02 Azure SQL DB & Azure Synapse Analytics

Azure SQL DB & Azure Synapse Analytics

Basic Level Questions (1-8)

1. What is Azure SQL Database and how does it differ from traditional on-premises SQL Server?

What to look for: Understanding of managed service benefits, automatic updates, built-in high availability, and scaling capabilities.

2. Explain the different purchasing models available in Azure SQL Database.

What to look for: DTU-based vs vCore-based models, understanding of compute and storage separation, and cost implications.

3. What is Azure Synapse Analytics and what are its main components?

What to look for: Data warehouse service, SQL pools, Spark pools, pipelines, and integration capabilities.

4. What is the difference between Azure SQL Database and Azure Synapse dedicated SQL pool?

What to look for: OLTP vs OLAP workloads, row-store vs column-store, scaling differences, and use case scenarios.

5. Explain the concept of Database Transaction Units (DTUs) in Azure SQL Database.

What to look for: Blended measure of CPU, memory, and I/O, performance levels, and when to use DTU vs vCore.

6. What are the main security features available in Azure SQL Database?

What to look for: Azure AD authentication, firewall rules, encryption at rest/transit, and threat detection.

7. What is a serverless compute tier in Azure SQL Database and when would you use it?

What to look for: Automatic scaling, pause/resume capabilities, cost optimization for intermittent workloads.

8. Explain what Data Warehouse Units (DWUs) are in Azure Synapse Analytics.

What to look for: Performance and cost scaling metric, compute resource allocation, and scaling flexibility.

Intermediate Level Questions (9-17)

9. How would you migrate a large on-premises SQL Server database to Azure SQL Database?

What to look for: Assessment tools (DMA, DMS), migration strategies, downtime considerations, and compatibility issues.

10. Explain the different backup and restore options available in Azure SQL Database.

What to look for: Automated backups, point-in-time restore, long-term retention, geo-restore capabilities.

11. How would you optimize query performance in Azure Synapse dedicated SQL pool?

What to look for: Distribution strategies, indexing (clustered columnstore), statistics, workload management, and query design.

12. What are the different distribution options in Azure Synapse and when would you use each?

What to look for: Hash distribution, round-robin distribution, replicated tables, and data skew considerations.

13. How would you implement disaster recovery for Azure SQL Database?

What to look for: Geo-replication, failover groups, backup strategies, and RTO/RPO requirements.

14. Explain how you would handle large data loads into Azure Synapse Analytics.

What to look for: PolyBase, COPY command, staging strategies, and performance optimization techniques.

15. What is Azure SQL Database Hyperscale and what are its benefits?

What to look for: Multi-tier architecture, rapid scaling, read replicas, and large database support (up to 100TB).

16. How would you monitor and troubleshoot performance issues in Azure SQL Database?

What to look for: Query Performance Insight, DMVs, Azure Monitor, intelligent insights, and automatic tuning.

17. Explain workload management in Azure Synapse Analytics and how you would configure it.

What to look for: Resource classes, workload groups, workload classification, and query prioritization.

Advanced/Difficult Level Questions (18-25)

18. Design a hybrid data architecture that integrates on-premises SQL Server with Azure Synapse Analytics.

What to look for: Data integration patterns, security considerations, network connectivity, and real-time vs batch processing.

19. How would you implement a real-time analytics solution using Azure SQL Database and streaming technologies?

What to look for: Change data capture, Azure Stream Analytics, Event Hubs integration, and lambda architecture patterns.

20. Explain how you would design and implement a multi-tenant SaaS application database architecture in Azure.

What to look for: Tenant isolation strategies, elastic pools, sharding patterns, and security considerations.

21. How would you optimize costs for a large-scale Azure Synapse Analytics implementation with varying workload patterns?

What to look for: Pause/resume strategies, workload management, resource allocation, and reserved capacity.

22. Design a data governance and security framework for sensitive data in Azure SQL Database and Synapse.

What to look for: Data classification, dynamic data masking, row-level security, always encrypted, and audit compliance.

23. How would you implement a CDC (Change Data Capture) solution to synchronize data between Azure SQL Database and a data lake?

What to look for: CDC mechanisms, delta processing, conflict resolution, and consistency guarantees.

24. Explain how you would design a high-performance data warehouse in Azure Synapse with complex ETL requirements.

What to look for: ELT vs ETL patterns, data pipeline orchestration, error handling, data quality, and performance optimization.

25. How would you implement a cross-region disaster recovery solution for a mission-critical data warehouse in Azure Synapse?

What to look for: Geo-redundancy strategies, automated failover, data consistency, testing procedures, and business continuity planning.

Technical Deep-Dive Follow-ups

For Azure SQL Database:

- "How would you handle connection pooling and deadlock resolution?"
- "Explain your approach to index maintenance in a high-transaction environment"
- "How would you implement temporal tables for audit trails?"

For Azure Synapse Analytics:

- "Walk me through designing fact and dimension tables with proper distributions"
- "How would you handle slowly changing dimensions (SCD) types 1, 2, and 3?"
- "Explain your strategy for handling late-arriving data in the warehouse"

Scenario-Based Questions:

- "Your data warehouse queries are running slowly during peak hours. Walk me through your troubleshooting approach."
- "A client needs to migrate 50TB of data with minimal downtime. What's your strategy?"
- "You need to ensure compliance with GDPR while maintaining query performance. How would you approach this?"

- "Design a solution that can handle both real-time dashboards and batch analytics on the same dataset."