

Informix DataStage

Operator's Guide

Version 3.6
October 1999
Part No. 000-6468

Published by Informix® Press

Informix Corporation
4100 Bohannon Drive
Menlo Park, CA 94025-1032

© 1999 Informix Corporation. All rights reserved. The following are trademarks of Informix Corporation or its affiliates, one or more of which may be registered in the United States or other jurisdictions:

Answers OnLine™; C-ISAM®; Client SDK™; DataBlade®; Data Director™; Decision Frontier™; Dynamic Scalable Architecture™; Dynamic Server™; Dynamic Server™, Developer Edition™; Dynamic Server™ with Advanced Decision Support Option™; Dynamic Server™ with Extended Parallel Option™; Dynamic Server™ with MetaCube®; Dynamic Server™ with Universal Data Option™; Dynamic Server™ with Web Integration Option™; Dynamic Server™, Workgroup Edition™; Dynamic Virtual Machine™; Enterprise Decision Server™; Formation™; Formation Architect™; Formation Flow Engine™; Frameworks for Business Intelligence™; Frameworks Technology™; Gold Mine Data Access®; i.Reach™; i.Sell™; Illustra®; Informix®; Informix® 4GL; Informix® COM Adapter™; Informix® Informed Decisions™; Informix® InquireSM; Informix® Internet Foundation.2000™; InformixLink®; Informix® Red Brick® Decision Server™; Informix Session Proxy™; Informix® Vista™; InfoShelf™; Interforum™; I-Spy™; Mediazation™; MetaCube®; NewEra™; Office Connect™; ON-Bar™; OnLine Dynamic Server™; OnLine/Secure Dynamic Server™; OpenCase®; Orca™; PaVER™; Red Brick® and Design; Red Brick® Data Mine™; Red Brick® Mine Builder™; Red Brick® Decisionscape™; Red Brick® Ready™; Red Brick Systems®; Regency Support®; Rely on Red BrickSM; RISQL®; Solution DesignSM; STARindex™; STARjoin™; SuperView®; TARGETindex™; TARGETjoin™; The Data Warehouse Company®; Universal Data Warehouse Blueprint™; Universal Database Components™; Universal Web Connect™; ViewPoint®; Visionary™; Web Integration Suite™. The Informix logo is registered with the United States Patent and Trademark Office. The DataBlade logo is registered with the United States Patent and Trademark Office.

GOVERNMENT LICENSE RIGHTS

Software and documentation acquired by or for the US Government are provided with rights as follows:

(1) if for civilian agency use, with rights as restricted by vendor's standard license, as prescribed in FAR 12.212;
(2) if for Dept. of Defense use, with rights as restricted by vendor's standard license, unless superseded by a negotiated vendor license, as prescribed in DFARS 227.7202. Any whole or partial reproduction of software or documentation marked with this legend must reproduce this legend.

Table of Contents

Preface

Organization of This Manual	vii
Documentation Conventions	viii
DataStage Documentation	viii

Chapter 1. Introducing DataStage

What Is a Data Warehouse?	1-1
Why Do I Need One?	1-1
What Does DataStage Do?	1-2
How Is DataStage Packaged?	1-2
DataStage Projects and Jobs	1-3

Chapter 2. DataStage Director

Starting the DataStage Director	2-1
The DataStage Director Window	2-3
Display Area	2-3
Menu Bar	2-4
Toolbar	2-5
Status Bar	2-5
Job Status View	2-6
Job States	2-6
Job Status Details	2-7
Shortcut Menus	2-8
Shortcut Menu in the Job Status View	2-8
Shortcut Menu in the Job Log View	2-9
Shortcut Menu in the Job Schedule View	2-9
Shortcut Menu in the Monitor Window	2-10

Filtering the Job Status or Job Schedule View	2-10
Examples of Filtering by Job Name	2-12
Finding Text	2-13
Sorting Columns	2-14
Printing the Current View	2-14
What Is in the Printout	2-15
Changing the Printer Setup	2-16
Director Options	2-16
General Page	2-17
Limits Page	2-18
View Page	2-19
Tracing Page	2-20
Priority Page	2-21
Choosing an Alternative Project	2-22
Viewing Jobs in Another Project	2-22
Viewing Jobs on a Different Server	2-23
Exiting the DataStage Director	2-23

Chapter 3. Running DataStage Jobs

Setting Job Options	3-1
Validating a Job	3-2
Running a Job	3-3
Stopping a Job	3-4
Resetting a Job	3-4
Job Scheduling	3-5
Job Schedule View	3-5
Viewing Details of a Job Schedule	3-7
Scheduling a Job	3-9
Unschedulering a Job	3-10
Reschedulering a Job	3-10
Deleting a Job	3-11
Job Administration	3-11
Cleaning Up Job Resources	3-12
Clearing a Job Status File	3-15

Chapter 4. Job Batches

What Is a Job Batch?	4-1
Creating a Job Batch	4-1
Running a Job Batch	4-3
Scheduling a Job Batch	4-4
Unscheduling a Job Batch	4-5
Rescheduling a Job Batch	4-5
Job Schedule Errors	4-6
Editing a Job Batch	4-6
Copying a Job Batch	4-6
Deleting a Job Batch	4-7

