SSIS Interview Questions

* While running a package using dtexec utility, an exit code of 5 is returned. What does it mean?
  1. The utility encountered an internal error of syntactic or semantic errors in the command line. Error code 6
  2. The package was canceled by the user
  3. The utility was unable to locate the requested package. The package could not be found.
  4. The **utility was unable to load the requested package**. The package could not be loaded. Error code 5

| Value | Description |
| --- | --- |
| 0 | The package executed successfully. |
| 1 | The package failed. |
| 3 | The package was canceled by the user. |
| 4 | The utility was unable to locate the requested package. The package could not be found. |
| 5 | The utility was unable to load the requested package. The package could not be loaded. |
| 6 | The utility encountered an internal error of syntactic or semantic errors in the command line. |

* Type method of component slowly changing dimensions, is adding additional set of the columns like “effective date” and “expiary date” in the tables to keep track of old data.
  1. Type3
  2. Type1
  3. Type2
  4. None of the options

SCD-1 (Current): Overwriting the old values in the dimension record.

SCD-2 (History and Current): Creating an additional dimension record at the time of the change with the new attribute values.

SCD-3 (Previous and Current): Creating new current fields and move the old attribute in a precedent field.

**SCD** means **S**lowly **C**hanging **D**imension. The prices of the product, address of the person, name of the city are few examples of SCD.

There are three methods to capture SCD.  
**Type 1: Overwriting the old values**  
As the old value is overwritten by the new (current) value, there is no way to find out what the old (historical) value(s) were.  In other words, it holds **ONLY** current value.  
**Type 2: Adding an additional record in the dimension table**  
In this case, an additional record is added in the dimension table.  This method is very useful for the ACCURATE reporting purpose as all the OLD VALUES are captured.  In this method, we need to change the table structure so that PK violation doesn't take place.  Generally speaking, addition of the columns like "Effective Date" and "Expiary Date" suffice the need.  
**Type 3: Addition of the column (old value) in the dimension table.**  
In this method, when the value of the dimension is updated (by the current value), the old value column of the table is updated by the most previous value of the dimension.  Thus we have the **CURRENT** value and the **MOST PREVIOUS** value of the dimension.  This method is also NOT EFFECTIVE when reporting is needed over the period of time when the value of the dimension has changed more than once.

* Which feature in SSIS 2008 allows you to allocate a certain percentage of your cache to remember rows that had no match in the database?
  1. Error output of Lookup transformation
  2. Mismatch transformation component
  3. There is no such feature as part of SSIS 2008
  4. Miss cache feature

New Miss Cache **feature** that **allows you** to **allocate** a **certain percentage** of **your cache** to **remembering rows** that **had no match** in the **database**

* Which of the following is not a part of Audit Transformation system variables?

1. Package name
2. Task name
3. **Container name**
4. Tiser name

* Exposes the data in a data flow by using the ADO.NET DataReader interface.
  1. DataReader Destination
  2. ADO Destination
  3. ADO NET DataReader Destination
  4. ADO NET Destination
* We cannot use the SQL Server profiler Log provider in a package that is running in 64 bit mode.

1. True
2. False

* Can the following query be used in Execute SQL task of SSIS 2008 which connects to SQL Server 2008?

MERGE inventorymaster AS in USING (select inventoryld, Descr FROM New inventory) AS src on im. Inventoryld = src. Inventoryld when MATCHED THEN update set im.Descr = sit. Descr when not MATCHED THEN insert (inventoryld, Descr) values (src. Inventoryld, src. Descr): -

* 1. True
  2. False
* Which is the only database maintenance task that does not encapsulate a transact-sql statement or a DBCC command?

