Azure Architect Technologies

AZ-301 Exam Preparation Deck

Fundamentals of Design

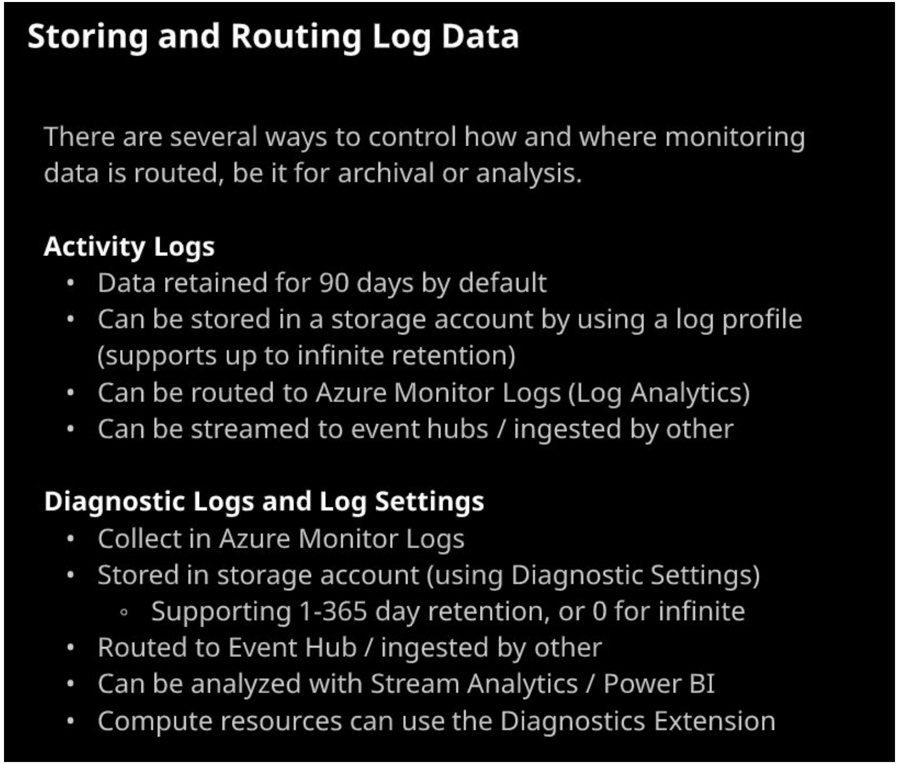
- Four fundamental design principles:
  - **Scalability**
    - Vertical scaling (scale up) - increase resource size.
    - Horizontal scaling (scale out) - add instances/resources.
  - **Reliability**
    - Availability - maintain uptime
    - Recoverability - restore services after failure
      - RPO / RTO
  - **Efficiency**
    - Cost effective/resource utilization/licensing
  - **Security**
    - Defense in depth / layered security
    - Shared security model
- In order to build a solution that correctly balances all four design fundamentals we need to ensure the **requirements** are accurately captured first and foremost.

https://docs.microsoft.com/en-us/azure/architecture/

Auditing & Monitoring Solutions

In Azure, there are four (4) main components that give us the ability to view and consume log and resource performance data within our environment. These are:

- Azure Monitor
  - Logs, metrics, health visualisation, and analytics.
- Application Insights (App Insights)
  - Used to gather in-depth analytical data about applications within Azure.
- Network Insights
- Integration
  - Relates to integrating with other monitoring tools; Power BI, Event Hubs, and SIEM solutions.
-



Storing and Routing Log Data

There are several ways to control how and where monitoring data is routed, be it for archival or analysis.

Activity Logs
- Data retained for 90 days by default
- Can be stored in a storage account by using a log profile (supports up to infinite retention)
- Can be routed to Azure Monitor Logs (Log Analytics)
- Can be streamed to event hubs / ingested by other

Diagnostic Logs and Log Settings
- Collect in Azure Monitor Logs
- Stored in storage account (using Diagnostic Settings)
  - Supporting 1-365 day retention, or 0 for infinite
- Routed to Event Hub / ingested by other
- Can be analyzed with Stream Analytics / Power BI
- Compute resources can use the Diagnostics Extension

Log Analytics Agent vs. Azure Diagnostics Extension

Azure Log Analytics Agent:

- Developed for comprehensive management across virtual machines in any cloud, on-premises machines, and those monitored by System Center Operations Manager.
- Windows and Linux agents send collected data from different sources to your Log Analytics workspace in Azure Monitor, as well as any unique logs or metrics as defined in a monitoring solution.
- Supports insights and other services in Azure Monitor.
- Log Analytics agent referred to as the Microsoft Monitoring Agent (MMA) or OMS Linux agent.
- Required for solutions, Azure Monitor for VMs, and other services such as Azure Security Centre.

Monitoring Log Agent Comparison:

https://docs.microsoft.com/en-us/azure/azure-monitor/platform/agents-overview

Azure Diagnostics Extension:

- **Can be used only with Azure virtual machines.**
- Azure Diagnostics extension sends data to **Azure Storage,** Azure Monitor Metrics (Windows only) and Event Hubs.

DETERMINE WORKLOAD REQUIREMENTS (10-15%)

Optimising Consumption

Purchasing Options:

- **PAYG** - Pay-as-you-go / standard purchase option for Azure. Only pay for services you use.
- **EA** - Enterprise Agreement. Requires a committed amount you will spend in Azure, paid upfront. Typically used for larger enterprises.
- **Dev/Test** - Access to discounted Azure services, only when used for **development and testing**. Includes Microsoft software licensing for virtual machines and access to Win 10.
- **CSP** - Cloud Solution Provider offer is where a Microsoft partner sells and bills the Azure services directly to customers. Agreement is with partner (CSP) not Microsoft.
- **Support** - Microsoft provides a range of support plans which largely differ based on response times and the level of advice provided.

- **Azure Reservations** - Save money by purchasing long-term service upfront.
- Services must be **prepaid** to receive discount.
- Discount at high as **72%***
- Reservation time varies - typically 1 or 3 years.
- Services available include:
  - Virtual Machines (VMs)
  - Azure SQL Data Warehouse
  - Azure Cosmos DB
  - SQL Database
  - Azure Databricks
  - App Service
  - A range of other software / OS licenses
- **Azure Hybrid Benefit** can provide up to 40% savings on Windows virtual machines and is available to those with **software assurance** for Windows Server licensing.
- Software assurance grants windows server licences to be used **on-premises AND in-cloud (Azure).**

Optimising Consumption (Cont.)

- Tools available to assist with Azure cost optimisation:
  - Azure Pricing Calculator
  - Total Cost of Ownership Calculator (TCO)

Service Level Agreements (SLAs)

Virtual Machines

- **99.99% when using Availability Zones**
- 99.9% when using Availability Sets

Virtual Machine Scale Sets

Availability Zones

- 99.99% SLA when availability zones are used.

Availability Sets

Application Gateways

- **99.95% with two or more medium or large instances**

Load-Balancer

- Standard SKU - 99.99% SLA
- **Basic is excluded** from SLA.

Azure Container Instance

- 99.9% SLA

https://azure.microsoft.com/en-us/support/legal/sla/

Traffic Manager

- 99.99% SLA for guaranteed response to DNS queries

Cosmos DB

- 99.99% SLA for all single region accounts, and all multi-region accounts with relaxed consistency.

Information Gathering & Requirements

Requirements gathering can take two main forms:

1. Technical requirements
    a. Ensuring to deploy the best Azure service for the job.
    b. Provide integration with existing solutions.
    c. Plan and undertake your migration.
    d. Technical questions - detailed in nature
        i. CPU, resources, architecture/patterns
        ii. Monolithic application or multi-tiered?
2. Non-technical requirements
    a. Meet or exceed service-level agreement targets.
    b. Secure data according to relevant legislation and requirements.
    c. Retain data according to relevant legislation and requirements.
    d. Train staff to maintain the solution.

