Microsoft Azure SQL

COSC 516 – Cloud Databases

Microsoft Azure

Azure is a cloud platform offered by Microsoft that has over 200 products and cloud services.

- Azure supports Microsoft products and open source technologies.

Azure supports multiple different database engines:

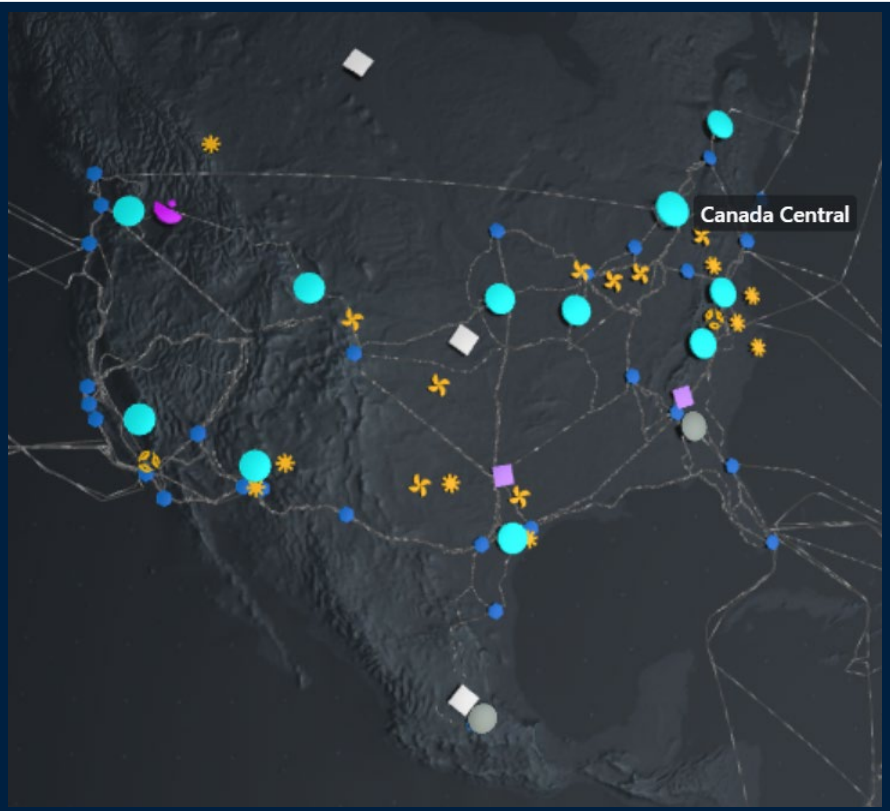
- SQL Server on Azure Virtual Machines
- PostgreSQL, MySQL, MariaDB (open-source relational)
- Azure SQL (scalable relational)
- CosmosDB (NoSQL)
- Azure Cache for Redis (in-memory)

Infrastructure

Azure global infrastructure consists of:

- ***Datacentre*** – unique physical building holding network servers
- ***Region*** – set of datacentres connected through low-latency network
- ***Geography*** – discrete market containing one or more regions with data residency and compliance boundaries
- ***Availability zone*** – unique physical locations within a region offering high availability to protect from datacentre failures (one or more datacentres)
- ***Global network*** – networking components connecting physical infrastructure

Azure Global Infrastructure



Azure Canada Regions

Canada Central: Toronto

Canada East: Quebec City

Source: Microsoft https://infrastructuremap.microsoft.com/explore

Microsoft SQL Server

Microsoft SQL Server is a relational DBMS designed for server deployments.

- One of the most widely deployed systems with a significant number of features.
- Usable for transactional databases and data warehousing/analytics.

SQL Server can be deployed on Azure as infrastructure as a service by creating compute instances that run the SQL Server database engine.

May also deploy as Azure SQL Managed Instance where infrastructure and server is managed for you.

Azure SQL

Azure SQL is a cloud-based database built upon SQL Server engine.

- It has many but not all the same features as SQL Server as Azure SQL is designed for the cloud rather than for server deployment.
- Latest optimizations appear on Azure first.