Chapter 5. Monitoring Jobs

The Monitor Window	5-1
Setting the Server Update Interval	5-2
Saving Monitor Window Settings	5-3
Showing Links	5-3
Showing CPU Usage	5-4
Viewing and Cleaning Up Job Resources	5-5
Switching Between Monitor Windows	5-5
The Stage Status Window	5-5

Chapter 6. The Job Log File

Job Log View	6-1
The Event Detail Window	6-3
Viewing Related Logs	6-4
Filtering the Job Log View	6-4
Purging Log File Entries	6-6
Purging Log Entries Immediately	6-7
Purging Log Entries Automatically	6-8

Glossary

Index

Preface

Informix DataStage is a powerful software suite that is used to develop and run DataStage jobs. A DataStage job can extract from different sources, and then cleanse, integrate, and transform the data according to your requirements. The clean data is ready to be imported into a data warehouse for analysis and processing by business information software.

This manual describes the DataStage Director, the DataStage component that is used to validate, schedule, run, and monitor DataStage jobs. To use this manual you should be familiar with the Windows 98 or Windows NT interface, but no other special skills or knowledge are required.

Organization of This Manual

This manual contains the following:

- Chapter 1 contains an overview of DataStage and its component parts.

- Chapter 2 describes DataStage Director and how to use it.

- Chapter 3 covers how to validate, run, delete, schedule, and administer DataStage jobs.

- Chapter 4 describes job batches.

- Chapter 5 describes how to monitor a running job.

- Chapter 6 describes the job log file and the Job Log view.

- The Glossary defines terms that have specific meaning in DataStage.

Documentation Conventions

This manual uses the following conventions:

Convention	Usage
Bold	In syntax, bold indicates commands, function names, keywords, and options that must be input exactly as shown. In text, bold indicates keys to press, function names, and menu selections.
<i>Italic</i>	In syntax, italic indicates information that you supply. In text, italic also indicates operating system commands and options, filenames, and pathnames.
Courier	Courier indicates examples of source code and system output.
Courier Bold	In examples, courier bold indicates characters that the user types or keys the user presses (for example, <Return>).
➤	A right arrow between menu options indicates you should choose each option in sequence. For example, “Choose File ➤ Exit ” means you should choose File from the menu bar, then choose Exit from the File pull-down menu.

DataStage Documentation

DataStage documentation includes the following:

DataStage Developer’s Guide: This guide describes the DataStage Manager and Designer, and how to create, design, and develop a DataStage application.

DataStage Administrator’s Guide: This guide describes DataStage setup, routine housekeeping, and administration.

DataStage Operator’s Guide: This guide describes the DataStage Director, and how to validate, schedule, run, and monitor DataStage applications.

These guides are also available online in PDF format. You can read them using the Adobe Acrobat Reader supplied with DataStage. See *DataStage Installation Instructions* in the DataStage CD jewel case for details on installing the manuals and the Adobe Acrobat Reader.

1

Introducing DataStage

Many organizations want to make better use of their data. But that data may be stored in different formats in different types of database. Some data sources may be dormant archives, others may be busy operational databases. Extracting and cleaning data from these varied sources has always been time-consuming and costly – until now. DataStage makes it simple to design and develop efficient applications that make data warehousing a reality where it was impossible before.

What Is a Data Warehouse?

A data warehouse is a central database containing copies of data from all the operational sources and archive systems in an organization. But the database does not have to be large. Instead of storing details of every transaction, order, or set of sales figures, the data warehouse stores totals, averages, area figures, and so on. This data is structured to make it easy to query and to generate reports.

Inside the data warehouse you can perform analyses that would be impractical on a working database. This means that anyone who needs access to the data gets all the information they want, and only the information they want. The data warehouse can be created or updated at any time, with minimum disruption to working systems.

Why Do I Need One?

Working databases are busy. It is hard to gain an accurate picture of the contents of the database at any time because it changes frequently. By transferring working data into a data warehouse, you can take snapshots of what is going on. Also, working databases contain dirty data – records in different formats, with key values missing or out of range, and so on. In a fast-moving working database it is difficult to trap mistakes or incomplete entries. Using DataStage, you can cleanse data before loading it into a data warehouse, ensuring that your business decisions are based only on valid information.

As well as a working database, you may have archive systems or incompatible data sources that you have inherited. These may be static, but inaccessible because their format is different from your working system. You can use DataStage to transform this data into compatible formats that can be stored in the data warehouse.

What Does DataStage Do?

DataStage comes in between your data and your data warehouse. DataStage jobs process the data to meet your needs, including:

- **Extraction** – DataStage takes data from indexed files, sequential files, networked databases, archives, and external data sources and stores it in the data warehouse.
- **Aggregation** – DataStage takes your working data, calculates totals and averages, then stores it in the data warehouse. This means you have to store much less data, which is quicker and easier to access.
- **Transformation** – DataStage converts inconsistent data into the required format and loads it into the data warehouse.

How Is DataStage Packaged?

DataStage is client/server software. The server holds the data while it is being processed. The client is the interface to DataStage that is used for designing and running jobs, or managing the data in the Repository. The client components include:

- **DataStage Designer**, for creating DataStage jobs, which are compiled into executable programs that are scheduled by the DataStage Director and run by the DataStage Server.
- **DataStage Director**, for running DataStage jobs.
- **DataStage Manager**, for viewing and editing the contents of the Repository.
- **DataStage Administrator**, for administering DataStage projects and conducting housekeeping on the server.

