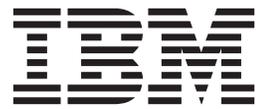IBM InfoSphere DataStage and QualityStage
Version 8 Release 7

Director Client Guide

IBM

IBM InfoSphere DataStage and QualityStage
Version 8 Release 7

Director Client Guide

IBM

Contents

Chapter 1. The InfoSphere DataStage and QualityStage Director client

The Director is the client component that validates, runs, schedules, and monitors jobs on the engine tier.

The Director client is the starting point for most of the tasks an IBM® InfoSphere® DataStage® operator needs to do in respect of jobs.

Note: InfoSphere DataStage mainframe jobs run on a mainframe computer, and use mainframe-specific tools. These jobs are not visible in the Director client. In this manual the term *job* therefore refers to server, sequence, and parallel jobs only. For information about running mainframe jobs, consult the documentation supplied with your mainframe software.

These topics describe the interface to the Director client and how to use it, including:

- Starting the Director
- Using the Director window
- Finding text in the Director window, sorting data, and printing out the display
- Setting options and defaults for the Director window and for jobs you want to run
- Switching between projects and exiting the Director

Starting the Director client

Start the IBM InfoSphere DataStage and QualityStage® Director client.

Procedure

1. Choose **Start** > **Programs** > **IBM InfoSphere Information Server** > **IBM InfoSphere DataStage and QualityStage Director** The Attach to Project dialog box appears.
2. Enter the name of the computer that is hosting the services tier in the **Host name of the services tier** field.
3. Enter your user name in the **User name** field. This is your user name on the application server system.
4. Enter your password in the **Password** field.
5. Enter the name of the project you want to use or choose one from the **Project** list, which displays all the projects installed on the engine tier.
6. Click **OK**. The Director window appears.

   Note: You can also start the Director from the Designer by clicking **Tools** > **Run Director**. You are automatically attached to the same project and you do not see the Attach to Project dialog box.

The Director client window

The Director client window opens when you start the Director.

Repository pane

The left pane of the Director client window is the repository pane.

The repository pane displays the repository tree, which lists folders and sub-folders that contain parallel, sequence, and server jobs. The jobs in the currently selected folder are listed in the display area. You can hide the repository pane by choosing **View** > **Show Folders**.

Display area

The display area is the main part of the Director window.

There are three views:

- **Job Status.** The default view, which appears in the right pane of the Director window. It displays the status of all jobs in the folder currently selected in the repository tree. If you hide the repository pane, the Job Status view includes a Folder column which shows the folder name, and displays the status of all jobs in the current project, regardless of what folder they are in. See "Job Status view" on page 4 for more information.
- **Job Schedule.** Displays a summary of scheduled jobs and batches in the currently selected folder. If the repository pane is hidden, the display area shows all scheduled jobs and batches, regardless of what folder they are in. See Chapter 5, "Designing job batches," on page 47 for a description of this view. To switch to the Job Schedule view, choose **View** > **Schedule**, or click the **Schedule** button on the toolbar.
- **Job Log.** Displays the log file for a job chosen from the Job Status view or the Job Schedule view. The repository pane is always hidden. See Chapter 4, "The job log," on page 39 for more details. To switch to this view, choose **View** > **Log**, or click the **Log** button on the toolbar.

Updating the display area

You can set how often the display area is updated from the repository by editing the Director client options.

About this task

For more information, see "Director Options".

You can also update the screen immediately by doing one of the following:

- Choose **View** > **Refresh**.
- Choose **Refresh** from the shortcut menus (see Shortcut Menus for more details).
- Press **Ctrl-R**.

Each entry in the display area represents a job, scheduled job, or event in the job log, depending on the current view. An icon is displayed by default for each entry. To hide the icons, choose **Tools** > **Options** ▸ **View**.

Note: You can improve refresh performance by organizing jobs within folders, so that you do not display more jobs than necessary in the display area. For information about how to organize jobs within folders, refer to *IBM InfoSphere DataStage Designer Client Guide*.

Menu bar

The menu bar has six pull-down menus that give access to all the functions of the Director.

- **Project.** Opens an alternative project and sets up printing.
- **View.** Displays or hides the toolbar, status bar, buttons, or repository pane, specifies the sorting order, changes the view, filters entries, shows further details of entries, and refreshes the screen.
- **Search.** Starts a text search dialog box.
- **Job.** Validates, runs, schedules, stops, and resets jobs, purges old entries from the job log file, deletes unwanted jobs, cleans up job resources (if the administrator has enabled this option), and allows you to set default job parameter values. You can also start the resource estimation tool.
- **Tools.** Monitors running jobs, manages job batches, and starts the Designer client, and allows you to manage data sets.
- **Help.** Invokes the Help system. You can also get help from any screen or dialog box in the Director.

Toolbar

The toolbar gives quick access to the main functions of the Director.

The toolbar is displayed by default, but can be hidden by choosing **View** > **Toolbar** or by changing the Director options. See Director Options for more details. To display ToolTips, let the cursor rest on a button in the toolbar.

From left to right the buttons are:
- Open project
- Print view
- Job status
- Job schedule
- Job log
- Find
- Sort - ascending
- Sort - descending
- Resource estimation
- Run a job
- Stop a job
- Reset a job
- Schedule a job
- Reschedule a job
- Help

Status bar

The status bar is displayed at the bottom of the Director window

The status bar displays the following information:
- The name of a job (if you are displaying the Job Log view).

- The number of entries in the display. If you look at the Job Status or Job Schedule view and use the `Filter Entries...` command, this panel specifies the number of lines that meet the filter criteria. If you have set a filter then (filtered) or (limited) is displayed.
- The date and time on the engine.

   Note: Under certain circumstances, the number of entries in the display is replaced by the last error message issued by the engine. The message disappears when the screen is refreshed.

Job Status view

The Job Status view is the default view displayed when you start the Director.

You can change the default view by editing the Director options (see Director Options for more details). You can also filter the jobs displayed in the view (see Filtering the Job Status or Job Schedule View).

The Job Status view shows the status of all the jobs in the currently selected folder, or, if the repository pane is hidden, in the current project. The view has the following columns:

Table 1. Job status view

Column	Description
Job name	The name of the job.
Status	The status of the job. See "Job States" for the possible job states and what they mean.
Started on date	The time and date a job was last started.
Last ran on date	The time and date the job was finished, stopped, or aborted. These columns are blank for jobs that have never been run.
Description	A description of the job, if available. This is the short description provided in the job design.

If you hover your mouse pointer over a job icon, you can view a tooltip that shows a sketch of the job. The sketch shows the job as it appears in the Designer client.

Job states

The Status column in the Job Status view displays the current status of the job.

The possible job states are given in the following table:

Table 2. Job States

Job State	Description
Job State	Description
Compiled	The job has been compiled but has not been validated or run since compilation.
Not compiled	The job is under development and has not been compiled successfully.
Running	The job is currently being run, reset, or validated.

Table 2. Job States (continued)

Job State	Description
Finished	The job has finished.
Finished (see log)	The job has finished but warning messages were generated or rows were rejected. View the log file for more details.
Stopped	The job was stopped by the operator.
Aborted	The job finished prematurely.
Validated OK	The job has been validated with no errors.
Validated (see log)	The job has been validated but warning messages were generated or rows were rejected. View the log file for more details.
Failed validation	The job has been validated, but an error was found.
Has been reset	The job has been successfully reset .

Job status details

About this task

To view more details about a job's status, select it in the display and do one of the following:

- Choose **View** > **Detail**.
- Right-click to display the shortcut menu and choose **Detail**.
- Double-click the job in the display.

The Job Status Detail dialog box appears. This dialog box contains details of the selected job's status as shown in the following table:

Table 3. Job status

This field...	Contains this information...
Server	The name of the computer that is hosting the engine tier..
Project	The name of the project to which the job belongs.
Status	The current status of the chosen job.
Job name	The name of the job
Invocation	You can run multiple versions of the same job on the Director. The Invocation number identifies the version of the job that you are viewing details of.
Wave Number	An internal number used by IBM InfoSphere DataStage when the job is run.
Status	The current status of the job
Description	A description of the job. This is the short description entered by the developer when the job was created. If the job has job parameters, this column also displays the values used in the run.

Use **Copy** to copy the whole window or selected text to the Clipboard for use elsewhere.

Click **Next** or **Previous** to display status details for the next or previous job in the list.

Click **Close** to close the dialog box.

Shortcut menus

The Director client has shortcut menus that are displayed when you right-click in the display area or repository pane.

The menu you see depends on the view or window you are using, and what is highlighted in the window when you click the mouse.

Shortcut menus in the Job Status view

This Job menu appears when you right-click on a selected job in the Job Status view.

About this task

If you right-click in the display area when the cursor is not over a job, a subset of this menu appears. From the Job menu you can:
* Add the job to the schedule (**Add to Schedule...**)
* Start a Monitor window (available only if an entry is selected) (**Monitor**)
* Set job parameter defaults (**Set Defaults...**)
* Display the Job Schedule or Job Log view (**View Schedule** and **View Log**)
* Use **Find** > **to search for text in the display area (Find...** and **Find Next)**
* Filter (limit) the jobs listed in the display area (**Filter...**)
* Refresh the display (**Refresh**)
* Display details of a log entry (available only if an entry is selected) (**Detail**)
* Delete the selected job (**Delete**)

Shortcut menus in the Job Log view

This menu appears when you right-click on an entry in the Job Log view.

About this task

If you right-click in the display area when the cursor is not over a log entry, a subset of this menu appears. From the full menu you can:
* Start a Monitor window (available only if a log entry is selected) (**Monitor**)
* Set job parameter defaults (**Set Defaults...**)
* Display the Job Status or Job Schedule view (**View Schedule** and **View Status**)
* Use **Find** > **to search for text in the display area (Find...** and **Find Next)**
* Filter (limit) log entries (**Filter...**)
* Refresh the display (**Refresh**)
* Display details of a log entry (available only if a log entry is selected) (**Detail**)
* Jump to the batch log from a job within a batch (**Related Log**)
* Delete an entry (**Delete**)

- Add rule to message handler

Shortcut menus in the Job Schedule view

This Job menu appears when you right-click on a job in the Job Schedule view.

About this task

If you right-click in the display area when the cursor is not over a job, a subset of this menu appears. From the Job menu you can:

- Schedule, reschedule, or unschedule the selected job (**Add to Schedule...**, **Reschedule**, **Unschedule**)
- Set job parameter defaults (**Set Defaults...**)
- Display the Job Status or Job Log view (**View Schedule** and **View Log**)
- Use **Find** > **to search for text in the display area (Find...** and **Find Next)**
- Filter (limit) the jobs listed in the display area (**Filter...**)
- Refresh the display (**Refresh**)
- Display details of an entry in the Job Schedule view (available only if an entry is selected) (**Detail**)
- Delete an entry (**Delete**)

Shortcut menus in the repository pane

This Job menu appears when you right-click on a job in the repository pane view.

