



Exam : 000-415

Title : IBM WebSphere IIS DataStage
Enterprise Edition v7.5

Ver : 12-18-2008

QUESTION 1:

Which three stages support the dynamic (runtime) definition of the physical column metadata? (Choose three.)

- A. the Sequential stage
- B. the Column Export stage
- C. the CFF stage
- D. the DRS stage
- E. the Column Import stage

Answer: A,B,E

QUESTION 2:

A Varchar(10) field named SourceColumn is mapped to a Char(25) field named TargetColumn in a Transformer stage. The APT_STRING_PADCHAR environment variable is set in Administrator to its default value. Which technique describes how to write the derivation so that values in SourceColumn are padded with spaces in TargetColumn?

- A. Include APT_STRING_PADCHAR in your job as a job parameter. Specify the C/C++ end of string character (0x0) as its value.
- B. Map SourceColumn to TargetColumn. The Transformer stage will automatically pad with spaces.
- C. Include APT_STRING_PADCHAR in your job as a job parameter. Specify a space as its value.
- D. Concatenate a string of 25 spaces to SourceColumn in the derivation for TargetColumn.

Answer: C

QUESTION 3:

Which three privileges must the user possess when running a parallel job? (Choose three.)

- A. read access to APT_ORCHHOME
- B. execute permissions on local copies of programs and scripts
- C. read/write permissions to the UNIX/etc directory
- D. read/write permissions to APT_ORCHHOME
- E. read/write access to disk and scratch disk resources

Answer: A,B,E

QUESTION 4:

In a Teradata environment, which stage invokes Teradata supplied utilities?

- A. Teradata API
- B. DRS Teradata
- C. Teradata Enterprise
- D. Teradata Multiload

Answer: D

QUESTION 5:

When importing a COBOL file definition, which two are required? (Choose two.)

- A. The file you are importing is accessible from your client workstation.
- B. The file you are importing contains level 01 items.
- C. The column definitions are in a COBOL copybook file and not, for example, in a COBOL source file.
- D. The file does not contain any OCCURS DEPENDING ON clauses.

Answer: A,B

QUESTION 6:

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

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

Answer: B,C

QUESTION 7:

Which three defaults are set in DataStage Administrator? (Choose three.)

- A. default prompting options, such as Autosave job before compile

- B. default SMTP mail server name
- C. project level default for Runtime Column Propagation
- D. project level defaults for environment variables
- E. project level default for Auto-purge of job log entries

Answer: C,D,E

QUESTION 8:

Which three are keyless partitioning methods? (Choose three.)

- A. Entire
- B. Modulus
- C. Round Robin
- D. Random
- E. Hash

Answer: A,C,D

QUESTION 9:

Which two must be specified to manage Runtime Column Propagation? (Choose two.)

- A. enabled in DataStage Administrator
- B. attached to a table definition in DataStage Manager
- C. enabled at the stage level
- D. enabled with environmental parameters set at runtime

Answer: A,C

QUESTION 10:

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

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

Answer: A,B,D

QUESTION 11:

A client requires that a database table be done using two jobs. The first job writes to a dataset. The second job reads the dataset and loads the table. The two jobs are connected in a Job Sequence. What are three benefits of this approach? (Choose three.)

- A. The time it takes to load the table is reduced.
- B. The database table can be reloaded after a failure without re-reading the source data.
- C. The dataset can be used by other jobs even if the database load fails.
- D. The dataset can be read if the database is not available.
- E. The data in the dataset can be archived and shared with other external applications.

Answer: B,C,D

QUESTION 12:

You are reading customer data using a Sequential File stage and transforming it using the Transformer stage. The Transformer is used to cleanse the data by trimming spaces from character fields in the input. The cleansed data is to be written to a target DB2 table. Which partitioning method would yield optimal performance without violating the business requirements?

- A. Hash on the customer ID field
- B. Round Robin
- C. Random
- D. Entire

Answer: B

QUESTION 13:

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

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

Answer: B,C,E

QUESTION 14:

An Aggregator stage using a Hash technique processes a very large number of rows during month end processing. The job occasionally aborts during these large runs with an obscure memory error. When the job is rerun, processing the data in smaller amounts corrects the problem. Which change would correct the problem?

- A. Set the Combinability option on the Stage Advanced tab to Combinable allowing the Aggregator to use the memory associated with other operators.
- B. Change the partitioning keys to produce more data partitions.
- C. Add a Sort stage prior to the Aggregator and change to a sort technique on the Stage Properties tab of the Aggregator stage.
- D. Set the environment variable APT_AGG_MAXMEMORY to a larger value.

Answer: C

QUESTION 15:

Which three actions are performed using stage variables in a parallel Transformer stage? (Choose three.)

- A. A function can be executed once per record.
- B. A function can be executed once per run.
- C. Identify the first row of an input group.
- D. Identify the last row of an input group.
- E. Lookup up a value from a reference dataset.

Answer: A,B,C

QUESTION 16:

The source stream contains customer records. Each record is identified by a CUSTID field. It is known that the stream contains duplicate records, that is, multiple records with the same CUSTID value. The business requirement is to add a field named NUMDUPS to each record that contains the number of duplicates and write the results to a target DB2 table.

Which job design would accomplish this?

- A. Send the incoming records to a Transformer stage. Use a Hash partitioning method with CUSTID as the key and sort by CUSTID. Use stage variables to keep a running count of the number of each new CUSTID. Add this count to a new output field named NUMDUPS then load the results into the DB2 table.
- B. Use a Modify stage to add the NUMDUPS field to the input stream then process the data via an Aggregator stage Group and CountRows options on CUSTID with the result of the sum operation sent to the NUMDUPS column in the Mapping tab for load into the

DB2 table.

C. Use a Copy stage to split the incoming records into two streams. One stream goes to an Aggregator stage that groups the records by CUSTID and counts the number of records in each group and outputs the results to the NUMDUPS field. The output from the Aggregator stage is then joined to the other stream using a Join stage on CUSTID and the results are then loaded into the DB2 table.

D. Use an Aggregator stage to group the incoming records by CUSTID and to count the number of records in each group then load the results into the DB2 table.

Answer: C

QUESTION 17:

A job contains a Sort stage that sorts a large volume of data across a cluster of servers. The customer has requested that this sorting be done on a subset of servers identified in the configuration file to minimize impact on database nodes.

Which two steps will accomplish this? (Choose two.)

A. Create a sort scratch disk pool with a subset of nodes in the parallel configuration file.

B. Set the execution mode of the Sort stage to sequential.

C. Specify the appropriate node constraint within the Sort stage.

D. Define a non-default node pool with a subset of nodes in the parallel configuration file.

Answer: C,D

QUESTION 18:

You have a compiled job and parallel configuration file. Which three methods can be used to determine the number of nodes actually used to run the job in parallel? (Choose three.)

A. within DataStage Designer, generate report and retain intermediate XML

B. within DataStage Designer, show performance statistics

C. within DataStage Director, examine log entry for parallel configuration file

D. within DataStage Director, examine log entry for parallel job score

E. within DataStage Director, open a new DataStage Job Monitor

Answer: C,D,E

QUESTION 19:

Which three features of datasets make them suitable for job restart points? (Choose

three.)

- A. They are indexed for fast data access.
- B. They are partitioned.
- C. They use datatypes that are in the parallel engine internal format.
- D. They are persistent.
- E. They are compressed to minimize storage space.

Answer: B,C,D

QUESTION 20:

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 Aggregator stage to a Transformer stage and use stage variables to accumulate the aggregations.
- 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 Sequential stage to use a Sort Merge collector on the aggregation keys.

Answer: C,D

QUESTION 21:

Which environment variable, when set to true, causes a report to be produced which shows the operators, processes and data sets in the job?

- A. APT_DUMP_SCORE
- B. APT_JOB_REPORT
- C. APT_MONITOR_SIZE
- D. APT_RECORD_COUNTS

Answer: A

QUESTION 22:

Which two system variables must be used in a parallel Transformer derivation to generate a unique sequence of integers across partitions? (Choose two.)

- A. @PARTITIONNUM
- B. @INROWNUM
- C. @DATE
- D. @NUMPARTITIONS

Answer: A,D

QUESTION 23:

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

Answer: A

QUESTION 24:

A job reads from a dataset using a DataSet stage. This data goes to a Transformer stage and then is written to a sequential file using a Sequential File stage. The default configuration file has 3 nodes. The job creating the dataset and the current job both use the default configuration file. How many instances of the Transformer run in parallel?

- A. 3
- B. 1
- C. 7
- D. 9

Answer: A

QUESTION 25:

Exhibit:

```

{
  node 'n1' {
    fastname "s1"
    pool "" "n1" "s1" "app2" "sort"
    resource disk "/orch/n1/d1" {}
    resource disk "/orch/n1/d2" {"bigdata"}
    resource scratchdisk "/temp" {"sort"}
  }
  node 'n2' {
    fastname "s2"
    pool "" "n2" "s2" "app1"
    resource disk "/orch/n2/d1" {}
    resource disk "/orch/n2/d2" {"bigdata"}
    resource scratchdisk "/temp" {}
  }
  node 'n3' {
    fastname "s3"
    pool "" "n3" "s3" "app1"
    resource disk "/orch/n3/d1" {}
    resource scratchdisk "/temp" {}
  }
  node 'n4' {
    fastname "s4"
    pool "n4" "s4" "app1"
    resource disk "/orch/n4/d1" {}
    resource scratchdisk "/temp" {} }
}

```

Which two hardware configurations are shown in the exhibit? (Choose two.)

- A. single processor
- B. grid
- C. cluster
- D. SMP

Answer: B,C

QUESTION 26:

What would require creating a new parallel Custom stage rather than a new parallel BuildOp stage?

- A. A Custom stage can be created with properties. BuildOp stages cannot be created with properties.
- B. In a Custom stage, the number of input links does not have to be fixed, but can vary, for example from one to two. BuildOp stages require a fixed number of input links.
- C. Creating a Custom stage requires knowledge of C/C++. You do not need knowledge of C/C++ to create a BuildOp stage.
- D. Custom stages can be created for parallel execution. BuildOp stages can only be built to run sequentially.

Answer: B

QUESTION 27:

In a Lookup operator, a key column in the stream link is VARCHAR and has Unicode set in the Extended attribute while the corresponding column in a reference link, also VARCHAR, does not.

What will allow correct comparison of the data?