The client and server components installed depend on the edition of DataStage you have purchased. DataStage is packaged in two ways:

- **Developer's Edition**, used by developers to design, develop, and create executable DataStage jobs. The Developer's Edition contains the server and all the client components described earlier. The developer's role, DataStage Designer, and DataStage Manager are described in *DataStage Developer's Guide*.
- **Operator's Edition**, used by operators to validate, schedule, run, and monitor DataStage jobs that have been developed elsewhere. The Operator's Edition contains DataStage Director, DataStage Administrator, and DataStage Server components only.

DataStage Projects and Jobs

You always enter DataStage through a DataStage project. When you start a DataStage client you are prompted to attach to a project. Each project contains DataStage jobs and the components required to develop or run them.

DataStage jobs are made up of individual stages. A stage represents a data source or a process. For example, one stage may extract data from a data source, while another transforms it. The data required at each stage and how it is handled is specified in the job design.

When a job design is compiled, an executable job is created. When the job is run, the processing described in the job design is performed. Variable parameters such as file names, dates, and so on, can be specified when the job is run. DataStage jobs can be exported for use on other DataStage systems.

2

DataStage Director

The DataStage Director is the client component that runs, schedules, and monitors jobs run by the DataStage Server. For DataStage operators, the DataStage Director is the starting point for most tasks you need to accomplish.

This chapter describes the interface to the DataStage Director and how to use it, including:

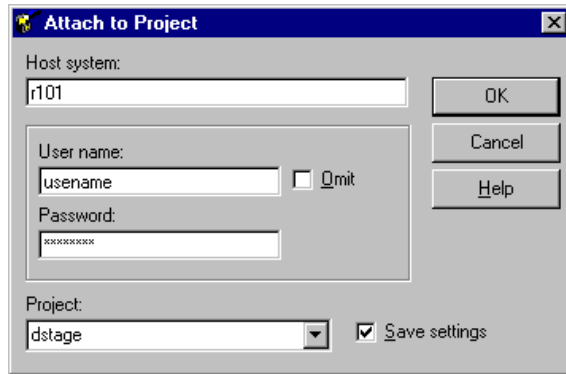
- Starting the DataStage Director
- A description of the DataStage Director window
- Finding text in the DataStage Director window, sorting data, and printing out the display
- Setting options and defaults for the DataStage Director window and for jobs you want to run
- Switching between projects and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

Starting the DataStage Director

To start the DataStage Director:

1. Choose **Start ► Programs ► Ardent DataStage ► DataStage Director**, or choose the appropriate program folder if you installed DataStage elsewhere.

The Attach to Project dialog box appears:



2. Enter the name of your host in the **Host system** field. This is the name of the system where the DataStage Server is installed.
3. Enter your user name in the **User name** field. This is your user name on the server system.
4. Enter your password in the **Password** field. If you are connecting to a Windows NT server via LAN Manager, you can select the **Omit** check box. The **User name** and **Password** fields gray out and you log on to the server using your current Windows account details.

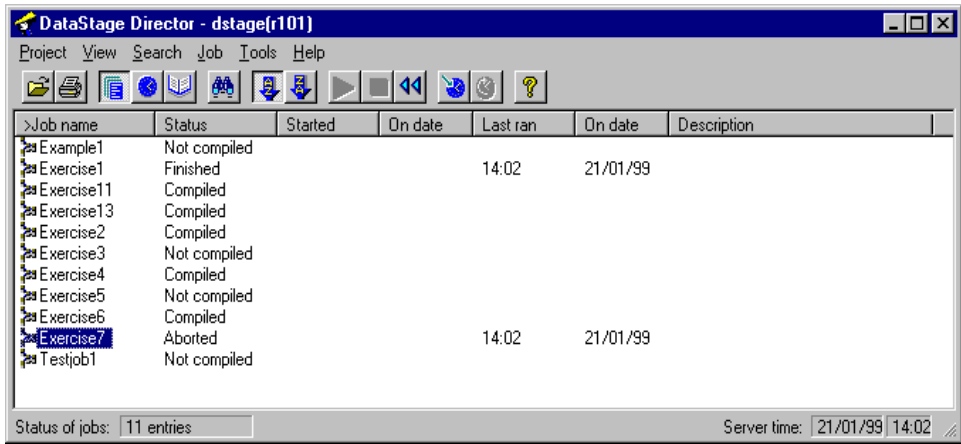
CAUTION: Think carefully before using the **Omit** option to log on to DataStage. If you use this option, note that:

- You cannot specify UNC pathnames in DataStage clients.
 - You must specify the host name in uppercase.
 - You cannot access remote files.
 - You may get errors if you import meta data from an ODBC data source if you are not logged on to the same domain as the server.
5. Enter the name of the project you want to use or choose one from the **Project** drop-down list, which displays all the projects installed on the server.
 6. Select the **Save settings** check box to save your logon settings. The next time you start the DataStage Director or any other DataStage client, you are presented with the saved settings (except for your password).
 7. Click **OK**. The DataStage Director window appears.

Note: You can also start the DataStage Director from the DataStage Designer or the DataStage Manager by choosing **Tools ► Run Director**. You are automatically attached to the same project and you do not see the Attach to Project dialog box. For more information about the DataStage Director, see *DataStage Developer's Guide*.

