

IBM InfoSphere DataStage Essentials v11.7

OEM: IBM • Duration: 4 Days (32 hrs) • Code: KM304G

COURSE MODULES & TOPICS

Objectives

- Describe the uses of DataStage, DataStage clients, and the DataStage workflow
- Describe the two types of parallelism exhibited by DataStage parallel jobs
- Describe what a deployment domain consists of, the different domain deployment options, and the installation process
- Create new users and groups
- Assign Suite roles and Component roles to users and groups
- Give users DataStage credentials
- Add a DataStage user on the Permissions tab and specify their role
- Specify DataStage global and project defaults
- List and describe important environment variables
- Navigate the DataStage Designer
- Import and export DataStage objects
- Design a parallel job in DataStage Designer
- Use the Row Generator, Peek, and Annotation stages in the job
- Compile, run, and monitor a job
- Create a parameter set and use it in a job
- Read and write to sequential files using the Sequential File stage
- Work with nulls in sequential files
- Read from multiple sequential files using file patterns
- Describe parallel processing architecture, pipeline parallelism, and partition parallelism
- Describe partitioning and collecting algorithms
- Describe the parallel job compilation process and how to use OSH (Orchestrate Shell Script)
- Explain the Score
- Combine data using the Lookup stage
- Combine data using the Merge, Join, and Funnel stages
- Sort data using in-stage sorts and the Sort stage
- Combine data using the Aggregator stage and the Remove Duplicates stage
- Use the Transformer stage in parallel jobs
- Define constraints and derivations
- Create a parameter set and use its parameters in constraints and derivations
- Perform a simple Find, Advanced Find, and an impact analysis

- Compare the differences between two table definitions and two jobs
- Import table definitions for relational tables
- Use ODBC and Db2 Connector stages in a job
- Use SQL Builder to define SQL SELECT and INSERT statements
- Use multiple input links into Connector stages to update multiple tables within a single transaction
- Use the DataStage job sequencer to build a job that controls a sequence of jobs
- Use Sequencer links and stages to control the sequence a set of jobs run in
- Pass information in job parameters from the master controlling job to the controlled jobs
- Handle errors and exceptions