- A. Convert both columns to CHAR, pad with spaces, and remove Unicode from the Extended attribute in Transformer operators prior to the Lookup operator.
- B. Convert the column in the reference link to the UTF-8 code page using the StringToUstring function in a Transformer operator prior to the Lookup operator, and set the Extended attribute of the column.
- C. Convert the column in the stream link to the default code page using the UstringToString function in a Transformer operator prior to the Lookup operator and remove Unicode from the Extended attribute of the column.
- D. Remove Unicode from the Extended attribute of the column from the beginning of the job to the Lookup operator and then set the Extended attribute of the column in the output mapping section of the Lookup operator.

Answer: B

QUESTION 28:

Exhibit:

```

{
  node "n1" {
    fastname "s1"
    pool "" "n1" "s1" "app2" "sort"
    resource disk "/orch/n1/d1" {}
    resource disk "/orch/n1/d2" {"bigdata"}
    resource scratchdisk "/temp" {"sort"}
  }
  node "n2" {
    fastname "s2"
    pool "" "n2" "s2" "app1"
    resource disk "/orch/n2/d1" {}
    resource disk "/orch/n2/d2" {"bigdata"}
    resource scratchdisk "/temp" {}
  }
  node "n3" {
    fastname "s3"
    pool "" "n3" "s3" "app1"
    resource disk "/orch/n3/d1" {}
    resource scratchdisk "/temp" {}
  }
  node "n4" {
    fastname "s4"
    pool "n4" "s4" "app1"
    resource disk "/orch/n4/d1" {}
    resource scratchdisk "/temp" {} }
}

```

In the exhibit, which item describes how you would make a stage run only one way parallel?

- A. constrain to node pool app2
- B. constrain to node pool app1
- C. constrain to node pool /temp
- D. constrain to node pool /orch/n4/d1

Answer: A

QUESTION 29:

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

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

metadata through Manager.

Answer: C

QUESTION 30:

Which task is performed by the DataStage JobMon daemon?

- A. writes the job's OSH script to the job log
- B. provides a snapshot of a job's performance
- C. graceful shutdown of DataStage engine
- D. automatically sets all environment variables

Answer: B

QUESTION 31:

What determines the degree of parallelism with which the Teradata Enterprise operator will read from a database?

- A. the value of the sessionsperplayer option found on the Additional Connections Options panel
- B. the number of Teradata AMPs
- C. the value of the Teradata MAXLOADTASKS parameter
- D. the number of nodes specified in the APT_CONFIG_FILE

Answer: B

QUESTION 32:

You need to move a DataStage job from a development server on machine A to a production server on machine B. What are two valid ways to do this? (Choose two.)

- A. Use the command line export tool to create a .dsx file on machine A, then move the .dsx file to machine B and use the command line import tool to load the .dsx file.
- B. Connect the Manager client to the source project on machine A and create a .dsx file of the job then connect the Manager client to the target project on machine B and import the .dsx file.
- C. Use the command line export tool to create a .dsx file on machine A then move the .dsx file to the client and use the Manager client to import it.
- D. Connect to machine A with the Manager client and create a .dsx file of the job, then move the .dsx file to machine B and use the command line import tool.

Answer: B,D

QUESTION 33:

Which two would cause a stage to sequentially process its incoming data? (Choose two.)

- A. The execution mode of the stage is sequential.
- B. The stage follows a Sequential File stage and its partitioning algorithm is Auto.
- C. The stage follows a Sequential File stage and the Preserve partitioning has been set to Clear.
- D. The stage has a constraint with a node pool containing only one node.

Answer: A,D

QUESTION 34:

Which two requirements must be in place when using the DB2 Enterprise stage in LOAD mode? (Choose two.)

- A. Tablespace cannot be addressed by anyone else while the load occurs.
- B. User running the job has dbadm privileges.
- C. Tablespace must be placed in a load pending state prior to the job being launched.
- D. Tablespace must be in read only mode prior to the job being launched.

Answer: A,B

QUESTION 35:

Which two statements are true for parallel shared containers? (Choose two.)

- A. When logic in a parallel shared container is changed, all jobs that use the parallel shared container inherit the new shared logic without recompiling.
- B. Within DataStage Manager, Usage Analysis can be used to build a multi-job compile for all jobs used by a given shared container.
- C. All container input and output links must specify every column that will be defined when the container is used in a parallel job.
- D. Parallel shared containers facilitate modular development by reusing common stages and logic across multiple jobs.

Answer: B,D

QUESTION 36:

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. Funnel
- B. Merge
- C. Lookup
- D. Transformer

Answer: B

QUESTION 37:

Your input rows contain customer data from a variety of locations. You want to select just those rows from a specified location based on a parameter value. You are trying to decide whether to use a Transformer or a Filter stage to accomplish this. Which statement is true?

- A. The Transformer stage will yield better performance because the Filter stage Where clause is interpreted at runtime.
- B. You cannot use a Filter stage because you cannot use parameters in a Filter stage Where clause.
- C. The Filter stage will yield better performance because it has less overhead than a Transformer stage.
- D. You cannot use the Transformer stage because you cannot use parameters in a Transformer stage constraint.

Answer: A

QUESTION 38:

You need to move a DataStage job from a development server on machine A to a production server on machine B.

What are two valid ways to do this? (Choose two.)

- A. Connect the Manager client to the source project on machine A and create a .dsx file of the job then connect the Manager client to the target project on machine B and import the .dsx file.
- B. Use the command line export tool to create a .dsx file on machine A then move the

.dsx file to the client and use the Manager client to import it.

- C. Use the command line export tool to create a .dsx file on machine A, then move the .dsx file to machine B and use the command line import tool to load the .dsx file.
- D. Connect to machine A with the Manager client and create a .dsx file of the job, then move the .dsx file to machine B and use the command line import tool.

Answer: A,D

QUESTION 39:

A dataset needs to be sorted to retain a single occurrence of multiple records that have identical sorting key values. Which Sort stage option can be selected to achieve this?

- A. Stable Sort must be set to True.
- B. Sort Group must include the sort key values.
- C. Allow Duplicates must be set to False.
- D. Unique Sort Keys must be set to True.

Answer: C

QUESTION 40:

The high performance ETL server on which DataStage EE is installed is networked with several other servers in the IT department with a very high bandwidth switch. A list of seven files (all of which contain records with the same record layout) must be retrieved from three of the other servers using FTP.

Given the high bandwidth network and high performance ETL server, which approach will retrieve and process all seven files in the minimal amount of time?

- A. In a single job, use seven separate FTP Enterprise stages the output links of which lead to a single Sort Funnel stage, then process the records without landing to disk.
- B. Setup a sequence of seven separate DataStage EE jobs, each of which retrieves a single file and appends to a common dataset, then process the resulting dataset in an eighth DataStage EE job.
- C. Use three FTP Plug-in stages (one for each machine) to retrieve the seven files and store them to a single file on the fourth server, then use the FTP Enterprise stage to retrieve the single file and process the records without landing to disk.
- D. Use a single FTP Enterprise stage and specify seven URI properties, one for each file, then process the records without landing to disk.

Answer: D

QUESTION 41:

Jobs that use the Sort stage are running slow due to the amount of data being processed. Which Sort stage property or environment variable can be modified to improve performance?

- A. Sort Stage Max Memory Limit property
- B. Sort Stage Restrict Memory Usage property
- C. APT_SORT_MEMORY_SIZE
- D. APT_AUTO_TRANSPORT_SIZE

Answer: B

QUESTION 42:

A credit card company has about 10 million unique accounts. The company needs to determine the outstanding balance of each account by aggregating the previous balance with current charges. A DataStage EE job with an Aggregator stage is being used to perform this calculation. Which Aggregator method should be used for optimal performance?

- A. Sort
- B. Group
- C. Auto
- D. Hash

Answer: A

QUESTION 43:

Your job reads from a file using a Sequential File stage running sequentially. The DataStage server is running on a single SMP system. One of the columns contains a product ID. In a Lookup stage following the Sequential File stage, you decide to look up the product description from a reference table. Which two partition settings would correctly find matching product descriptions? (Choose two.)

- A. Hash algorithm, specifying the product ID field as the key, on both the link coming from the Sequential File stage and the link coming from the reference table.
- B. Round Robin on both the link coming from the Sequential File stage and the link coming from the reference table.
- C. Round Robin on the link coming from the Sequential File stage and Entire on the link coming from the reference table.
- D. Entire on the link coming from the Sequential File stage and Hash, specifying the product ID field as the key, on the link coming from the reference table.

Answer: A,C

QUESTION 44:

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 server platform in text mode and capture the metadata through Manager.
- C. FTP the COBOL copybook to the client workstation in binary and capture the metadata through Manager.
- D. FTP the COBOL copybook to the client workstation in text mode and capture the copybook with Manager.

Answer: D

QUESTION 45:

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

- A. No special settings are required. XML Input stage automatically detects the repetition element from the XPath expression.
- B. Set the "Key" property for the column on the output link to "Yes".
- C. Check the "Repetition Element Required" box on the output link tab.
- D. Set the "Nullable" property for the column on the output link to "Yes".

Answer: B

QUESTION 46:

Which two job design techniques can be used to give unique names to sequential output files that are used in multi-instance jobs? (Choose two.)

- A. Use the DataStage DSJobInvocationId macro to prepend/append the Invocation Id to the file name.
- B. Use parameters to identify file names.
- C. Use a Transformer Stage variable to generate the name.
- D. Use the Generate Unique Name property of the Sequential File Stage.

Answer: A,B

QUESTION 47:

A job design consists of an input fileset followed by a Peek stage, followed by a Filter stage, followed by an output fileset. The environment variable APT_DISABLE_COMBINATION is set to true, and the job executes on an SMP using a configuration file with 8 nodes defined. Assume also that the input dataset was created with the same 8 node configuration file.

Approximately how many data processing processes will this job create?