DESIGN FOR IDENTITY & SECURITY (20-25%)

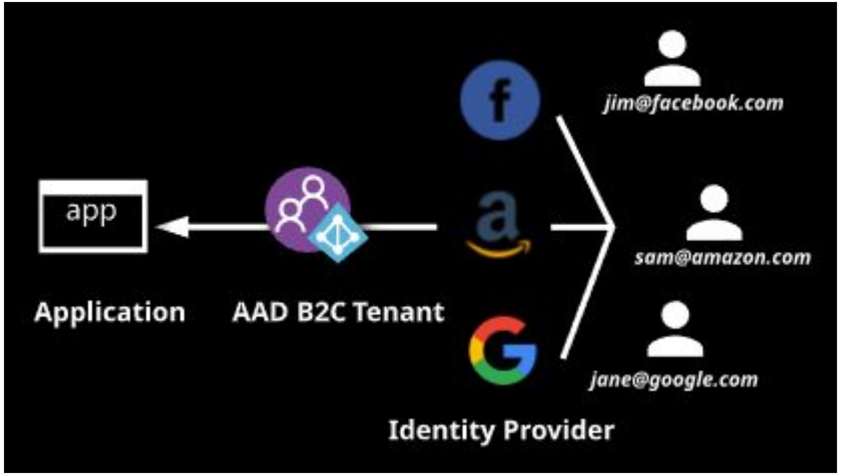
Azure AD Business-to-Business (B2B)

- Azure AD B2B allows you to invite users from external or partner organisations to collaborate.
- Guests/partners **use their own identity management solution**.
- Guest users sign-in **using their own work, school, social identities**.
- Simplified/streamlined process for federating your business identity with an external partner's.
- Main features include:
  - Guest Access - easily invite users as **guests** to securely share access to your resources.
  - Organisation Relationships - simplified federation with partner identities.
  - Terms of Use - custom welcome emails with terms-of-use sent for guest invitations.
  - Licensing - automatic licensing provided to guest users (calculated at 1:5 ratio).

Note: Guest accounts can be assigned to access controls.

Azure AD Business-to-Consumer (B2C)

- Azure AD B2C helps simplify the identity management for developers who need to secure access to their applications.
- Allows developers to use Azure AD as their identity system.
- Supports customer sign-in from existing identities like:
    - Facebook
    - Amazon
    - Google



Azure AD Domain Services (DS)

- The Azure AD DS managed domain is a stand-alone domain. It **IS NOT** an extension of an on-premises domain.
- BackEnd active directory virtual machines are managed by Microsoft.
- Integrates with Azure AD (using synchronisation).
- Supports a range of typical AD features, such as:
  - LDAP binding and LDAP read
  - Secure LDAP (LDAPS)
  - Group Policy (GPO)
  - DNS Management
- Important limitations:
  - **Does not support LDAP write.**
  - **Cannot support AD domain or forest trusts.**
  - **Not all objects are synchronised (on-prem GPOs, computer objects, and organisation units (OUs).**
- One-way sync from Azure AD to Azure AD DS.
- You can create resources in the managed domain, but they **are not** synchronised back to Azure AD.

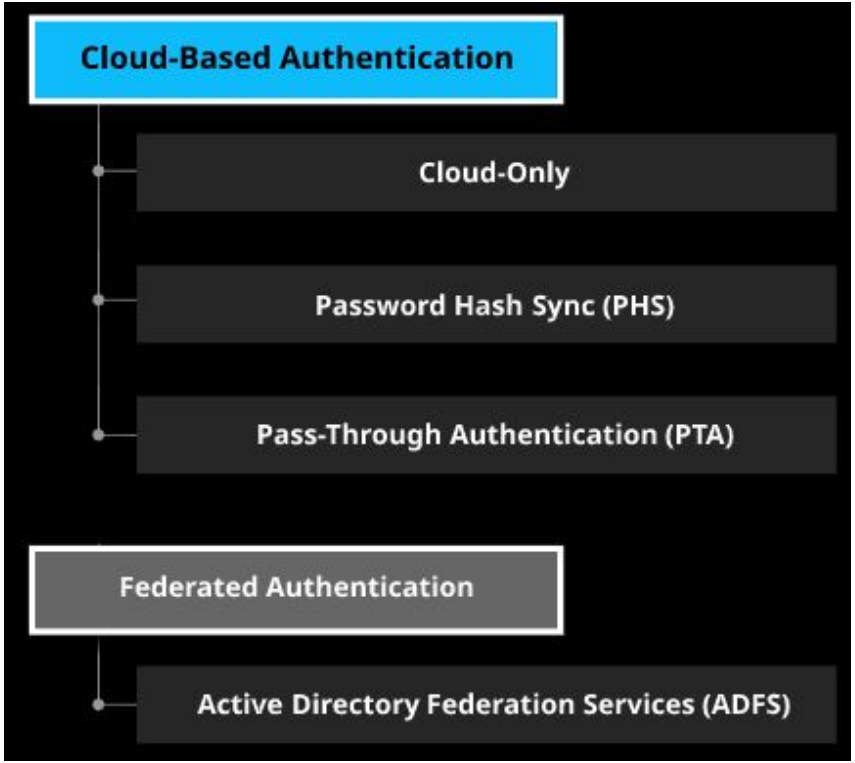
- https://docs.microsoft.com/en-us/azure/active-directory-domain-services/overview
- Also supports **domain join** functionality for Azure virtual machines.
- **Cannot** be used to form domain forests and trusts.

Management Groups

- Provide a top-level of control and governance over Azure subscriptions.
- Once management groups are created, the default top-level management group is called the “**Tenant Root Group**”
- TNG contains all Azure subscriptions beneath it.
- Uses Role-Based Access Control (RBAC)
- A management group tree can support up to six levels of depth.
  - This limit **doesn’t include the ROOT LEVEL or the SUBSCRIPTION LEVEL.**
- Each management group and subscription can only support **ONE PARENT**.
- Each management group can have many children.
- 10,000 management groups can be supported in a single directory.

- https://docs.microsoft.com/en-us/azure/governance/management-groups/overview

Azure AD Authentication Options


Cloud-Based Authentication
Cloud-Only
Password Hash Sync (PHS)
Pass-Through Authentication (PTA)
Federated Authentication
Active Directory Federation Services (ADFS)

Important things to note:
- Cloud authentication ONLY occurs in the cloud.
- On-premises enterprise management is improving, with products such as:
  - Azure AD DS
  - Microsoft InTune and enterprise device management
  - Azure AD Join

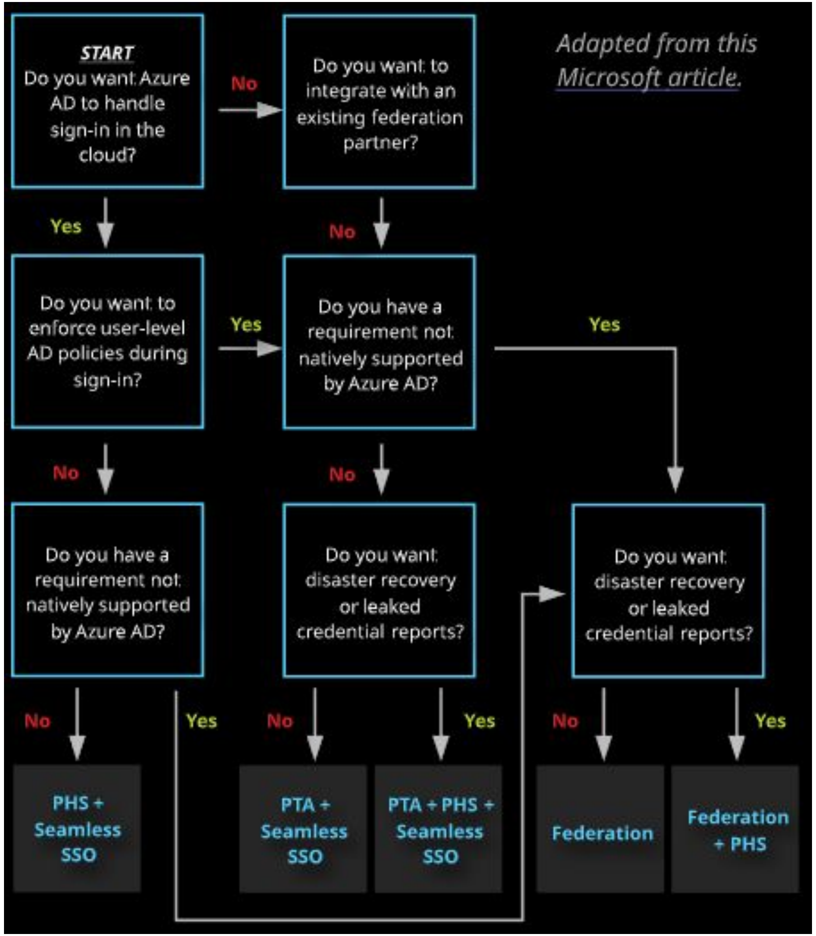
- **Password Hash Sync:**
  - Stores hash of a hash of the password in the cloud (double hash). May be an issue for regulatory/compliance reasons.
- Improves user experience with seamless single sign-on (SSSO).
- Co-exists with ADFS or pass-through authentication.
- **Important limitations of PHS**:
  - Account changes not immediately enforced.
  - **On-prem account policies do not apply.**

Azure AD Authentication Options (Cont).

