Database Administrator Technical Test

Section 1 – General Admin

1. Explain the differences between a database and an instance  
     
   *A database is a collection of related database objects such as tables, indexes or stored procedures contained in related physical OS files (mdf, ndf). The instance is a set of databases related by a common set of system tables (model, master, msdb, tempdb)*
2. What are the names of the system databases?  
     
   *model, master, msdb and tempdb (plus Distribution for Replication )*
3. What is the role of the SQL Server Agent?  
     
   *Runs scheduled jobs and handles alerting to operators*
4. Describe how you would create a SSIS package to carry out a simple population of a table in one database from a table in another.

*Either Use the Wizard (Tasks -> Import /Export Data) or Create through BIDS:  
Create a new package  
Create a connection for the source and target databases  
Add a transform data task between the connections specifying a copy column transformation for each field.*

1. A SSIS package is required to be run at 09:30, 10:30, 11:30, 20:00 and 21:00. Describe how you would create a job to run a SSIS package to this schedule

*There are multiple ways to achieve this, simplest is to:   
Create a SQL Job to execute an SSIS package  
Amend the schedule to run hourly from 09:30 to 11:31  
Create a second schedule within the same job to run hourly from 20:00 to 21:01*

1. How do you restart SQL Server in single user mode? How do you start SQL Server in minimal configuration mode?

*SQL Server can be started from command line, using the SQLSERVR.EXE. This EXE has some very important parameters with which a DBA should be familiar with. -m is used for starting SQL Server in single user mode and -f is used to start the SQL Server in minimal configuration mode. Check out SQL Server books online for more parameters and their explanations*

1. As a part of your job, what are the DBCC commands that you commonly use for database maintenance?

*DBCC CHECKDB, DBCC CHECKTABLE, DBCC CHECKCATALOG, DBCC CHECKALLOC, DBCC SHOWCONTIG, DBCC SHRINKDATABASE, DBCC SHRINKFILE etc. But there are a whole load of DBCC commands which are very useful for DBAs. Check out SQL Server books online for more information*

1. What's the difference between DELETE TABLE and TRUNCATE TABLE commands?

*DELETE TABLE is a logged operation, so the deletion of each row gets logged in the transaction log, slower than truncate. TRUNCATE TABLE also deletes all the rows in a table, but it won't log the deletion of each row, instead it logs the de-allocation of the data pages of the table, which makes it faster. Of course, TRUNCATE TABLE can be rolled back.*

1. Can you rollback a TRUNCATE TABLE command

*Yes – common misconception that nothing is logged and therefore cannot be rolled back*

1. What is RAID and what are different types of RAID configurations you are aware of?

*RAID stands for Redundant Array of Inexpensive/Independent Disks, used to provide fault tolerance to (database) servers. Typically set up with either 1 (mirror), 1+0 (Mirrored Striping) 5 (Block Level striping with Distributed Parity), 0 (stripe set)*

1. What is the difference between varchar and nvarchar?  
     
   *varchar is one byte per character. nvarchar is two bytes per character to allow for Unicode language sets such as Japanese.*
2. What are triggers? How many triggers you can have on a table?

*Triggers are special kind of stored procedures that get executed automatically when an INSERT, UPDATE or DELETE operation takes place on a table. They are Event driven stored procedures*

*No restrictions on number in 2008 but different versions of SQL have different restrictions*

1. How can you move the data or log files for tempdb

*Several ways but here’s one……………*

|  |  |
| --- | --- |
| *use master* | |
| *go* |

|  |  |
| --- | --- |
| *Alter database tempdb modify file (name = tempdev, filename = 'e:\tempdb\tempdb.mdf')* | |
| *go* |

|  |  |
| --- | --- |
| *Alter database tempdb modify file (name = templog, filename = 'e:\tempdb\templog.ldf')* | |
| *Go*  *The files are moved when SQL restarts* |