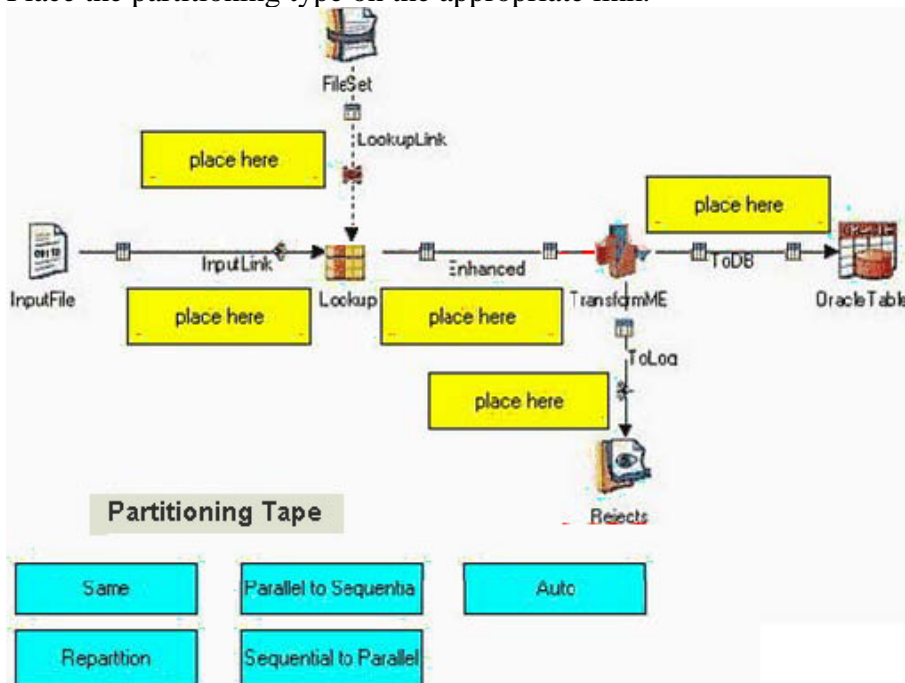
- A. 32
- B. 8
- C. 16
- D. 1

Answer: A

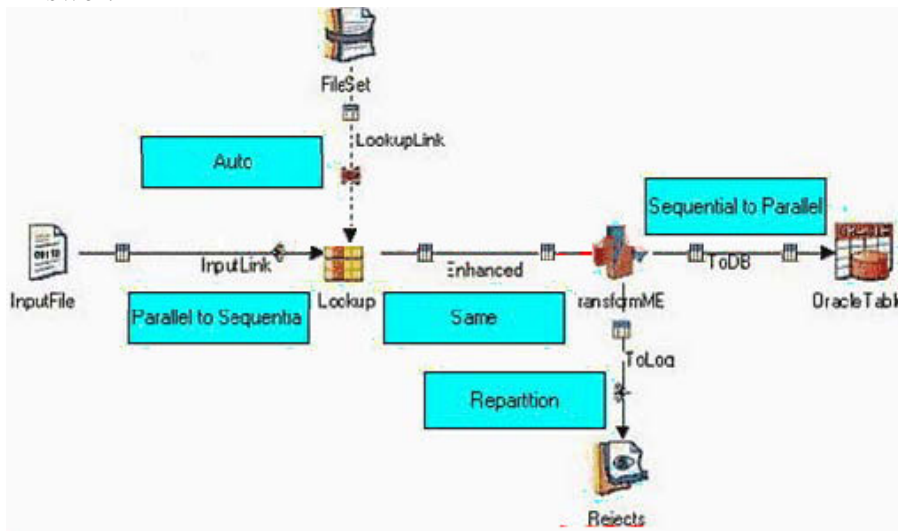
QUESTION 48:

DRAG DROP

Place the partitioning type on the appropriate link.



Answer:



QUESTION 49:

In a Transformer you add a new column to an output link named JobName that is to contain the name of the job that is running. What can be used to derive values for this column?

- A. a DataStage function
- B. a link variable
- C. a system variable
- D. a DataStage macro

Answer: D

QUESTION 50:

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

- A. Orchestrate schema column data types are the same as those used in DataStage stages.
- B. Examples of Orchestrate schema column data types are varchar and integer.
- C. Examples of Orchestrate schema column data types are int32 and string [max=30].
- D. OSH import operators are needed to convert data read from sequential files into schema types.

Answer: C,D

QUESTION 51:

Data volumes have grown significantly in the last month. A parallel job that used to run well is now using unacceptable amounts of disk space and running slowly. You have reviewed the job and explicitly defined the sorting and partitioning requirements for each stage but the behavior of the job has not changed.

Which two actions improve performance based on this information? (Choose two.)

- A. Change the sort methods on all sorts to "Don't Sort - Previously Grouped".
- B. Enable the environment variable APT_NO_SORT_INSERTION in the job.
- C. Increase the value of environment variable APT_PARTITION_NUMBER to increase the level of parallelism for the sorts.
- D. Enable the environment variable APT_SORT_INSERTION_CHECK_ONLY in the job.

Answer: B,D

QUESTION 52:

Which two statements are true about usage of the APT_DISABLE_COMBINATION environment variable? (Choose two.)

- A. Locks the job so that no one can modify it.
- B. Disabling generates more processes requiring more system resources and memory.
- C. Must use the job design canvas to check which stages are no longer being combined.
- D. Globally disables operator combining.

Answer: B,D

QUESTION 53:

Which technique would you use to abort a job from within the Transformer stage?

- A. Call the DSLogFatal function from a stage variable derivation.
- B. Call the DSStopJob function from a stage or column derivation.
- C. Create a dummy output link with a constraint that tests for the condition to abort on - set the "Abort After Rows" property to 1.
- D. Use the SETENV function to set the environmental APT_PM_KILL_NOWAIT.

Answer: C

QUESTION 54:

Your business requirement is to read data from three Oracle tables that store historical

sales data from three regions for loading into a single Oracle table. The table definitions are the same for all three tables, the only difference is that each table contains data for a particular region.

Which two statements describe how this can be done? (Choose two.)

- A. Create a job with a single Oracle Enterprise stage that executes a custom SQL statement with a UNION ALL operator that outputs the data to an Oracle Enterprise stage.
- B. Create a job with three Oracle Enterprise stages to read from the tables and output to a Collector stage which in turn outputs the data to an Oracle Enterprise stage.
- C. Create a job with three Oracle Enterprise stages to read from the tables and output to a Funnel stage which in turn outputs the data to an Oracle Enterprise stage.
- D. Create a job with a single Oracle Enterprise stage that executes a custom SQL statement with a FETCH ALL operator that outputs the data to an Oracle Enterprise stage.

Answer: A,C

QUESTION 55:

You have a series of jobs with points where the data is written to disk for restartability. One of these restart points occurs just before data is written to a DB2 table and the customer has requested that the data be archivable and externally shared. Which storage technique would optimize performance and satisfy the customer's request?

- A. DB2 tables
- B. filesets
- C. sequential files
- D. datasets

Answer: B

QUESTION 56:

Which environment variable controls whether performance statistics can be displayed in Designer?

- A. APT_NO_JOBMON
- B. APT_PERFORMANCE_DATA
- C. APT_PM_SHOW_PIDS
- D. APT_RECORD_COUNTS

Answer: A

QUESTION 57:

Which statement describes how to add functionality to the Transformer stage?

- A. Create a new parallel routine in the Routines category that specifies the name, path, type, and return type of a function written and compiled in C++.
- B. Create a new parallel routine in the Routines category that specifies the name, path, type, and return type of an external program.
- C. Create a new server routine in the Routines category that specifies the name and category of a function written in DataStage Basic.
- D. Edit the C++ code generated by the Transformer stage.

Answer: A

QUESTION 58:

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. easing use of captured job statistics (eg. row counts in an XML file) by processes or users outside of DataStage
- D. improving developer productivity and quality by distinguishing link names within stage editors

Answer: C,D

QUESTION 59:

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

Answer: A,B

QUESTION 60:

Data volumes have grown significantly in the last month. A parallel job that used to run well is now using unacceptable amounts of disk space and running slowly. You have reviewed the job and explicitly defined the sorting and partitioning requirements for each stage but the behavior of the job has not changed. Which two actions improve performance based on this information? (Choose two.)

- A. Change the sort methods on all sorts to "Don't Sort - Previously Grouped".
- B. Increase the value of environment variable APT_PARTITION_NUMBER to increase the level of parallelism for the sorts.
- C. Enable the environment variable APT_SORT_INSERTION_CHECK_ONLY in the job.
- D. Enable the environment variable APT_NO_SORT_INSERTION in the job.

Answer: C,D

QUESTION 61:

You must create a job that extracts data from multiple DB2/UDB databases on a mainframe and AS400 platforms without any database specific client software other than what is included with DataStage. Which two stages will let you parameterize the options required to connect to the database and enable you to use the same job if the source metadata matches all source tables? (Choose two.)

- A. DB2/UDB Enterprise stage
- B. ODBC Enterprise stage
- C. DB2/UDB API stage
- D. Dynamic RDBMS stage

Answer: B,D

QUESTION 62:

You have set the "Preserve Partitioning" flag for a Sort stage to request that the next stage preserves whatever partitioning it has implemented. Which statement describes what will happen next?

- A. The job will compile but will abort when run.
- B. The job will not compile.

- C. The next stage can ignore this request but a warning is logged when the job is run depending on the stage type that ignores the flag.
- D. The next stage disables the partition options that are normally available in the Partitioning tab.

Answer: C

QUESTION 63:

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

- A. Cleanup resources from a failed DataStage job.
- B. Start and stop the DataStage engine.
- C. Provide read access to a DataStage EE configuration file.
- D. Report DataStage client connections.

Answer: B

QUESTION 64:

To encourage users to update the short description for a job, how can you make the short description visible and updateable on the main canvas?

- A. Add a Description Annotation field to the job canvas and select the Short Description property.
- B. Right-click on the job canvas and choose Show Job Short Description in the submenu.
- C. Click the Show Job Short Description option in the Job Properties.
- D. Add an Annotation stage to the job canvas and copy and paste in the short description.

Answer: A

QUESTION 65:

Which feature in DataStage will allow you determine all the jobs that are using a Shared Container?

- A. Extended Job view
- B. Reporting Assistant
- C. Usage Analysis
- D. Impact Analysis

Answer: C

QUESTION 66:

Which two statements regarding the usage of data types in the parallel engine are correct?
(Choose two.)

- A. The best way to import RDBMS data types is using the ODBC importer.
- B. The parallel engine will use its interpretation of the Oracle meta data (e.g, exact data types) based on interrogation of Oracle, overriding what you may have specified in the Columns tabs.
- C. The best way to import RDBMS data types is using the Import Orchestra Schema Definitions using orchdbutil.
- D. The parallel engine and server engine have exactly the same data types so there is no conversion cost overhead from moving data between the engines.

Answer: B,C

QUESTION 67:

What is the default data type produced by the Aggregator stage?

- A. integer
- B. double precision floating point
- C. decimal [38,9]
- D. single precision floating point

Answer: B

QUESTION 68:

Which two describe a DataStage EE installation in a clustered environment? (Choose two.)