The DataStage Director Window

The DataStage Director window appears when you start the Director:



This section describes the features of the DataStage Director window including:

- The display area
- The menu bar
- The toolbar
- The status bar

Display Area

The display area is the main part of the DataStage Director window. There are three views:

- **Job Status.** The default view which displays the status of all jobs in the current project. See "Job Status View" on page 2-6 for more information.
- **Job Schedule.** Displays a summary of scheduled jobs and batches. See Chapter 3, "Running DataStage Jobs," for a description of this view. To switch to this view, choose **View ► Schedule**, or click the **Schedule** icon on the toolbar.

- **Job Log.** Displays the log file for a chosen job. See Chapter 6, “The Job Log File,” for more details. To switch to this view, choose **View ► Log**, or click the **Log** icon on the toolbar.

Updating the Display

You can set how often the display area is updated from the server by editing the Director options. For more information, see “Director Options” on page 2-16. 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” on page 2-8 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**.

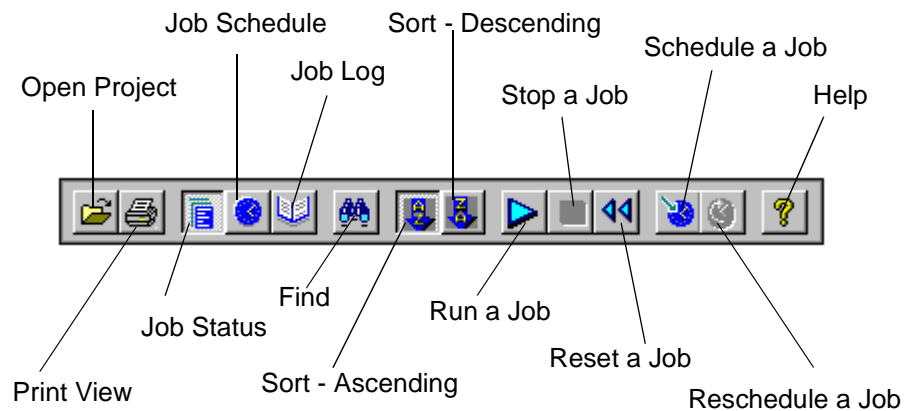
Menu Bar

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, or icons, 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.
- **Tools.** Monitors running jobs, manages job batches, and starts the DataStage Designer or the DataStage Manager (if these components are installed on the system) and other custom software.
- **Help.** Invokes the Help system. You can also get help from any screen or dialog box in the DataStage Director.

Toolbar

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



The toolbar is displayed by default, but can be hidden by choosing **View ► Toolbar** or by changing the Director options. See “Director Options” on page 2-16 for more details. To display ToolTips, let the cursor rest on an icon in the toolbar.

Status Bar

The status bar appears at the bottom of the DataStage Director window and 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...** option, this panel specifies the number of lines that meet the filter criteria.
- The date and time on the DataStage server.
- The number of lines in the display. If you look at the Job Log view and use the **Filter Entries...** option, this panel specifies the number of lines that meet the filter criteria.

Under certain circumstances, the number of lines in the display is replaced by the last error message issued by the server. The message disappears when the screen is refreshed.

Job Status View

The Job Status view is the default view displayed when you start the DataStage Director (see “The DataStage Director Window” on page 2-3). You can change the default view by editing the Director options (see “Director Options” on page 2-16 for more details). You can also filter the jobs displayed in the view (see “Filtering the Job Status or Job Schedule View” on page 2-10).

The Job Status view contains the status of all the jobs in the current project and has the following columns:

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 started. These fields are only filled in for a job with a status of Running.
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.

Job States

The Status column in the Job Status view displays the current status of the job. The possible job states are as follows:

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.
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.

Job State	Description
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 reset with no errors.

Job Status Details

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:

This field...	Contains this information...
Project	The name of the project and the DataStage server.
Status	The current status of the chosen job.
Wave #	An internal number used by DataStage when the job is run.
Job name	The name of the job.
Started at	The date and time the job was started. This field is used only for a job with a status of Running.

This field...	Contains this information...
---------------	------------------------------

Last run at	The date and time the job was last run.
-------------	---

Description	A description of the job. This is the 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.

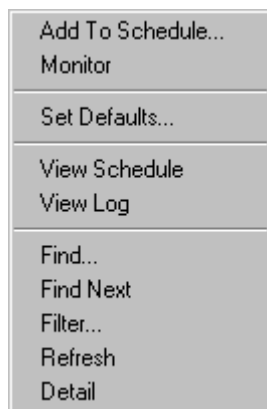
Click **Close** to close the dialog box.