1. What is the difference between Clustered and Non-clustered Indexes?

*Non-clustered Index*

*The data is present in random order, but the logical ordering is specified by the index. The data rows may be randomly spread throughout the table. The non-clustered index tree contains the index keys in sorted order, with the leaf level of the index containing the pointer to the page and the row number in the data page. In non-clustered index:*

*\* The physical order of the rows is not the same as the index order.*

*\* Typically created on column used in JOIN, WHERE, and ORDER BY clauses.*

*\* Good for tables whose values may be modified frequently.*

*Microsoft SQL Server creates non-clustered indices by default when CREATE INDEX command is given. There can be more than one non-clustered index on a database table. It also creates a clustered index on a primary key by default.*

*Clustered Index*

*Clustering alters the data block into a certain distinct order to match the index, resulting in the row data being stored in order. Therefore, only one clustered index can be created on a given database table. Clustered indices can greatly increase overall speed of retrieval, but usually only where the data is accessed sequentially in the same or reverse order of the clustered index, or when a range of items is selected.*

*Since the physical records are in this sort order on disk, the next row item in the sequence is immediately before or after the last one, and so fewer data block reads are required. The primary feature of a clustered index is therefore the ordering of the physical data rows in accordance with the index blocks that point to them. Some databases separate the data and index blocks into separate files, others put two completely different data blocks within the same physical file(s). Create an object where the physical order of rows is same as the index order of the rows and the bottom(leaf) level of clustered index contains the actual data rows.*

Section 2 – Security and Permissions

1. When trying to grant permissions to a database that has been imported from another instance you get an error stating that the user already exists in the database you are granting permissions to.  
     
   Why does this happen and how would you resolve the issue?  
     
   *Database users are stored locally within the database. Server logins are system level entities stored in master. When you move a database across instances users are moved but logins aren’t. To resolve the problem, remove the DB user then reassign permissions to the login*

OR

*It may indicate the logins need remapping if the database has been restored from another server – this is done with sp\_change\_users\_login*

1. What is the system function to get the current user's user id?

*USER\_ID(). Related functions - USER\_NAME(), SYSTEM\_USER, SESSION\_USER, CURRENT\_USER, USER, SUSER\_SID(), HOST\_NAME().*

1. Stored procedure USP\_get\_hist is owned by a developer with login domain\smithj; the owner needs to ‘dbo’. What command would you use to make this change?  
     
   *exec sp\_changeobjectowner (‘domain\smithj.USP\_get\_hist, ‘dbo’)*
2. What login methods are available to SQL Server
   1. Windows Authentication / Integrated Auth
   2. SQL Auth
3. What are the steps you will take, if you are tasked with securing an SQL Server?

*Very open ended question. Here are some things you could talk about: Preferring NT authentication, using server, databse and application roles to control access to the data, securing the physical database files using NTFS permissions, using an unguessable SA password, restricting physical access to the SQL Server, renaming the Administrator account on the SQL Server computer, disabling the Guest account, enabling auditing, using multiprotocol encryption, setting up SSL, setting up firewalls, isolating SQL Server from the web server etc.*

1. What port does SQL server run on?

*1433 is default but can be changed*

Section 3 – SQL Clustering & High Availability

1. Explain difference between Active/Active and Active/Passive cluster configurations.

Active/Active – In the Microsoft world, this means man separate resource groups live on the cluster at any time (either across multiple nodes or on the same node

Active/Passive – 1 application resource group only therefore live on only 1 node

1. What is a quorum disk?

*Cluster lock disk*

1. You are running SQL Server in a clustered environment - how do you stop the Instance without failing over?  
     
   *Start the cluster administrator, go to active resources, locate SQL Server, right click and take the resource* offline
2. What versions of SQL Server 2008 can be implemented in an Active/Passive Microsoft clustering environment?

Datacenter, Enterprise & Standard

1. How can clusters help to provide Highly Available Systems?

Node failover prevents single point of failure for server / OS / Network but doesn’t prevent against disk problems (if shared) data issues or instance issues.

