



Object Arena Software Solutions
hr@objectarena.com



IBM WebSphere DataStage v8.0.1 Training/Workshop

Object Arena Career Development Centre offers the unique opportunity to aspiring fresher's and experienced professionals to get real time experience in **ETL – IBM DataStage**. We also cover the training for the IBM Certified Solution Developer – WebSphere IIS DataStage certification exam.

Course Description

With this training the professional will get critical knowledge and skills, to design and develop an efficient and scalable DataStage solution using minimum number of components. Also the professional can control data partitioning and provide minimum data skew to a complex enterprise level business problem.

The expert should be able to tune a parallel application to determine where bottlenecks exist and how to eliminate them. The trained professional can work with complex data import and export scenarios and understand the designing of parallel solutions without relying on automated insertion of partitioning and sort operations. The candidate clearly understands what an inserted buffer operation is, can enable environment variables to collect performance statistics and can determine the appropriate response to the collected data.

Support

- ✓ Placement assistance through our Placement Cell.
- ✓ Trained by IBM certified professionals with real time experience.
- ✓ IBM Certification assistance.
- ✓ Resume and Interview preparation assistance.

Highlights

- Very reasonable charges and value for your money.
- Industry standard course materials.
- Hands on Lab with real-time scenarios.
- Accommodation assistance for outstation students.

Duration: 40 Hours

Customizable: Yes

Pre-requisites: Students should have strong knowledge in the following topics

- RDBMS
- Data Warehousing

IBM WebSphere DataStage v8.0.1 Training/Workshop

Curriculum

Module: 01

Introduction

- DataStage Architecture.
- DataStage Clients
 - Designer
 - Director
 - Administrator

DataStage Workflow

Module: 02

Types of DataStage Job

- Parallel Jobs
- Server Jobs
- Job Sequences

Setting up DataStage Environment

- DataStage Administrator Properties
- Defining Environment Variables
- Importing Table Definitions

Module: 03

Creating Parallel Jobs

- Design a simple Parallel job in Designer
- Compile your job
- Run your job in Director
- View the job log
- Command Line Interface (dsjob)

Module: 04

Accessing Sequential Data

- Sequential File stage
- Data Set stage
- Complex Flat File stage
- Create jobs that read from and write to sequential files
- Read from multiple files using file patterns
- Use multiple readers
- Null handling in Sequential File Stage

Module: 05

Platform Architecture

- Describe parallel processing architecture Describe pipeline & partition parallelism
- List and describe partitioning and collecting algorithms
- Describe configuration files
- Explain OSH & Score

Module: 06

Combining Data

- Combine data using the Lookup stage
- Combine data using merge stage
- Combine data using the Join stage
- Combine data using the Funnel stage

Module: 07

Sorting and Aggregating Data

- Sort data using in-stage sorts and Sort stage
- Combine data using Aggregator stage
- Remove Duplicates stage

Module: 08

Transforming Data

- Understand ways DataStage allows you to transform data
- Create column derivations using user-defined code and system functions
- Filter records based on business criteria
- Control data flow based on data conditions

Module: 09

Repository Functions

- Perform a simple Find
- Perform an Advanced Find
- Perform an impact analysis
- Compare the differences between two Table Definitions and Jobs.

Module: 10

Working with Relational Data

- Import Table Definitions for relational tables.
- Create Data Connections.
- Use Connector stages in a job.
- Use SQL Builder to define SQL Select statements.
- Use SQL Builder to define SQL Insert and Update statements.
- Use the DB2 Enterprise stage.

Module: 11

Metadata in Parallel Framework:

- Explain schemas.
- Create schemas.
- Explain Runtime Column Propagation (RCP).
- Build a job that reads data from a sequential file using a schema.
- Build a shared container.

Module: 12

Job Control:

- 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.
- Use Sequencer triggers and stages to control the conditions under which jobs run.
- Pass information in job parameters from the master controlling job to the controlled jobs.
- Define user variables.
- Enable restart.
- Handle errors and exceptions.