Shortcut Menus

DataStage has shortcut menus that appear when you right-click in the display area. 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 Menu in the Job Status View

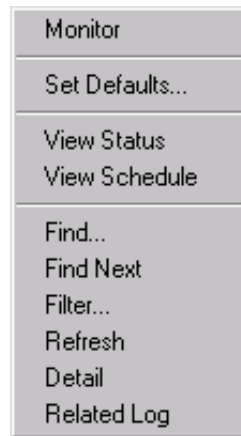
This menu appears when you right-click in the display area of the Job Status view. Some options are not available if no log entry is selected. From this menu you can:



- Add the job to the schedule
- Start a Monitor window (available only if an entry is selected)
- Set job parameter defaults
- Display the Job Schedule or Job Log view
- Use **Find** to search for text in the display area
- Filter (limit) the jobs listed in the display area
- Refresh the display
- Display details of a log entry (available only if an entry is selected)

Shortcut Menu in the Job Log View

This menu appears when you right-click in the display area of the Job Log view. Some options are not available if no log entry is selected. From this menu you can:



- Start a Monitor window (available only if a log entry is selected)
- Set job parameter defaults
- Display the Job Status or Job Schedule view
- Use **Find** to search for text in the display area
- Filter (limit) log entries
- Refresh the display
- Display details of a log entry (available only if a log entry is selected)
- Jump to the batch log from a job within a batch

Shortcut Menu in the Job Schedule View

This menu appears when you right-click in the display area of the job schedule view. Some options are not available if no entry is selected. From this menu you can:



- Schedule, reschedule, or unschedule the selected job
- Set job parameter defaults
- Display the Job Status or Job Log view
- Use **Find** to search for text in the display area
- Filter (limit) the jobs listed in the display area
- Refresh the display
- Display details of an entry in the Job Schedule view (available only if an entry is selected)

Shortcut Menu in the Monitor Window

This menu appears when you right-click in the Monitor window (described in Chapter 5). From this menu you can:



- Refresh the display
- Display details of a selected stage
- Show link information in the Monitor window
- Show CPU usage in the Monitor window
- Clean up job resources
- Save or clear the Monitor window settings
- Display information about the DataStage release

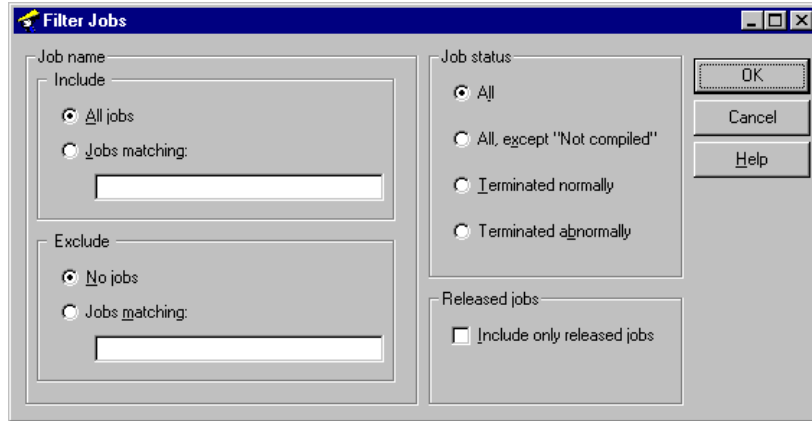
Filtering the Job Status or Job Schedule View

By default, the Job Status view or Job Schedule view displays information about all the jobs in a project. On a large project, the display may include more information than you need. If you want to focus on specific categories of job, based either on their name or their status, you can filter the view so that it displays only those jobs.

To filter the jobs in the Job Status view or the Job Schedule view:

1. If the view you want to filter is not already displayed, choose **View ► Status** or **View ► Schedule**, as appropriate.
2. Start the **Filter** option 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 wildcards, character lists, and character ranges.

Wildcard/Pattern	Description
?	Matches any single character.
*	Matches zero or more characters.
#	Matches a single digit.
[<i>charlist</i>]	Matches any single character in <i>charlist</i> .
[! <i>charlist</i>]	Matches any single character not in <i>charlist</i> .
[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 **Exclude** 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. If you want to restrict the display to released jobs, select the **Include only released jobs** check box in the **Released jobs** area.
 7. 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

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

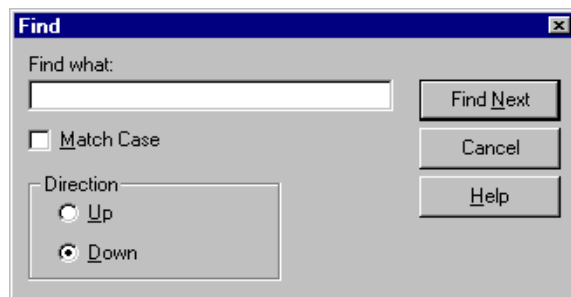
Job Names	"Include" Filter	Job View
A3tires	[A-E]3*	A3tires
A3valves		A3valves
B3tires		B3tires
B3valves		B3valves
F3tires		
F3valves		

Finding Text

If there are many entries in the display area, you can use **Find** to search for a particular job or event. You start **Find** in one of three ways:

- Choose **Search ► Find...**
- Choose **Find...** from the shortcut menu.
- Click the **Find** icon on the toolbar.

