified workspace for notebook, cluster, data, access management and collaboration
- Based on Apache Spark
- Integrates very well with common Azure data services

What is Artificial Intelligence?

Artificial Intelligence (**AI**) is the simulation of human intelligence & capabilities by computer software.

What is Machine Learning?

Machine Learning is a subcategory of AI where a computer software is "**taught**" to **draw conclusions** and **make predictions from data**.

Azure Machine Learning

- Cloud-based platform for creating, managing and publishing machine learning models
- Platform as a Service (PaaS)
- Machine Learning Workspace – top level resource
- Machine Learning Studio – web portal for end-2-end development
- Features
    - Notebooks – using Python and R
    - Automated ML – run multiple algorithms/parameters combinations, choose the best model

    - Designer – graphical interface for no-code development
    - Data & Compute – management of storage and compute resources
    - Pipelines – orchestrate model training, deployment and management tasks

What is Serverless?

Serverless computing is cloud-hosted execution environment that allows customers to **run their applications** in the cloud while **completely abstracting underlying infrastructure**.

Azure Functions

- Serverless coding platform (Functions as a Service, FaaS)
- Designed for nano-service architectures and event-based applications
- Scales up and down very quickly
- Highly scalable
- Supports popular languages and frameworks (.NET & .NET Core, Java, Node.js, Python, PowerShell, etc.)

Azure Logic Apps

- Serverless enterprise integration service (PaaS)
- 200+ connectors for popular services
- Designed for orchestration of
    - business processes,
    - integration workflows for applications, data, systems and services
- No-code solution

Azure Event Grid

- Fully managed serverless event routing service
- Uses publish-subscribe model
- Designed for event-based and near-real time applications
- Supports dozens of built-in events from most common Azure services

What is DevOps?

DevOps is a set of practices that combine both development (**Dev**) and operations (**Ops**).

DevOps aims to **shorten the development life cycle** by providing **continuous integration** and **delivery** (CI/CD) capabilities while **ensuring high quality** of deliverables.

Azure DevOps

- **Collection of services** for building solutions using DevOps practices
- Services included
    - **Boards** – tracking work
    - **Pipelines** – building CI/CD workflows (build, test and deploy apps)
    - **Repos** – code collaboration and versioning with Git
    - **Test Plans** – manual and exploratory testing
    - **Artifacts** – manage project deliverables
- Extensible with **Marketplace** – over 1000 of available apps
- Evolved from **TFS** (Team Foundation Server), through **VSTS** (Visual Studio Team Services)

Azure DevTest Labs

- Service for creation of **sandbox environments** for developers/testers (PaaS)
- Quick setup of **self-managed virtual machines**
- **Preconfigured templates** for VMs
- Plenty of additional **artifacts** (tools, apps, custom actions)
- Lab **policies** (quotas, sizes, auto-shutdowns)
- **Share** and **automate** labs via custom images
- Premade plugins/API/tools for **CI/CD pipeline automation**

Describe Azure management tools

Azure Portal

- Public web-based interface for management of Azure platform
- Designed for self-service
- Customizable
- Simple tasks

Azure PowerShell

- PowerShell and module
- Designed for automation
- Multi-platform with PowerShell Core
- Simple to use
    - Connect-AzAccount – log into Azure
    - Get-AzResourceGroup – list resource groups
    - New-AzResourceGroup – create new resource group
    - New-AzVm – create virtual machine

Azure CLI

- Command Line Interface for Azure
- Designed for automation
- Multi-platform (Python)
- Simple to use
    - az login – log into Azure
    - az group list – list resource groups
    - az group create – create new resource group
    - az vm create – create virtual machine
- Native OS terminal scripting

Azure Cloud Shell

- Cloud-based scripting environment
- Completely free
- Supports both Azure PowerShell and Azure CLI
- Dozen of additional tools
- Multiple client interfaces
    - Azure Portal integration (portal.azure.com)
    - Shell Portal (shell.azure.com)
    - Visual Studio Code Extension
    - Windows Terminal
    - Azure Mobile App
    - Microsoft Docs integration

Azure Advisor

- **Personalized consultant** service
- Designed to provide **recommendations** and **best practices** for
    - **Cost** (SKU sizes, idle services, reserved instances, etc.)
    - **Security** (MFA settings, vulnerability settings, agent installations, etc.)
    - **Reliability** (redundancy settings, soft delete on blobs, etc.)
    - **Performance** (SKU sizes, SDK versions, IO throttling, etc.)
    - **Operational** Excellence (service health, subscription limits, etc.)
- **Actionable** recommendations
- **Free**!

Azure file movement options

AzCopy
it is a command-line tool that can be used to upload and download data to from Azure Blob storage. AzCopy can be used to transfer data within Azure storage accounts or between storage accounts. It supports both block blobs and page blobs.

Azure storage Explorer
It is a graphical user interface (GUI) tool that you can use to manage your Azure Storage resources. With Storage Explorer, you can create and manage storage accounts, blobs, queues, tables, and files. You can also monitor your storage account metrics and access your stored data through the various Storage Explorer features.

Azure Files
It is a cloud file storage service from MS Azure that operates like a traditional Service Message Block (SMB) or Network File System (NFS) file server. This makes Azure File directly accessible by windows, Linux, macOS clients.