- A. The C++ compiler must be installed on all cluster nodes.
- B. Transform operators must be copied to all nodes of the cluster.
- C. The DataStage parallel engine must be installed or accessible in the same directory on all machines in the cluster.
- D. A remote shell must be configured to support communication between the conductor and section leader nodes.

Answer: C,D

QUESTION 69:

Using FTP, a file is transferred from an MVS system to a LINUX system in binary transfer mode. Which data conversion must be used to read a packed decimal field in the file?

- A. treat the field as a packed decimal
- B. packed decimal fields are not supported
- C. treat the field as ASCII
- D. treat the field as EBCDIC

Answer: A

QUESTION 70:

Which two statements describe the properties needed by the Oracle Enterprise stage to operate in parallel direct load mode. (Choose two.)

- A. Only index organized tables allow you to use parallel direct load mode and you need to set both DIRECT and PARALLEL properties to True, otherwise the load will be executed in parallel but will not use direct mode.
- B. Set the Write Method to Load.
- C. The table can have any number of indexes but the job will run slower since the indexes will have to be updated as the data is loaded.
- D. The table must not have any indexes, or you must include the rebuild Index Mode property unless the only index on the table is the primary key in which case you can use the Disable Constraints property.

Answer: B,D

QUESTION 71:

When a sequential file is read using a Sequential File stage, the parallel engine inserts an operator to convert the data to the internal format. Which operator is inserted?

- A. import operator
- B. copy operator
- C. tsort operator
- D. export operator

Answer: A

QUESTION 72:

A client requires that any job that aborts in a Job Sequence halt processing. Which three activities would provide this capability? (Choose three.)

- A. Nested Condition Activity
- B. Exception Handler
- C. Sequencer Activity
- D. Sendmail Activity
- E. Job trigger

Answer: A,B,E

QUESTION 73:

Exhibit:



In the exhibit, Sequential_File reads a single file stored on an unshared file system. The Oracle stage includes a Partition Table argument. The job runs with a configuration file representing a four node cluster of machines.

Which three components in the job can be run on all members of the cluster? (Choose three.)

- A. the Sequential File stage
- B. the Lookup stage
- C. the Oracle stage
- D. the Server Shared Container
- E. the Join stage

Answer: B,C,E

QUESTION 74:

Which partitioning method would yield the most even distribution of data without duplication?

- A. Entire
- B. Round Robin

- C. Hash
- D. Random

Answer: B

QUESTION 75:

Which two statements describe the operation of the Merge stage? (Choose two.)

- A. Duplicate records should always be removed from the master data set and from the update data sets if there is more than one update data set.
- B. Merge stages can only have one reject link.
- C. Duplicate records should always be removed from the master data set and from the update data set even if there is only one update data set.
- D. Merge stages can have multiple reject links.

Answer: A,D

QUESTION 76:

Which three accurately describe the differences between a DataStage server root installation and a non-root installation? (Choose three.)

- A. A non-root installation enables auto-start on reboot.
- B. A root installation must specify the user "dsadm" as the DataStage administrative user.
- C. A non-root installation inherits the permissions of the user who starts the DataStage services.
- D. A root installation will start DataStage services in impersonation mode.
- E. A root installation enables auto-start on reboot.

Answer: C,D,E

QUESTION 77:

Which two statements are true about the Join stage? (Choose two.)

- A. All the inputs to the Join stage must be sorted by the Join key.
- B. Join stages can have reject links that capture rows without matches.
- C. The Join stage supports inner, left outer, and right outer joins.
- D. The Join stage uses more memory than the Lookup stage.

Answer: A,C

QUESTION 78:

Which type of file is both partitioned and readable by external applications?

- A. fileset
- B. Lookup fileset
- C. dataset
- D. sequential file

Answer: A

QUESTION 79:

Which three statements about the Enterprise Edition parallel Transformer stage are correct? (Choose three.)

- A. The Transformer allows you to copy columns.
- B. The Transformer allows you to do lookups.
- C. The Transformer allows you to apply transforms using routines.
- D. The Transformer stage automatically applies 'NullToValue' function to all non-nullable output columns.
- E. The Transformer allows you to do data type conversions.

Answer: A,C,E

QUESTION 80:

Exhibit:

Revolution	Citizen	Citizen	Exchange
1789	Left	M_B_Dextrous	Nasdaq
1776	M_B_Dextrous	Righty	NYSE

Table A Table B

In the exhibit, which statement is true about performing a Full Outer Join on the two tables where the primary key is Citizen?

- A. Since this is an outer join, only rows from Table A whose key columns contain equal values are transferred to the output link.
- B. The default non-key value for unmatched records is always null.
- C. Two new columns (leftRec_Citizen and rightRec_Citizen) are created on the output.
- D. All the columns from Table A and all the non-key columns from Table B are transferred to the output link.

Answer: C

QUESTION 81:

Your job reads from a file using a Sequential File stage running sequentially. You are using a Transformer following the Sequential File stage to format the data in some of the columns. Which partitioning algorithm would yield optimized performance?

- A. Hash
- B. Random
- C. Round Robin
- D. Entire

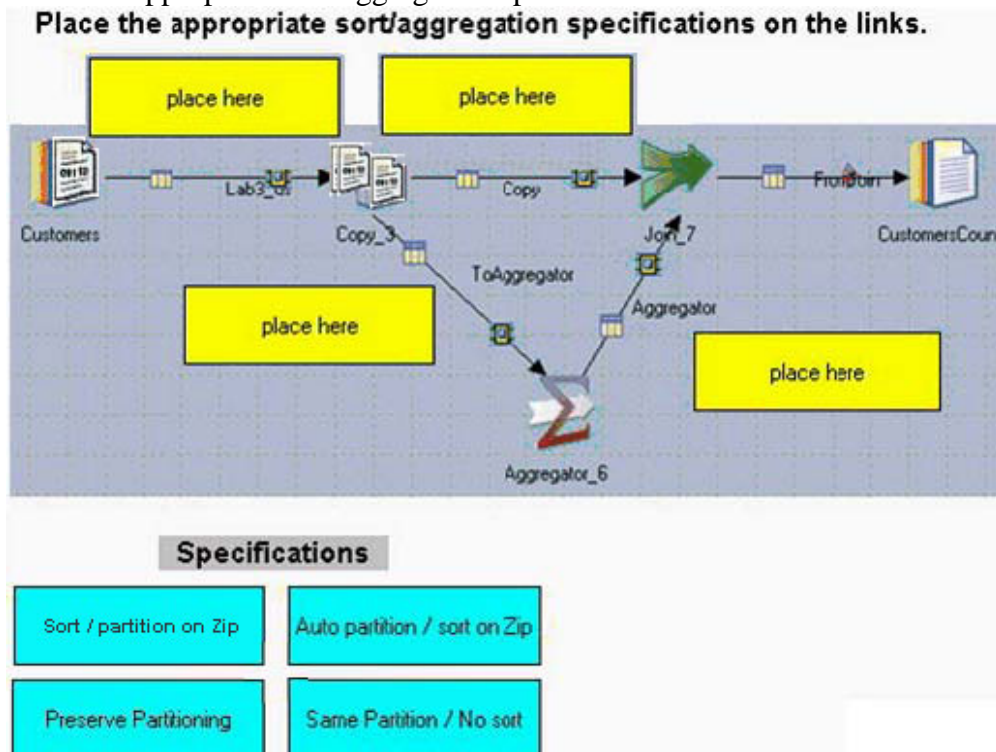
Answer: C

QUESTION 82:

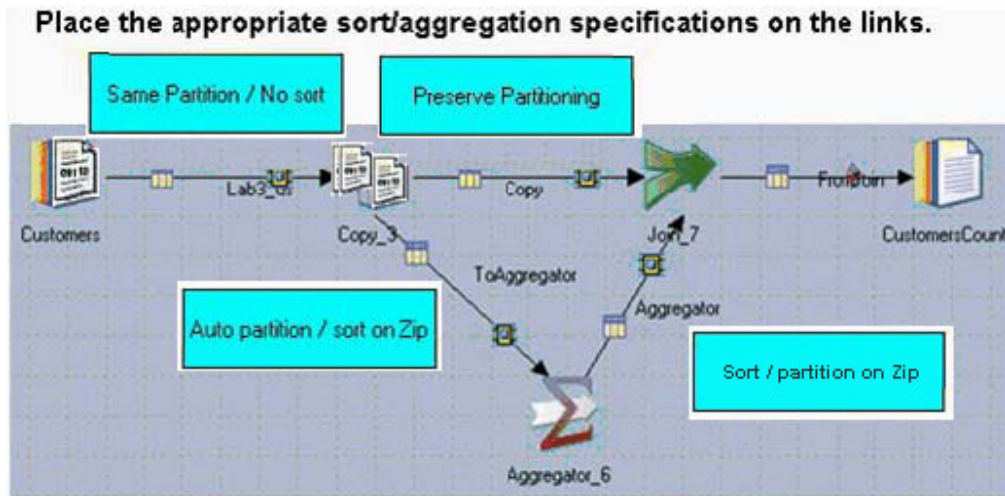
DRAG DROP

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.



Answer:



QUESTION 83:

Which two statements are true about XML Meta Data Importer? (Choose two.)

- A. XML Meta Data Importer is capable of reporting syntax and semantic errors from an XML file.
- B. XPATH expressions that are created during XML metadata import cannot be modified.
- C. XML Meta Data Importer can import Table Definitions from only XML documents.
- D. XPATH expressions that are created during XML metadata import are used by XML Input stage and XML Output stage.

Answer: A,D

QUESTION 84:

In the exhibit, what is "bigdata"?

```
{
  node 'n1' {
    fastname "s1"
    pool "" "n1" "s1" "app2" "sort"
    resource disk "/orch/n1/d1" {}
    resource disk "/orch/n1/d2" {"bigdata"}
    resource scratchdisk "/temp" {"sort"}
  }
  node 'n2' {
    fastname "s2"
    pool "" "n2" "s2" "app1"
    resource disk "/orch/n2/d1" {}
    resource disk "/orch/n2/d2" {"bigdata"}
    resource scratchdisk "/temp" {}
  }
  node 'n3' {
    fastname "s3"
    pool "" "n3" "s3" "app1"
    resource disk "/orch/n3/d1" {}
    resource scratchdisk "/temp" {}
  }
  node 'n4' {
    fastname "s4"
    pool "n4" "s4" "app1"
    resource disk "/orch/n4/d1" {}
    resource scratchdisk "/temp" {} }
}
```

- A. a resource pool
- B. a resource
- C. a node
- D. a node pool

Answer: A

QUESTION 85:

Which two stages allow field names to be specified using job parameters? (Choose two.)