1. What are the high availability solutions offered by SQL server?

Replication

Database Mirroring

Clustering

Log Shipping

Optional - VM Ware

1. What is database replication? What are the different types of replication you can set up in SQL Server?

*Replication is the process of copying/moving data between databases on the same or different servers. SQL Server supports the following types of replication scenarios:*

* *Snapshot replication*
* *Transactional replication (with immediate updating subscribers, with queued updating subscribers)*
* *Merge replication*

1. In terms of Replication, describe the following – Article, Publication, Distributor, Subscription

*Article Thing to be published, e.g. a table*

*Publication The collection of articles*

*Distributor The server which ‘distributes’ the replicated data*

*Subscription On the target server, the construct to receive the replicated data*

Section 4 – Performance

1. How do you attempt to improve the performance of a poorly performing query

*Using query plan to identify problem areas such as Table Scans instead of Index Scans & Seeks. Checking for appropriate & up to date statistics. Check for missing indexes, blocking, excess recompilations of stored procedures, procedures and triggers without SET NOCOUNT ON, poorly written query with unnecessarily complicated joins, too much normalization, excess usage of cursors and temporary tables.*

*Also, check Windows Perfromace Monitor and Activity monitor for things like Memory Usage, CPU usage, Disk bottlenecks, etc*

1. What is a deadlock and what is a livelock? How will you go about resolving deadlocks?

*Deadlock is a situation when two processes, each having a lock on one piece of data, attempt to acquire a lock on the other's piece. Each process would wait indefinitely for the other to release the lock, unless one of the user processes is terminated. SQL Server detects deadlocks and terminates one user's process.  
  
A livelock is one, where a request for an exclusive lock is repeatedly denied because a series of overlapping shared locks keeps interfering. SQL Server detects the situation after four denials and refuses further shared locks. A livelock also occurs when read transactions monopolize a table or page, forcing a write transaction to wait indefinitely.  
  
Check out SET DEADLOCK\_PRIORITY and "Minimizing Deadlocks" in SQL Server books online. Also check out the article Q169960 from Microsoft knowledge base*

1. What is blocking and how would you troubleshoot it?

*Blocking happens when one connection from an application holds a lock and a second connection requires a conflicting lock type. This forces the second connection to wait, blocked on the first.*

1. How would you find out the fragmentation level of the indexes on a table?  
     
   *dbcc showcontig(‘<tablename>’) [Optional -with fast, tableresults]*
2. What command would you use to defragment a fragmented index?  
     
   *dbcc indexdefrag(<db>, <table>, <index>)*
3. What are statistics, what might make them go out of date, and how do you update them?

*Statistics determine the selectivity of the indexes. If an indexed column has unique values then the selectivity of that index is more, as opposed to an index with non-unique values. Query optimizer uses these indexes in determining whether to choose an index or not while executing a query.   
  
Some situations under which you should update statistics:  
1) If there is significant change in the key values in the index  
2) If a large amount of data in an indexed column has been added, changed, or removed (that is, if the distribution of key values has changed), or the table has been truncated using the TRUNCATE TABLE statement and then repopulated  
3) Database is upgraded from a previous version  
  
Look up SQL Server books online for the following commands: UPDATE STATISTICS, STATS\_DATE, DBCC SHOW\_STATISTICS, CREATE STATISTICS, DROP STATISTICS, sp\_autostats, sp\_createstats, sp\_updatestats*

Section 5 – Backup and Recovery

1. Explain different types of BACKUPs available in SQL Server?

*Types of backups you can create in SQL Sever 7.0+ are Full database backup, differential database backup, transaction log backup, filegroup backup. Check out the BACKUP and RESTORE commands in SQL Server books online. Be prepared to write the commands in your interview. Books online also has information on detailed backup/restore architecture and when one should go for a particular kind of backup*