- Pass-through Authentication:
  - Passwords/hashes are NOT stored in the cloud.
  - Supports seamless single sign-on (SSSO)
  - Slightly more complex than password hash sync, but less complex than ADFS.
  - Is able to enforce on-premises account policy at the time of sign-in (e.g. disabled accounts for logon hour restrictions).
- Important limitations for PTA include:
  - Does not include access to advanced features like Azure AD Identity Protection / leaked password protection.
  - Using multi-factor auth requires Azure Multi-Factor Authentication.

- Active Directory Federation Services (ADFS)
  - All authentication occurs on-premises.
  - No support for seamless single sign-on (SSSO).
  - Requires complex on-premises infrastructure.
  - Does support additional features such as:
    - Authentication using smartcards and certificates.
    - On-prem multi-factor authentication servers.
    - Sign-in that requires an sAMAccountName (e.g. MYDOMAIN\jason) instead of a User Principal Name (UPN) (e.g. jason@mydomain.com).

Azure AD Authentication Options (Cont).



START
Do you want Azure AD to handle sign-in in the cloud?
No
Do you want to integrate with an existing federation partner?
Adapted from this Microsoft article.
Yes
No
Do you want to enforce user-level AD policies during sign-in?
Yes
Do you have a requirement not natively supported by Azure AD?
Yes
No
No
Do you have a requirement not natively supported by Azure AD?
Do you want disaster recovery or leaked credential reports?
Do you want disaster recovery or leaked credential reports?
No
Yes
No
Yes
No
Yes
PHS + Seamless SSO
PTA + Seamless SSO
PTA + PHS + Seamless SSO
Federation
Federation + PHS

Azure AD Connect Health

Azure AD Connect Health provides monitoring and alerts for synchronisation and infrastructure health for hybrid identities using Azure AD Connect.

Key features:

- **Synchronisation monitoring**
  - Alerts for sync issues.
  - Sync insights (sync latency, object changes).
  - Object-level synchronisation error reporting.
- **AD FS monitoring**
  - Monitoring and alerts for infra health.
  - Extranet lockout trends.
  - Failed sign-ins reporting.
  - Usage analytics (top apps, performance metrics).
- **Azure AD DS monitoring**
  - Monitoring and alerts for infra health.
  - Domain controller dashboard.
  - Replication status dashboard.

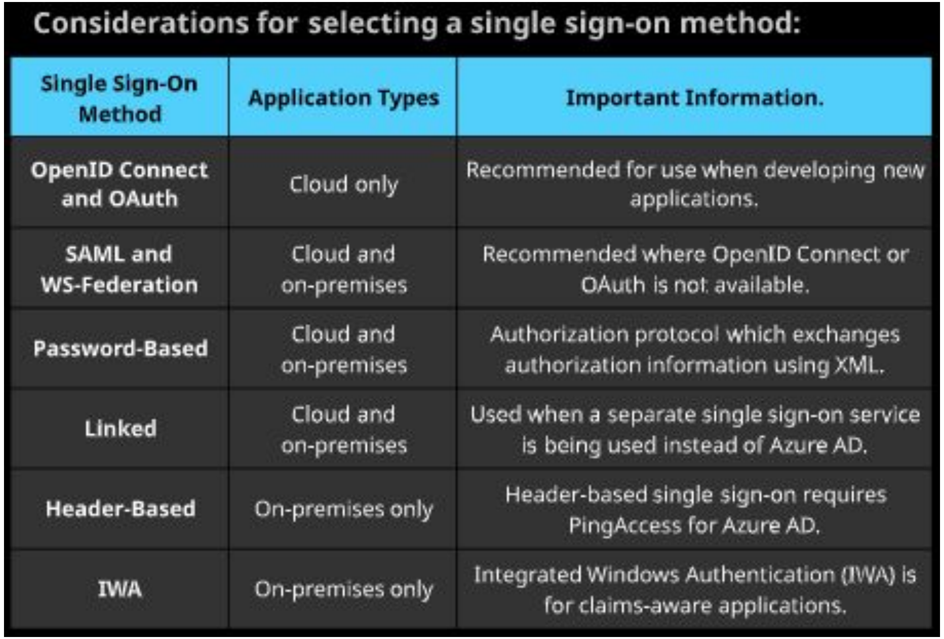
Important considerations:

- Requires Azure AD Premium licensing (P1 or greater).
  - The first Connect Health Agent requires **one license**.
  - Each **additional agent requires 25 additional licenses**.
- Requires agent installation on each targeted server.

Types of sync errors Azure AD Connect Health can report on:

- Duplicate attributes
- Data mismatch
- Data validation failures
- Large attribute
- Federated domain controllers
- Existing admin role

Application Single Sign-On (SSO)

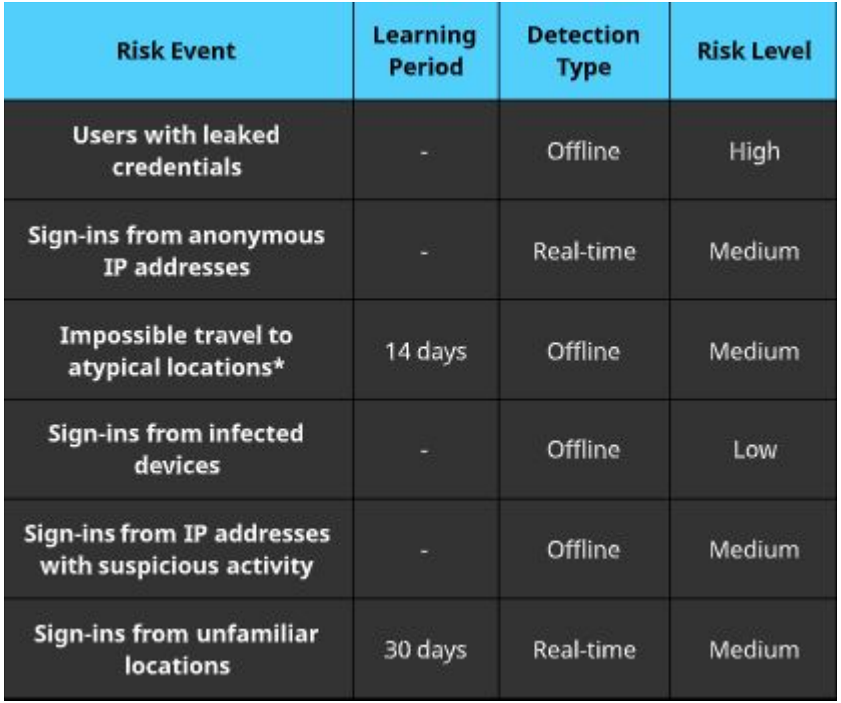


Considerations for selecting a single sign-on method:

Single Sign-On Method	Application Types	Important Information.
OpenID Connect and OAuth	Cloud only	Recommended for use when developing new applications.
SAML and WS-Federation	Cloud and on-premises	Recommended where OpenID Connect or OAuth is not available.
Password-Based	Cloud and on-premises	Authorization protocol which exchanges authorization information using XML.
Linked	Cloud and on-premises	Used when a separate single sign-on service is being used instead of Azure AD.
Header-Based	On-premises only	Header-based single sign-on requires PingAccess for Azure AD.
IWA	On-premises only	Integrated Windows Authentication (IWA) is for claims-aware applications.

Azure AD Identity Protection

- Supports three (3) main types of policies:
  - MFA registration policy - enforce MFA registration
  - Sign-in risk policy (mitigation) - require MFA to sign-in based on **real-time risk.**
  - User risk policy (remediation) - block access or require password changes for users identified as risky (after the fact - offline detection).
- **Key takeaways:**
  - **Sign-in risk policies occur in real-time.**
  - **User risk policies mitigate AFTER** / offline.
- **Requires Azure AD Premium P2 licensing.**
- **Each administrator or affected user requires a license.**



Risk Event	Learning Period	Detection Type	Risk Level
Users with leaked credentials	-	Offline	High
Sign-ins from anonymous IP addresses	-	Real-time	Medium
Impossible travel to atypical locations*	14 days	Offline	Medium
Sign-ins from infected devices	-	Offline	Low
Sign-ins from IP addresses with suspicious activity	-	Offline	Medium
Sign-ins from unfamiliar locations	30 days	Real-time	Medium

Azure AD Conditional Access Policies

- Azure AD Conditional Access Policies can integrate with other solutions, such as:
  - Identity Protection: to enable risk based policies*
  - Intune: modern workplace management (MDM)
  - Cloud App Security: cloud access security broker
- Conditional Access licensing requirements:
  - Core functionality - Azure AD Premium P1 licensing.
  - Risk-based policies* - Azure AD Premium P2 licensing.

