

1.What is a control task?

A control task is used to alter the normal processing of a workflow by stopping, aborting, or failing a workflow or worklet.

2.What is a pipeline partition and how does provide a session with higher performance?

Within a mapping, a session can break apart different source qualifier to target pipelines into their own reader/transformation/writer thread(s). This allows the Integration Service to run the partition in parallel with other pipeline partitions in the same mapping. The parallelism creates a higher performing session.

3.What is the maximum number of partitions that can be defined for in a single pipeline?

You can define up to 64 partitions at any partition point in a pipeline.

4.Pipeline partitions is designed to increase performance, however list one of it's disadvantages?

Increasing the number of partitions increases the load on the node. If the node does not contain enough CPU bandwidth, you can overload the system.

5.What is a dynamic session partition?

A dynamic session partition is where the Integration Service scales the number of session partitions at runtime. The number of partitions is based on a number of factors including number of nodes in a grid or source database partitions.

6.List three dynamic partitioning configurations that cause a session to run with one partition

1. You set dynamic partitioning to the number of nodes in the grid, and the session does not run on a grid.
2. You create a user-defined SQL statement or a user-defined source filter.

3. You use dynamic partitioning with an Application Source Qualifier.

7.What is pushdown optimization?

Pushdown optimization is a feature within Informatica PowerCenter that allows us to push the transformation logic in a mapping into SQL queries that are executed by the database. If not all mapping transformation logic can be translated into SQL, then the Integration Service will process what is left.

8.List the different types of pushdown optimization that can be configured?

1. Source-side pushdown optimization – The Integration Service pushes as much transformation logic as possible to the source database.

2. Target-side pushdown optimization – The Integration Service pushes as much transformation logic as possible to the target database.

3. Full pushdown optimization – The Integration Service attempts to push all transformation logic to the target database. If the Integration Service cannot push all transformation logic to the database, it performs both source-side and target-side pushdown optimization.

Question #9

What databases are we able to configure pushdown optimization?

IBM DB2

Microsoft SQL Server

Netezza

Oracle

Sybase ASE

Teradata

Databases that use ODBC drivers

Question #10

List several transformations that work with pushdown optimization to push logic to the database?

Aggregator, Expression, Filter, Joiner, Lookup, Router, Sequence Generator, Sorter, Source Qualifier, Target, Union, Update Strategy

Question #11

What is real-time processing?

Data sources such as JMS, WebSphere MQ, TIBCO, webMethods, MSMQ, SQP, and webservices can publish data in real-time. These real-time sources can be leveraged by Informatica PowerCenter to process data on-demand. A session can be specifically configured for real-time processing.

Question #12

What types of real-time data can be processed with Informatica PowerCenter

1. Messages and message queues. Examples include WebSphere MQ, JMS, MSMQ, SAP, TIBCO, and webMethods sources.
2. Web service messages. Example includes receiving a message from a web service client through the Web Services Hub.
3. Change data from PowerExchange change data capture sources.

13.What is a real-time processing terminating condition?

A real-time processing terminating condition determines when the Integration Service stops reading messages from a real-time source and ends the session.

14.List three real-time processing terminating conditions?

1. Idle time – Time Integration Service waits to receive messages before it stops reading from the source.

2. Message count – Number of messages the Integration Service reads from a real-time source before it stops reading from the source.

3. Reader time limit – Amount of time in seconds that the Integration Service reads source messages from the real-time source before it stops reading from the source

15. What is real-time processing message recovery?

Real-time processing message recovery allows the Integration Service to recover unprocessed messages from a failed session. Recovery files, tables, queues, or topics are used to recover the source messages or IDs. Recovery mode can be used to recover these unprocessed messages.

16. What factors play a part in determining a commit point?

1. Commit interval
2. Commit interval type
3. Size of the buffer blocks

17. List all configurable commit types?

1. Target-based commit – Data committed based on the number of target rows and the key constraints on the target table.
2. Source-based commit – Data committed based on the number of source rows.
3. User-defined commit – Data committed based on transactions defined in the mapping properties.

18. What performance concerns should you be aware of when logging error rows?

Session performance may decrease when logging row errors because the Integration Service processes one row at a time instead of a block of rows at once.

19.What functionality is provided by the Integration Service when error logging is enabled?

Error logging builds a cumulative set of error records in a error log file or error table created by the Integration Service.

20. What is the difference between stopping and aborting a workflow session task?

A **stop command** tells the Integration Service to stop reading session data, but will continue writing and committing data to targets.

A **abort command** works exactly like the stop command, however it will tell the Integration to stop processing and committing data to targets after 60 seconds. If all processes are not complete after this time out period, the session gets terminated.

21.What is a concurrent workflow?

A concurrent workflow is a workflow that can run as multiple instances concurrently. Concurrent workflows can be configured in one of two ways:

1. Allow concurrent workflows with the same instance name.
2. Configure unique workflow instances to run concurrently.

22.What is Informatica PowerCenter grid processing and its benefits?

Grid processing is a feature of PowerCenter that enables workflows and sessions to be run across multiple domain nodes. PowerCenter grid's parallel processing provides increased performance and scalability.

23.List the types of parameters and variables that can be defined within a parameter file?

Service variables, service process variables, workflow and worklet variables, session parameters, and mapping parameters and variables.

24. With PowerCenter, what two locations can one specify a parameter file?

1. Within the session task. 2. Within the workflow.

25. How are maplet parameters and variables defined within a parameter file different?

Maplet parameters and variables are different because they must be preceded with the maplet name they were defined within. For example, a parameter by name of MyParameter, defined within maplet MyMaplet, would be set to a value of 10 in a related parameter file by using syntax:
MyMaplet.MyParameter=10.

26. What is an SQL Transformation in Informatica?

The SQL transformation is active, passive, and connected. It allows for runtime SQL processing. It allows data to be retrieved, inserted, updated, and deleted midstream in a mapping pipeline. SQL transformations have two modes, script mode and query mode. Script mode allows for external located script files to be called to execute SQL. Query mode allows for SQL to be placed within the transformation's editor to execute SQL logic.

Question #27

What is dynamic lookup cache?

Dynamic lookup cache is a cache that has been built from the first lookup request. Each subsequent row that passes through the lookup will query the cache. As these rows are processed or inserted into the lookup's target table, the lookup cache is also updated dynamically.

Question #28

What is a Unstructured Data transformation?

The Unstructured Data transformation is active, passive, and connected. It leverages the Data Transformation application to transform unstructured, semi-structured, and structured file formats such as messaging formats, HTML pages, PDF documents, ACORD, HIPAA, HL7, EDI-X12, EDIFACT, and SWIFT. Once data has been transformed by Data Transformation, it can be returned in a mapping pipeline and further transformed and/or loaded to an appropriate target.