About this task

When you right-click in the repository pane, this menu appears, from which you can:

- Filter (limit) the jobs listed in the display area (**Filter...**)
- Refresh the repository pane (**Refresh**)

Shortcut menu in the Monitor window

This menu appears when you right-click in the Monitor window.

About this task

From this menu you can:

- Refresh the display (**Refresh**)
- Display details of a selected stage (**Detail**)
- Show link information in the Monitor window (**Show links**)
- Show CPU usage in the Monitor window (**Show %CP**)
- Clean up job resources (**Cleanup Resources**)
- Save or clear the Monitor window settings (**Clear window settings**)
- Display information about the IBM InfoSphere DataStage release (**About...**)

Filtering the Job Status or Job Schedule view

By default, the Job Status view or Job Schedule view displays information about all the jobs in the currently selected folder.

About this task

If you hide the repository pane, the view displays information about all the jobs in a project. On a large project, the display might include more information than you need. If you want to focus on specific types of job, based either on their name or their status, you can filter the view so that it displays only those jobs.

Procedure

1. If the view you want to filter is not already displayed, choose **View** > **Status** > **or** > **View** > **Schedule** > **, as appropriate.**

2. Start the **Filter** facility by doing one of the following:
   - Choose **View** > **Filter Entries...** from the menu bar.
   - Choose **Filter...** from the shortcut menu.
   - Press **Ctrl-T.**

   The Filter Jobs dialog box appears.

3. Choose which jobs to include in the view by clicking either the **All jobs** or **Jobs matching** option button in the **Include** area.

   If you select **Jobs matching**, enter a string in the **Jobs matching** field. Only jobs that match this string will be displayed. The string can include wildcard, character lists, and character ranges.

   Wildcard/Pattern
   > **Description**

   ? Matches any single character.

   * Matches zero or more characters.

   # Matches a single digit.

   [*charlist***]**
   > Matches any single character in *charlist*.

   [*!charlist***]**
   > Matches any single character not in *charlist*.

   [a-z] Matches any single character in the range a-z.

4. Choose which jobs to exclude from the view by clicking either the **No jobs** or **Jobs matching** option button in the **Include** area. If you select **Jobs matching**, enter a string in the **Jobs matching** field. Only jobs that match this string will be excluded. The string definition is the same as in step 3.

5. Specify the status of the jobs you want to display by clicking an option button in the **Job status** area.
   - **All** lists jobs that have any status.
   - **All, except "Not compiled"** lists jobs with any status except Not compiled.
   - **Terminated normally** lists jobs with a status of Finished, Validated, Compiled, or Has been reset.
   - **Terminated abnormally** lists jobs with a status of Aborted, Stopped, Failed validation, Finished (see log), or Validated (see log).

6. Click **OK** to activate the filter. The updated view displays the jobs that meet the filter criteria. The status bar indicates that the entries have been filtered.

Examples of filtering by job name

The following examples show how to use the **Jobs Matching** field in the Filter Jobs dialog box to filter the Job Status view or the Job Schedule view.

Example 1

Table 4. Example 1

Job Names	"Include" Filter	Job View
job2input	job2*	job2input
job2output		job2output
job3input		
job3output		

Example 2

Continuing Example 1, if you also specify ***input** as an "Exclude" filter, the Job Status view shows only job2output.

Example 3

Table 5. Example 3

Job Names	"Include" Filter	Job View
A3tires	[A-E]3*	A3tires
A3valves		A3valves
B3tires		B3tires
B3valves		B3valves
F3tires		
F3valves		

Finding text

If there are many entries in the display area, you can use **Find** to search for a particular job or event.

About this task

You start **Find** in one of three ways:
- Choose **Search** > **Find...** .
- Choose **Find...** from the shortcut menu.
- Click the **Find** button on the toolbar.

The Find dialog box appears.

The **Look in** field shows the currently selected folder. If the repository pane is hidden, and the display area lists jobs from all folders, the **Look in** field specifies All Folders. You cannot edit this field.

Procedure

1. Enter text in the **Find what** field. This could be a date, time, status, or the name of a job.

   Note: If the text entered matches any portion of the text in any column, this constitutes a match.

2. If the displayed entry must match the case of the text you entered, select the **Match Case** check box. The default setting is cleared.
3. Choose the search direction by clicking the **Up** or **Down** option button. The default setting is **Down**.
4. Click **Find Next**. The display columns are searched to find the entered text. The first occurrence of the text is highlighted in the display. The text can appear in any column or row of the display area.
5. Click **Find Next** again to search for the next occurrence of the text.
6. Click **Cancel** to close the Find dialog box.

Note: You can also use **Search** > **Find Next** to search for an entry in the display. If there is a search string in the Find dialog box, **Find Next** acts in the same way as the **Find Next** button on the Find dialog box. If there is no search string in the Find dialog box, this option displays the Find dialog box where you must enter a search string.

Sorting columns

You can organize the entries in the display area by sorting the columns in ascending or descending order.

About this task

The column currently being used for sorting is indicated by a symbol in the column title:
- > indicates the sort is in ascending order.
- < indicates the sort is in descending order.

To sort a column do any one of the following:
- Click the column title. This selects the column for sorting toggles between ascending and descending.
- Click the **Ascending** or **Descending** button on the toolbar.
- Choose **View** > **Ascending** or **View** > **Descending**.

If you choose a column that contains a date or a time, both the date and time columns are sorted together.

Printing the current view

You can print information from the current view. The content of the printout depends on the view you are using and the options you choose in the Print dialog box.

Procedure

1. Do one of the following to display the Print dialog box:
    - Choose **Project** > **Print...** .
    - Click the **Print** button on the toolbar.
      The Print dialog box contains the name of the printer to use. By default, this is the default Windows printer. For information on how to specify an alternative printer, see Changing the Printer Setup.
2. Choose the range of items to print by clicking an appropriate option button in the **Range** area:

- **All entries** prints all entries in the current view.
- **Current® page** prints all entries visible in the display area for the current view.
- **Current item** prints the selected item only.

3. Choose what to print by clicking an appropriate option button in the **Print what** area:
   - **Summary only** prints a summary for each item.
   - **Full details** prints detailed information for each item.

4. Specify the print quality to use by choosing an appropriate setting from the **Print quality** list:
   - **High** (the default setting)
   - **Medium**
   - **Low**
   - **Draft**

   Note: This setting is ignored if **Print to file** is selected.

5. Select the **Print to file** check box if you want to print to a file only.

6. Click **OK**. If you are printing to file, the Print to file dialog box appears. Enter the name of a text file to use. The default is *DSDirect.txt* in the client directory.

What Is in the Printout?

The content of the printout depends on the view you are using.

- For the Job Status view, the printout contains the current status for each job in the project, and the date and time the job was last run.
- For the Job Schedule view, the printout contains an entry for each scheduled job in the project specifying when the job is scheduled to run.
- For the Job Log view, the printout can include information about each event in the job log file. For more information about the job log file, see The Job Log File.

Changing the printer setup

Printouts from the Director are normally output to the default Windows printer. You can choose an alternative printer or specify other printer settings.

Procedure

1. Choose **Project ▸ Print Setup...** . The Print Setup dialog box appears. (This dialog box is also displayed if you click **Setup...** in the Print dialog box.)
2. Change any of the settings as required or choose a different printer.
3. Click **OK** to save the settings and close the dialog box.

DataStage Director options

When you start the Director, the current settings for the Director options determine what is displayed in the Director window.

The settings include:
- The Director window position and size
- The time interval between updates from the engine
- The number of data rows processed before a job is stopped (only applies to server jobs)

- The number of warnings that are permitted before a job is aborted
- Whether the toolbar, status bar, or icons are displayed
- The filter criteria
- The process priority level for the Director

You can modify the settings for the Director options by choosing **Tools** >
Options.... The Options dialog box appears.

The four pages in the dialog box are described in the following sections.

General page

From the General page you can:
- Enable/Disable automatic engine update
- Change the engine update interval
- Compare the client and engine tier host times
- Save window settings

The engine update interval
The **Refresh Interval (seconds)** is the time, in seconds, between updates from the
engine.

Click the arrow buttons to increase or decrease the value in the box. The default
setting is 4.

Note that if you choose a long refresh time, the status displayed in the Director
window might not represent what is happening on the engine. For example, if you
start a run, the job status might not update to Running until a whole refresh
interval has elapsed. Conversely, if you choose a refresh time that is too short, the
Director requests information from the engine at a rate that is too frequent and
unproductive. You must find a value between these extremes that meets your
update requirements.

If you have a large project, you might wish to disable automatic refresh altogether
by clearing the **Enabled** check box. You can refresh manually by using **View** >
Refresh.

Comparing client and server times
If you select the **Compare client and server system times** check box, when the
Director client first attaches to a project, it checks that the system times on the
client and the engine tier host are within five minutes of each other.

If they are not, a warning message appears. This check box is selected by default.

Saving window settings
The **Save on exit** area contains two check boxes that determine the settings that are
saved on exit from the Director client.
- If **Main window size and position** is selected, the Director is restarted at the
  same screen coordinates as when it exited.
- If **Filter settings** is selected, the current filter settings are saved and used when
  the Director is started.

Both check boxes are selected by default.

Limits page

The Limits page sets the maximum number of rows to process in a job run, and the maximum number of warning messages to allow before a job aborts.

The row limit applies to all server jobs in the current session. You can override the settings for an individual job when it is validated, run, or scheduled.

Setting the maximum number of rows

The option buttons set the maximum number of rows to be processed by the job.

Click **No limit** to process all rows or **Stop stages after *n* rows** to specify a number of rows.

Enter a value 1 thru 99999 or click the arrow buttons to increase or decrease the value. The default value is 1000.

Setting the maximum number of warnings

The option buttons set the maximum number of warning messages allowed before a job is aborted.

Select **No limit** to log all error messages or **Abort job after** to specify a number of warning messages. Enter a value 1 thru 99999 or click the arrow buttons to increase or decrease the value. The default value is 50.

The abort is requested when the warning limit is reached. The stages in the job will continue to process data until they actually receive the abort request, so extra warnings can be logged before the job terminates. It is even possible for the job to complete if the warning limit is reached just before the end of the job run.

View page

The options on the View page determine what is displayed in the Director client window.

The check boxes in the **Show** area are selected by default:

Toolbar
> Displays the toolbar.

Status bar
> Displays the status bar.