The Find dialog box appears:



To use **Find**:

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 unchecked.
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. 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:

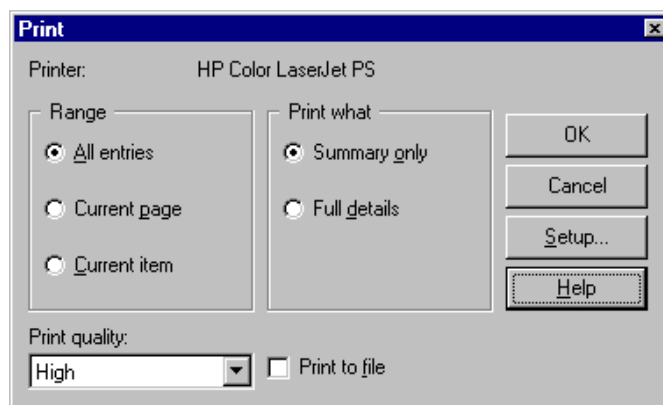
- Click the column title. The sort toggles between ascending and descending.
- Click the **Ascending** or **Descending** icon on the toolbar.
- Choose **View ► Ascending** or **View ► Descending**.

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. To print the current view:

1. Do one of the following to display the Print dialog box:
 - Choose **Project ► Print...**
 - Click the **Print** icon on the toolbar.



This 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” on page 2-16.

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** drop-down list box:
 - **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 *Ardent\DataStage* directory.

What Is in the Printout

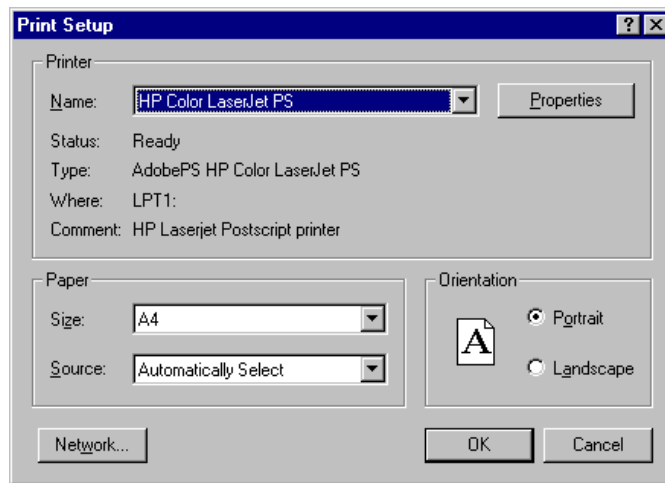
The content of the printout depends on the view you are using:

- 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 Chapter 6.

Changing the Printer Setup

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

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 to close the dialog box.

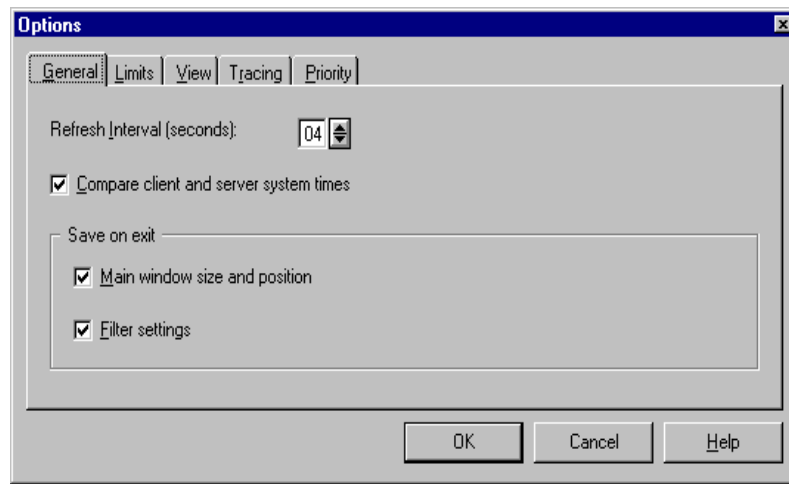
Director Options

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

- The DataStage Director window position and size
- The time interval between updates from the server
- The number of data rows processed before a job is stopped
- The number of unexpected warnings that are permitted before a job is aborted
- Whether the toolbar, status bar, or icons are displayed
- The filter criteria

- Tracing levels for a chosen job
- The process priority level for the DataStage Director

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



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

General Page

From the **General** page you can:

- Change the server update interval
- Save window settings
- Compare the client and server times

Changing the Server Update Interval

The **Refresh** area has a box labelled **Every *n* seconds**. This is the time, in seconds, between updates from the server. Click the arrow buttons to increase or decrease the value in the box. The default setting is 4, the minimum is 2, and the maximum is 65.

Note that if you choose a long refresh time, the status displayed in the DataStage Director window may not represent what is happening on the server. For example, if you start a run, the job status may not update to Running until a whole refresh interval has elapsed. Conversely, if you choose a refresh time that is too short, the

DataStage Director requests information from the server at a rate that is too frequent and unproductive. You must find a value between these extremes that meets your update requirements.

Saving Window Settings

The **Save on exit** area contains two check boxes that determine the settings that are saved on exit from the DataStage Director.

- If **Main window size and position** is selected, the DataStage 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 DataStage Director is started.