Azure AD Password Protection

- Azure AD Password Protection provides two important features to help users protect their identities:
    - **Banned Passwords**
        - Protects Azure AD and on-premises AD users
        - Over 500 of the most commonly used passwords in ban list.
        - Over 1 million character substitution variations.
    - **Smart Lockout**
        - Intelligent lock-out system enabled by default.
        - Protects against brute-force and password spray.
        - Supports hybrid identity password hash sync or pass-through authentication with limitations.

LICENSING:

- Password Protection licensing requirements:
    - No license required for cloud-only organisations (when custom lists are not required).
    - Full functionality requires **Azure AD Premium P1**.
- Smart lockout licensing requirements:
    - Included in all versions of Azure AD.
    - Customisation requires **Azure AD Premium P1**.

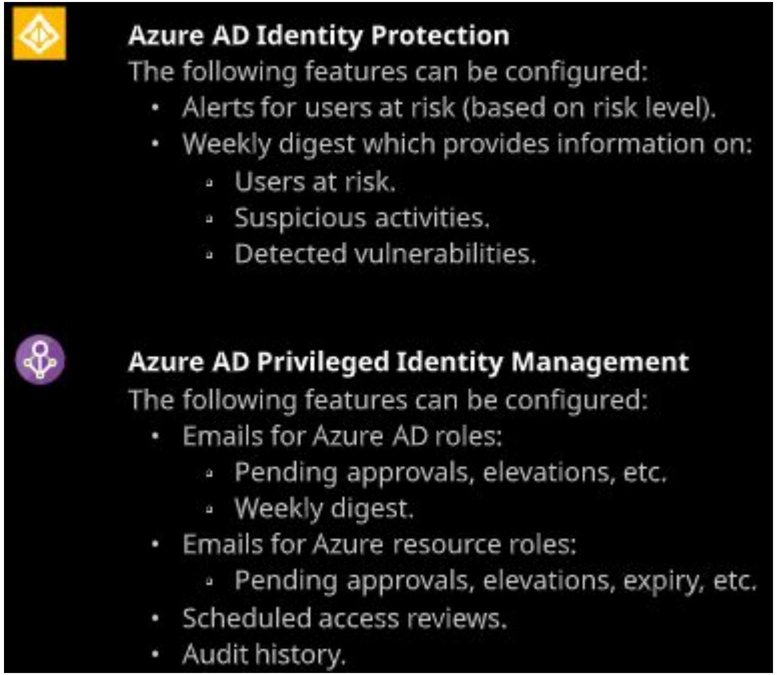
Azure Licensing

https://azure.microsoft.com/en-us/pricing/details/active-directory/

Azure AD Monitoring

- Key monitoring features available within Azure AD:
  - Security reports
    - Users flagged for risk
    - Risky sign-ins
  - Activity reports
    - Audit logs
    - Sign-ins (**requires Azure AD Premium**)
    - Usage and Insights (**requires Azure AD Premium**)
  - Standard monitoring experience
    - Route to storage account for retention (30 days by default). Use storage accounts for longer retention (e.g. 6 months or more).
    - Stream to Event Hubs for integration.
    - Use Azure Monitor Logs for analytics and alerts.

Monitoring for Other Products includes:



Azure AD Identity Protection

The following features can be configured:

- Alerts for users at risk (based on risk level).
- Weekly digest which provides information on:
  - Users at risk.
  - Suspicious activities.
  - Detected vulnerabilities.

Azure AD Privileged Identity Management

The following features can be configured:

- Emails for Azure AD roles:
  - Pending approvals, elevations, etc.
  - Weekly digest.
- Emails for Azure resource roles:
  - Pending approvals, elevations, expiry, etc.
- Scheduled access reviews.
- Audit history.

DESIGN A DATA PLATFORM SOLUTION (15-20%)

Data Management Platforms

There are three (3) main types of data management platforms available within Azure, there are:

- **Relational and non-relational databases:**
  - Where we can house structured and unstructured data.
- **Data warehouses:**
  - Structured data repositories designed for business intelligence and analytics.
  - Used for frequent reads/writes
- **Data lakes:**
  - Storage for big data analytics, machine learning, and data science.
  - Primarily stores **unstructured data**.

Relational Databases:

- Schema (structure) is well-defined and fixed.
- **Generally scales UP and DOWN**.
- Examples of relational databases can be:
  - Student management system
  - Product management
  - Bank transaction system
  - Library book and loan management

Non-relational Databases:

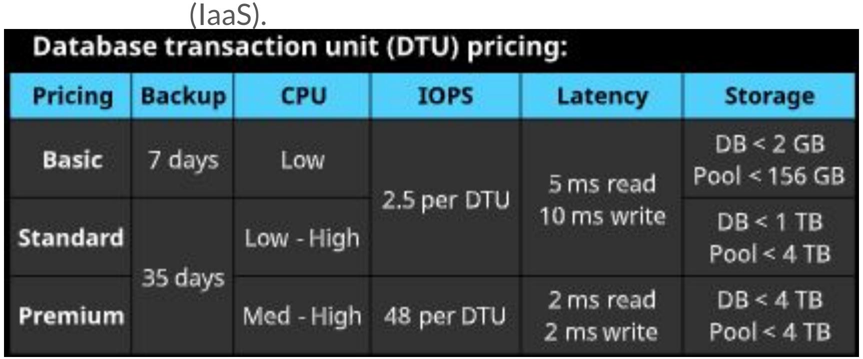
- Generally used when:
  - There are large amounts of data
  - The relationship between data isn't important
  - Data may change over time
- Database is more flexible than relational DB.
- **Generally scales IN and OUT.**
- Example use cases:
  - Storing social media information (tweets, pictures).
  - Internet of Things (IoT) or Content management.

Azure SQL

Houses structured data.
Relational database management system (DBMS).
For transactions (OLTP) - Online Transaction Platform

Azure SQL Deployment Options (3):

- **Databases**. Fully managed (backups, patching...etc)
- **Managed Instances** - lift-and-shift ready.
  - More of a managed service offering - PaaS
- **SQL virtual machines** - lift-and-shift ready
  - More control over the SQL server - dedicated VM (IaaS).



Database transaction unit (DTU) pricing:

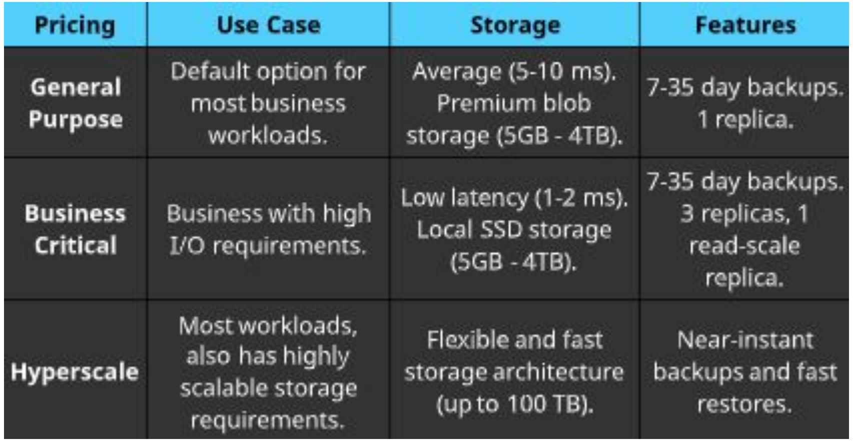
Pricing	Backup	CPU	IOPS	Latency	Storage
Basic	7 days	Low	2.5 per DTU	5 ms read 10 ms write	DB < 2 GB Pool < 156 GB
Standard	35 days	Low - High			DB < 1 TB Pool < 4 TB
Premium		Med - High	48 per DTU	2 ms read 2 ms write	DB < 4 TB Pool < 4 TB

SQL Elastic Pools - economic resource sharing

- Provides a way to group multiple databases together so as to better utilise a pool of resources.
- Suitable for housing several databases with:
  - Low average utilisation.
  - Relatively infrequent utilisation spikes.
- **Important info about elastic pools:**
  - You can create multiple pools on a server, **but you CANNOT add databases from different servers into the same pool.**
  - There is no per-database charge for elastic pools. **You are billed for each hour a pool exists** at the highest eDTU or vCores, regardless of usage or whether the pool was active for less than an hour (one hour).
- In general, not more than ⅔ (or 67%) of the databases in the pool should simultaneously peak to their resource limit.