Date and time
> Displays the date and time (of the engine tier host) on the status bar.

Icons   Displays the icons in the views.

Specify the view to display when the Director is started, by clicking the appropriate option button:
- Status of jobs (the default setting)
- Schedule
- Log for last job

Priority page

When jobs are running, the performance of the Director client might be noticeably slower.

You can improve the performance by changing the priority of the Director process.

Setting the process priority

The option buttons in the **Set process priority** area allow you to change the priority of the Director process.

About this task

- Click **High always** to increase the priority of the Director, regardless of where the client and engine components are installed. Note that if you choose this option and the client and engine components are installed on different computers, you might not see an improvement in the performance of the Director.
- Click **High if project on local system** to increase the priority of the Director process if the client and engine components reside on the same machine. This is the default setting.
- Click **Normal always** if you do not want to change the priority setting for the Director process.

Displaying the current state

The **Current state** area displays the current priority setting.

About this task

Note: If you choose a high priority setting, it might take longer for a job to run. This is because processor cycles are directed toward monitoring jobs rather than running them.

Click **OK** to save the settings.

Choosing an alternative project

When you start the Director client, the project chosen in the Attach to Project dialog box is opened.

You can view the jobs in another IBM InfoSphere DataStage project without exiting the Director.

Viewing jobs in another project

You can view jobs in another project.

Procedure

1. Do one of the following to display the Open Project dialog box:
   - Choose **Project** > **Open...** .
   - Click the **Open Project** button on the toolbar.
2. Choose the project you want to open from the **Projects** list box. This list box contains all the IBM InfoSphere DataStage projects on the engine tier host specified in the **DataStage Server** list, which is the engine tier that you initially attached to.
3. Click **OK** to open the chosen project. The updated Director window displays the jobs in the new project.

Viewing jobs on a different engine tier

You can view jobs on a different engine tier.

Procedure

1. Do one of the following to display the Open Project dialog box:
   - Click **Project** > **Open...** .
   - Click the **Open Project** button on the toolbar.
2. Choose the computer that hosts the engine tier from the **DataStage server** list. The list contains all the engine tier hosts that are registered to the services tier. On a typical installation, only one engine tier is registered.
3. Choose the project you want to open from the **Projects** list box.
4. Click **OK** to open the chosen project. The updated Director window displays the jobs in the new project.

   Note: If you have Monitor windows open when you choose an alternative project, you are prompted to confirm that you want to change projects. If you click **Yes**, the Monitor windows are closed before the new project is opened. See Chapter 3, "Monitoring jobs," on page 33 for more details.

Exiting the IBM InfoSphere DataStage Director client

To exit the Director, choose **Project** > **Exit**.

About this task

Any open windows (for example, Monitor windows) are automatically closed on exit.

Chapter 2. Running InfoSphere DataStage jobs

When you run a job, the extraction, loading, and transforming tasks that were built into the job designs are actually implemented.

These topics describe how to run IBM InfoSphere DataStage jobs, including the following topics:
- Setting job options
- Validating jobs
- Starting, stopping, and resetting a job run
- Deleting jobs from a project
- Cleaning up the resources of jobs that have hung or aborted
- Creating multiple job invocations
- Setting tracing for parallel jobs
- Generating operational metadata
- Disabling message handlers for parallel jobs

These tasks are performed from the Job Status view in the Director window. To switch to this view, choose **View** > **Status**, or click the **Status** button on the toolbar.

Setting job options

About this task

Each time you validate, run, or schedule a job you can:
- Change the job parameters associated with the job, as appropriate, and set values for environment variables that have been defined as job parameters. (You can also set up default values for job parameters).
- Override any default limits for row processing (server jobs and job sequences only) and warning messages that are set for the job run.
- Assign Invocation IDs to create multiple job invocations. You can create as many invocations as you want. For more information on creating Multiple Job Invocations see Multiple Job Invocations.
- Set tracing options for server jobs and job sequences. For more information, see Setting Tracing Options.
- Override the project default setting for generating operational metadata.

You do this from the Job Run Options dialog box which appears automatically when you run, validate, or schedule a job.

Some parameters have default values assigned to them. You can use the default or enter another value. You can reinstate the default values by clicking **Set to Default** or **All to Default**. In the example screen, the default values are encrypted strings which are shown as asterisks. You can also set your own defaults for job parameters from the Director.

Some job parameters hold variable information such as dates or file names that you need to enter for each job run. You must enter appropriate values in all the fields before you can continue.

If the job's designer included help text for the job parameters, you can get help by selecting the parameter and clicking **Property Help**.

You can also use this dialog box to set values for environment variables that affect parallel job runs. When you design the job, you can add environment variables to the list of job parameters, this dialog box will then ask you to supply values for those variables for this run. Environment variables are identified by a $ sign. When setting a value for an environment variable, you can specify one of the following special values:

- $ENV the value for the environment variable is retrieved from the environment in which the job is run.
- $PROJDEF the current project default as defined in the Administrator client is used.
- $UNSET the environment variable is unset.

Note: The dialog box displays a Parameters page only if the job has parameters.

Validating a job

You can check that a job or job invocation will run successfully by validating it.

About this task

Jobs should be validated before running them for the first time, or after making any significant changes to job parameters.

When a server job is validated, the following checks are made without actually extracting, converting, or writing data:

- Connections are made to the data sources or data warehouse.
- SQL SELECT statements are prepared.
- Files are opened. Intermediate files in Hashed File, UniVerse, or ODBC stages that use the local data source are created, if they do not already exist.

When a parallel job is validated, the job is run in `check only' mode so data is not affected.

Procedure

1. Select the job or job invocation you want to validate in the Job Status view.
2. Choose **Job** > **Validate...** . The Job Run Options dialog box appears. See Setting Job Options.
3. Fill in the job parameters as required.
4. Click **Validate**. Click **OK** to acknowledge the message. The job is validated and the job's status is updated to **Running**.

   Note: It might take some time for the job status to be updated, depending on the load on the engine tier host and the refresh rate for the client.

Results

Once validation is complete, the updated job's status displays one of these status messages:

- Validated OK. You can now schedule or run the job.

- Failed validation. You need to view the job log file for details of where the validation failed. For more details, see "The Job Log File."

If you want to monitor the validation in progress, you can use a Monitor window. For more information, see "Monitoring Jobs."

Running a job

You can run a job immediately or you can schedule a job to run at a later time or date.

About this task

You can run a job in two ways:
- Immediately.
- By scheduling it to run at a later time or date. See Job Scheduling for how to do this.

If you run a job immediately, you must ensure that the data sources and data warehouse are accessible, and that other users on your system will not be affected by the job run.

Procedure

1. Select the job or job invocation in the Job Status view.
2. Do one of the following:
   - Choose **Job** > **Run Now...** .
   - Click the **Run** button on the toolbar.
     The Job Run Options dialog box appears. See Setting Job Options.
3. Fill in the job parameters and check warning and row limits for the job, as appropriate.
4. Optionally, click **Validate** to validate the job.
5. Click **Run**. The job is scheduled to run with the current date and time and the job's status is updated to **Running**.

   Note: It might take some time for the job status to be updated, depending on the load on the engine tier host and the refresh rate for the client.

Results

All types of job in your project (other than mainframe jobs) are displayed in the Director repository pane. This can include Web services jobs, which are distinguished by having a separate icon and being grayed out.

If you attempt to run an Web service enabled job from within the Director, you will be warned with a message asking whether you really want to run the job from the Director client (that is, not from the web services console).

Stopping a job

You can stop a job that is currently running.

Procedure

1. Select the job to be stopped in the Job Status view.
2. Do one of the following:
   - Click **Job** > **Stop**.
   - Click the **Stop** button on the toolbar.

     The job or invocation is stopped, regardless of the stage currently being processed, and the status of the job is updated to `Stopped`.

     Note: It might take some time for the status of the job to be updated, depending on the load on the engine tier host and the refresh rate for the client.

Resetting a job

If a job has stopped or aborted, it is difficult to determine whether all the required data was written to the target data tables. You can reset a job or job invocation.

About this task

When a job has a status of Stopped or Aborted, you must reset it before running the job again.

By resetting a job, you set it back to a runnable state and, optionally (on server jobs only), return your target files to the state they were in before the job was run.

Note: You can only reinstate server job sequential files and hashed files to a pre-run state if the backup option has been chosen on the corresponding stage in the job design.

If you want to undo the updates performed during a successful job, you can also use the **Reset** command for jobs with a status of Finished. The **Reset** command is not available for jobs with a status of Not compiled or Running.

Procedure

1. Select the job or invocation you want to reset in the Job Status view.
2. Choose **Job** > **Reset** or click the **Reset** button on the toolbar. A message box appears.
3. Click **Yes** to reset the tables. All the files in the job are reinstated to the state they were in before the job was run. The job's status is updated to Has been reset.

   Note: It might take some time for the job status to be updated, depending on the load on the engine tier host and the refresh rate for the client.

Restarting job sequences

About this task

If a sequence is restartable (tht is, is recording checkpoint information) and has one of its jobs fail during a run, then the following status appears in the Director:

`Aborted/restartable`

In this case you can take one of the following actions:

- **Run Job**. The sequence is re-executed, using the checkpoint information to ensure that only the required components are re-executed.
- **Reset Job**. All the checkpoint information is cleared, ensuring that the whole job sequence will be run when you next specify run job.

Note: If, during sequence execution, the flow diverts to an error handling stage, IBM InfoSphere DataStage does not checkpoint anything more. This is to ensure that stages in the error handling path will not be skipped if the job is restarted and another error is encountered.

Setting default job parameters

You can set default values for any parameters a job has.

About this task

These will override an defaults set in the job design (although note that, if you recompile the job, the defaults will be reset to the design ones, similarly if you upgrade IBM InfoSphere DataStage).

Procedure

1. Select the job in the display area.
2. Choose **Job** > **Set Defaults...**. The Set Job Parameter Defaults dialog box appears.
3. If defaults have been set in the Designer for this job, they will be displayed. Edit them to override them.

Job scheduling

About this task

You can schedule a job to run in a number of ways:
- Once today at a specified time
- Once tomorrow at a specified time
- On a specific day and at a particular time
- Daily at a particular time
- On the next occurrence of a particular date and time

Each job can be scheduled to run on any number of occasions, by using different job parameters if necessary. For example, you can schedule a job to run at different times on different days. The scheduled jobs are displayed in the Job Schedule view.

Note: Windows restricts job scheduling to administrators, therefore you need to be logged on as a Windows administrator in order to use the IBM InfoSphere DataStage scheduling features. Microsoft has published a workaround to this restriction - visit http://support.microsoft.com/directory/ and look up article Q124859 for details.