- A. Transformer stage
- B. Funnel stage
- C. Modify stage
- D. Filter stage

Answer: C,D

QUESTION 86:

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

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

Answer: B,C

QUESTION 87:

When importing a COBOL file definition, which two are required? (Choose two.)

- A. The file you are importing is accessible from your client workstation.
- B. The file you are importing contains level 01 items.
- C. The column definitions are in a COBOL copybook file and not, for example, in a COBOL source file.
- D. The file does not contain any OCCURS DEPENDING ON clauses.

Answer: A,B

QUESTION 88:

The parallel dataset input into a Transformer stage contains null values. What should you do to properly handle these null values?

- A. Convert null values to a valid values in a stage variable.
- B. Convert null values to a valid value in the output column derivation.
- C. Null values are automatically converted to blanks and zero, depending on the target data type.
- D. Trap the null values in a link constraint to avoid derivations.

Answer: A

QUESTION 89:

Which two statements are correct about XML stages and their usage? (Choose two.)

- A. XML Input stage converts XML data to tabular format.
- B. XML Output stage converts tabular data to XML hierarchical structure.
- C. XML Output stage uses XSLT stylesheet for XML to tabular transformations.
- D. XML Transformer stage converts XML data to tabular format.

Answer: A,B

QUESTION 90:

Which Oracle Enterprise Stage property can be set using the DB Options group to tune the performance of your job with regard to the number of network packets transferred during the execution of the job?

- A. memsize
- B. blocksize
- C. arraysize
- D. transactionsize

Answer: C

QUESTION 91:

Which two statements are true of the Merge stage? (Choose two.)

- A. The Merge stage supports inner, left outer, and right outer joins.
- B. All the inputs to the Merge stage must be sorted by the key column(s).
- C. The Merge stage supports only two input links.
- D. The columns that are used to merge the incoming data must be identically named.

Answer: B,D

QUESTION 92:

Your business requirement is to read data from three Oracle tables that store historical sales data from three regions for loading into a single Oracle table. The table definitions are the same for all three tables, the only difference is that each table contains data for a particular region.

Which two statements describe how this can be done? (Choose two.)

- A. Create a job with a single Oracle Enterprise stage that executes a custom SQL statement with a FETCH ALL operator that outputs the data to an Oracle Enterprise stage.
- B. Create a job with a single Oracle Enterprise stage that executes a custom SQL statement with a UNION ALL operator that outputs the data to an Oracle Enterprise stage.
- C. Create a job with three Oracle Enterprise stages to read from the tables and output to a Collector stage which in turn outputs the data to an Oracle Enterprise stage.

D. Create a job with three Oracle Enterprise stages to read from the tables and output to a Funnel stage which in turn outputs the data to an Oracle Enterprise stage.

Answer: B,D

QUESTION 93:

In the exhibit, how many parallel partitions will each DataStage EE operator run by default?

```
{
  node 'n1' {
    fastname "s1"
    pool "" "n1" "s1" "app2" "sort"
    resource disk "/orch/n1/d1" {}
    resource disk "/orch/n1/d2" {"bigdata"}
    resource scratchdisk "/temp" {"sort"}
  }
  node 'n2' {
    fastname "s2"
    pool "" "n2" "s2" "app1"
    resource disk "/orch/n2/d1" {}
    resource disk "/orch/n2/d2" {"bigdata"}
    resource scratchdisk "/temp" {}
  }
  node 'n3' {
    fastname "s3"
    pool "" "n3" "s3" "app1"
    resource disk "/orch/n3/d1" {}
    resource scratchdisk "/temp" {}
  }
  node 'n4' {
    fastname "s4"
    pool "n4" "s4" "app1"
    resource disk "/orch/n4/d1" {}
    resource scratchdisk "/temp" {} }
}
```

- A. four
- B. five
- C. three
- D. two

Answer: C

QUESTION 94:

Which two would require the use of a Transformer stage instead of a Copy stage?
(Choose two.)

- A. Drop a column.
- B. Send the input data to multiple output streams.
- C. Trim spaces from a character field.
- D. Select certain output rows based on a condition.

Answer: C,D

QUESTION 95:

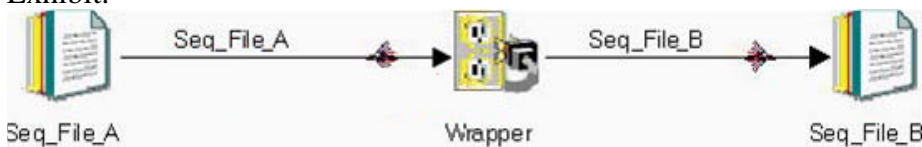
Which two requirements must be in place when using the DB2 Enterprise stage in LOAD mode? (Choose two.)

- A. Tablespace must be in read only mode prior to the job being launched.
- B. User running the job has dbadm privileges.
- C. Tablespace must be placed in a load pending state prior to the job being launched.
- D. Tablespace cannot be addressed by anyone else while the load occurs.

Answer: B,D

QUESTION 96:

Exhibit:



In the exhibit, which sequence of events is true?

- A. Import, Export
- B. Import, Export, Export
- C. Import, Export, Import, Export
- D. Export, Import, Export

Answer: C

QUESTION 97:

Detail sales transaction data is received on a monthly basis in a single file that includes CustID, OrderDate, and Amount fields. For a given month, a single CustID may have multiple transactions.

Which method would remove duplicate CustID rows, selecting the largest transaction amount for a given CustID?

- A. Hash partition on CustID. Perform a non-stable, unique Sort on CustID and Amount (descending).
- B. Use Auto partitioning on all links. Perform a non-unique Sort on CustID and Amount (ascending) followed by a RemoveDuplicates on CustID and Amount (duplicateToRetain=Last).
- C. Use Auto partitioning. Perform a unique Sort on CustID and Amount (ascending).
- D. Hash partition on CustID. Perform a non-unique Sort on CustID and Amount (descending). Use same partitioning, followed by a RemoveDuplicates on CustID (duplicateToRetain=First).

Answer: D

QUESTION 98:

A file created by a mainframe COBOL program is imported to the DataStage parallel engine. The resulting record schema contains a fixed length vector of the three most recent telephone numbers called by each customer. The processing requirements state that each of the three most recent calls must be processed differently. Which stage should be used to restructure the record schema so that the Transformer stage can be used to process each of the three telephone numbers found in the vector?

- A. the Make Vector stage followed by a Split Subrecord stage
- B. the Split Subrecord stage followed by a Column Export stage
- C. the Promote Subrecord stage
- D. the Split Vector stage

Answer: D

QUESTION 99:

Which two statements are correct when using the Change Capture and Change Apply stages together on the same data? (Choose two.)

- A. You must apply a Differences stage to the output of the Change Capture stage before passing the data into the Change Apply stage.
- B. A Compare stage must be used following the Change Capture stage to identify changes to the change_code column values.
- C. The input to the Change Apply stage must have the same key columns as the input to the prior Change Capture stage.
- D. Both inputs of the Change Apply stage are designated as partitioned using the same partitioning method.

Answer: C,D

QUESTION 100:

Which three UNIX kernel parameters have minimum requirements for DataStage installations? (Choose three.)

- A. MAXUPROC - maximum number of processes per user
- B. NOFILES - number of open files
- C. MAXPERM - disk cache threshold
- D. NOPROC - no process limit
- E. SHMMAX - maximum shared memory segment size

Answer: A,B,E

QUESTION 101:

What is the default data type produced by the Aggregator stage?

- A. decimal [38,9]
- B. integer
- C. single precision floating point
- D. double precision floating point

Answer: D

QUESTION 102:

Which partitioning method requires specifying a key?

- A. Random
- B. DB2
- C. Entire
- D. Modulus

Answer: D

QUESTION 103:

Which "Reject Mode" option in the Sequential File stage will write records to a reject link?

- A. Output
- B. Fail

- C. Drop
- D. Continue

Answer: A

QUESTION 104:

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

Answer: D

QUESTION 105:

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

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

Answer: A

QUESTION 106:

A single sequential file exists on a single node. To read this sequential file in parallel, what should be done?

- A. Set the Execution mode to "Parallel".
- B. A sequential file cannot be read in parallel using the Sequential File stage.
- C. Select "File Pattern" as the Read Method.
- D. Set the "Number of Readers Per Node" optional property to a value greater than 1.

Answer: D

QUESTION 107:

How does a Join stage process an Inner join?

- A. 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.
- B. It transfers records from the input data sets whose key columns contain equal values to the output data set.
- C. 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.
- 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.

Answer: B

QUESTION 108:

When a sequential file is written using a Sequential File stage, the parallel engine inserts an operator to convert the data from the internal format to the external format. Which operator is inserted?

- A. export operator
- B. copy operator
- C. import operator
- D. tsort operator

Answer: A

QUESTION 109:

How are transaction commit operations handled by the DB2/UDB Enterprise stage?

- A. Commit operations can only be defined by the number of rows since the start of a transaction.
- B. Transaction commits can be controlled by defining the number of rows per transaction or by a specific time period defined by the number of seconds elapsed between commits.
- C. Commit operations can only be defined by the number of seconds since the start of a transaction.
- D. Commit operations can be defined globally by setting APT_TRANSACTION_ROWS variable.

Answer: B

QUESTION 110:

In which situation should a BASIC Transformer stage be used in a DataStage EE job?

- A. in a job containing complex routines migrated from DataStage Server Edition
- B. in a job requiring lookups to hashed files
- C. in a large-volume job flow
- D. in a job requiring complex, reusable logic

Answer: A

QUESTION 111:

To encourage users to update the short description for a job, how can you make the short description visible and updateable on the main canvas?

- A. Click the Show Job Short Description option in the Job Properties.
- B. Right-click on the job canvas and choose Show Job Short Description in the submenu.
- C. Add an Annotation stage to the job canvas and copy and paste in the short description.
- D. Add a Description Annotation field to the job canvas and select the Short Description property.

Answer: D

QUESTION 112:

Which two statements describe the properties needed by the Oracle Enterprise stage to operate in parallel direct load mode. (Choose two.)

- A. The table can have any number of indexes but the job will run slower since the indexes will have to be updated as the data is loaded.
- B. The table must not have any indexes, or you must include the rebuild Index Mode property unless the only index on the table is the primary key in which case you can use the Disable Constraints property.
- C. Only index organized tables allow you to use parallel direct load mode and you need to set both DIRECT and PARALLEL properties to True, otherwise the load will be executed in parallel but will not use direct mode.
- D. Set the Write Method to Load.