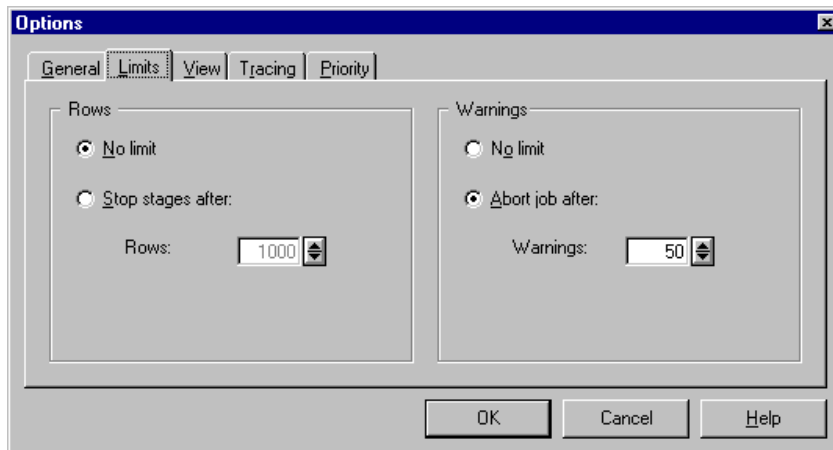
Both check boxes are selected by default.

Comparing Client and Server Times

If you select the **Compare client and server system times** check box, when the Director first attaches to a project, it checks that the system times on the client and the server are within five minutes of each other. If they are not, a warning message appears. This check box is 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 is aborted. The limits apply to all 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.

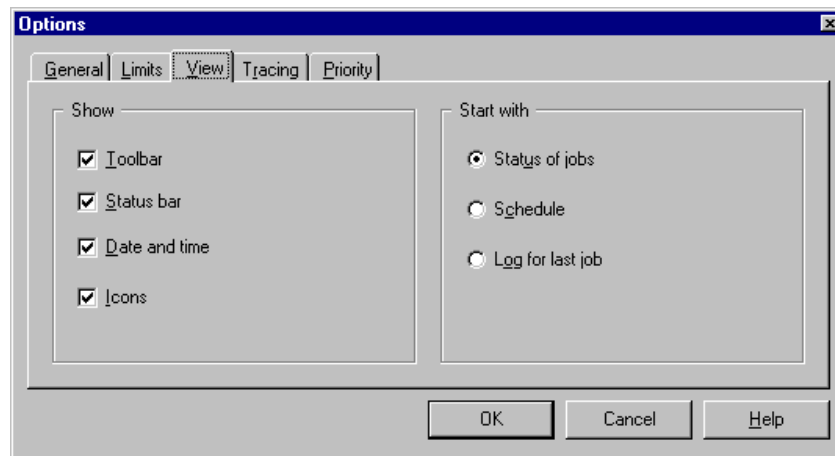
Enter a value 1 through 99999 or click the arrow buttons to increase or decrease the value. The default value is 1000.

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 through 99999 or click the arrow buttons to increase or decrease the value. The default value is 50.

View Page

The options on the **View** page determine what is displayed in the DataStage Director 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 server) on the status bar. |
| Icons | Displays the icons in the views. |

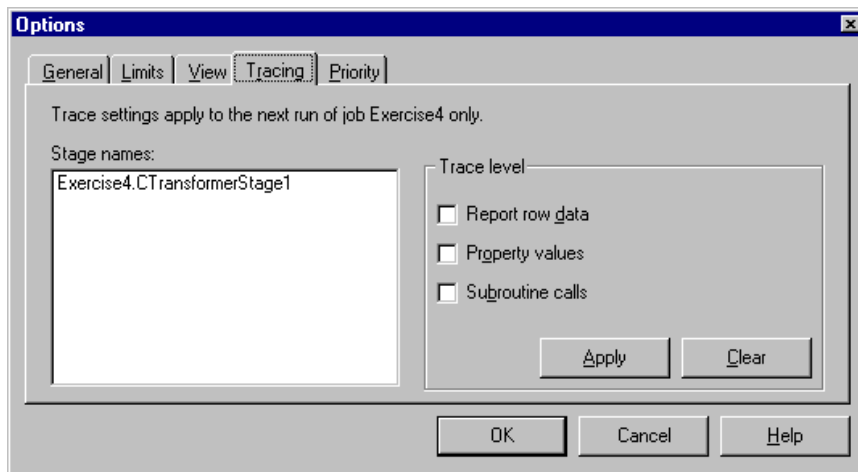
Specify the view to display when the Director is started, by clicking the appropriate option button:

- **Status of jobs** (the default setting)
- **Schedule**
- **Log for last job**

Tracing Page

The **Tracing** page is included to help Informix analysts during troubleshooting and appears only for compiled jobs.

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, press **Ctrl-Alt** and click to highlight stages in the **Stage names** list box, then select any of the following check boxes:

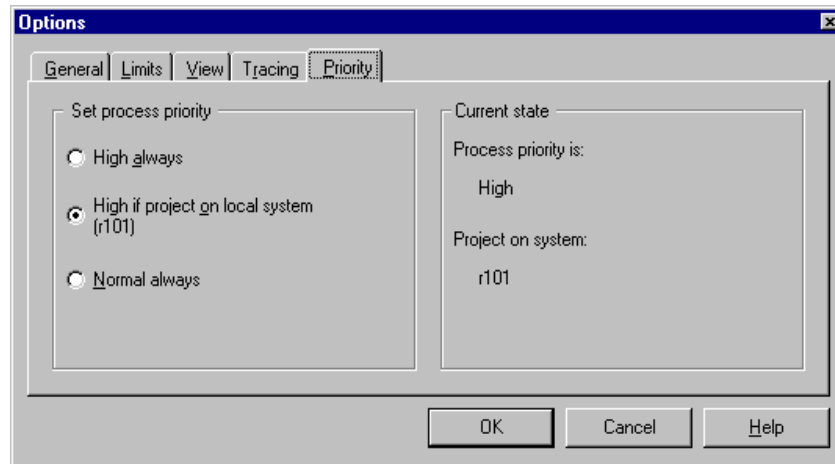
- **Report row data.** Records an entry for every data row read on input and written on output. This check box is cleared by default.
- **Property values.** Records an entry for every input and output opened and closed. This check box is cleared by default.
- **Subroutine calls.** Records an entry for every BASIC subroutine used. This check box is cleared by default.