1. Maintenance cleanup Task
2. History cleanup Task
3. Update Statistics Task
4. Notify Operator Task

* Which transformation is used to remove duplicate values :-
  1. Merge
  2. Sort
  3. Multicast
  4. Unionatl

We had Sort **transformation** with **delete duplicate** option selected **in SSIS**

* Which of the components perform row transformation?
  1. Multicast
  2. Look up
  3. All the options
  4. Derived column

**Row Transformations** - **Derived Column, Data Conversion, Multicast, and Lookup**.

* Which of the components perform partially blocking transformations
  + Merge join, merge, union all
  + Merge, union all
  + Merge, look up, sort
  + Merge, merge join, aggregate

**Partially blocking transformations: -** SSIS include **Merge, Merge Join, and Union All**.

**Blocking transformations** - SSIS include **Aggregate and Sort. -- srinivas**

* When transactionoption property is set to “supported”, what does it mean?
  1. Use existing transaction
  2. If a transaction exists join it ease start a new one
  3. If a transaction exists join it
  4. Do not join an existing transaction

**Supported** - The container/task does not create a separate transaction, but if the parent object has already initiated a transaction then participate in it

**Required** - The container/task creates a new transaction irrespective of any transaction initiated by the parent object

**NotSupported** - The container/task neither creates a transaction nor participates in any transaction initiated by the parent object

* \_\_\_\_is a managed set of APIs that allows us to write custom applications, command Line utilities etc.
  1. Integration Services Data flow
  2. Integration Services Object model
  3. Integration Services Service
  4. Integration Services Runtime

SSIS object model **is a managed set of APIs that allows us to write custom applications**, **command line utilities etc**

* Which of the following task checks the allocation and structural integrity of all the objects in the specified database?
  1. Check Database Integrity Task
  2. Maintenance cleanup Task
  3. History Cleanup Task
  4. Update Statistics Task
* Which of the following is an example of text mining scenario?

1. Key term extraction from texts
2. Automatic categorization of texts
3. Analysis of texts
4. Combining text mining data with tabular data
5. A & b
6. C & d
7. A, b, & c
8. A, b, c & d

* What Are The New Features Available In SSIS 2008

1. Lookup enhancement
2. Data profiling task
3. Pipeline memory limiter
4. Option ii and iii only
5. Option i and iii only
6. Option iii only
7. All the options

This article series will discuss the following new features in SSIS 2008:

* Lookup Enhancement
* Data Profiling Task
* Pipeline Memory Limiter
* Pipeline Performance – Multicast Transform and Threading in SQL Server 2008
* VSTA support for Script Task and Script Component
* Import & Export Wizard Enhancements
* Brief explanation of leveraging SQL Server 2008 features in SSIS viz. CDC and MERGE statement
* A variable can be created within the scope of

1. Event Handler
2. All the options
3. Package
4. Container

A variable is created within the scope of a package or within the scope of a container, task, or event handler in the package

* The names of user-defined and system variables are case sensitive.

1. Partially true
2. False
3. True

* Aggregate transformation supports which of the following operations?

1. Count
2. All the options
3. Sum
4. Minimum

* Which of the following is not a feature of dimension processing destination?
  + Full processing
  + Partial processing
  + Update processing
  + Incremental processing
* Converts columns into rows.

1. Unpivot
2. Row sampling
3. Pivot
4. Slowly changing dimension

* Unique Conditions Exist For Each Breakpoint

1. 10
2. 8
3. 11
4. 12

* Which of the following is a scripting engine In SSIS 2008?

1. VTSA
2. VSTA
3. VSA
4. VSAT

* Utility loads the package

1. DTEXE
2. DTEXEC
3. MSDTC
4. DTUtility

* Which one of the following is used to link executables, containers, and tasks in packages into a control flow?

1. Sequence containers
2. Tasks
3. Precedence constraints
4. Connection managers

* Is a stopping point in the code?
  + Breakpoints
  + Pausepoints
  + Failurepoints
  + Checkpoints
* Which transform would you use to split your data based on conditions you define?
  + Conditional split
  + Sort
  + Merge
  + Multicast
* In which order does the dtexec utility process during the configuration phase?

1. Options that set package flags, variables and properties.
2. Options that configure the run time behavior of the utility, such as reporting
3. Options that verify the package version and build
4. B,C,A
5. A,C,B
6. A,B,C
7. A , B only

Options that set package flags, variables, and properties.

Options that verify the package version and build.

Options that configure the run-time behavior of the utility, such as reporting.