Answer: B,D

QUESTION 113:

Which requirement must be satisfied to read from an Oracle table in parallel using the Oracle Enterprise stage?

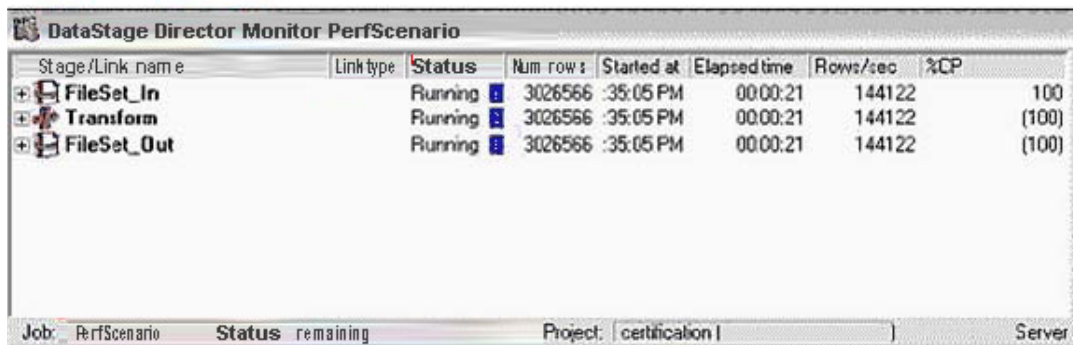
- A. Set the environment variable \$ORA_IGNORE_CONFIG_FILE_PARALLELISM.
- B. Configure the source table to be partitioned within the source Oracle database.
- C. Oracle Enterprise stage always reads in parallel.
- D. Specify the Partition Table option in the Oracle source stage.

Answer: D

QUESTION 114:

The exhibit shows the result in the DataStage Director Monitor on the first three components in a larger flow that reads data from a source before temporarily checkpointing the data.

What would improve performance of the monitored parallel job?



Stage/Link name	Link type	Status	Num rows	Started at	Elapsed time	Rows/sec	%CP
FileSet_In		Running	3026566	:35:05 PM	00:00:21	144122	100
Transform		Running	3026566	:35:05 PM	00:00:21	144122	(100)
FileSet_Out		Running	3026566	:35:05 PM	00:00:21	144122	(100)

Job: PerfScenario Status: remaining Project: certification I Server t

- A. Split the Logic performed in the Transformer into two separate Transformer stages.
- B. Replace target FileSet stage with target Sequential File stage.
- C. Set APT_BUFFERING_POLICY=NO_BUFFERING.
- D. Replace target FileSet stage with Data Set stage.

Answer: D

QUESTION 115:

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

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

Answer: A,C

QUESTION 116:

A DataStage EE job is sourcing a flat file which contains a VARCHAR field. This field needs to be mapped to a target field that is a date. Which task will accomplish this?

- A. Perform a datatype conversion using DateToString function inside the Transformer stage.
- B. Use a Modify stage to perform the type conversion.
- C. DataStage automatically performs this type conversion by default.
- D. Use a Copy stage to perform the type conversion.

Answer: B

QUESTION 117:

What are two causes for records being incorrectly identified as an Edit in the Change Capture stage? (Choose two.)

- A. The before and after input datasets are not in the same sort order.
- B. At least one field containing unique values for all records is specified as a change value.
- C. The key fields are not unique.
- D. A key field contains null values in some records.

Answer: A,C

QUESTION 118:

Using a second (reject) output link from a Sequential File read stage, which two methods will allow the number of records rejected by the Sequential File stage to be captured and logged to a file? (Choose two.)

(Note: The log file should contain only the number of records rejected, not the records themselves. Assume the file is named reject_count.stage_10.)

- A. Send the rejected records to a dataset named rejects_stage_10.ds, and define an After Job Subroutine to execute the command 'dsrecords rejects_stage_10.ds > reject_count.stage_10'.
- B. Send the rejected records to a Column Generator stage to generate a constant key value, use the Aggregator stage with the generated key value to count the records, then write the results to the log file using another Sequential File stage.

- C. Send the rejected records to a Peek stage and define an After Job Subroutine to execute the command 'dsjob -report peek_rejects > reject_count.stage_10' with the appropriate -jobid and -project arguments.
- D. Send the rejected records to a Change Capture stage, use a Sort Funnel stage to sort it all out, then write the resulting record count to the log file using another Sequential File stage.

Answer: A,B

QUESTION 119:

What does setting an environment variable, specified as a job parameter, to PROJDEF do?

- A. Populates the environment variable with the value of PROJDEF.
- B. Explicitly unsets the environment variable.
- C. Uses the value for the environment variable as shown in the DataStage Administrator.
- D. Uses the current setting for the environment variable from the operating system.

Answer: C

QUESTION 120:

A bank receives daily credit score updates from a credit agency in the form of a fixed width flat file. The monthly_income column is an unsigned nullable integer (int32) whose width is specified as 10, and null values are represented as spaces. Which Sequential File property will properly import any nulls in the monthly_income column of the input file?

- A. Set the record level fill char property to the space character (' ').
- B. Set the null field value property to a single space (' ').
- C. Set the C_format property to "%d. 10".
- D. Set the null field value property to ten spaces (' ').

Answer: D

QUESTION 121:

Which three can be identified from a parallel job score? (Choose three.)

- A. components inserted by DataStage parallel engine at runtime (such as buffers, sorts).
- B. runtime column propagation was enabled to support dynamic job designs.
- C. whether a particular job was run in a clustered or SMP hardware environment.

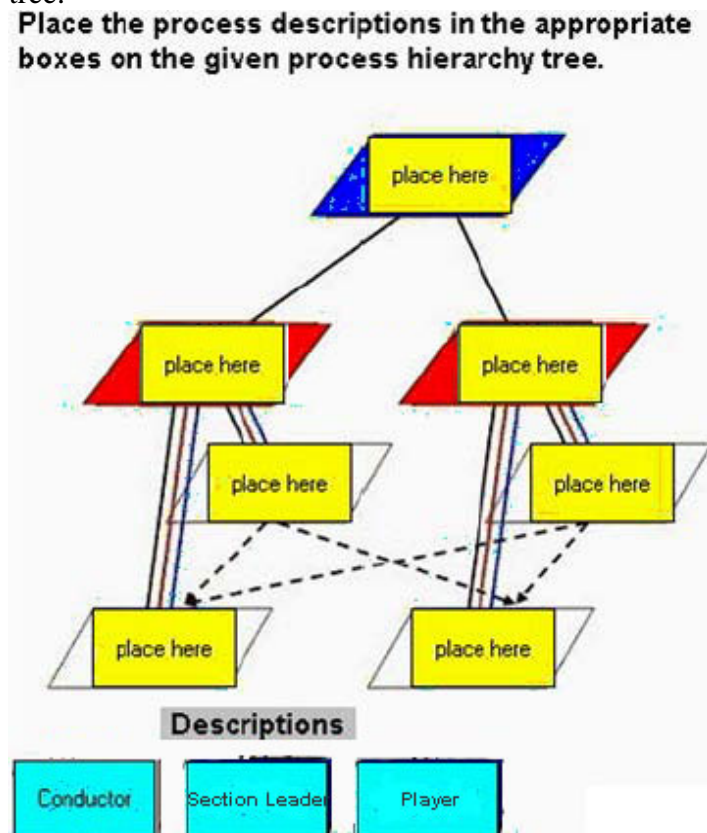
- D. stages whose logic was combined together to minimize the number of physical processes used to execute a given job.
- E. actual partitioning method used at runtime for each link defined with "Auto" partitioning.

Answer: A,D,E

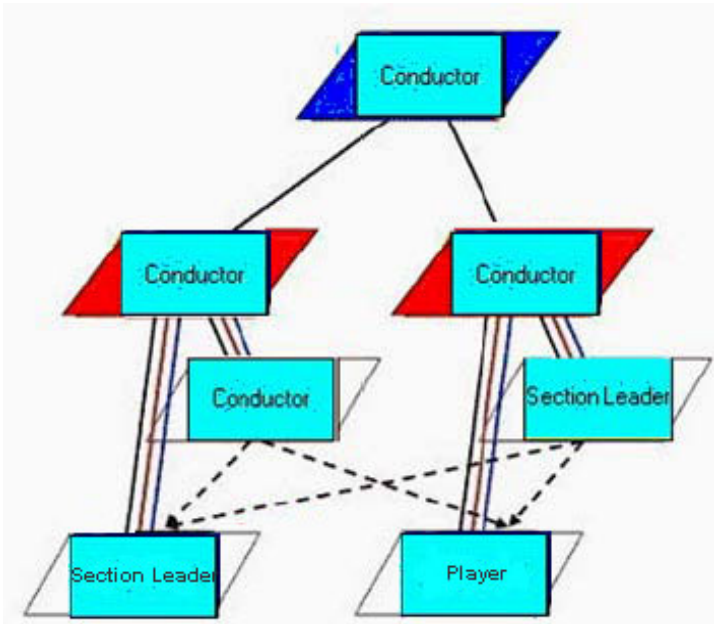
QUESTION 122:

DRAG DROP

Place the process descriptions in the appropriate boxes on the given process hierarchy tree.



Answer:



QUESTION 123:

A job requires that data be grouped by primary keys and then sorted within each group by secondary keys to reproduce the results of a group-by and order-by clause common to relational databases. The designer has chosen to implement this requirement with two Sort stages in which the first Sort stage sorts the records by the primary keys. Which set of properties must be specified in the second Sort stage?

- A. Specify both sets of keys, with a Sort Key Option of "Don't Sort (Previously Sorted)" on the primary keys, and "Don't Sort (Previously Grouped)", and the Stable Sort option set to True.
- B. Specify only the secondary keys, with a Sort Key Option of "Don't Sort (Previously Sorted)" and the Stable Sort option set to True.
- C. Specify both sets of keys, with a Sort Key Option of "Don't Sort (Previously Grouped)" on the primary keys, and "Sort" on the secondary keys.
- D. Specify only the secondary keys, with a Sort Key Option of "Sort" and the Stable Sort option set to True.

Answer: C

QUESTION 124:

HOTSPOT

Given the following DataStage Director log, click on the entry that contains the parallel job score for the running job.