Job schedule view

The Job Schedule view displays details of all scheduled and unscheduled jobs and batches in the currently selected job folder.

If the repository pane is hidden, the Job Schedule view displays details of all scheduled and unscheduled jobs and batches in the project, regardless of the folder they are in.

To display the job schedule, choose **View** > **Schedule** or click the **Schedule** button on the toolbar.

You can filter the view to display specific types of job, based on their name or status (see Filtering the Job Status or Job Schedule View). The icon on the left of the **Job name** column indicates that a job is scheduled.

The **To be run** column shows when the job is scheduled to run, as shown in the following table:

To be run...
> **Means this...**

Every *n*
> *n* is a number representing the date. For example, *Every 12&27* means the job is scheduled to run on the 12th and 27th day of each month.

Every *x*
> *x* represents the day of the week:
>
> M = Monday
>
> T = Tuesday
>
> W = Wednesday
>
> Th = Thursday
>
> F = Friday
>
> S = Saturday
>
> Su = Sunday
>
> For example, *Every Th&F* means the job is scheduled to run every Thursday and Friday.

Every *n***&***x*
> *n* is a date and *x* is a day of the week (as above). For example, Every 10&Su means the job is scheduled to run on every 10th day of the month and every Sunday.

Today The job is run today at the specified time.

Tomorrow
> The job will run tomorrow at the specified time.

Next *n*
> *n* is a date (as above). For example, *Next 28* means the job is run on the next 28th of the month.

Next *x* *x* is a day of the week. For example, *Next W* means the job is run the next Wednesday in the month.

Next *n***&***x*
> *n* is a number and *x* is a day of the week. For example, Next 5&12&T means the job is scheduled to run on the next 5th and 12th day of the month, and the next Tuesday.

The **At time** column lists the time at which the job will run. This is displayed in the system's current time format: 12-hour or 24-hour clock.

The **Parameters/Description** column lists the parameters required to run the job. Each job has built-in job parameters which must be entered when you schedule or run a job. The entered values are displayed in this column in the following format:

parameter1 name = value, parameter2 name = value, ...

A brief description appears here if there is a short description defined and there are no job parameters.

Viewing the details of a job schedule
About this task

To view more details about a scheduled job or batch, select it in the display area and do one of the following:
- Choose **View** > **Detail**.
- Choose **Detail** from the shortcut menu.
- Double-click the job or batch in the display.

If you choose a batch, it has the same effect as choosing **Tools** > **Batch**, as described in Creating a Job Batch.

If you choose a job, the Job Schedule Detail dialog box appears. This dialog box contains a summary of the job details and all the settings used to schedule the job.

This field...
> **Contains this information...**

Project
> The name of the project and the computer that hosts the engine tier.

Schedule #
> The schedule number the job has been assigned.

Occurrences
> The number of times the job will be run using this schedule. A value of Repeats means that the job is continuously rescheduled.

Job name
> The name of the job.

Run time
> The time the job is set to run, in 24-hour format.

Run date
> The date the job is set to run.

Job parameters
> The job parameters. Each entry in this field is in the format *parameter name=value*.

Note: The parameter name displayed here is the name used internally by the job, not the descriptive parameter name you see when you enter job parameter values.

Use **Copy** to copy the schedule details and job parameters to the Clipboard for use elsewhere.

Click **Next** or **Previous** to display schedule details for the next or previous job in the list. These buttons are only active if the next or previous job is scheduled to run.

Click Close to close the window.

Scheduling a job

You can schedule a job or job invocation.

Procedure

1. Select the job or job invocation you want to schedule in the Job Status or Job Schedule view.

   Note: You cannot schedule a job with a status of Not compiled or a web service-enabled job.

2. Do one of the following to display the Add to schedule dialog box:
   - Choose **Job** > **Add to Schedule...** .
   - Choose **Add To Schedule...** from the appropriate shortcut menu.
   - Click the **Schedule** button on the toolbar.Choose when to run the job by clicking the appropriate option button:

     Today runs the job today at the specified time (in the future).

     Tomorrow runs the job tomorrow at the specified time.

     Every runs the job on the chosen day or date at the specified time in this month and repeats the run at the same date and time in the following months.

     Next runs the job on the next occurrence of the day or date at the specified time.

     Daily runs the job every day at the specified time.

3. If you selected **Every** or **Next** in step 3, choose the day to run the job by doing one of the following:
   - Choose an appropriate day or days from the **Day** list.
   - Choose a date from the calendar.

     Note: If you choose an invalid date, for example, 31 September, the behavior of the scheduler depends upon the operating system of the computer that hosts the engine tier, and you might not receive a warning of the invalid date. Refer to your documentation for the engine tier host for further information.

4. Choose the time to run the job. There are two time formats:
   - 12-hour clock. Click either **AM** or **PM**.
   - 24-hour clock. Click **24H Clock**.

     Click the arrow buttons to increase or decrease the hours and minutes, or enter the values directly.

5. Click **OK**. The Add to schedule dialog box closes and the Job Run Options dialog box appears.

6. Fill in the job parameter fields and check warning and row limits, as appropriate.

7. Click **Schedule**. The job is scheduled to run and is added to the Job Schedule view.

Unscheduling a job

If you want to prevent a job from running at the scheduled time, you must unschedule it.

Procedure

1. Select the job you want to unschedule in the Job Schedule view.
2. Do one of the following:
   - Choose **Job** > **Unschedule**.
   - Choose **Unschedule** from the Job shortcut menu.

Results

If the job is not scheduled to run at another time, the job status is updated to **Not scheduled** in the **To be run** column, and is not run again until you add it to the schedule.

Rescheduling a job

If you have a job scheduled to run, but you want to change the frequency, day, or time it is run, you can reschedule it.

Procedure

1. Select the job you want to reschedule in the Job Schedule view.
2. Do one of the following to display the Add to schedule dialog box:
   - Choose **Job** > **Reschedule...** .
   - Choose **Reschedule...** from the Job shortcut menu.
   - Click the **Reschedule** button on the toolbar.

   The current settings for the job are shown in the dialog box.
3. Edit the frequency, day, or time you want the job to run.
4. Click **OK**. The Add to schedule dialog box closes and the Job Run Options dialog box appears.
5. Fill in the job parameters and check warning and row limits as appropriate.
6. Click **Reschedule**. The job is rescheduled and the **To be run** column in the Job Schedule view is updated.

Deleting a job

You can remove unwanted or old versions of jobs from your project.

Procedure

1. Select the job or job invocation in the Job Status view. You can make multiple selections.
2. Choose **Job** > **Delete**. A message confirms that you want to delete the chosen job, or jobs.
3. Click **Yes** to delete the jobs. A message confirms they have been deleted.
4. Click **OK**. The job design and all the associated components used at run time are deleted, including the files and records used by the Job Log view and the Monitor window.
5. If you delete a job that is part of a batch, edit the batch to remove the deleted job to prevent the batch from failing. See Chapter 5, "Designing job batches," on page 47

Job administration

From the Administrator client, the administrator can enable job administration commands in the Director client that let you clean up the resources of a job that has hung or aborted.

These commands help you return the job to a state in which you can rerun it after the cause of the problem has been fixed. You should use them with care, and only after you have tried to reset the job and you are sure it has hung or aborted.

There are two job administration commands:
* Cleanup Resources
* Clear Status File

Cleaning up job resources
About this task

This facility only applies to server jobs. The `Cleanup Resources` command lets you:
* View and end job processes
* View and release the associated locks

To run this command, do one of the following:
* Choose **Job** > **Cleanup Resources** from the menu bar.
* Choose **Cleanup Resources** from the Monitor window shortcut menu.

The Job Resources dialog box appears, from which you can view and clean up the resources of the selected job.

In the **Processes** area:

This column...
        Displays this information...

PID #   The process identification number.

Context
        The process context. In a job with more than one active stage, a PID might be reused during a job run. In that case the context field might have entries for more than one active stage. The context will always be listed as Unavailable if the **Show All** option button is selected.

User Name
        The identity of the user whose job started the process.

Last Command Processed
        The command executed most recently by the process.

In the **Locks** area:

This column...
        Displays this information...

PID/User #
        The identification number of the process associated with the lock.

Lock Type
        The type of lock: File, Record, or Group.

Item Id

> The identity of the item (record) locked by the process. For a Group lock, this column is left blank.

Viewing processes and locks

The default is for the Job Resources dialog box to display all the processes and locks associated with the job currently selected in the Director client.

About this task

You can, however, filter the display by using the option buttons in the **Processes** and **Locks** areas of the Job Resources dialog box.

In the **Processes** area:
- **Show All** displays all current IBM InfoSphere DataStage processes.
- **Show by job** (the default) displays all processes for the selected job.

In the **Locks** area:
- **Show All** displays all the current locks.
- **Show by job** (the default) displays all locks for the selected job.
- **Show by process** displays all the locks associated with the process that you have selected in the **Processes** area in the Job Resources dialog box.

Ending job processes
Procedure

1. From the Job Resources dialog box, choose the range of processes to list by using the **Show All** or **Show by job** option buttons in the **Processes** area.
2. To end all the processes associated with a job, click **Logout All**. (This button is disabled if you have clicked the **Show All** option button.)

   To end a particular process, select the process in the **Processes** list box, then click **Logout**.
3. Wait for the processes to end (log out) and the display to update.

   You can refresh the display manually at any time by clicking **Refresh**.

   If this procedure fails to end a process that you believe is causing a job to hang, try the following steps:
4. Log out of all IBM InfoSphere DataStage clients.
5. Try to end the process by using the Windows Task Manager or kill the process in UNIX.
6. Stop and restart the InfoSphere DataStage Server Engine.
7. Reset the job from the Director (see Resetting a Job).

Results

If there is a problem with a job, you can also release locks (see the next section), or clear the job status file (see Clearing a Job Status File).

Releasing locks
Before you begin

On UNIX or Linux systems, you must be user dsadm or root in order to release locks.

Procedure

1. From the Job Resources dialog box, choose the range of locks to list by doing either of the following:
   - Click the **Show by job** option button in the **Locks** area.
   - Select a process in the **Processes** area, then click the **Show by process** option button.

     Note: You cannot release locks if you have clicked the **Show All** option button in the **Locks** area.
2. Click **Release All**. Each of the displayed locks is unlocked and the display updates automatically. (You cannot select or release individual locks.)

Results

You can refresh the display manually at any time by clicking **Refresh**.

Clearing a job status file

When you clear a job's status file you reset the status records associated with all stages in that job.

About this task

You should therefore use this option with great care, and only when you believe a job has hung or aborted.