Azure SQL (Cont.)

Pricing by vCore:



Pricing	Use Case	Storage	Features
General Purpose	Default option for most business workloads.	Average (5-10 ms). Premium blob storage (5GB - 4TB).	7-35 day backups. 1 replica.
Business Critical	Business with high I/O requirements.	Low latency (1-2 ms). Local SSD storage (5GB - 4TB).	7-35 day backups. 3 replicas, 1 read-scale replica.
Hyperscale	Most workloads, also has highly scalable storage requirements.	Flexible and fast storage architecture (up to 100 TB).	Near-instant backups and fast restores.

- Retention:
  - **BASIC, STANDARD, PREMIUM all support long-term database backup retention (up to 10 years).**

Important scaling considerations:

- Managed instances and single databases:
  - Will scale within limits of the pricing or plan.
  - Cannot truly autoscale.
  - Can be scaled manually.
- Alternative scaling methods:
  - Read scale-out: leverage read replicas to distribute load.
  - **Sharding:** an architectural pattern whereby identically structured data is distributed across many databases.
- Automatic tuning:
  - Automatic additions of indexes for databases.
  - Helps improve SQL query performance.
- SQL Server Integration Services (SSIS) is NOT available within Azure SQL.
  - We can use Azure Data Factory (ADF) as an alternative.

Azure SQL (Cont.)

- Azure SQL Server and SQL Managed Instance both support **cross-database queries**.
- Azure SQL Stretch Database is used for migrating cold data transparently and securely to Azure cloud.
  - https://docs.microsoft.com/en-us/sql/sql-server/stretch-database/stretch-database?view=sql-server-ver15
-

Data Migration Assistant (DMA)

- Helps upgrade to a modern data platform by detecting compatibility issues that can impact database functionality in your new version of SQL Server or Azure SQL database.
- DMA recommends performance and reliability improvements for your target environment.
- **For large migrations (in terms of number and size of databases) Microsoft recommends using Azure Database Migration Service - for migrating databases at scale.**
- Assess on-premises SQL server instance(s) migrating to Azure SQL database(s).
- **ONLY supports Microsoft SQL Server (sources and targets).**
- Supports migrations to Azure SQL Database single database, and Azure SQL managed instance.
- https://docs.microsoft.com/en-us/sql/dma/dma-overview?view=sql-server-ver15

Azure SQL Managed Instance

Azure SQL Managed Instance is designed for customers looking to migrate a large number of apps from an on-premises or IaaS, self-built, or ISV provided environment to a fully managed PaaS cloud environment, with as low a migration effort as possible.

- Uses **Azure Data Migration Service** to perform migration.
- Native VNET support - isolation of customer DB instances.
- No hardware purchasing, and management.
- Automated patching and version upgrade.
- **99.99% uptime SLA**
- Built-in high availability.
- User-initiated backups.
- Supports AAD authentication, single sign-on support.
- Default deployment - **SQL endpoint is exposed only through a private IP address.**

- SQL Managed Instance is available in two service tiers:
    - General Purpose
        - Typical performance and I/O latency requirements.
        - High-performance Azure Blob storage (8TB).
    - Business Critical
        - Super-fast SSD storage (up to 1TB on Gen4, and up to 4TB on Gen5).
        - Built-in additional read-only database replica that can be used for reporting and other read-only workloads.
- https://docs.microsoft.com/en-us/azure/azure-sql/managed-instance/sql-managed-instance-paas-overview

Azure SQL DB Availability

Current protection solutions to protect Azure SQL databases from disruption and provide high-availability are:

- **Active geo-replication:**
    - Disaster recovery for regional disasters or outages.
    - Supports up to 4 secondary regional replicas.
    - Requires manual or application-initiated failover.
    - Does **NOT** support managed instances.
- **Failover Groups:**
    - Extends active geo-replication functionality.
    - Adds DNS management and failover policies.
    - https://docs.microsoft.com/en-us/azure/azure-sql/database/auto-failover-group-overview?tabs=azure-powershell
- **Automated backups:**
    - Full database backups each week.
    - Differential DB backups every 12 hours.
    - Transaction log backups every 5 - 10 minutes.
    - Point in time restores (PITR) up to 35 days (**standard and premium pricing tiers only**).

Long-term retention:

- Extended automated backups.
- Supports up to **10 years - Basic, Standard, and Premium pricing tiers.**
    - **No long-term support for Azure SQL DataWarehouse (SQL DW).**
- Retention options: weekly, monthly, week-of-year.

Azure SQL Monitoring

- SQL Analytics - Cloud-based SQL monitoring solution
  - Can monitor performance
  - Performance degradation
  -
- Intelligent Insights (Preview) SQLInsights logs can be exported to the following locations:
  - Log Analytics
  - Event Hub
  - Azure Storage
- Azure SQL Database and Azure SQL Managed Instance support performance diagnostics log generated by Intelligent Insights.
  - Identifies performance issues.

- https://docs.microsoft.com/en-us/azure/azure-sql/database/monitor-tune-overview
- https://docs.microsoft.com/en-us/azure/azure-monitor/insights/azure-sql

Cosmos DB

- Multi-model (key-value, column-family, document, graph)
- Suitable for NoSQL (non-relational) data
- Designed for massive scale (scaling out)

Primary use cases:

- For applications that require major scale
- Used extensively within Microsoft's e-commerce solutions
- Web and mobile (social media, personalisation, etc.)

Cosmos DB API Options:

1. SQL (Core)
2. Cassandra
3. MongoDB - **Currently ONLY supported migration option**.
4. Gremlin
5. Table

Selecting the right approach to partitioning is important:

- The **partition key** controls the logical division of items in a container.
- A partition key should consider read/write distribution, and storage requirements. (See az-300 prep deck for more information on Cosmos DB, particularly **partition keys**).

CosmosDB Data Migration Tool

- Dedicated tool for performing data migrations into CosmosDB
- Supports migration from:
    - SQL (SQL API)
    - MongoDB
    - Table
    - Cassandra
    - JSON files, CSV files, Azure CosmosDB collections.

Cosmos DB (Cont.)

Azure Cosmos DB uses two types of keys to authenticate users and provide access to its data and resources:

- **Master keys**
  - Used for administrative resources:
    - Database accounts, databases, users, and permissions.
  - **Cannot be used to provide granular access to containers and documents**.
- **Resource tokens**
  - Used for application resources:
    - Containers, documents, attachments, stored procedures, triggers, and UDFs
  - Uses a hash resource token specifically constructed for the user, resource, and permission
  - Typically a mid-tier web app service will **REQUEST** permission to **receive a resource token from the Cosmos DB account**.

Resource token generation and management is handled by the native Cosmos DB client libraries.

- Resource token are also time bound with a customisable validity period.
  - Default time span is **one hour**.
- https://docs.microsoft.com/en-us/azure/cosmos-db/secure-access-to-data#resource-tokens-

Azure SQL Data Warehouse (DW)

- Designed for business intelligence and analytics.
- Uses massively parallel processing (MPP) to perform complex queries and provide value to businesses.
- Data is first prepared and then stored in relation tables. This reduces storage costs, and improves query performance.

Primary use cases:

- As a replacement for traditional on-premises enterprise data warehouse (which is otherwise costly to maintain).
- As part of a big-data analytics solution.

Pricing is measured using two key components:

1. Data Warehouse Units (DWUs)
    a. Gen 1: Data Warehouse Units (DWUs)
    b. Gen 2: Compute Data Warehouse Units (cDWUs)
    c. Ranges from 100 DWU to 6000 DWUs
2. Consumed Storage

Consumed storage:

- Measured in GBs consumed.
- Data is sharded into **distributions** across Azure storage.
- Maximum database storage:
    - 240TB rowstore
    - Unlimited columnstore tables (gen 2).

