

IBM

Exam 000-303

IBM InfoSphere DataStage v9.1

Version: 6.0

[Total Questions: 134]

Topic 1, Volume A**Question No : 1 - (Topic 1)**

In your ETL application design you have found several areas of common processing requirements in the mapping specifications. These common logic areas found include: code validation lookups and name formatting. The common logic areas have the same logic, but the jobs using them would have different column metadata. Choose the action that gives you the best reusability design to effectively implement these common logic areas in your ETL application?

- A. Create parallel routines for each of the common logic areas and for each of the unique column metadata formats.
- B. Create separate jobs for each layout and choose the appropriate job to run within a job sequencer.
- C. Create parallel shared containers and define columns combining all data formats.
- D. Create parallel shared containers with Runtime Column Propagation (RCP) ON and define only necessary common columns needed for the logic.

Answer: D

Question No : 2 - (Topic 1)

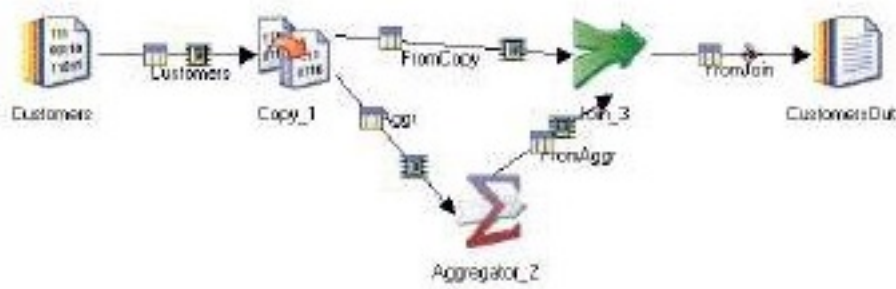
When optimizing a job, Balanced Optimization will NOT search the job for what pattern?

- A. Links
- B. Stages
- C. Sequencers
- D. Property Settings

Answer: C

Question No : 3 - (Topic 1)

You are asked to optimize the fork join job design in the exhibit. This job uses the sort aggregator and a left outer join on the ZIP code column. Currently all partitioning is set to "Auto" and automatic sort insertion is allowed.



Which change will reduce the cost of partitioning that occurs in this job?

- A. Use Entire partitioning on the input links to the Aggregator and Join stages.
- B. Hash partition and sort on ZIP code column on the input links to the Aggregator and Join stages.
- C. Hash partition and sort on ZIP code column prior to the Copy stage and use entire partitioning on the Aggregator and Join stages.
- D. Hash partition and sort on ZIP code column prior to the Copy stage, and use same partitioning on the Aggregator and Join stages.

Answer: D

Question No : 4 - (Topic 1)

You have a parallel job that based on operational recoverability requirements needs to be broken up into two separate parallel jobs. You have decided to use the Data Set stage to support this job design change. What two characteristics of Data Sets make them a good design consideration in your jobs design change? (Choose two.)

- A. They sort the data in a staging area.
- B. They automatically convert data types.
- C. They persist the parallelism of the job creating them.
- D. They use the same data types as the parallel framework.
- E. They persist parallelism into a temporary repository table.

Answer: C,D

Question No : 5 - (Topic 1)

What two binding types are supported by Information Services Director (ISD) for a parallel

job that is designed to be used as a service? (Choose two.)

- A. EJB
- B. SQL
- C. HDFS
- D. SOAP
- E. STREAMS

Answer: A,D

Question No : 6 - (Topic 1)

You are assigned to correct a job from another developer. The job contains 20 stages sourcing data from two Data Sets and many sequential files. The annotation in the job indicates who wrote the job and when, not the objective of the job. All link and stage names use the default names. One of the output columns has an incorrect value which should have been obtained using a lookup. What could the original developer have done to make this task easier for maintenance purposes?

- A. Name all stage and links the same.
- B. Name all stages and links based on what they do.
- C. Indicate all stage names within the job annotation.
- D. Name all stage and links with column names and ideas.

Answer: B

Question No : 7 - (Topic 1)

You are asked by management to document all jobs written to make future maintenance easier. Which statement is true about annotations?

- A. The short job description can be identified within the Description Annotation stage.
- B. The Description Annotation stage contains both the short and full descriptions for the job.
- C. The background for the Description Annotation stage can be changed for each unique stage.
- D. The Description Annotation stage can be added several times at different locations to identify business logic.