Procedure

1. Try to reset the job (see Resetting a Job).
2. Ensure that all the job's processes have ended (see Ending Job Processes).

Results

To clear the job status file, choose **Job** > **Clear Status File** from the menu bar. The job status changes to Compiled and no evidence will remain that the job has ever run.

If there is a problem with a job, you can also release the locks (see Releasing Locks).

Multiple job invocations

You can create multiple invocations of a server job, parallel job, or job sequence, with each invocation starting with different parameters to process different data sets.

A job invocation can be invoked regardless of the state of other invocations which are processing different data sets.

If you run a job in the Director without giving an invocation Id then you cannot create any new invocations of that job until the job has finished. If you want to run several invocations of the same job at the same time, you must give an invocation Id for the first invocation.

The job designer should ensure that the job is suitable to have multiple invocations run. For example, an unsuitable job might have different invocations running concurrently and writing to the same table. An unsuitable job might also adversely affect job performance.

Parallel job invocations resulting from a decision to invoke multiple invocations of a job should not be confused with the several instances of the same job that you get when running a partitioned job across several processors. In the latter case the partitioning and collecting built-in to the job will handle the situation where several processes want to read or write to the same data source.

Creating multiple job invocations

If the designer of the job has enabled the feature, you can run several invocations of the same job.

Procedure

1. From the Job Status view, select the job, and choose **Job** > **Validate**. The Job Run Options window opens.
2. Enter an Id into the **Invocation Id** field. This ID will be suffixed to the job name to create the invocation. For example if the job name is `Exercise5' and you enter an Invocation Id of `Test1', then job will appear as `Execrcise5.Test1'.
3. Click **Validate**. The Job Status view now shows the new job invocation.`

Running a job invocation

Running an invocation is done in the same way as you would run other jobs.

Procedure

1. From the Job Status view select the invocation from the list and choose **Job** > **Run now**. The Job Run Option window opens.
2. Fill in any Parameters, Rows and Warning Limits as required. Click **Run**.

   Note: Another way to run an invocation is to choose the job from the list, enter the Invocation Id in the text field and then click **Run**.

Viewing the job log for an invocation

Viewing the job log for job invocations is similar to viewing other job logs.

About this task

The status view shows a list of jobs and job invocations. For example, for a job named myjob you might see the following items in the status view:

```
myjob
myjob.invk1
myjob.invk2
```

If you select myjob and switch to the log view, you will see the log for myjob and both of the invocations of myjob. A column named **Invocation** in the log identifies the invocation that generated each log entry.

If you select myjob.invk1 or myjob.invk2 and switch to the log view, you will see only the job log for that particular invocation of the job.

Setting tracing options

You can generate tracing information and performance statistics for server jobs.

About this task

The Tracing page is included in the Job Run Options dialog box to help analysts during troubleshooting.

The options on this page determine the amount of diagnostic information generated the next time a job is run. Diagnostic information is generated only for the active stages in a chosen job. To specify the job, highlight it in the Job Status or Job Schedule view before choosing **Tools** > **Options...** .

The **Stage names** list box contains the names of the active stages in the job in the format *jobname.stagename*. To set a trace level, highlight the stages in the **Stage names** list box, then select any of the following check boxes:

- **Report row data**. Records an entry for every data row read on input and written on output. This check box is cleared by default.
- **Property values**. Records an entry for every input and output opened and closed. This check box is cleared by default.
- **Subroutine calls**. Records an entry for every BASIC subroutine used. This check box is cleared by default.
- **Performance statistics**. When performance tracing is turned on a special log entry is generated immediately after the stage completion message. This contains performance statistics about the selected stages statistics in a tabular form. (For more information see *IBM InfoSphere DataStage Administrator Client Guide*)

When the job runs, a file is created for each active stage in the job. The files are named using the format *jobname.stagename.trace* and are stored in the &PH& subdirectory of your server engine installation directory. If you need to view the content of these files, contact your local Customer Support center for assistance.

Enabling operational metadata at the job level (parallel and server jobs)

If operational metadata is not enabled at the project level, you can enable it for individual jobs in the project by selecting **Generate operational metadata** on the **General** tab of the Job Run Options dialog box in the Director client.

If operational metadata is enabled at the project level, you can disable it for individual jobs by clearing the selection.

Disabling message handlers

When you run a parallel job you can disable any message handler that has been defined for that job.

About this task

For parallel jobs, you can define message handlers, which specify that particular types of informational or warning messages are suppressed from the log. Message handlers can be assigned to jobs at the project level (from the Administrator), the

Job Design level (from the Designer), or for individual job runs (from the Director). You can define a handler from the Director, or from the Designer. For more details see Message Handlers.

Message handlers are disabled on the General page of the **Job Run Options** dialog box. To disable message handling:

• Select **Disable project-level message handling** to disable the handler that has been defined in the Administrator to apply to all jobs in the project.

• Select **Disable compiled-in job-level message handling** to disable a local handler that has been compiled into this particular job.

The message handlers are only disabled for this particular job run.

Chapter 3. Monitoring jobs

Processing details that occur at each active stage of the job are displayed in a Monitor window.

These topics describe how to monitor running jobs by using the **Monitor** command. These details include:

- The name of the stages performing the processing
- The status of each stage
- The number of rows processed
- The time taken to complete each stage

For parallel jobs, where the job is partitioned and different instances are running on different processors, the monitor can provide information for each instance if you request this option.

The Monitor window

To display a Monitor window, select a job in the Job Status view and do one of the following:

- Choose **Tools** > **New Monitor**.
- Choose **Monitor** from the shortcut menu.

(You can also start the monitor from the **Tools** menu of the Designer).

The monitor opens, displaying a tree structure containing stages in a job and associated links. For server jobs active stages are shown (active stages are those that perform processing rather than ones reading or writing a data source). For parallel jobs, all stages are shown.

The window displays the following information when you select a stage in the tree:

Table 6. The Monitor window: stage selected

This column...	Contains this information...
Status	The status of each stage. The possible states are:
State	**Meaning**
Aborted	The processing terminated abnormally at this stage.
Finished	All data has been processed by the stage.
Ready	The stage is ready to process the data.
Running	Data is being processed by the stage.
Starting	The processing is starting.
Stopped	The processing was stopped manually at this stage.
Waiting	The stage is waiting to start processing.

Table 6. The Monitor window: stage selected (continued)

This column...	Contains this information...
Num Rows	The number of rows of data processed so far by each stage on its primary input.
Started at	The time the processing started on the engine.
Elapsed time	The elapsed time since processing started.
Rows/sec	The number of rows processed per second.
%CP	The percentage of CPU the stage is using (you can turn the display of this column on and off from the shortcut menu - see Showing CPU Usage).

If you are monitoring a parallel job, and have not chosen to view instance information (see below), the monitor displays information for Parallel jobs as follows:

- If a stage is running in parallel then x N is appended to the stage name, where N gives how many instances are running.
- If a stage is running in parallel then the Num Rows column shows the total number of rows processed by all instances. The Rows/sec is derived from this value and shows the total throughput of all instances.
- If a stage is running in parallel then the %CP might be more than 100 if there are multiple CPUs on the engine tier host. For example, on a machine with four CPUs, %CP could be as high as 400® where a stage is occupying 100% of each of the four processors, alternatively, if the stage is occupying only 25% of each processor the %CP would be 100%.

To monitor instances of parallel jobs individually, choose **Show Instances** from the shortcut menu. The monitor will then show each instance of a stage as a sub-branch under the `parent' stage, The monitor displays the information for all stage instances under the `parent' stage. Only relevant information is shown for each stage instance.

Table 7. The Monitor window: job instance selected

This column...	Contains this information...
Status	The status of each stage. The possible states are:
State	**Meaning**
Aborted	The processing terminated abnormally at this stage.
Finished	All data has been processed by the stage.
Ready	The stage is ready to process the data.
Running	Data is being processed by the stage.
Starting	The processing is starting.

Table 7. The Monitor window: job instance selected (continued)

This column...	Contains this information...
Stopped	The processing was stopped manually at this stage.
Waiting	The stage is waiting to start processing.
Num Rows	The number of rows of data processed so far by each stage on its primary input.
%CP	The percentage of CPU the instance is using (you can turn the display of this column on and off from the shortcut menu - see Showing CPU Usage).

If you select a link in the tree (either under a stage or a stage instance) the following information is shown:

Table 8. The Monitor window: link selected

This column...	Contains this information...
Link Type	The type of link as follows:
Type	**Meaning**
<<Pri	primary input link
<Ref	input link
>Out	output link
>Rej	output link for rejected rows
Num Rows	The number of rows of data processed so far by each stage on its primary input.
Rows/sec	The number of rows processed per second.

The status bar at the bottom of the window displays the name and status of the job, the name of the project and the engine tier host, and the current time on the engine tier host.

Monitor shortcut menu

The Monitor window shortcut menu allows you to perform the following tasks:

* **Refresh**. Refresh the monitor with the most recent information (this overrides the interval setting which normally decides how often the monitor is refreshed - see Setting the Engine Update Interval).
* **Detail**. Select this to display the Stage Status window - see The Stage Status Window.
* **Show Instances**. Select this to show details for each instance of a parallel job that has been partitioned between processors - see Monitoring instances of parallel jobs.
* **Show %CP**. Select this to display the %CP column in the monitor - see Showing CPU Usage.

Showing CPU usage

While a stage or stage instance is processing data rows, it can use a large portion of the CPU. You can use the Monitor window to check what percentage of the computer processor is being used by an active stage.

To show the percentage of CPU usage, choose **Show %CP** from the Monitor shortcut menu. The updated Monitor window includes the **%CP** column.

The number displayed in the **%CP** column updates at regular intervals. The **%CP** column is empty if a stage is not processing data or has finished. If several running stages share a process, only one of the stages displays a value in the **%CP** column; the other stages display the same value in parentheses.

To hide the **%CP** column, choose **Show %CP** from the shortcut menu again.

Setting the engine update interval

The Monitor window display is updated with new information from the engine at regular intervals.

You can set how often the updates occur by setting a time (in seconds) in the **Interval** field. Click the arrow buttons to increase or decrease the value, or enter the value directly. The default setting is 10, the minimum is 5, and the maximum is 65.

Switching between monitor windows

You can display a Monitor window before you start the job or while a job is running.

About this task

If you want to monitor more than one job at a time, you can display more than one Monitor window.

If you have several Monitor windows open, they can become hidden by other windows on the screen.

You can bring a Monitor window to the front of the screen.

Procedure

1. Choose **Tools** > **View Monitor**.
2. Select the job you want to view from the submenu.

The Stage Status window

As well as viewing the information in the Monitor window, you can obtain more information about each stage from the Stage Status window.