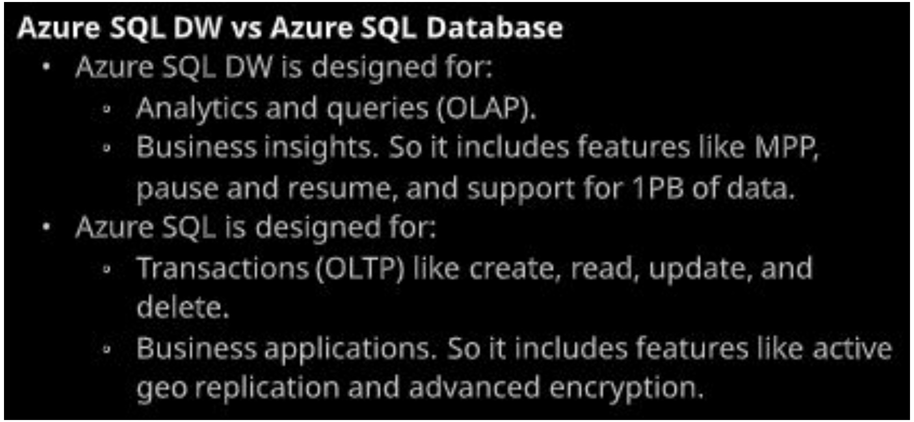
Scale and Performance:

- Compute can be scaled and paused manually.
- **Cannot be configured for autoscaling without additional code (such as Azure functions).**

Massively Parallel Processing (MPP) architecture

- Azure SQL DW uses scale-out architecture to distribute processing across multiple nodes.
- Data is **sharded** into **distributions** to Azure Storage.
- Distributions are mapped to compute nodes for processing.

Azure SQL Data Warehouse (DW) (Cont.)



Azure SQL DW vs Azure SQL Database

- Azure SQL DW is designed for:
  - Analytics and queries (OLAP).
  - Business insights. So it includes features like MPP, pause and resume, and support for 1PB of data.
- Azure SQL is designed for:
  - Transactions (OLTP) like create, read, update, and delete.
  - Business applications. So it includes features like active geo replication and advanced encryption.

- Currently **no native support for long-term retention**.
  - **The maximum limit is for 7 days** whether for automatic backups or for user-defined restore points.

Azure Data Lake Storage (ADLS)

Azure Data Factory (ADF)

- https://docs.microsoft.com/en-us/azure/data-factory/create-azure-ssis-integration-runtime
- ADF supports migration of Azure SQL workloads that are using SQL Service Integration Services (SSIS).
-

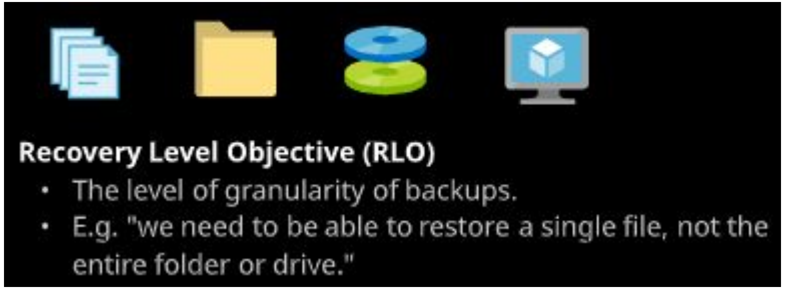
Azure SQL DB: Security

DESIGN A BUSINESS CONTINUITY STRATEGY (15-20%)

Site Recovery

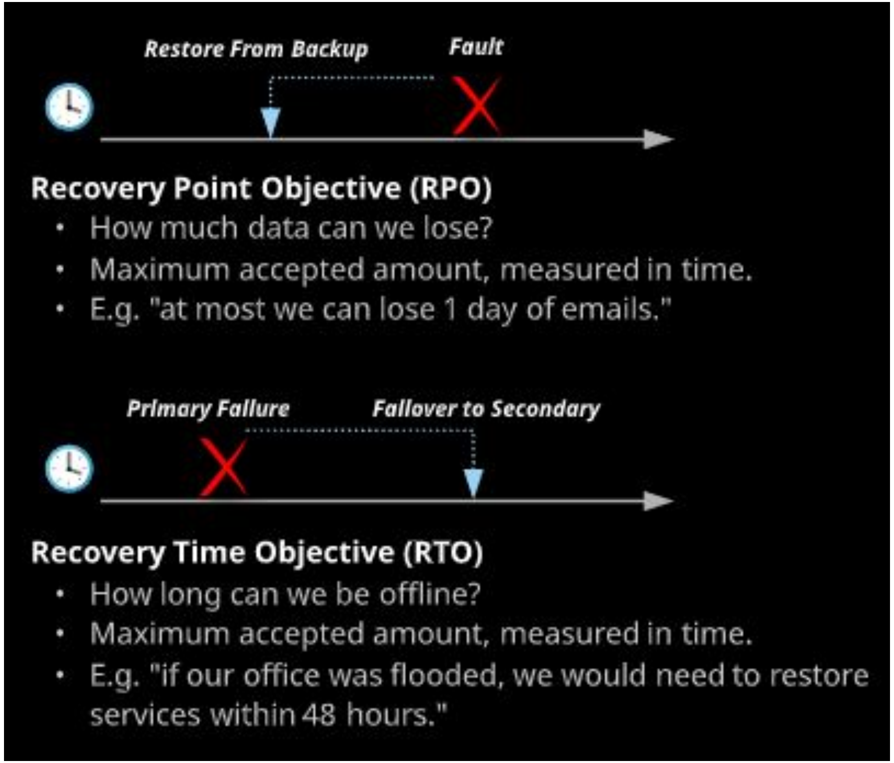
Design principle of resilience

- Availability: protect against minor faults and outages.
- Recoverability: protect against loss of data or service.



Recovery Level Objective (RLO)

- The level of granularity of backups.
- E.g. "we need to be able to restore a single file, not the entire folder or drive."



Recovery Point Objective (RPO)

- How much data can we lose?
- Maximum accepted amount, measured in time.
- E.g. "at most we can lose 1 day of emails."

Recovery Time Objective (RTO)

- How long can we be offline?
- Maximum accepted amount, measured in time.
- E.g. "if our office was flooded, we would need to restore services within 48 hours."

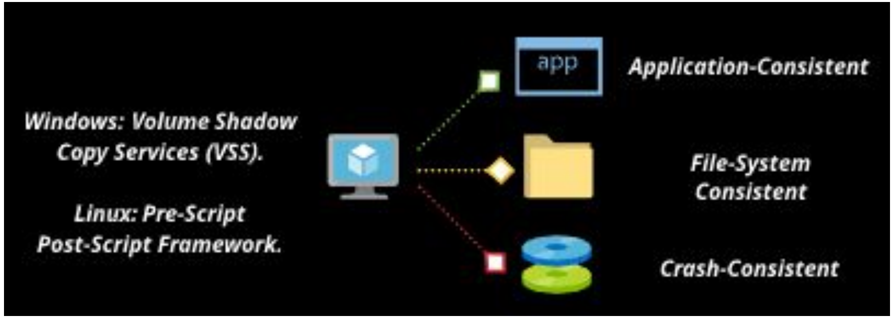
Backup or Snapshot Consistency

There are three (3) consistency levels:

1. **Application-consistent**
   a. For supported application workloads, backups ensure that ALL DATA (pending I/O, data from memory, etc.) is written to the backup.
2. **File-system consistent**
   a. The virtual machine (VM) has been backed up successfully. The VM file system is not corrupted and there is no data loss. Application data may be in a corrupt or inconsistent state.
3. **Crash-consistent**
   a. The VM will likely boot, however a disk-check process is required to identify and fix any file-system errors. Application data may be corrupt.

Important note:

Setting an application to perform an application-consistent snapshot can cause some performance degradation.