Key features:

- Platform as a service (PaaS) handles DBMS upgrading, backups, and monitoring.
- Supports high-performance in-memory technologies and intelligent query processing.
- Built-in high availability

Azure SQL Deployment Models

Single database – fully managed, isolated database

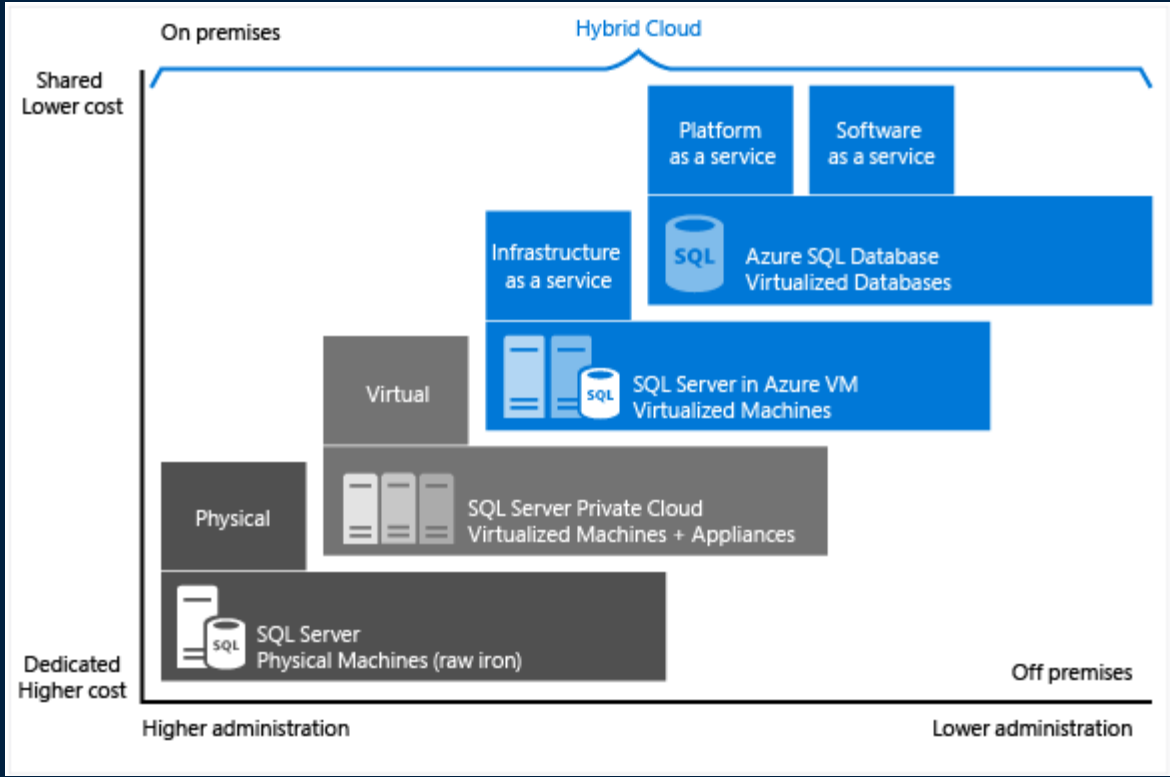
Elastic pool – collection of single databases with shared CPU/memory resources. Databases can be added/removed from pool.

- Multi-tenant SaaS

Azure SQL Managed Instance (PaaS) is best for cloud migrations.

- Similar to instance of Microsoft SQL Server
- Full SQL Server access and near 100% compatibility

SQL Server Service Comparison



Source: Microsoft https://learn.microsoft.com/en-ca/azure/azure-sql/azure-sql-iaas-vs-paas-what-is-overview

SQL Server versus Azure SQL

Features in common:

- Uses SQL Server engine. Similar support for SQL syntax and data types.

Key differences:

- Azure does not support common language runtime (CLR)
  - No system views, stored procedures, triggers, or user-defined functions
- Azure does not use mirroring or failover clustering as built-in HA on platform.

Use SQL Server versus Azure SQL when:

- Have existing application built on SQL Server that requires OS-level access or CLR.
- Require 100% compatibility with SQL Server and changes for porting may be extensive.

Azure SQL Purchasing Models

vCore-based: select # of vCores, memory, and storage.

- Microsoft allows companies that already have license to transfer to Azure.

DTU-based: three service tiers with different amount of compute, memory, I/O.

- A database transaction unit (DTU) represents a blended measure of CPU, memory, reads, and writes.

Azure SQL Database Service Tiers

General Purpose/Standard: common workloads, budget-oriented

Business Critical/Premium: designed for OLTP with high transaction rates and low latency I/O requirements. High resilience to failures using several isolated replicas.