To display the Stage Status window, select a stage in the Monitor window and do one of the following:

- Choose **Detail** from the shortcut menu.
- Double-click the stage in the Monitor window.

The fields and parameters displayed in the Stage Status window depend on whether you have selected a stage, instance, or a link in the monitor window tree.

The following screen shows the Stage Status window for a selected stage.

If you are looking at an instance of a parallel job stage the stage name will be given as Stage.Instance, for example CTransformerSatge1xz.2. The fields in the window cannot be edited.

This field...
> **Displays this information...**

Project
> The name of the project and the computer that hosts the engine tier.

Job name
> The name of the job.

Stage The name of the stage (or stage and instance)

Status The status of the stage. This is one of the states described in The Monitor Window.

Started at
> The date and time (on the engine) the processing started at this stage.

Ended at
> The date and time (on the engine) the processing ended. This is set to 00:00:00 or is left blank if processing is still taking place.

Row count
> The number of rows processed.

Control
> An internal number assigned by IBM InfoSphere DataStage.

Wave #
> An internal number assigned by InfoSphere DataStage.

User The name or process number of the user who ran the job.

Retrieved
> The date and time (on the engine) the stage retrieved the data to process.

The same fields described earlier are displayed in this window except:
- **Stage** is replaced by **Stage.Link** (or **Stage.Instance.Link**). This field contains the name of the stage followed by the name of the link.
- **Control** is replaced by **Link type**. This field contains the type of link. There are four possible settings:

> **Primary**
>> A primary input link

> **Reference**
>> A reference input link

> **Output**
>> An output link

> **Reject** An output link marked as "Rejected Rows"

Chapter 4. The job log

The job log is updated when you validate, run, or reset a job. You can use the log to troubleshoot jobs that fail during validation, or that are aborted when they are run.

For parallel jobs, you can also define message handlers, which specify that particular types of informational or warning messages are suppressed from the log. Message handlers can be assigned to jobs at the project level (from the Administrator client), the Job Design level (from the Designer client), or for individual job runs (from the Director client). You can define a handler from the Director client, or from the Designer client. For more details see "Message handlers" on page 44.

You can make information from job logs available to other components in the IBM InfoSphere Information Server suite by writing logs to the operational repository. For more information, see http://publib.boulder.ibm.com/infocenter/iisinfsv/ v8r7/topic/com.ibm.swg.im.iis.ds.admin.doc/topics/Enabling_OR_logging.html.

Viewing job results

The log file for a job is displayed in the Job Log view.

About this task

To display the log, select the job or job invocation in the Job Status view, then choose **View** > **Log** or click the **View Log** toolbar button. The display is updated to list the job log.

Each log contains entries describing events that occurred during the last (or previous) runs of the job. The most recent run is shown in black, others are shown in blue.

Entries are written to the log when:
- A job or batch starts or finishes
- A stage starts or finishes
- Rejected rows are output (server jobs and some stage types in parallel jobs)
- Errors are generated

The columns in the Job Log view are described in this table:

Table 9. Job log columns

Column name	Description
Occurred	The time the event occurred.
On date	The date the event occurred.
Type	The type of event.

Table 9. Job log columns (continued)

Column name	Description
Event	A message describing the event. The first line of the message is displayed. If a message has an ellipsis (...) at the end, it contains more than 1 line. You can view the full message in the Event Detail window.
Invocation	If the log is for a job invocation, gives the invocation ID, otherwise is blank.

The types of events that can occur are described in this table:

Table 10. Types of events

Event type	Cause of the event
Info	Stages start and finish. No action required.
Warning	An error occurred. You might need to investigate the cause of the warning, as it might cause the job to abort.
Fatal	A fatal error occurred. You need to investigate the cause of this error.
Control	The job starts and finishes.
Reject	Rejected rows are output.
Reset	A job is reset.

The Event Detail window

The full message for an event can be viewed in the Event Detail window.

About this task

To display the Event Detail window, select an event in the Job Log view and do one of the following:

- Choose **View** > **Detail**.
- Choose **Detail** from the shortcut menu.
- Double-click the event line.
- Press **Ctrl-D**.

This window contains a summary of the job and event details. The fields are described in the following table:

Table 11. Event Detail window

This field...	Contains this information...
Project	The name of the project and the computer that hosts the engine tier.
Event #	The ID assigned to the event during the job run.

Table 11. Event Detail window (continued)

This field...	Contains this information...
Event type	The type of event. The event types are: • Info • Warning • Fatal • Control • Reject • Reset
Job name	The name of the job.
Invocation	If the log entry is for a job invocation, gives the invocation ID, otherwise is blank.
Timestamp	The date and time the event occurred.
User	The name of the user who reset, validated, or ran the job.
Message	The full event message. The text contains names of stages in the job which were processed during this event. If you are viewing an event with an event type of Fatal or Warning, this text describes the cause of the event. The message also displays data associated with the event, unless the Administrator has restricted the operator's view of the log, in which case the data is visible to users with Developer rights only. You can specify that the tab character is used to delimit columns when logging for certain stage types. Depending on the length of the preceding data, the tab character appears as anything between 1 to 8 spaces.

You can use **Copy** to copy the event details and message or selected text to the Clipboard for use elsewhere.

Click **Next** or **Previous** to display details for the next or previous event in the list.

Click **Close** to close the window.

If you are running a parallel job, and your administrator has enabled the **Generated OSH Visible** option in the Administrator, the Event Detail window contains the OSH that was run for the job. This appears after the 'Parallel Job Initiated' message (usually event 3). This facility is intended for expert users.

Viewing related logs

You can examine the log entries for a job that was run as part of a batch.

Procedure
1. Switch to the log entries for the batch itself by selecting a log entry.
2. Do one of the following:
    - Choose **View** > **Related Log**.

- Choose **Related Log** from the shortcut menu.

  You see the first log entry for the batch that the job belongs to.

Filtering the Job Log view

By default, the information displayed in the Job Log view includes all the events that have occurred for the chosen job, regardless of how many times the job has been run.

About this task

On an established system, the display might include more information than you need.

If you want specific information, for example, a list of all the events after a particular date, you can use the **Filter** facility.

To start the **Filter** facility, do one of the following:
- Choose **View** > **Filter Entries...** from the menu bar.
- Choose **Filter...** from the shortcut menu.
- Press **Ctrl-T**.

The Filter dialog box appears.

You can choose to display only those entries between particular dates and times. You can also further reduce the entries by displaying only those entries with a particular event type.

Procedure

1. Choose where to start the filter by clicking the appropriate option button:
   - **Oldest** displays all the entries from the oldest event in the log file.
   - **Start of last run** displays the entries from the start of the last run.
   - **Day** and **Time** displays all the entries from the specified date and time. Enter the date and time or click the arrow buttons. The format of the date depends on how your Windows system is set up, for example, $dd\backslash mm\backslash yyyy$ or $mm\backslash dd\backslash yyyy$. The time is always in 24-hour format.
2. Choose where to end the filter by clicking the appropriate option button:
   - **Newest** displays entries up to the most recent date and time.
   - **Day** and **Time** displays all the entries up to the specified date and time.
3. Choose what to display from the filtered selection by clicking the appropriate option button:
   - **Select all entries** displays all the entries between the chosen start and end points.
   - **Last** displays the last n number of entries, where n is a number. The default setting is 100. Click the arrow buttons to increase or decrease the value, or enter a value 1 through 9999.
4. Choose the event types you want to display by selecting the appropriate check boxes:
   - **Information**
   - **Warning**
   - **Fatal**

- **Reject**
- **Other**. All event types other than those listed above.

5. Click **OK** to activate the filter. The updated Job Log view displays the entries that meet the filter criteria. The status bar indicates that the entries have been filtered.

Results

The Job Log view uses the filter settings until you change them or redisplay the whole log file. To display all the entries, choose **View** > **Show** > **All**.

Purging log file entries

You must purge job log entries routinely so that the log files do not become too large.

There are three ways to purge log entries:
- Automatically for a whole project. This is set up in the Administrator client.
- Automatically for each job. You can override any project-wide purging policy and set the purging policy for individual jobs.
- Manually for individual jobs.

You purge log entries for a job from the Clear Log dialog box. To display this dialog box, select the job in the Job Status view, then choose **Job** > **Clear Log...** .

For job shows the name of the job. **Oldest entry** shows the date and time of the oldest entry in the log file. (The format of the date and time might look different on your screen depending on your Windows settings.)

Purging log entries immediately
About this task

Click **Immediate purge** to select the **Immediate** area. You set the purging criteria by clicking one of the following:
- **Clear all entries.** Deletes all entries in the log file.
- **Up to last run.** Deletes all entries except the last one.
- **Before date.** Deletes all entries before the specified date. Specify the date by clicking the arrow buttons or entering the values directly.

The log file is purged after you click **OK**.

Purging log entries automatically
About this task

Click **Auto-purge** to select the **Auto-purge** area. You set the purging criteria by clicking one of the following:
- **Up to previous (job runs).** Purges old log entries, leaving the specified number of recent job run entries in the file.
- **Older than (days).** Purges all log entries older than the specified number of days.

Specify the number of job run entries or days by clicking the arrow buttons or entering the value directly.

Click **OK** to close the dialog box. The log file is purged at the next job run, and at subsequent job runs, according to your settings.

The aim of auto-purge is to purge the log regularly so that the storage space allocated to it does not keep growing, but remains at a constant size, so the storage area is not completely cleared. To completely empty a log, use the manual purge facilities to clear the whole log, then set up auto-purge to keep it at a manageable size.

Note that, if you are running multiple invocations of the same job, all invocations share the same job log. This means that if you, for example, choose to purge entries for all but the last four runs, and the job is invoked 6 times, then the entries for only 4 of the invocations will be kept. When using invocations to run jobs it might be better to auto-purge based on the number of days entries to be kept.

Message handlers

When you run a parallel job, any error messages and warnings are written to the job log and can be viewed from the Director client.

You can choose to handle specified errors in a different way by creating one or more message handlers.

A message handler defines rules about how to handle messages generated when a parallel job is running. You can, for example, use one to specify that certain types of message should not be written to the log.

You can edit message handlers in the Designer or in the Director. The recommended way to create them is by using the **Add Rule to Message Handler** feature in the Director.

You can specify message handler use at different levels:
- **Project Level**. You define a project level message handler in the Administrator, and this applies to all parallel jobs within the specified project.
- **Job Level**. From the Designer you can specify that any existing handler should apply to a specific job. When you compile the job, the handler is included in the job executable as a local handler (and so can be exported to other systems if required).

You can also add rules to handlers when you run a job from the Director (regardless of whether it currently has a local handler included). This is useful, for example, where a job is generating a message for every row it is processing. You can suppress that particular message.