Windows: Volume Shadow Copy Services (VSS).
Linux: Pre-Script Post-Script Framework.
app
Application-Consistent
File-System Consistent
Crash-Consistent

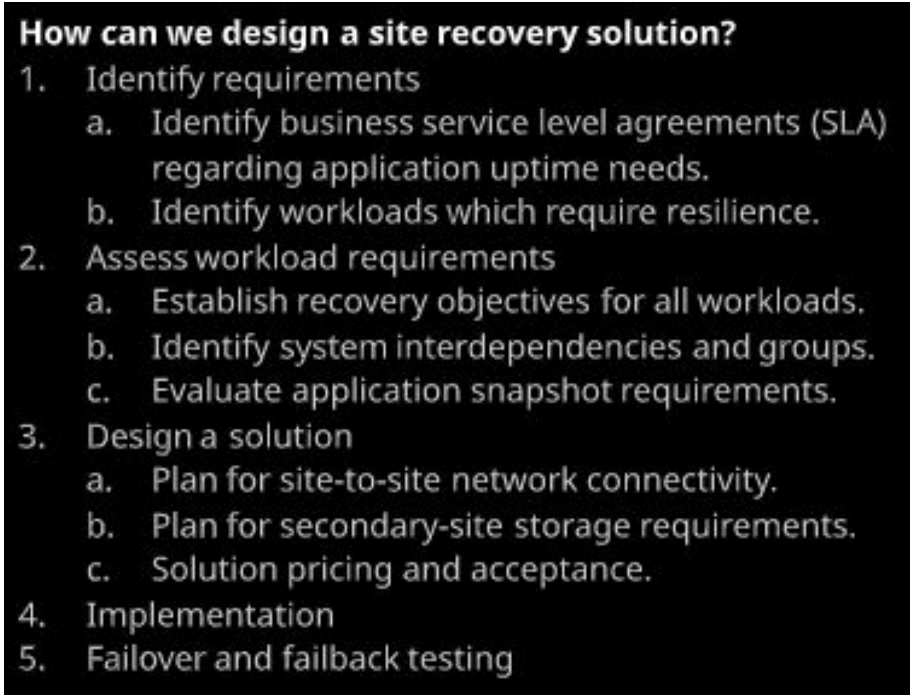
Architecting for Site Recovery

Common strategies:

- Backups:
  - Local short-term backups
  - Long-term offsite backups
- Replication:
  - Replication from a primary site to a secondary site.

Important things to consider:

- Backup and recovery (RPO/RTO/RLO) values typically differ.
- Availability strategies can positively influence our recovery strategies.
  - Having replicated systems/site infrastructure can allow for a smaller RPO/RTO.
  - Designing for more frequent application-consistent snapshots can decrease RPO but application performance may suffer.



How can we design a site recovery solution?

1. Identify requirements
   a. Identify business service level agreements (SLA) regarding application uptime needs.
   b. Identify workloads which require resilience.
2. Assess workload requirements
   a. Establish recovery objectives for all workloads.
   b. Identify system interdependencies and groups.
   c. Evaluate application snapshot requirements.
3. Design a solution
   a. Plan for site-to-site network connectivity.
   b. Plan for secondary-site storage requirements.
   c. Solution pricing and acceptance.
4. Implementation
5. Failover and failback testing

Azure Site Recovery

Replicated Items:

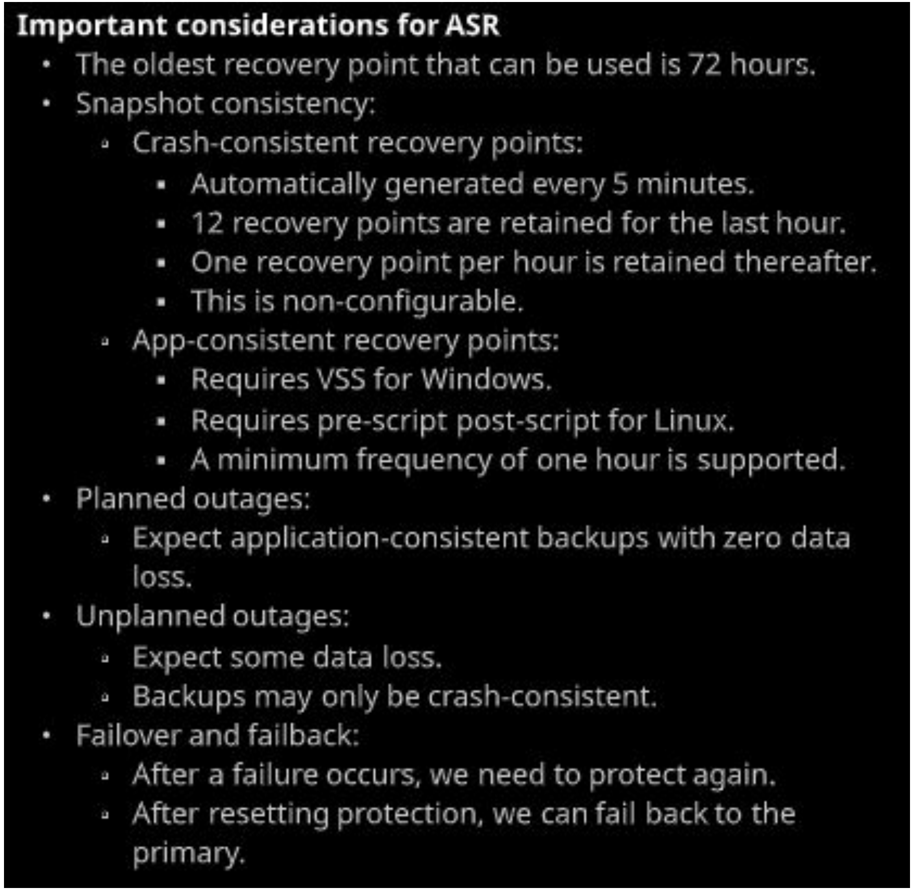
- Actual workloads replicated using ASR
- Workloads can be grouped (**multi-VM consistency**)

Replication Policies:

- Focused on recovery point objective (RPO).
- Recovery point retention (0-72 hours).
- Frequency of app-consistent snapshots (1-12 hours).
- Supports multi-VM consistency (replication groups).

Important considerations for ASR:

1. Protects against region failure (azure to azure).
2. Supports on-premises to Azure (site failure)
3. Supports Azure, Hyper-V, VMware, and Physical servers.
4. Supports Windows and Linux.



Important considerations for ASR

- The oldest recovery point that can be used is 72 hours.
- Snapshot consistency:
  - Crash-consistent recovery points:
    - Automatically generated every 5 minutes.
    - 12 recovery points are retained for the last hour.
    - One recovery point per hour is retained thereafter.
    - This is non-configurable.
  - App-consistent recovery points:
    - Requires VSS for Windows.
    - Requires pre-script post-script for Linux.
    - A minimum frequency of one hour is supported.
- Planned outages:
  - Expect application-consistent backups with zero data loss.
- Unplanned outages:
  - Expect some data loss.
  - Backups may only be crash-consistent.
- Failover and failback:
  - After a failure occurs, we need to protect again.
  - After resetting protection, we can fail back to the primary.

Azure Backup

- Uses Recovery Services Vaults (RSV) for short-term and long-term data retention.
- Uses geo-replicated storage (GRS) option by default.
-

LINKS:

- https://docs.microsoft.com/en-us/azure/backup/backup-support-matrix
- https://docs.microsoft.com/en-us/azure/backup/backup-azure-recovery-services-vault-overview
-

Azure Migrations

Main steps of a migrations:

1. Design
   a. Identify workloads for migration.
   b. **Assess compatibility.**
   c. **Determine the migration strategy.**
   d. Design the solution architecture.
   e. Design base architecture.
2. Deployment of solution
   a. Base architecture implementation.
   b. Solution infrastructure deployment.
3. Migration
   a. Configuration and testing (UAT).
   b. Migrate data as required, depending on scope and scale.
   c. Service cutover.
4. Integration
5. Ongoing maintenance and operations

The Four **R's** of migration:

1. **Rehost**
2. **Refactor**
3. **Rearchitect**
4. **Rebuild**

DESIGN FOR DEPLOYMENT, MIGRATION, AND INTEGRATION (10-15%)

Azure Automation

1. Automation and configuration as a service, providing:
    - Process automation (graphical, PowerShell, Python).
    - Configuration management (inventory and change).
    - Update management (integrated with Log Analytics).
2. Support for hybrid environments
    - Leverages Hybrid Runbook Worker.
    - Supports AWS, on-premises, physical, and virtual.
3. Supports for Windows and Linux.

- Build and deploy resources across hybrid environments
- Configure virtual machines (VMs):
    - Extend PowerShell Desired State Configuration (DSC).
    - Ensure VMs and compliant with configuration standards.