* SSIS 2008 provides built in support for spatial data types.
  + false
  + true
* Which type of join does the lookup transformation performs during lookups?
  + Equi join
  + Merge join
  + Both a. And b.
  + None
* On data Load from one source to its destination, in order to maintain the identity constraint what can be done
  + Create identity constraint on the destination column
  + In the OLE DB Source component window, check Keep identity.
  + Both a. and b.
  + None of the options
* Fuzzy Lookup transformation returns — or more close matches from the reference table.
  + 0
  + 1

Fuzzy Lookup transformation uses fuzzy matching to return one or more close matches from the reference table

* Variables in SSIS packages are scope oriented.
  + False
  + True
* True or False: SSIS has a default means to Log all records updated, deleted or inserted on a per table basis.
  + false
  + true
* Different options for dynamic configurations in SSIS?

1. Use an xml file
2. Use custom variables
3. Use a database per environment with the variables
4. Use a centralized database with all variables
   1. I & II
   2. I, Il & Ill
   3. I, III & IV
   4. All are correct

* Is a windows service, monitors running SSIS packages and manages the storage of packages.
  + Integration Services Service
  + Integration Services Data Flow
  + Integration Services Runtime
  + Integration Services object Model
* Converts columns into rows.
  + Pivot
  + Row sampling
  + Unpivot
  + Slowly changing dimension
* Which Are The Various Protection Levels In SSIS Packages :-
  1. EncryptAllwithPassword
  2. EncryptSensitiveWithPassword
  3. DontSaveSensitive
  4. All the options
  5. EncryptALLwithUserKey
* To combine values from the FirstName and LastName columns into a single column named FullName, which transformation can be used?
  + Import column
  + Derived Column
  + Data Column
  + Export column
* Which of the following statements are true for optimizing the performance, during SSIS package design?
  1. In OLEDB source remove the columns which are not used downstream
  2. In flat file source for transforms that have all columns with the default column type and size, consider using the ‘suggest type’ functionality.
  3. For OLEDB destinations/sql server destinations that perform bulk insert into a database, verify that the logging mode is appropriate for performance.
  4. A & b
  5. B & c
  6. A , b & c
  7. A only
* Which of the following is not an element of SSIS expression?
  + Expression parser
  + Expression builder
  + Expression evaluator
  + Expression grammar
* Is the detection of a dimensional natural key during fact table source loading, prior to the assignment of a corresponding primary key or surrogate key in the dimension table?
  + Inferred dimension
  + Early-arriving dimension
  + Early-arriving fact
* The tasks and containers in a package can be enabled for logging even if the package itself is not
  + Partially true
  + True
  + False
* Which is not a connection manager provider?
  + OLEDB
  + WMI
  + FTP
  + JIL
* Which one of the following is used to link executables, containers, and tasks in packages into a control flow?
  + Precedence constraints
  + Sequence containers
  + Tasks
  + Connection managers
* A for loop container can have only one evaluation expression.
  + False
  + True
* Which of the following task is used to access data by using other technologies that are not supported by built-in connection types?
  + Web service task
  + Script task
  + Ftp task
  + Xml task
* Property expressions can be updated in which of the following ways?

1. User-defined variables can be included in package configurations and then updated when the package is deployed.
2. Date and time functions are evaluated at run time and provide the updated values to property expressions.
3. Variables in expressions can be updated by the scripts that the Script task and the Script component run.
4. I & II
5. II only
6. II & III
7. I, II & III
8. User-defined variables can be included in package configurations and then updated when the package is deployed. At run time, the property expression is evaluated using the updated variable value.
9. System variables that are included in expressions are updated at run time, which changes the results of the property evaluation.
10. Date and time functions are evaluated at run time and provide the updated values to property expressions.
11. Variables in expressions can be updated by the scripts that the Script task and the Script component run

* Name the container that groups the tasks and containers into control flows, which are subsets of the package control flow
  + For loop container
  + For each loop container
  + Sequence container
  + Task host container
* Which of the following is a scripting engine in SSIS 2008?
  + VSAI
  + VTSA
  + VSTA
  + VSAT
* Which of the following is not configurable?
  + Execution tree
  + Variables
  + Connection managers
  + None of the options
* Which task can be used to re-create fact and dimension tables before loading data into them?
  + Execute package task
  + Execute process task
  + All the options
  + Execute sql task
* Which of the following is not a feature/tool for troubleshooting the data flows in a package?
  + Progress Reporting
  + Row Counts
  + Data Grid
  + Data viewers

Troubleshoot the data flows in an Integration Services package.