Answer: A

Question No : 8 - (Topic 1)

A job design consists of an input Row Generator stage, a Filter stage, followed by a Transformer stage and an output Sequential File stage. The job is run on an SMP machine with a configuration file defined with three nodes. The \$APT_DISABLE_COMBINATION variable is set to True. How many player processes will this job generate?

- A. 8
- B. 10
- C. 12
- D. 16

Answer: A

Question No : 9 - (Topic 1)

Which partitioning method requires a key?

- A. Same
- B. Modulus
- C. Sort Merge
- D. Round Robin

Answer: B

Question No : 10 - (Topic 1)

A job design consists of an input Row Generator stage, a Sort stage, followed by a Transformer stage and an output Data Set stage. The job is run on an SMP machine with a configuration file defined with four nodes. The \$APT_DISABLE_COMBINATION variable is set to True. How many player processes will this job generate?

- A. 7
- B. 16
- C. 13

D. 16

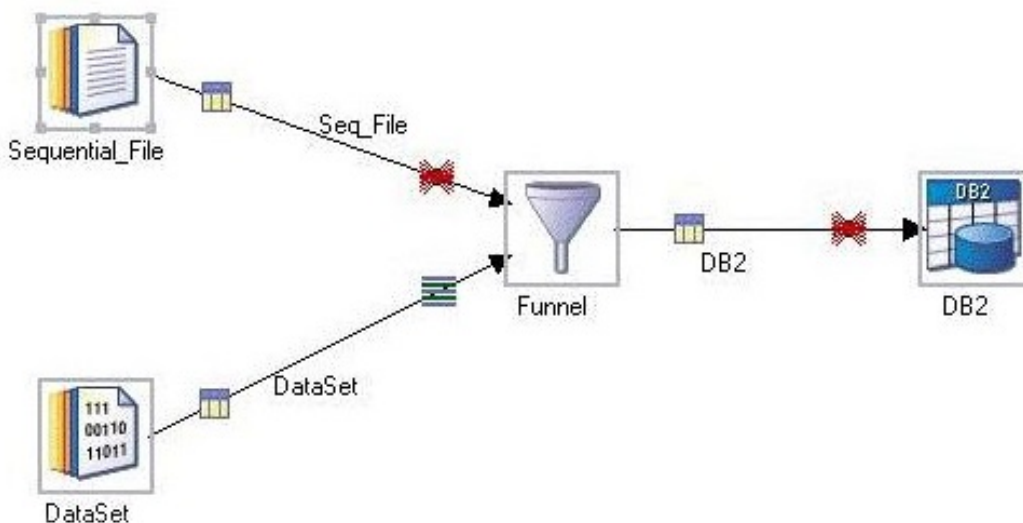
Answer: C
Question No : 11 - (Topic 1)

The data going into the target Sequential Files stage is range-partitioned and sorted. Which technique method would be the most efficient to create a globally sorted target sequential file?

- A. Select an in-stage sort in the final Sequential File stage.
- B. Select the Ordered collector method for the final Sequential File stage.
- C. Select the Sort Merge collector method for the final Sequential File stage.
- D. Insert a Funnel stage before the final Sequential File stage and select Sequence as Funnel Type.

Answer: B
Question No : 12 - (Topic 1)

In the exhibit, a Funnel stage has two input links. Input 1 (Seq_File) comes from a Sequential File stage with "Readers per Node" set to "2". Input 2 (DataSet) comes from a dataset created with 3 partitions. In the Funnel stage, the funnel type is set to "Sequence".



The parallel configuration file contains 4 nodes. How many instances of the Funnel stage

run in parallel?

- A. 1
- B. 2
- C. 4
- D. 6

Answer: C

Question No : 13 - (Topic 1)

Your job sequence must be restartable. It runs Job1, Job2, and Job3 serially. It has been compiled with "Add checkpoints so sequence is restartable". Job1 must execute every run even after a failure. Which two properties must be selected to ensure that Job1 is run each time, even after a failure? (Choose two.)

- A. Set the Job1 Activity stage to "Do not checkpoint run".
- B. Set trigger on the Job1 Activity stage to "Unconditional".
- C. In the Job1 Activity stage set the Execution action to "Run".
- D. In the Job1 Activity stage set the Execution action to "Reset if required, then run".
- E. Use the Nested Condition Activity with a trigger leading to Job1; set the trigger expression type to "Unconditional".