1. What are the different ways of moving data/databases between servers and databases in SQL Server?

*There are lots of options available; you have to choose your option depending upon your requirements. Some of the options you have are: BACKUP/RESTORE, detaching and attaching databases, replication, DTS, BCP, logshipping, INSERT...SELECT, SELECT...INTO, creating INSERT scripts to generate data.*

1. How would you typically recover a SQL database that is suspect

*First choice should be to restore a full backup, diff and the transaction logs since the last full/diff backup.  If these are unavailable you can use single user mode and then emergency mode and then running a dbcc checkdb.  Here is the T-SQL you should use.  NOTE: You should be in the master database when you run this.*

*Alter database mydb set Single\_User*

*Alter database mydb set Emergency*

*DBCC CheckDB ('mydb') -- This will tell you the Repair level to use*

*DBCC CheckDB ('mydb', { REPAIR\_ALLOW\_DATA\_LOSS | REPAIR\_FAST | REPAIR\_REBUILD })*

Alter database mydb set Multi\_User

*After a successful DBCC CheckDB the database is back in a consistent state and its status is now online*

1. A Database has become corrupted and you need to restore from backup. You check the backups and there are lots of full, differential and Log backups available. What steps would you take to get the data back up to the latest point available

*Try to do a final transaction log backup*

*Restore: Most recent Full*

*Most recent diff*

*All logs since the last diff*

1. A data volume has been lost from the main ERP system’s database server, the log volumes remain intact. How would you ensure you do not lose any transaction from this system?

*Attempt to Backup the transaction log*

1. How would you mount a database in emergency mode? Give a scenario when this may need to be done.

*ALTER DATABASE foo SET EMERGENCY, used to make data available when a Databases is suspect*

Section 6 – Querying

The questions in this section are based on the following tables:

CREATE TABLE Sites

(

SiteID INT IDENTITY(1,1) PRIMARY KEY,

SiteName VARCHAR(25) NOT NULL,

SiteLocation VARCHAR(25) NOT NULL,

)

CREATE TABLE Employee

(

Empid INT IDENTITY(1,1) PRIMARY KEY,

FirstName VARCHAR(25) NOT NULL,

Secondaname VARCHAR(25) NOT NULL,

Manager INT,

JobTitle VARCHAR(25)NOT NULL,

DOB SMALLDATETIME NOT NULL,

SiteID INT FOREIGN KEY (SiteID) REFERENCES Sites(SiteID)

)

1. Write the SQL to list all Employees and which sites they work at

SELECT FirstName, SecondName, SiteName, SiteLocation

FROM Employee EMP

JOIN Sites S ON EMP.SiteID = S.SiteID

1. Write the SQL to Give everyone who does not work at any specific site

SELECT FirstName, SecondName

FROM Employee EMP

LEFT OUTER JOIN Sites S ON EMP.SiteID = S.SiteID

WHERE S.SiteID is NULL

--or--

SELECT FirstName, SecondName

FROM Employee EMP

WHERE SiteID IS NULL

--or--

You could use a sub query but not very efficient in this case

1. Write the a procedure to selects the Site Name and location for all or part of an employee surname

CREATE PROC USP\_GetLocation (@Name VARCHAR(25))

AS

SELECT S.SiteName, S.SiteLocation

FROM Sites S

JOIN Employee EMP ON S.SiteID = EMP.SiteID

WHERE EMP.SecondName LIKE @Name + '%'

1. Write the SQL to select the youngest employee and their age in years

SELECT FirstName, SecondName, DATEDIFF(YY,DOB,GETDATE())

FROM EMployee

WHERE DOB = (SELECT MAX(DOB) FROM Employee)

1. Write the SQL to list the number of employees at each site

SELECT S.SiteName, COUNT(\*) AS NumResidents

FROM Sites S

JOIN Employee EMP ON S.SiteID = EMP.SiteID

GROUP BY SiteName