* SSIS Designer provides data viewers.
* SSIS Designer and Integration Services transformations provide row counts.
* SSIS Designer provides progress reporting at run time.
* Does SSIS support direct and indirect configurations? True or false?
  + True
  + False
* Various types of variables that can define in SSIS:-
  + User Defined
  + System Defined
  + Process Defined
  + Both a. and b.
* Demonstrate how your package uses buffers and threads.
  + Event handlers
  + Execution trees
  + Package trees
  + Package explorer
* How do you upgrade an SSIS package?
  + Recode the package based on the functionality in sql server dts
  + Use the migrate dts 2000 package wizard in bids then recode any portion of the package that is not accurate
  + SSIS wizard
  + Both a. and b.
* Which of the following transformation is used for text mining?
  + - 1. Lookup
      2. Fuzzy Lookup
      3. Term Lookup
      4. Text Lookup
      5. Term Extraction
      6. a, b, c, d & e -- some what doubt
      7. a, b & c
      8. b, c & d
      9. a & c
      10. c & e
      11. a & b
* Is it possible to run SSIS 2005 and SSIS 2008 packages on the same server at the same time?
  + No it’s not possible.
  + Yes it is possible.
  + Maybe
* Consists of one or more tasks and containers that execute when the package runs.
  + Control Flow
  + Data Flow
  + Both a & b
* Which setup is not needed while setting up a checkpoint file in SSIS? checkpoint
  + FaiLParentPackage
  + Savecheckpoints
  + checkpointFileName
* Which is not a type of Extraction?
  + Push
  + None
  + Pull
  + Dribble
* \_\_\_\_\_ is a managed set of APIs that allows us to write custom applications, command Line utilities etc.
  + Integration Services Object Model
  + Integration Services Data Flow
  + Integration Services Service
  + Integration Services Runtime
* Which of the following data flow source reads data more quickly than other sources?
  + Raw File Source
  + Flat File Source
  + None
  + OLEDB Source
* SQL Server Profiler Log provides the following function
  + Both a. and b.
  + executes traces in SQL Server Profiler
  + Write traces which can be viewed using SQL Server Profiler.

| Log provider | ProgID | ClassID |
| --- | --- | --- |
| Text file | DTS.LogProviderTextFile | {0A039101-ACC1-4E06-943F-279948323883} |
| SQL Server Profiler | DTS.LogProviderSQLProfiler | {E93F6300-AE0C-4916-A7BF-A8D0CE12C77A} |
| SQL Server | DTS.LogProviderSQLServer | {94150B25-6AEB-4C0D-996D-D37D1C4FDEDA} |
| Windows Event Log | DTS.LogProviderEventLog | {071CC8EB-C343-4CFF-8D58-564B92FCA3CF} |
| XML File | DTS.LogProviderXMLFile | {440945A4-2A22-4F19-B577-EAF5FDDC5F7A} |

* While scheduling SSIS packages, what are the numerous ways in which we can start SQL Server Agent Services?

1. using SQL Server Management studio
2. Using Command Line
3. Using services.msc console
4. All the options
5. using SQL Server configuration manager

* Includes managed application programming interfaces for creating custom components for use in packages, or custom applications that create, load, run, and manage packages.
  1. Data flow engine
  2. Object model
  3. Runtime engine
  4. Integration services service
* **Which of the following statements are true for optimizing the performance, during SSIS package design?** 
  1. In OLEDB source remove the columns which are not used downstream
  2. In flat file source for transforms that have all columns with the default column type and size, consider using the ‘suggest type’ functionality.
  3. For OLEDB destinations/sql server destinations that perform bulk insert into a database, verify that the logging mode is appropriate for performance.
  4. A & B
  5. B & C
  6. A,B & C
  7. A ONLY
* While loading the dimension data, it is found that we don't have the fact data available. How do we handle such kind of situations

1. Load dims: populate an unknown load dims :populate dim data
2. Load facts: join llookup to dim's, and if no match, set as 'unknown' dimension record
3. I
4. II
5. I and II -- some what doubt
6. None of the options

----------------