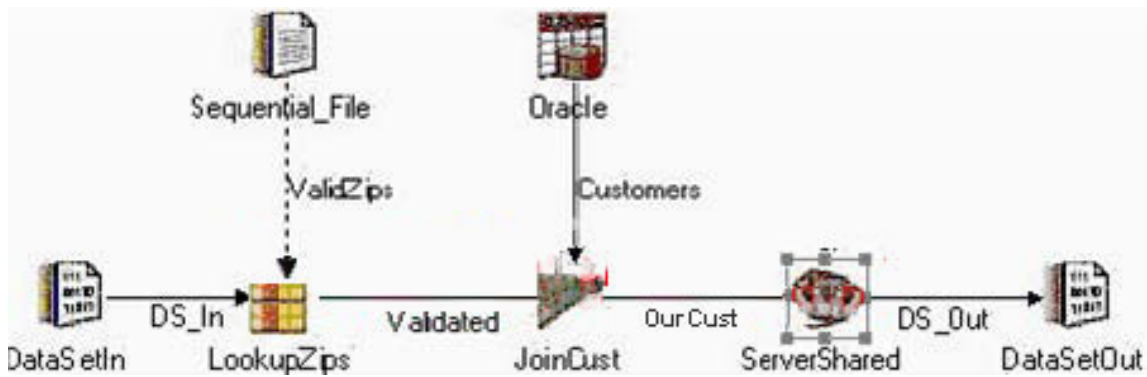
Ascential DataStage Director			
Project View Search Job Tools Help			
<div> </div>			
> Occured	/Off date	Type	Every
6:27:20 AM	10/12/2005	Control	Starting Job cpJoinPartEx1a. (...)
6:27:20 AM	10/12/2005	Info	Environment variable settings: (...)
6:27:20 AM	10/12/2005	Info	Parallel job initiated (...)
6:27:05 AM	10/12/2005	Info	main_program: Ascential DataStage(tm) Enterprise Edition 7.5.1A (...)
6:27:05 AM	10/12/2005	Info	main_program: orchgeneral: loaded(...)
6:27:05 AM	10/12/2005	Info	main_program: APT configuration file: /opt/Ascential/DataStage/Conf
6:27:06 AM	10/12/2005	Info	main_program: This step has no datasets. (...)
6:27:20 AM	10/12/2005	Info	main_program: This step has 10 datasets: (...)
6:27:20 AM	10/12/2005	Info	Customer_Expert_D.0: Field export complete. 50000 records converted s
6:27:20 AM	10/12/2005	Info	Customer_Expert_T.0: Field export complete. 1 records converted succe
6:27:20 AM	10/12/2005	Info	Sequential_File_5.0: Export complete; 100002 records exported succe
6:27:20 AM	10/12/2005	Info	main_program: Step execution finished with status = OK.
6:27:20 AM	10/12/2005	Info	main_program: Startup time, 0:10; production run time, 0:05.
6:27:20 AM	10/12/2005	Info	Parallel job reports successful completion
6:27:20 AM	10/12/2005	Control	Finished Job cpJoinPartEx1a.
Log for job: cpJoinPartEx1a 23 entries (filtered) Server time: 10/12/2005 06:27 AM			

Answer:

Ascential DataStage Director			
Project View Search Job Tools Help			
<div> </div>			
> Occured	/Off date	Type	Every
6:27:10 AM	10/12/2005	Control	Starting Job cpJoinPartEx1a. (...)
6:27:04 AM	10/12/2005	Info	Environment variable settings: (...)
6:27:04 AM	10/12/2005	Info	Parallel job initiated (...)
6:27:05 AM	10/12/2005	Info	main_program: Ascential DataStage(tm) Enterprise Edition 7.5.1A (...)
6:27:05 AM	10/12/2005	Info	main_program: orchgeneral: loaded(...)
6:27:05 AM	10/12/2005	Info	main_program: APT configuration file: /opt/Ascential/DataStage/Conf
6:27:06 AM	10/12/2005	Info	main_program: This step has no datasets. (...)
6:27:20 AM	10/12/2005	Info	main_program: This step has 10 datasets: (...)
6:27:20 AM	10/12/2005	Info	Customer_Expert_D.0: Field export complete. 50000 records converted s
6:27:20 AM	10/12/2005	Info	Customer_Expert_T.0: Field export complete. 1 records converted succe
6:27:20 AM	10/12/2005	Info	Sequential_File_5.0: Export complete; 100002 records exported succe
6:27:20 AM	10/12/2005	Info	main_program: Step execution finished with status = OK.
6:27:20 AM	10/12/2005	Info	main_program: Startup time, 0:10; production run time, 0:05.
6:27:20 AM	10/12/2005	Info	Parallel job reports successful completion
6:27:20 AM	10/12/2005	Control	Finished Job cpJoinPartEx1a.
Log for job: cpJoinPartEx1a 23 entries (filtered) Server time: 10/12/2005 06:27 AM			

QUESTION 125:

Exhibit:



In the exhibit, Sequential_File reads a single file stored on an unshared file system. The Oracle stage includes a Partition Table argument. The job runs with a configuration file representing a four node cluster of machines.

Which three components in the job can be run on all members of the cluster? (Choose three.)

- A. the Lookup stage
- B. the Oracle stage
- C. the Server Shared Container
- D. the Join stage
- E. the Sequential File stage

Answer: A,B,D

QUESTION 126:

Which two statements are true about the Lookup stage? (Choose two.)

- A. The Lookup stage supports one and only one lookup table per stage.
- B. A reject link can be specified to capture input rows that do not have matches in the lookup tables.
- C. The source primary, input link must be sorted.
- D. The Lookup stage uses more memory than the Merge and Join stages.

Answer: B,D

QUESTION 127:

Which statement about job parameter usage is true?

- A. You can change job parameters while a job is running and the changes will immediately be applied mid-job.
- B. You can use environment variables to set parameter values linked to the Job Sequence.
- C. You can change the parameter values in an initialization file linked to a Job Sequence

.ini file.

D. Changes to the job parameters in the Designer do not require a recompile to be applied to the job.

Answer: B

QUESTION 128:

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

- A. Set the "Nullable" property for the column on the output link to "Yes".
- B. Set the "Key" property for the column on the output link to "Yes".
- C. Check the "Repetition Element Required" box on the output link tab.
- D. No special settings are required. XML Input stage automatically detects the repetition element from the XPath expression.

Answer: B

QUESTION 129:

During a sequential file read, you experience an error with the data. What is a valid technique for identifying the column causing the difficulty?

- A. Set the "data format" option to text on the Record Options tab.
- B. Enable tracing in the DataStage Administrator Tracing panel.
- C. Enable the "print field" option at the Record Options tab.
- D. Set the APT_IMPORT_DEBUG environmental variable.

Answer: C

QUESTION 130:

Which two would determine the location of the raw data files for a parallel dataset? (Choose two.)

- A. the orchadmin tool
- B. the Data Set Management tool
- C. the DataStage Administrator
- D. the Dataset stage

Answer: A,B

QUESTION 131:

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

Answer: C,D

QUESTION 132:

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. A database function that returns the current value of an Oracle object is required as part of the result set.
- B. When the number of input rows is significantly smaller than the number of rows in the lookup table.
- C. When the Oracle database is on the same system as the DataStage server.
- D. When the number of input rows is significantly larger than the number of rows in the lookup table.

Answer: A,B

QUESTION 133:

On which two does the number of data files created by a fileset depend ? (Choose two.)

- A. the size of the partitions of the dataset
- B. the number of CPUs
- C. the schema of the file
- D. the number of processing nodes in the default node pool

Answer: A,D

QUESTION 134:

You are reading data from a Sequential File stage. The column definitions are specified by a schema. You are considering whether to follow the Sequential File stage by either a Transformer or a Modify stage. Which two criteria require the use one of these stages instead of the other? (Choose two.)

- A. You want to dynamically specify the name of an output column based on a job parameter, therefore you select a Modify stage.
- B. You want to replace NULL values by a specified constant value, therefore you select a Modify stage.
- C. You want to add additional columns, therefore you select a Transformer stage.
- D. You want to concatenate values from multiple input rows and write this to an output link, therefore you select a Transformer stage.

Answer: A,D

QUESTION 135:

Which three are the critical stages that would be necessary to build a Job Sequence that: picks up data from a file that will arrive in an directory overnight, launches a job once the file has arrived, sends an email to the administrator upon successful completion of the flow? (Choose three.)

- A. Sequencer
- B. Notification Activity
- C. Wait For File Activity
- D. Job Activity
- E. Terminator Activity

Answer: B,C,D

QUESTION 136:

Detail sales transaction data is received on a monthly basis in a single file that includes CustID, OrderDate, and Amount fields. For a given month, a single CustID may have multiple transactions.

Which method would remove duplicate CustID rows, selecting the most recent transaction for a given CustID?

- A. Use Auto partitioning. Perform a unique Sort on CustID and OrderDate (descending).
- B. Hash partition on CustID. Perform a non-unique Sort on CustID and OrderDate (ascending). Use same partitioning, followed by a RemoveDuplicates on CustID (duplicateToRetain=Last).
- C. Hash partition on CustID. Perform a unique Sort on CustID and OrderDate (ascending).

D. Use Auto partitioning on all links. Perform a non-unique Sort on CustID and OrderDate (ascending) followed by a RemoveDuplicates on CustID and OrderDate (duplicateToRetain=Last).

Answer: B

QUESTION 137:

You have three output links coming out of a Transformer. Two of them (A and B) have constraints you have defined. The third you want to be an Otherwise link that is to contain all of the rows that do not satisfy the constraints of A and B. This Otherwise link must work correctly even if the A and B constraints are modified. Which two are required? (Choose two.)

- A. The Otherwise link must be first in the link ordering.
- B. A constraint must be coded for the Otherwise link.
- C. The Otherwise link must be last in the link ordering.
- D. The Otherwise check box must be checked.

Answer: C,D

QUESTION 138:

Which two statements are true about DataStage Parallel Buildop stages? (Choose two.)

- A. Unlike standard DataStage stages they do not have properties.
- B. They are coded using C/C++.
- C. They are coded using DataStage Basic.
- D. Table Definitions are used to define the input and output interfaces of the BuildOp.

Answer: B,D

QUESTION 139:

Which environment variable controls whether performance statistics can be displayed in Designer?

- A. APT_RECORD_COUNTS
- B. APT_PM_SHOW_PIDS
- C. APT_NO_JOBMON
- D. APT_PERFORMANCE_DATA

Answer: C

QUESTION 140:

Which two field level data type properties are schema properties that are standard SQL properties? (Choose two.)