When the job runs it will look in the local handler (if one exists) for each message to see if any rules exist for that message type. If a particular message is not handled locally, it will look to the project-wide handler for rules. If there are none there, it writes the message to the job log.

Note that message handlers do not deal with fatal error messages, these will always be written to the job log.

Note: You cannot add message rules to jobs from an earlier release of IBM InfoSphere DataStage without first re-running those jobs.

Adding rules to message handlers

You can add rules to message handlers 'on the fly' from within the Director client.

About this task

Using this method, you can add rules to handlers that are local to the current job, to the project default handler, or to any previously-defined handler.

To add rules in this way, highlight the message that you want to add a rule about in the job log and choose **Add rule to message handler...** from the job log shortcut menu or from the **Job** menu on the menu bar. The Add rule to message handler dialog box appears.

Procedure

1. Choose an option to specify which handler you want to add the new rule to. Choose between the local runtime handler for the currently selected job, the project-level message handler, or a specific message handler. If you want to edit a specific message handler, select the handler from the **Message Handler** drop-down list. Choose **(New)** to create a new message handler.
2. Choose an **Action** from the drop down list. Choose from:
   - **Suppress from log**. The message is not written to the job's log as it runs.
   - **Promote to Warning**. Promote an informational message to a warning message.
   - **Demote to Informational**. Demote a warning message to become an informational one.

     The **Message ID**, **Message type** and **Example of message text** fields are all filled in from the log entry you have currently selected. You cannot edit these.
3. Click **Add Rule** to add the new message rule to the chosen handler.
4. Click **Previous** or **Next** to move through the messages in the job log and add further rules to the selected handler.

Results

If you click **Edit Handler...**, the **Edit Message Handlers** dialog box appears, enabling you to manage the message handlers.

Managing message handlers

You can view, edit, or delete message handlers from the Message Handler Manager.

About this task

You can also define new handlers if you are familiar with the message IDs (although note that IBM InfoSphere DataStage will not know whether such messages are warnings or informational). The preferred way of defining new handlers is by using the add rule to message handler feature.

To open the Message Handler Manager, choose **Tools** > **Message Handlers** (you can also open the manager from the Add rule to message handler dialog box). The

Edit Message Handlers dialog box appears.

To edit an existing handler:
Procedure

1. Choose an option to specify whether you want to edit the local runtime handler for the currently selected job, edit the project-level message handler, or edit a specific message handler. If you want to edit a specific message handler, select the handler from the drop-down list. The settings for whichever handler you have chosen to edit appear in the grid.

2. Edit the grid as required. You can:
   - Choose a new Action for a particular message. Select a new action from the drop-down list. Possible Actions are:

     Suppress from log. The message is not written to the job's log as it runs.

     Promote to Warning. Promote an informational message to a warning message.

     Demote to Informational. Demote a warning message to become an informational one.

   - Delete a message. Select the message in the grid, right-click and choose Delete Row from the shortcut menu.

   - Add a new message. Right-click in the grid and choose **Insert Row** from the shortcut menu. You can then type in details of the message you want to add to the handler.

     When you are done with your edits, click Save and choose Save Message Handler to save the handler with its current name or Save Message Handler As to save it as a new handler.

To delete a handler:
Procedure

1. Choose an option to specify whether you want to delete the local runtime handler for the currently selected job, delete the project-level message handler, or delete a specific message handler. If you want to delete a specific message handler, select the handler from the drop-down list. The settings for whichever handler you have chosen to edit appear in the grid.

2. Click the **Delete** button.

To define a new handler:
Procedure

1. Choose **Edit chosen message handler** and select **(New)** from the drop-down list. The grid clears, allowing new message rules to be added.

2. Type in the message ID for the first message you want to add to the handler.

3. Choose an **Action** from the drop-down list. Possible Actions are:
   - **Suppress from log**. The message is not written to the job's log as it runs.
   - **Promote to Warning**. Promote an informational message to a warning message.
   - **Demote to Informational**. Demote a warning message to become an informational one.

4. Repeat this process until you have added all the required messages.

5. When you have defined all the required messages for the handler, click Save and choose **Save Message Handler As** from the menu. A dialog box opens and prompts you for the name of the new handler.

6. Type in the name for your new message handler.

Chapter 5. Designing job batches

You can create job batches, which run a number of jobs in sequence.

These topics describe how to create and run job batches, including the following topics:

- What is a job batch
- Creating and running job batches
- Scheduling, unscheduling and rescheduling job batches
- Deleting batches from a project

Job batches are supported to provide continuity for existing DataStage users. If you want to create sequences of jobs, you should design sequence jobs in the Designer client rather than defining job batches in the Director client. Job sequences are more versatile and flexible than job batches.

What is a job batch?

A job batch is a group of jobs or separate invocations of the same job (with different job parameters) that you want to run in sequence.

IBM InfoSphere DataStage treats a batch as though it were a single job. If any job fails to complete successfully, the batch run stops. You can follow the progress of jobs within a batch by examining the log files of each job or of the batch itself. These contain messages about the progress of the batch, as well as the job. You can create, schedule, edit, or delete job batches.

Creating a job batch

You can create a job batch.

Procedure

1. Choose **Tools** > **Batch** > **New...** . The Create New Batch dialog box appears.
2. Select the repository folder in which you want to store the batch job. The path and folder name are shown in the Folder Path field.
3. Enter a name for the new batch in the Item name field.
4. Click **OK**. The Job Properties dialog box appears.
5. Select the jobs to add to the batch from the **Add Job** list on the Job control page (note that you cannot add an RTI-enabled job to a job batch). This list displays all the server and parallel jobs in the project. You are prompted to enter parameter values, and row and warning limits for each job in the batch. As each job is added to the batch, the control information is added to the **Job** control page.
6. Click the **General** tab. The General page appears at the front of the Job Properties dialog box. Optionally, enter a brief description of the batch in the **Short job description** field. This description appears in the **Parameters/Description** column in the Job Schedule view.
7. Optionally, enter a more detailed description of the batch in the **Full job description** field.

8. Click the **Parameters** tab. The **Parameters** page appears at the front of the Job Properties dialog box.

9. Define any parameters that you want to specify for the batch. For example, a user name and password to prompt for when the batch is run.

10. When you have defined the batch, click **OK**. The batch is compiled and appears in the Job Status view.

   Note: The Dependencies page allows you to specify the dependencies of the batch job. You only need to do this if you intend to export the batch job for deployment on another system by using the Designer. Information on dependencies and exporting jobs is in *IBM InfoSphere DataStage Designer Client Guide.*

   Note: When you create a batch job you are, in effect, creating a server job whose function is to run other jobs (these can be other server jobs or parallel job). The Job Properties dialog box give access to the same performance improvement facilities of ordinary server jobs. The Performance page allows you to improve the performance of the job by specifying the way the system divides jobs into processes. For a full explanation of this, see Chapter 2 of *InfoSphere DataStage Server Job Developer Guide*.

Running a job batch

You can run a job batch immediately or you can schedule a job batch to run at a later time or date.

About this task

You can run a job batch in the same way as a standard job:

- Immediately.
- By scheduling it to run at a later time or date. See Scheduling a Job Batch for how to do this.

If you run a batch immediately, you must ensure that the data sources and data warehouse are accessible, and that other users on your system will not be affected by the job run.

To run a batch immediately:

Procedure

1. Select the batch in the Job Status view.
2. Do one of the following:
   - Choose **Job** > **Run Now...** .
   - Click the **Run** button on the toolbar.
     The Job Run Options dialog box appears. See Setting Job Options.
3. Fill in the job parameters and check warning and row limits for the batch, as appropriate.
4. Optionally, click **Validate** to validate the job.
5. Click **Run**. The batch is started and its status is updated to **Running**.

   Note: It might take some time for the job status to be updated, depending on the load on the engine tier host and the refresh rate for the client.

Scheduling a job batch

You can schedule a job batch.

Procedure

1. Display the Job Schedule view. Job batches are displayed in the Job Schedule view in the format Batch::*batchname*. *batchname* is the name entered when the batch was created.
2. Select the batch in the list and do one of the following to display the Add to schedule dialog box:
   - Choose **Job** > **Add to Schedule...** .
   - Choose **Add To Schedule...** from the Job shortcut menu.
   - Click the **Schedule** button on the toolbar.
3. Choose when to run the batch by clicking the appropriate option button:

   Today runs the batch today at the specified time (in the future).

   Tomorrow runs the batch tomorrow at the specified time.

   Every runs the batch on the chosen day or date at the specified time in this month and repeats the run at the same date and time in the following months.

   Next runs the batch on the next occurrence of the day or date at the specified time.

   Daily runs the batch every day at the specified time.
4. If you selected **Every** or **Next** in step 3, choose the day to run the batch from the **Day** list or a date from the calendar.

   Note: If you choose an invalid date, for example, 31 September, the behavior of the scheduler depends upon the operating system of the computer hosting the engine tier, and you might not receive a warning of the invalid date. Refer to your computer documentation for further information.
5. In the **Time** area, select one option from **AM**, **PM, or 24H Clock**. Then click the arrow buttons to increase or decrease the hours and minutes, or enter the values directly.
6. Click **OK**. The Add to schedule dialog box closes and the Job Run Options dialog box appears.
7. Click **Schedule**. The batch is scheduled to run and is added to the Job Schedule view. The job parameters entered when the batch was created are used when the batch is run.

Unscheduling a job batch

If you want to prevent a batch from running at the scheduled time, you must unschedule it.

About this task

To unschedule a batch, select the batch in the Job Schedule view and do one of the following:

- Choose **Job** > **Unschedule**.
- Choose **Unschedule** from the Job shortcut menu.

The batch status is updated to **Not scheduled** in the **To be run** column and is not run again until you schedule it again.

Rescheduling a job batch

If you have a batch scheduled to run, but you want to change the frequency, day, or time it is run, you can reschedule it.

Procedure

1. Select the batch in the Job Schedule view and do one of the following:
   - Choose **Job** > **Reschedule...** .
   - Choose **Reschedule...** from the Job shortcut menu.
   - Click the **Reschedule** button on the toolbar.

   The Add to schedule dialog box appears with the current settings for the batch.
2. Edit the frequency, day, or time you want the batch to run.
3. Click **OK**. The Add to schedule dialog box closes, the batch is rescheduled, and the **To be run** column in the Job Schedule view is updated.

Job schedule errors

If you get an error when you try to schedule jobs, ask your system administrator to check that:

- The Windows Schedule service is running on the computer that hosts the engine tier.
- You belong to a Windows group that has permission to use the Schedule service.

Editing a job batch

To add or remove jobs from a job batch or change any job parameters, you must edit the job batch.