DESIGN AN INFRASTRUCTURE STRATEGY (15-20%)

Azure Storage Solutions

1. **Block-level Storage**
   a. Accessed directly by OS - low level access for storing data.
   b. iSCSI access protocol
   c. Used by Azure StorSimple
2. **File-level Storage**
   a. File system-like storage (NTFS, FAT32, EXT4)
   b. Use case would be Azure Files for replacing traditional on-premises file servers.
   c. Provides granular access control.
   d. **Not suitable** for large volumes of data, and unstructured data.
3. **Object-level Storage**
   a. Modern storage for cloud-based applications.
   b. Capable of massive scale and availability.
   c. Ideal for:
      i. Unstructured data.
      ii. Backups and archive data.
      iii. Static content that doesn’t change often.

Azure Batch

- Used to run large-scale parallel and high-performance computing (HPC) batch jobs efficiently in Azure.
- Creates and manages a pool of compute nodes (virtual machines).
- Configure and manage cluster/nodes with Batch APIs and tools, command-line scripts, or the Azure portal.
- Works well with **intrinsically parallel workloads**.
  - Workloads that can run independently.
  - Each instance completing part of the work.
- Can also be used to run **tightly coupled workloads,** where the applications you run need to communicate with each other rather than independently.
- Tightly coupled applications normally use the **Message Passing Interface (MPI) API.**
- Works with Azure Storage Accounts to receive input and send output from completed batch jobs.

Some examples of intrinsically parallel workloads you can bring to Batch:

- Financial risk modelling using Monte Carlo simulations
- VFX and 3D imaging rendering
- Image analysis and processing
- Media transcoding
- Data ingestion and, processing, and ETL operations.

Some examples of tightly coupled workloads:

- Finite element analysis
- Fluid dynamics
- Multi-node AI training.

Parallel Task Execution - used to maximise resource utilisation.
Pricing:

- https://azure.microsoft.com/en-in/pricing/details/batch/

Containers on Azure https://azure.microsoft.com/en-au/product-categories/containers/

Azure Container Registry (ACR)

- Managed container image platform for storing multiple container images based on Docker Registry.
- Supports DC/OS, Docker, Swarm, Kubernetes, and more.

Azure Container Instances (ACI)

- Simple service for easily running containers in Azure.
- Both Windows and Linux-based containers are supported.
- **Covered by 99.9% SLA**

Azure Kubernetes Service (AKS)

- Managed Kubernetes cluster (only pay for worker notes).
- Full Kubernetes orchestration and scalability.
- Supports service discovery, RBAC, and integrated logging and monitoring, and pod scaling.
- Open source solution
- https://docs.microsoft.com/en-au/azure/aks/intro-kubernetes

Azure Service Fabric (ASF)

- Deploy to Azure or to on-premises data centers that run Windows or Linux with zero code changes.
- Supports stateless and stateful microservices.
- Microsoft's container solution - used in-house before releasing to public as a microservices product.
- Uses Service Fabric programming model
  - Guest executables
  - Reliable Services - lightweight framework for writing services that integrate with the Service Fabric platform.
  - ASP.NET Core - open-source, cross-platform framework.
  - https://docs.microsoft.com/en-us/azure/service-fabric/service-fabric-choose-framework
- https://docs.microsoft.com/en-us/azure/service-fabric/service-fabric-overview

Compute Solutions - Virtual Machines

Design considerations:

- Useful for lift-and-shift of legacy applications or solutions.
- Supports full-access to operating system.
- Provides support for virtually any application.
- Supports up to **99.99% SLA when using Availability Zones (AZ)**
- Supports up to **99.9% SLA** when using **Availability Sets (AS)**

Virtual machine series

Parking Lot

1. ~~Azure API Management~~
   a. ~~Transform policies to strip HTTP headers.~~
   b. ~~Rate-limiting policies.~~
   c. ~~IP whitelisting access to backEnd API's.~~
2. Azure B2B - Guest Access
3. Azure B2C
4. Licensing - Azure Premium P1 vs P2 - E3 vs E5
5. **Azure Automation**
   a. **Hybrid Runbook Worker**
6. Disk Types;
   a. https://docs.microsoft.com/en-us/azure/virtual-machines/disks-types
7. Storage Service Encryption vs. Azure Disk Encryption
   a. https://devblogs.microsoft.com/premier-developer/azure-storage-encryption-and-azure-disk-encryption-demystified/
8. Azure Traffic Manager - Routing methods
9. Azure Container Instances (Container Groups)
10. Azure Service Fabric
    a. https://docs.microsoft.com/en-us/azure/service-fabric/service-fabric-cluster-capacity
11. **Azure Sentinel**

NICE TO KNOW

- **Customer-managed Encryption Keys**
- https://docs.microsoft.com/en-us/azure/storage/common/storage-service-encryption#about-encryption-key-management
- You can use **customer-managed encryption keys** to encrypt data in the following two Azure storage services:
  - **Azure Blob storage**
  - **Azure Files**
- **Azure Redis Cache:**
  - https://docs.microsoft.com/en-us/azure/azure-cache-for-redis/cache-overview
  - **Stores/caches data close to the application** for super low-latency access to frequently used data.
- **Azure Content Delivery Network (CDN)**
  - Distributed network of servers that can efficiently deliver web content to users.
  - Stores caches content on edge servers in point-of-presence (POP) locations.

On-premises Data Gateway

- Provides secure data transfer between on-premises data sources and your Azure Analysis Services servers in the cloud.
- https://docs.microsoft.com/en-us/azure/analysis-services/analysis-services-gateway
- https://docs.microsoft.com/en-us/data-integration/gateway/service-gateway-onprem
- Supports secure transfer of data from on-premises to the following Azure cloud services:
  - Power BI,
  - Power Apps,
  - Power Automate,
  - Azure Analysis Services,
  - Azure Logic Apps.

NICE TO KNOW (Cont.)

Always Encrypted vs. **Transparent Data Encryption (TDE)**

- https://azure.microsoft.com/en-in/blog/transparent-data-encryption-or-always-encrypted/
- **Always Encrypted supported on SQL server 2016 and above.**

Virtual Network Peering

- You **CANNOT** peer two virtual networks created through the classic deployment model.
-

ACRONYMS

Acronyms

- OLTP = Online Transaction Platform (SQL Database)
- OLAP = Online Analytics Platform (SQL Data Warehouse)
- RDBMS = Relational Database Management System
- SSIS = SQL Server Integration Service
- MPP = Massively Parallel Processing
- ADLS = Azure Data Lake Storage (Gen 2 storage)
- ADLA = Azure Data Lake Analytics (Uses Gen1 storage)
- RDMA = Remote Direct Memory Access
- MPI = Message Passing Interface
- HPC = High Performance Compute
- PTE = Parallel Task Execution (Running tasks across the same (single) HPC node to save cost of running multiple nodes for small tasks).
-

CRAMMING NOTES

CRAMMING

ROUTING:

- You **CANNOT** specify user-defined routes (UDR) when using Azure ExpressRoute circuits. **BGP must be used**.
  - https://docs.microsoft.com/en-us/azure/virtual-network/virtual-networks-udr-overview#custom-routes

Application Gateway:

- https://docs.microsoft.com/en-us/azure/application-gateway/application-gateway-autoscaling-zone-redundant

Azure Data Factory (ADF)

- **Supports "Gantt views" in your data pipeline runs.**
- Select "Gantt View" in "Pipeline Runs screen inside the **Monitor & Manage** tile within Azure Data Factory.

BACKUPS/REPLICATION & STORAGE:

- Azure Backup uses Recovery Services Vault (RSV) to store backup data.
- Azure Blob storage can retain data for 7 years.
  - Archive tier is the cheapest possible option.
  - Data being retrieved from archive tier needs to be "**Rehydrated**" and can take hours.
- Cool tier blob storage can still be accessed **immediately** and is a cheaper storage cost option compared to hot storage.
  - Cool storage does have a higher data-retrieval cost however.
  - **Data must remain in cool tier for at least 30 DAYS**
- https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-storage-tiers?tabs=azure-portal

CRAMMING - Continued...

Azure App Service:

- In order to access resources that are accessible across ExpressRoute or VPN connections (in Azure) use **VNET Integration.**
- https://docs.microsoft.com/en-us/azure/app-service/networking-features
- App Service Plan comparison and limits:
    - https://docs.microsoft.com/en-us/azure/azure-resource-manager/management/azure-subscription-service-limits#app-service-limits

Azure Containers and Azure Container Registry (ACR):

- Load-balancing is NOT supported for containers.
- Azure containers use Azure Files for storage mapping and d not support Azure Disks.
- Azure Container Instances do NOT support service discovery.

Azure Kubernetes Service (AKS):