- A. the character used to mark the end of a record
- B. values to be generated for a field
- C. whether a field is nullable
- D. the value that is to be written to a sequential file when the field is NULL

Answer: A,C

QUESTION 141:

In a Lookup operator, a key column in the stream link is VARCHAR and has Unicode set in the Extended attribute while the corresponding column in a reference link, also VARCHAR, does not.

What will allow correct comparison of the data?

- A. Convert the column in the stream link to the default code page using the UstringToString function in a Transformer operator prior to the Lookup operator and remove Unicode from the Extended attribute of the column.
- B. Convert the column in the reference link to the UTF-8 code page using the StringToUstring function in a Transformer operator prior to the Lookup operator, and set the Extended attribute of the column.
- C. Convert both columns to CHAR, pad with spaces, and remove Unicode from the Extended attribute in Transformer operators prior to the Lookup operator.
- D. Remove Unicode from the Extended attribute of the column from the beginning of the job to the Lookup operator and then set the Extended attribute of the column in the output mapping section of the Lookup operator.

Answer: B

QUESTION 142:

You have a parallel job that does not scale beyond two nodes. After investigation you find that the data has been partitioned on gender and you have an Aggregator stage that is accumulating totals based on gender using a sort method. Which technique should you use to allow the job to scale?

- A. Change the preserve partitioning option on the stage ahead of the aggregator to clear partitioning.

- B. Change the aggregation method to hash to eliminate the blocking sort operation.
- C. Add an additional column for partitioning to result in additional data partitions.
- D. Change the partitioning key to one that results in more partitions, add a second Aggregator stage that re-aggregates based on gender.

Answer: D

QUESTION 143:

HOTSPOT

Given the excerpt from a parallel job score shown in the exhibit, click on the four components that are inserted at runtime by the DataStage parallel engine. (Choose four.)

```
It has 8 operators:
op0[1p] {(sequential Sequential_File)
  on nodes {
    node1[op0,p0]
  }
}
op1[2p] {(parallel APT_Transform OperatorImpl VO SI_opJoin PastEx1b_Transformer
  on nodes {
    node1[op1,p0]
    node2[op1,p1]
  }
}
op2[2p] {(parallel inserted tsort operator (key={value=joinkey, subArgs={asc}}){0}}
  on nodes {
    node1[op2,p0]
    node2[op2,p1]
  }
}
op3[2p] {(parallel inserted tsort operator (key={value=joinkey, subArgs={asc}}))
  on nodes {
    node1[op3,p0]
    node2[op3,p1]
  }
}
op4[2p] {(parallel buffer{0}}
  on nodes {
    node1[op4,p0]
    node2[op4,p1]
  }
}
op5[2p] {(parallel buffer{1}}
  on nodes {
    node1[op5,p0]
    node2[op5,p1]
  }
}
op6[2p] {(parallel APT Join6 Operator in Joint
  on nodes {
    node1[op6,p0]
    node2[op6,p1]
  }
}
op7[2p] {(parallel Peek)
  on nodes {
    node1[op7,p0]
    node2[op7,p1]
  }
}
It runs 15 processes on 2 nodes.
```

Answer:

```

It has 8 operators:
op0[1p] {sequential Sequential_File}
  on node: {
    node1[op0,p0]
  }
}

op1[2p] {parallel APT_TransformOperatorImplVQSI_opJoinParallel_Transformer}
  on node: {
    node1[op2,p0]
    node2[op2,p1]
  }
}

op2[2p] {parallel inserted sort operator (key={value=junkkey, subArgs={a:c}})[0]}
  on node: {
    node1[op2,p0]
    node2[op2,p1]
  }
}

op3[2p] {parallel inserted sort operator (key={value=junkkey, subArgs={a:c}})}
  on node: {
    node1[op3,p0]
    node2[op3,p1]
  }
}

op4[2p] {parallel buffer[0]}
  on node: {
    node1[op4,p0]
    node2[op4,p1]
  }
}

op5[2p] {parallel buffer[1]}
  on node: {
    node1[op5,p0]
    node2[op5,p1]
  }
}

op6[2p] {parallel APT_JoinSubOperator in Join}
  on node: {
    node1[op6,p0]
    node2[op6,p1]
  }
}

op7[2p] {parallel Peek}
  on node: {
    node1[op7,p0]
    node2[op7,p1]
  }
}

It runs 15 processes on 2 nodes.

```

QUESTION 144:

DRAG DROP

Place the process descriptions in the appropriate boxes on the given process hierarchy tree.

```

graph TD
    C[Conductor] --> SL1[Section Leader]
    C --> SL2[Section Leader]
    SL1 --> P1[Player]
    SL1 --> P2[Player]
    SL1 --> P3[Player]
    SL2 --> P4[Player]
    SL2 --> P5[Player]
    SL2 --> P6[Player]
  
```

Descriptions

- Conductor
- Section Leader
- Player

- A. This cannot be accomplished in a single job.
- B. Use a Copy stage to split incoming records into two streams. One stream uses an Aggregator stage to count the number of duplicates. Join the Aggregator stage output back to the other stream using a Join stage.
- C. Use an Aggregator stage to group incoming records by customer ID and to count the number of records in each group. Output the results to the target DB2 table.

D. Use stage variables in a Transformer stage to keep a running count of records with the same customer ID. Add this count to a new output field and write the results to the DB2 table.

Answer: B

QUESTION 146:

A developer has separated processing into multiple jobs in order to support restartability. The first job writes to a file that the second job reads. The developer is trying to decide whether to write to a dataset or to a sequential file. Which factor favors using a dataset?

- A. I/O performance
- B. ability to archive
- C. memory usage
- D. disk space usage

Answer: A

QUESTION 147:

Which two statements describe functionality that is available using the dsjob command? (Choose two.)

- A. dsjob can be used to get a report containing job, stage, and link information.
- B. dsjob can be used to add a log entry for a specified job.
- C. dsjob can be used to compile a job.
- D. dsjob can be used to export job executables.

Answer: A,B

QUESTION 148:

Using FTP, a file is transferred from an MVS system to a LINUX system in binary transfer mode. Which data conversion must be used to read a packed decimal field in the file?

- A. treat the field as EBCDIC
- B. treat the field as a packed decimal
- C. packed decimal fields are not supported
- D. treat the field as ASCII

Answer: B

QUESTION 149:

Which two statements about shared containers are true? (Choose two.)

- A. You can make a local container a shared container but you cannot make a shared container a local container.
- B. Shared containers can be used to make common job components that are available throughout the project.
- C. Changes to a shared container are automatically available in all dependant jobs without a recompile.
- D. Server shared containers allow for DataStage Server Edition components to be placed in a parallel job.

Answer: B,D

QUESTION 150:

What are two ways to delete a persistent parallel dataset? (Choose two.)

- A. standard UNIX command rm
- B. orchadmin command rm
- C. delete the dataset Table Definition in DataStage Manager
- D. delete the dataset in Data Set Manager

Answer: B,D

QUESTION 151:

A parts supplier has a single fixed width sequential file. Reading the file has been slow, so the supplier would like to try to read it in parallel.

If the job executes using a configuration file consisting of four nodes, which two Sequential File stage settings will cause the DataStage parallel engine to read the file using four parallel readers? (Choose two.)

(Note: Assume the file path and name is /data/parts_input.txt.)

- A. Set the read method to specific file(s), set the file property to '/data/parts_input.txt', and set the number of readers per node option to 2.
- B. Set the read method to specific file(s), set the file property to '/data/parts_input.txt', and set the read from multiple nodes option to yes.
- C. Set read method to file pattern, and set the file pattern property to '/data/(@PART_COUNT)parts_input.txt'.
- D. Set the read method to specific file(s), set the file property to '/data/parts_input.txt',

and set the number of readers per node option to 4.

Answer: B,D

QUESTION 152:

A new column is required as part of a sorted dataset to be set to 1 when the value of sort key changes and to 0 when the value of sort key is the same as the prior record. Which statement is correct?

- A. This can be handled entirely within the Sort stage by setting the Create Key Change Column to True.
- B. This can be handled within the Sort stage by including a new column name in the output tab of the stage and by placing an "if, then, else" expression in the Derivation field in the stage Mapping tab to generate the value for the column.
- C. This can be handled entirely within the Sort stage by setting the Create Cluster Key Change Column to True.
- D. This cannot be handled entirely within the Sort stage.

Answer: A

QUESTION 153:

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 Hash method, the Aggregator stage will partition sort the data by the group key before building a hash table in memory.
- 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 Sort method, the Aggregator stage requires the data to have been partition sorted by the group key.

Answer: D

QUESTION 154:

What is the purpose of the Oracle Enterprise Stage Exception Table property?

- A. Enables you to specify a table which is used to capture selected column data that meets user defined criteria for debugging purposes. The table needs to exist before the

job is run.

- B. Enables you to specify a table which is used to record the ROWID information on rows that violate constraints during upsert/write operations.
- C. Enables you to specify a table which is used to record ROWID information on rows that violate constraints when constraints are re-enabled after a load operation.
- D. Enables you to specify that a table should be created (if it does not exist) to capture all exceptions when accessing Oracle.

Answer: C

QUESTION 155:

Which statement is true when Runtime Column Propagation (RCP) is enabled?

- A. DataStage Manager does not import meta data.
- B. DataStage Director does not supply row counts in the job log.
- C. DataStage Designer does not enforce mapping rules.
- D. DataStage Administrator does not allow default settings for environment variables.

Answer: C

QUESTION 156:

A customer wants to migrate only the performance-critical jobs of an existing DataStage Server Edition environment to a clustered deployment of DataStage Enterprise Edition. Which migration approach will allow the new parallel jobs to run on all nodes of the cluster using DataStage Enterprise Edition?

- A. Create a Server Shared Container containing the key stages for each job to be migrated, and use this container in a parallel job flow with native parallel source and target stages.
- B. Using requirements for the existing Server Edition jobs, create new parallel jobs using native parallel stages that implement the required source to target logic.
- C. Configure a DataStage Server Edition environment on each node in the DataStage Enterprise Edition cluster, import the existing .dsx export file to each node, compile and run the jobs in parallel using a parallel configuration file.
- D. Export the Server Edition jobs to a .dsx file, import them into the new DataStage Enterprise Edition environment, and re-compile.

Answer: B

QUESTION 157:

You are reading customer data using a Sequential File stage and Sorting it by customer ID using the Sort stage. The data is to be written to a sequential file in sorted order. Which collection method is more likely to yield optimal performance without violating the business requirements?

- A. Sort Merge on customer ID
- B. Auto
- C. Round Robin
- D. Ordered

Answer: A