About this task

Do not edit a job batch while it is running.

Procedure

1. Select the batch and choose **Tools** > **Batch** > **Properties**. The Job Properties dialog box appears.
2. Add further jobs by selecting them in the **Add Job** list, or remove jobs from the batch by selecting them in the Job control page and by using the **Cut** button or the **Delete** key.
3. Click the **General** tab and edit the batch description, if required.
4. Click the **Parameters** page and edit the batch parameters, if required.
5. Click **OK** to save the edits.

Copying a job batch

You can copy a job batch.

Procedure

1. In the Job Status view, select the batch and choose **Tools** > **Batch** > **Save As...** . The **Save Batch As** dialog box appears.
2. Choose the repository folder where you want the copied batch job to be located.

3. Enter a new name in the **Item Name** field or edit the existing name. You must specify a new name even if you are saving the copy to a different folder.
4. Click **OK** to copy the batch.

Deleting a job batch

If you do not need a job batch, you can delete it.

Procedure

1. Select the batch, or batches, in the Job Status view.
2. Choose **Tools** > **Batch** > **Delete**. A message appears asking you to confirm the deletion.
3. Click **Yes** to delete the batch or batches.

Product accessibility

You can get information about the accessibility status of IBM products.

The IBM InfoSphere Information Server product modules and user interfaces are not fully accessible. The installation program installs the following product modules and components:

- IBM InfoSphere Business Glossary
- IBM InfoSphere Business Glossary Anywhere
- IBM InfoSphere DataStage
- IBM InfoSphere FastTrack
- IBM InfoSphere Information Analyzer
- IBM InfoSphere Information Services Director
- IBM InfoSphere Metadata Workbench
- IBM InfoSphere QualityStage

For information about the accessibility status of IBM products, see the IBM product accessibility information at http://www.ibm.com/able/product_accessibility/index.html.

Accessible documentation

Accessible documentation for InfoSphere Information Server products is provided in an information center. The information center presents the documentation in XHTML 1.0 format, which is viewable in most Web browsers. XHTML allows you to set display preferences in your browser. It also allows you to use screen readers and other assistive technologies to access the documentation.

IBM and accessibility

See the IBM Human Ability and Accessibility Center for more information about the commitment that IBM has to accessibility.

Accessing product documentation

Documentation is provided in a variety of locations and formats, including in help that is opened directly from the product client interfaces, in a suite-wide information center, and in PDF file books.

The information center is installed as a common service with IBM InfoSphere Information Server. The information center contains help for most of the product interfaces, as well as complete documentation for all the product modules in the suite. You can open the information center from the installed product or from a Web browser.

Accessing the information center

You can use the following methods to open the installed information center.

- Click the **Help** link in the upper right of the client interface.

  Note: From IBM InfoSphere FastTrack and IBM InfoSphere Information Server Manager, the main Help item opens a local help system. Choose **Help > Open Info Center** to open the full suite information center.

- Press the F1 key. The F1 key typically opens the topic that describes the current context of the client interface.

  Note: The F1 key does not work in Web clients.

- Use a Web browser to access the installed information center even when you are not logged in to the product. Enter the following address in a Web browser: http://host_name:port_number/infocenter/topic/ com.ibm.swg.im.iis.productization.iisinfsv.home.doc/ic-homepage.html. The host_name is the name of the services tier computer where the information center is installed, and port_number is the port number for InfoSphere Information Server. The default port number is 9080. For example, on a Microsoft® Windows® Server computer named iisdocs2, the Web address is in the following format: http://iisdocs2:9080/infocenter/topic/ com.ibm.swg.im.iis.productization.iisinfsv.nav.doc/dochome/ iisinfsrv_home.html.

A subset of the information center is also available on the IBM Web site and periodically refreshed at http://publib.boulder.ibm.com/infocenter/iisinfsv/v8r7/ index.jsp.

Obtaining PDF and hardcopy documentation

- A subset of the PDF file books are available through the InfoSphere Information Server software installer and the distribution media. The other PDF file books are available online and can be accessed from this support document: https://www.ibm.com/support/docview.wss?uid=swg27008803&wv=1.

- You can also order IBM publications in hardcopy format online or through your local IBM representative. To order publications online, go to the IBM Publications Center at http://www.ibm.com/e-business/linkweb/publications/ servlet/pbi.wss.

Providing feedback about the documentation

You can send your comments about documentation in the following ways:

- Online reader comment form: www.ibm.com/software/data/rcf/
- E-mail: comments@us.ibm.com

Links to non-IBM Web sites

This information center may provide links or references to non-IBM Web sites and resources.

IBM makes no representations, warranties, or other commitments whatsoever about any non-IBM Web sites or third-party resources (including any Lenovo Web site) that may be referenced, accessible from, or linked to any IBM site. A link to a non-IBM Web site does not mean that IBM endorses the content or use of such Web site or its owner. In addition, IBM is not a party to or responsible for any transactions you may enter into with third parties, even if you learn of such parties (or use a link to such parties) from an IBM site. Accordingly, you acknowledge and agree that IBM is not responsible for the availability of such external sites or resources, and is not responsible or liable for any content, services, products or other materials on or available from those sites or resources.

When you access a non-IBM Web site, even one that may contain the IBM-logo, please understand that it is independent from IBM, and that IBM does not control the content on that Web site. It is up to you to take precautions to protect yourself from viruses, worms, trojan horses, and other potentially destructive programs, and to protect your information as you deem appropriate.

Notices and trademarks

This information was developed for products and services offered in the U.S.A.

Notices

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785 U.S.A.

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web

sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003 U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to

IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. _enter the year or years_. All rights reserved.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at www.ibm.com/legal/copytrade.shtml.

The following terms are trademarks or registered trademarks of other companies:

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office

UNIX is a registered trademark of The Open Group in the United States and other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

The United States Postal Service owns the following trademarks: CASS, CASS Certified, DPV, LACS^Link, ZIP, ZIP + 4, ZIP Code, Post Office, Postal Service, USPS and United States Postal Service. IBM Corporation is a non-exclusive DPV and LACS^Link licensee of the United States Postal Service.

Other company, product or service names may be trademarks or service marks of others.

Contacting IBM

You can contact IBM for customer support, software services, product information, and general information. You also can provide feedback to IBM about products and documentation.

The following table lists resources for customer support, software services, training, and product and solutions information.

Table 12. IBM resources

Resource	Description and location
IBM Support Portal	You can customize support information by choosing the products and the topics that interest you at www.ibm.com/support/ entry/portal/Software/ Information_Management/ InfoSphere_Information_Server
Software services	You can find information about software, IT, and business consulting services, on the solutions site at www.ibm.com/ businesssolutions/
My IBM	You can manage links to IBM Web sites and information that meet your specific technical support needs by creating an account on the My IBM site at www.ibm.com/account/
Training and certification	You can learn about technical training and education services designed for individuals, companies, and public organizations to acquire, maintain, and optimize their IT skills at http://www.ibm.com/software/sw-training/
IBM representatives	You can contact an IBM representative to learn about solutions at www.ibm.com/connect/ibm/us/en/

Providing feedback

The following table describes how to provide feedback to IBM about products and product documentation.

Table 13. Providing feedback to IBM

Type of feedback	Action
Product feedback	You can provide general product feedback through the Consumability Survey at www.ibm.com/software/data/info/ consumability-survey

Table 13. Providing feedback to IBM (continued)

Type of feedback	Action
Documentation feedback	To comment on the information center, click the Feedback link on the top right side of any topic in the information center. You can also send comments about PDF file books, the information center, or any other documentation in the following ways: • Online reader comment form: www.ibm.com/software/data/rcf/ • E-mail: comments@us.ibm.com

Index

A
Active stages 33
Add to schedule dialog box 49
alternative printer 11
alternative project 14
Attach to Project dialog box 1

C
changing the printer setup 11
choosing an alternative printer 11
cleaning up job resources 26
clearing
    job log file 43
    job status file 28
columns, sorting 10
CPU usage, showing 36
Create New Batch dialog box 47
creating job batches 47
customer support
    contacting 63

D
DataStage Director
    exiting 15
    starting 1
DataStage Director options 11
    display settings 13
    Filter settings 12
    limits 13, 15
    Main window size and position 12
    Options dialog box 12
    priority 14
    Refresh setting 12
    save settings 12
DataStage Director window 2
    display area 4
    menu bar 2
    shortcut menus 6
    status bar 3
    toolbar 3
DataStage engine tier 1
default display options 11
deleting
    job batches 51
    jobs 25
dependencies, specifying 48
displaying
    Monitor window 33
    Stage Status window 36
    ToolTips 3

E
editing job batches 50
ending job processes 26
entering job parameters 17
Event Detail window 40

examples
filtering views 9
Job Log view 39
Job Schedule view 22
Job Status view 4
exiting DataStage Director 15

F
Filter dialog box 42
Filter facility 8, 42
Filter Jobs dialog box 8
filtering
    examples 9
    Job Schedule view 8
    Job Status view 8
Find dialog box 9
Find facility 9

H
Help system, starting 3
hiding
    icons 2
    repository pane 2

I
icons, hiding 2
instances, job 28

J
job administration 26
job batches 47
    copying 50
    creating 47
    deleting 51
    editing 50
    related log entries 41
    rescheduling 50
    running 48
    scheduling 48, 49
    unscheduling 49
job instances 28
job log file
    purging entries 43
Job Log view 2, 39
    Event Detail window 40
    filtering the display 42
    shortcut menus 6
job parameters 17
    setting defaults 21
job processes
    ending 26
    viewing 26
Job Properties Job control page 47
Job Resources dialog box 26
job resources, cleaning up 26

(continued)
job results, Viewing 39
Job Run Options dialog box 17
Job Schedule Detail window 23
Job Schedule view 2, 22
    filtering the display 8
    shortcut menus 7
Job Status Detail dialog box 5
job status file, clearing 28
Job Status view 2, 4
    filtering the display 8
    shortcut menus 6
jobs
    deleting 25
    rescheduling 25
    resetting 20
    running 19
    scheduling 21
    stopping 19
    unscheduling 25
    validating 18

L
legal notices 59
limiters
    row 13
    warning 13
locks
    releasing 26
    viewing 26

M
mainframe jobs 1
menus
    pull-down 3
    shortcut 6
message handlers 44
    adding rules to 45
    disabling for job runs 30
    managing 45
Monitor window
    % CPU 36
    displaying 33
    shortcut menu 7
    switching between 36
multiple job instances 28

N
non-IBM Web sites
    links to 57

O
options, DataStage Director 11

© Copyright IBM Corp. 1997, 2012

65

IBM ®

Printed in USA