



**Exam Code:** 000-415

**Exam Name:** IBM WebSphere IIS DataStage Enterprise

Edition v7.5

**Vendor:** IBM

**Version:** DEMO

## Part: A

1: The last two steps of a job are an Aggregator stage using the Hash method and a Sequential File stage with a Collector type of Auto that creates a comma delimited output file for use by a common spreadsheet program. The job runs a long time because data volumes have increased. Which two changes would improve performance? (Choose two.)

- A.Change the Sequential stage to use a Sort Merge collector on the aggregation keys.
- B.Change the Sequential stage to a Data Set stage to allow the write to occur in parallel.
- C.Change the Aggregator stage to use the sort method. Hash and sort on the aggregation keys.
- D.Change the Aggregator stage to a Transformer stage and use stage variables to accumulate the aggregations.

**Correct Answers: A C**

2: Which three are valid ways within a Job Sequence to pass parameters to Activity stages? (Choose three.)

- A.Routine Activity stage
- B.Sequencer Activity stage
- C.UserVariables Activity stage
- D.ExecCommand Activity stage
- E.Nested Condition Activity stage

**Correct Answers: A C D**

3: Which two tasks will create DataStage projects? (Choose two.)

- A.Install the DataStage engine.
- B.Copy a project in DataStage Administrator.
- C.Add new projects from DataStage Administrator.
- D.Export and import a DataStage project from DataStage Manager.

**Correct Answers: A C**

4: Which statement about job parameter usage is true?

- A.You can use environment variables to set parameter values linked to the Job Sequence.
- B.You can change the parameter values in an initialization file linked to a Job Sequence .ini file.
- C.Changes to the job parameters in the Designer do not require a recompile to be applied to the job.
- D.You can change job parameters while a job is running and the changes will immediately be applied mid-job.

**Correct Answers: A**

5: Which statement describes a process for capturing a COBOL copybook from a z/OS system?

- A.Select the COBOL copybook using the Browse button and capture the COBOL copybook with Manager.
- B.FTP the COBOL copybook to the client workstation in binary and capture the metadata through Manager.
- C.FTP the COBOL copybook to the client workstation in text mode and capture the copybook

with Manager.

D. FTP the COBOL copybook to the server platform in text mode and capture the metadata through Manager.

**Correct Answers: C**

6: Which two statements are true of the column data types used in Orchestra schemas? (Choose two.)

A. Examples of Orchestra schema column data types are varchar and integer.

B. Examples of Orchestra schema column data types are int32 and string [max=30].

C. Orchestra schema column data types are the same as those used in DataStage stages.

D. OSH import operators are needed to convert data read from sequential files into schema types.

**Correct Answers: B D**

7: What is the purpose of the uv command in a UNIX DataStage server?

A. Start and stop the DataStage engine.

B. Report DataStage client connections.

C. Cleanup resources from a failed DataStage job.

D. Provide read access to a DataStage EE configuration file.

**Correct Answers: A**

8: A job has two input sources that need to be combined. Each input source exceeds available physical memory. The files are in the same format and must be combined using a key value. It is guaranteed that there will be at least one match.

Given the above scenario, which stage would consume the least amount of physical memory?

A. Merge

B. Funnel

C. Lookup

D. Transformer

**Correct Answers: A**

9: In which two scenarios should a sparse lookup be used in place of a normal lookup to retrieve data from an Oracle database? (Choose two.)

A. When the Oracle database is on the same system as the DataStage server.

B. When the number of input rows is significantly larger than the number of rows in the lookup table.

C. When the number of input rows is significantly smaller than the number of rows in the lookup table.

D. A database function that returns the current value of an Oracle object is required as part of the result set.

**Correct Answers: C D**

10: Which two statements about performance tuning a DataStage EE environment are true? (Choose two.)

A. Overall job design has a minimal impact in actual real-world performance.

- B.Only adjust buffer tuning parameters after examining other performance factors.
- C.A single, optimized configuration file will yield best performance for all jobs and be easier to administer.
- D.Performance tuning is an iterative process - adjust one item at a time and examine the results in isolation.

