Introduction to SQL Server

What is SQL Server

RDBMS from Microsoft

Unmanaged box product – you are responsible for running it

SQL Server features:

SQL Server Editions:

Express Edition-

Max 1.4GB RAM

1 socket or 4 cores, whichever is lesser

No analysis services

No backup compression

No integration with Python & R

Web Edition

Standard Edition-

Max 128 GB RAM

4 sockets or 24 cores

Only basic integration for Python & R

Enterprise Edition

Unlimited RAM i.e., all the memory that OS can support

Unlimited CPU

Full integration of Python & R

There are 2 subcategories of enterprise edition if needed:

Developer Edition

Evaluation Edition

Accessing SQL Server

Authentication

You can authenticate a user using the following options:

SQL Server Authentication

Active Directory

Azure Active Directory – Requires Azure Arc-enabled SQL Server 2022

Protocols used:

TCP/IP (Default port: 1433)

Shared memory

Named pipes

Disk configuration

SQL Server can in theory, run on a single disk that has all the DBs, SQL Server, OS etc. But in a production environment there a multiple disks that could have one disk for OS + App, one for data, one for logs, one for TempDB etc. that sits on multiple servers.

Database consists of 2 main components:

Data

Logs – Transactional logs.

Recovery models – the type of recovery model used will determine the restore granularity

Full – every single transaction is logged. The log size could increase rapidly but the DB can be restored to any pt in time

Bulk Logged

Simple – the log size won’t grow too fast.

SQL Server agent – orchestrates tasks & helps in automation i.e, runs time-specific tasks, multiple-step jobs, etc.

Another way of doing automation is using triggers.

SQL Server Client Tools

Some common tools for accessing SQL Server are:

SQL Server Management Studio (SSMS) –

Graphical tool

Windows only

Primarily for Admins

Advanced support for Replication & High Availability (HA)

Azure Data Studio (ADS)

Graphical tool

Cross Platform

Primarily for Developers

Advanced support for Notebooks & Extensions

sqlcmd

Command line tool

Cross platform

Primarily for automation & scripting scenarios

In most cases we use a combination of these tools as they cater for different services.

What is T-SQL

Language used to communicate with SQL Server

Consists of DML(Data Manipulation Language) & DDL(Data Definition Language)

Comes with full support for flow control, error handling, variables, transactions, procedures & functions

Functions:

Scalar functions –

Table valued functions – returns a table

Aggregate functions –

Stored Procedures – pre-compiled collection of SQL statements & procedural local logic that allows Code reusability

Improves performance

Added security (we can grant access of a procedure to a user without having to grant them access to the underlying tables)

Modularity & Encapsulation

Transaction Control

Maintenance & Versioning

sys-tables – System tables (not generated by user) are present in every SQL database. There are sys.tables, sys.columns, sys.types, etc. These can be used to get SQL Server metadata.

View – Instead of using a query every time, we can create a view instead. The view stores the query & works on the underlying tables so we don’t need to update views once, they are created. We can then work on these views like we do with tables.

SQL Server Analysis Services (SSAS)

Has 2 modes:

Multi-dimensional mode – uses MDX language

Tabular mode – uses DAX language

BI tools consume data from SSAS which in-turn is obtained from a data-source which is usually SQL Server. Therefore, SSAS acts as a data mart which takes data from a data warehouse & provides it to a Visualization/Analytics layer.

SQL Server Integration Services (SSIS)

Built-in ETL service

SQL Server Reporting Services (SSRS) – Built-in visualization layer. But nowadays, we’re switching to Power BI & SSRS is losing its relevancy.

Arc Enablement – SQL Server metadata is stored in Azure cloud, allowing us to manage multiple SQL Servers via cloud. It also provides the following features:

Best practice assessment

Metrics

Logs

Centralized update management

Azure Active Directory Authentication

Azure Synapse Link for SQL Server

Bring OLTP tables from SQL Server to Azure Synapse workspace in real-time.

For this we install a self-hosted integration runtime which makes use of an Azure Data Lake Storage to feed a dedicated SQL pool in Azure Synapse.

Managed Instance Link

An on-premise SQL server & a managed instance in the cloud can be connected via a Managed Instance Link. The users normally use the on-premise SQL Server but the Managed Instance in the cloud can act as a stand-in whenever there’s a failure in the on-premises server. Whenever the on-premises server comes back up again, a FAILBACK happens & the users will automatically start using the on-premises server again.