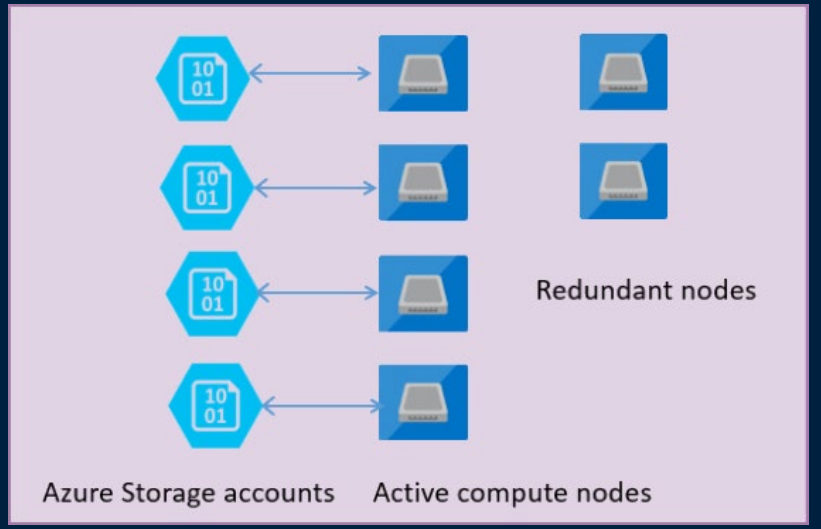
Hyperscale: for most business workloads. High performance with scalable compute and storage resources.

Serverless compute: Within vCore-based purchasing model and general purpose service tier automatically scales compute based on workload demand, and bills for the amount of compute used per second.

General Purpose Tier Architecture

Two layers:

- 1) Stateless compute layer running the sqlservr.exe process.
  - Only has transient data.
- 2) stateful data layer with database files (.mdf/.ldf) stored in Azure Blob storage.
  - Guarantees no data loss of any record placed in database file.
  - Built-in data availability/redundancy.



Storage latency between 5 and 10 ms.

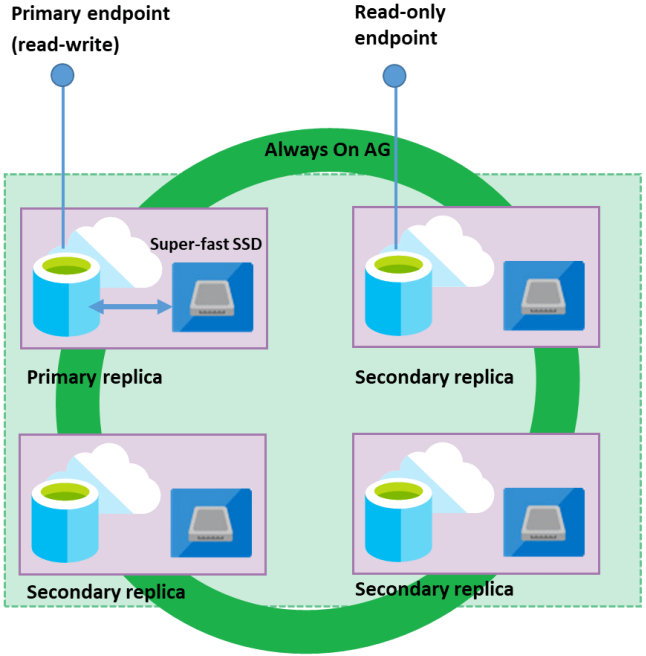
Business Critical Service Tier

Compute and storage integrated on each node. Database files on same node as database engine.

High availability by replication between database engines on each node of a four node cluster (one primary and three secondaries), with each node using ***locally attached SSD as data storage***. Similar to SQL Server Always On availability groups.

Failover handled by promoting secondary replica to be new primary. Read-only queries can be performed on secondary replicas.

Designed for low-latency applications.



Business Critical service tier: collocated compute and storage

Manageability

Database monitoring: Using Azure portal for performance metrics and database configuration.

Security

Data Encryption: Data is encrypted during communications using SSL/TLS. Data and backups are encrypted on storage.

Backups

Point-in-time Restore (PITR): Backups are copied to Azure storage automatically.

- Configurable retention period 0 to 35 days.
- Storage size increases dynamically as new backups are created.
- Backup storage amount equal to the maximum data size is provided at no extra charge.

Long-term Retention (LTR): May retain full backups for up to 10 years.

- Stored in Azure Blob storage automatically
- May select difference retention periods to satisfy compliance requirements

Scalability

Compute Scaling: Adjust compute and memory resources up or down within minutes. Maximum of 40 vCores and 120 GB of RAM.

Storage Scaling: Storage size up to 4 TB (general purpose tier).

Read Scale-out: Allows one or more read replicas of a primary server.

- Only available in Business critical service tier.
- Propagation latency in range from tens of milliseconds to single-digit seconds. No fixed upper bound on data propagation latency.

In-Memory Technologies

In-memory technologies keep data memory-resident for faster query response times. Useful for applications:

- Transactional/OLTP with many small read/update requests.
- Analytic (OLAP) with expensive computational queries.
- Mixed (hybrid transaction/analytical processing (HTAP)) where both OLTP and OLAP queries are executed on the same set of data.