**Correct Answers: B D**

11: Job run details for a specific invocation of a multi-instance job can be viewed by which two clients? (Choose two.)

- A.dsjob
- B.dsjobinfo
- C.DataStage Director
- D.DataStage Manager

**Correct Answers: A C**

12: Which statement is true about Aggregator Sort and Hash methods when the APT\_NO\_SORT\_INSERTION environment variable is set to TRUE?

- A.If you select the Hash method, the Aggregator stage requires the data to have the partition sorted by the group key.
- B.If you select the Sort method, the Aggregator stage requires the data to have been partition sorted by the group key.
- C.If you select the Sort method, the Aggregator stage will partition sort the data by the group key before performing the aggregation.
- D.If you select the Hash method, the Aggregator stage will partition sort the data by the group key before building a hash table in memory.

**Correct Answers: B**

13: How does a Join stage process an Inner join?

- A.It transfers records from the input data sets whose key columns contain equal values to the output data set.
- B.It transfers all values from the left data set but transfers values from the right data set and intermediate data sets only when key columns match.
- C.It transfers all values from the right data set and transfers values from the left data set and intermediate data sets only where key columns match.
- D.It transfers records in which the contents of the key columns are equal from the left and right input data sets to the output data set. It also transfers records whose key columns contain unequal values from both input data sets to the output data set.

**Correct Answers: A**

14: An XML file is being processed by the XML Input stage. How can repetition elements be identified on the stage?

- A.Set the "Key" property for the column on the output link to "Yes".
- B.Check the "Repetition Element Required" box on the output link tab.
- C.Set the "Nullable" property for the column on the output link to "Yes".

D.No special settings are required. XML Input stage automatically detects the repetition element from the XPath expression.

**Correct Answers: A**

15: Click the Task button.

You have been asked to improve the performance of a job. A consultant has convinced your manager that sorts and partitions are major contributors to the performance issues being experienced since they are specified as "auto". Both the aggregation and join are done on Zip (postal) code.

Place the appropriate sort/aggregation specifications on the links.

**Place the appropriate sort/aggregation specifications on the links.**

**Specifications**

Sort / partition on Zip	Auto partition / sort on Zip
Preserve Partitioning	Same Partition / No sort

Done

**Correct Answers:**

**Place the appropriate sort/aggregation specifications on the links.**

**Specifications**

Sort / partition on Zip      Auto partition / sort on Zip

Preserve Partitioning      Same Partition / No sort

Done

16: Which command can be used to execute DataStage jobs from a UNIX shell script?

- A.osh
- B.dsjob
- C.DSRunJob
- D.DSExecute

**Correct Answers: B**

17: Which three are valid trigger expressions in a stage in a Job Sequence? (Choose three.)

- A.Unconditional
- B.Equality(Conditional)
- C.Custom(Conditional)
- D.Difference(Conditional)
- E.Return Value(Conditional)

**Correct Answers: A C E**

18: Which two can be implemented in a Job Sequence using job parameters? (Choose two.)

- A.All options of the Start Loop Activity stage.
- B.Name of a job to be executed by a Job Activity stage.
- C.A command to be executed by a Routine Activity stage.
- D.The body of the email notification activity using the user interface.

**Correct Answers: A C**

19: Establishing a consistent naming standard for link names is useful in which two ways? (Choose two.)

- A.using less memory at job runtime
- B.specifying link order without having to specify stage properties to generate correct results
- C.improving developer productivity and quality by distinguishing link names within stage editors
- D.easing use of captured job statistics (eg. row counts in an XML file) by processes or users outside of DataStage

**Correct Answers: C D**

20: A customer is interested in selecting the right RDBMS environment to run DataStage Enterprise Edition to solve a multi-file and relational database data merge. The customer realizes the value of running in parallel and is interested in knowing which RDBMS stage will match the internal data partitioning of a given RDBMS. Which RDBMS stage will satisfy the customer's request?

- A.ODBC Enterprise
- B.Oracle Enterprise
- C.Sybase Enterprise
- D.DB2/UDB Enterprise

**Correct Answers: D**