Answer: A,D

Question No : 14 - (Topic 1)

Which two actions are available when editing a message handler? (Choose two.)

- A. Abort job
- B. Demote to warning
- C. Suppress from job log
- D. Demote to informational
- E. Suppress from the project

Answer: C,D

Question No : 15 - (Topic 1)

What is the result of running the following command: dsjob -report DSProject ProcData

- A. Generates a report about the ProcData job, including information about its stages and links.
- B. Returns a report of the last run of the ProcData job in a DataStage project named DSProject.
- C. Runs the DataStage job named ProcData and returns performance information, including the number of rows processed.
- D. Runs the DataStage job named ProcData and returns job status information, including whether the job aborted or ran without warnings.

Answer: B

Question No : 16 - (Topic 1)

You would like to pass values into parameters that will be used in a variety of downstream activity stages within a job sequence. What are two valid ways to do this? (Choose two.)

- A. Use local parameters.
- B. Place a parameter set stage on the job sequence.
- C. Add a Transformer stage variable to the job sequence canvas.
- D. Check the "Propagate Parameters" checkbox in the Sequence Job properties.
- E. Use the UserVariablesActivity stage to populate the local parameters from an outside source such as a file.

Answer: A,E

Question No : 17 - (Topic 1)

On the DataStage development server, you have been making enhancements to a copy of a DataStage job running on the production server. You have been asked to document the changes you have made to the job. What tool in DataStage Designer would you use?

- A. Compare Against
- B. diffapicmdline.exe
- C. DSMakeJobReport
- D. Cross Project Compare

Answer: D

Question No : 18 - (Topic 1)

You are working on a project that contains a large number of jobs contained in many folders. You would like to review the jobs created by a former developer of the project. How can you find these jobs?

- A. Sort the jobs by date in the Repository window.
- B. Use Advanced Find feature in the Designer interface.
- C. Select the top project folder and choose Find Dependencies (deep).
- D. Right-click on jobs folder in the Repository window and select Filter jobs.

Answer: B

Question No : 19 - (Topic 1)

Your customer is using Source Code Control Integration for Information server and have tagged artifacts for version 1. You must create a deployment package from the version 1. Before you create the package you will have to ensure the project is up to date with version 1. What two things must you do to update the meta-data repository with the artifacts tagged as version 1? (Choose two.)

- A. Right-click the asset and click the Deploy command.
- B. Right-click the asset and click the Team Import command.
- C. Right-click the asset and click Update From Source Control Workspace.
- D. Right-click the asset and click Replace From Source Control Workspace.
- E. Right-click the asset and click the Team command to update the Source Control Workspace with the asset.

Answer: D,E

Question No : 20 - (Topic 1)

You want to find out which table definitions have been loaded into a job, and specifically which stages of the job they has been loaded into? How will you determine this?

- A. Select the job, right-click, then click the Find where used (deep) command.

- B. Select the job, right-click, then click the Find dependencies (deep) command.
- C. Select the job, right-click, then click the Find where used command. Then right-click and select "Show the dependency path from the job".
- D. Select the job, right-click, then click the Find dependencies command. Then right-click and select "Show the dependency path from the job".

Answer: D

Question No : 21 - (Topic 1)

You are responsible for deploying objects into your customers production environment. To ensure the stability of the production system the customer does not permit compilers on production machines. They have also protected the project and only development machines have the required compiler. What option will enable jobs with a parallel transformer to execute in the customers production machines?

- A. Add \$APT_COMPILE_OPT=-portable
- B. Set \$APT_COPY_TRANSFORM_OPERATOR
- C. Use protected projects in the production environment.
- D. Create a package with Information Server Manager and select the option to include executables.

Answer: D

Question No : 22 - (Topic 1)

What two features distinguish the Operations Console from the Director job log? (Choose two.)

- A. Jobs can be started and stopped in Director, but not in the Operations Console.
- B. The Operations Console can monitor jobs running on only one DataStage engine.
- C. Workload management is supported within Director, but not in the Operations Console.
- D. The Operations Console can monitor jobs running on more than one DataStage engine.
- E. The Operations Console can run on systems where the DataStage clients are not installed.

Answer: D,E