Optimizations:

- Memory-resident data
- Native compilation of queries
- Batch processing and SIMD instructions

Temporal Tables

Temporal tables track and analyze full history of data changes in the table.

- Start and end time when data record is valid.

Provides efficient time-based analysis.

Simplifies implementation compared to manually tracking data changes.

Cost

Cost is based on usage.

Reserved instances allow for a discount by reserving a instance for a one or three year term.

Cost depends on deployment model, service tier, and licensing.

- Different pricing for customers bringing existing Microsoft licenses.

Free tier: Azure SQL Database 250 GB S0 instance with 10 database transaction units

Azure Free Account

An Azure free account provides free access to 40+ services always and free services to others for the first 12 months.

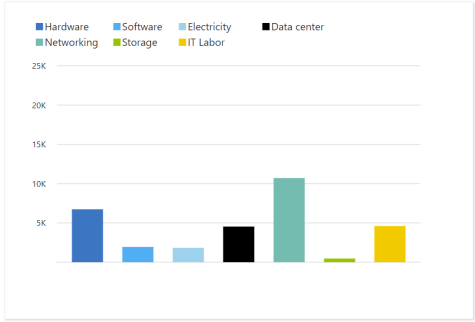
- https://azure.microsoft.com/en-ca/free/
- SQL Server Developer Edition – free always
- Azure Cosmos DB – free always
- Azure Database for MySQL or PostgreSQL – free first 12 months
- Azure SQL Database – free 250 GB S0 instance for first 12 months

Signup requires a credit card and phone number for validation.

Azure vs SQL Server On-Premise Cost

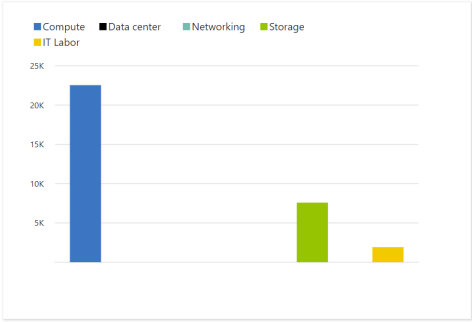
Total on-premises cost breakdown

In Azure, several of the cost categories from the on-premises environment are consolidated and decrease with the efficiency that comes with the cloud.



Total Azure cost breakdown

In Azure, several of the cost categories from the on-premises environment are consolidated and decrease with the efficiency that comes with the cloud.



SQL Server with 16 GB RAM, 1 CPU, 1 TB Storage

US$59,329
Cost over 5 year(s)

US$32,044
Cost over 5 year(s)

On-premises cost breakdown summary

Category	Cost
Compute	US$39,014.80
Hardware	US$6,736.00
Software	US$1,944.00
Electricity	US$1,822.80
Database	US$28,512.00
Data Center	US$4,534.05
Networking	US$10,713.30
Storage	US$467.20
IT Labor	US$4,600.00
Total	**US$59,329.00**

Azure cost breakdown summary

Category	Cost
Compute	US$22,541.76
Data Center	US$0.00
Networking	US$0.60
Storage	US$7,584.91
IT Labor	US$1,917.05
Total	**US$32,044.00**

Azure vs AWS

Azure is competing aggressively with AWS on cost.

- Claims AWS costs 5x more when running SQL Server on Windows server.
- Price matching on common instances (e.g. EC2, S3 storage).

Azure advantages:

- Integrated support for Microsoft products especially useful for organizations with existing on-premise Microsoft systems.
- Focus on support for hybrid, on-premise, and edge deployments

Both offer wide range of services and support high availability and security.

Conclusion

Azure supports multiple database services:

- SQL Server on Azure Virtual Machines
- PostgreSQL, MySQL, MariaDB (open-source relational)
- Azure SQL (scalable relational)
- CosmosDB (NoSQL)
- Azure Cache for Redis (in-memory)

Provides different deployment options for customers with existing SQL Server databases:

- SQL Server instances (IaaS)
- Azure SQL (PaaS)

Competes aggressively with AWS especially on Microsoft products.

Objectives

- List the database engines supported by Azure.
- Define: datacenter, region, geography, availability zone, global network
- How do these components combine into global infrastructure?
- Compare/contrast different deployment models.
- How does SQL Server and Azure SQL differ? How are those differences related to their architectures?
- When would you select vCore-based versus DTU-based cost model?
- When would you select either of the service tiers (general purpose, business critical, hyperscale)?
- For backups, what is the difference between Point-in-time Restore (PITR) and Long-term Retention (LTR)?
- When would in-memory technologies be useful? What are key considerations with this technology?
- How do temporal tables save developer effort when tracking data histories?
- Describe some of the factors for Azure SQL costs.


UBC
THE UNIVERSITY OF BRITISH COLUMBIA