Click **Apply** to set these options for the chosen stage. To clear options for a stage, highlight it in the **Stage names** list box and click **Clear**.

The next time the job is run, 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 DataStage server installation directory.

Priority Page

The **Priority** page is included for sites where the DataStage client and server components are installed on the same computer. When jobs are running, the performance of the DataStage Director may be noticeably slower. You can improve the performance by changing the priority of the DataStage Director process.



Setting the Process Priority

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

- Click **High always** to increase the priority of the DataStage Director, regardless of where the client and server components are installed. Note that if you choose this option and the client and server components are installed on different computers, you may not see an improvement in the performance of the DataStage Director.
- Click **High if project on local system** to increase the priority of the DataStage Director process if the client and server 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 DataStage Director process.

Displaying the Current State

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

Note: If you choose a high priority setting, it may 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. Settings on the **Tracing** page are applied to the next run of the specified job, whenever it occurs.

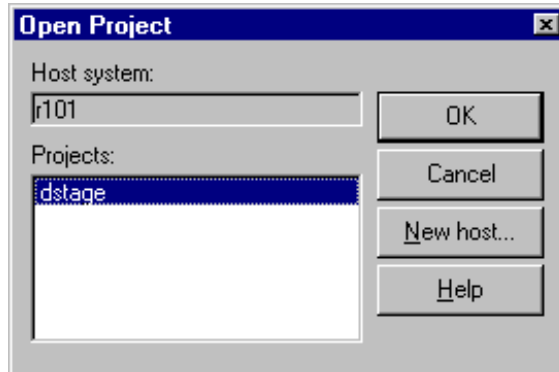
Choosing an Alternative Project

When you start the DataStage Director, the project chosen in the Attach to Project dialog box is opened. You can view the jobs in another DataStage project without exiting the DataStage Director.

Viewing Jobs in Another Project

To view the jobs in another project:

1. Do one of the following to display the Open Project dialog box:
 - Choose **Project ► Open...**
 - Click the **Open Project** icon on the toolbar.



2. Choose the project you want to open from the **Projects** list box. This list box contains all the DataStage projects on the server specified in the **Host system** field, which is the server you initially attached to.
3. Click **OK** to open the chosen project. The updated DataStage Director window displays the jobs in the new project.

Viewing Jobs on a Different Server

To open a project on a different DataStage server:

1. Click **New host...** . The Attach to Project dialog box appears.
2. Enter the name of the DataStage server and your logon details before choosing the project from the **Projects** drop-down list.
3. Click **OK** to open the chosen project. The updated DataStage 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 5, "Monitoring Jobs," for more details.

Exiting the DataStage Director

To exit the Director, choose **Project ► Exit**. Any open windows (for example, Monitor windows) are automatically closed on exit.

3

Running DataStage Jobs

This chapter describes how to run 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

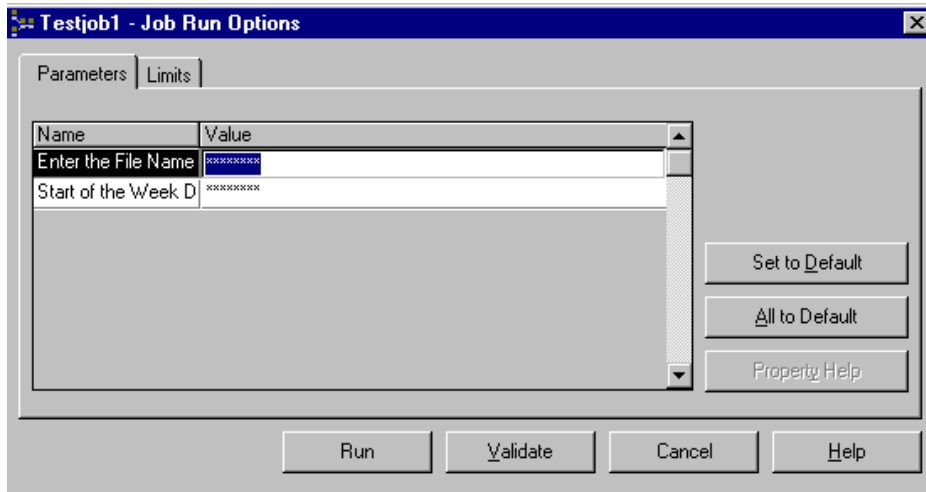
These tasks are performed from the Job Status view in the DataStage Director window. To switch to this view, choose **View ► Status**, or click the **Status** icon on the toolbar.

Setting Job Options

Each time you validate, run, or schedule a job you can:

- Change the job parameters associated with the job, as appropriate
- Override any default limits for row processing and warning messages that are set for the job run

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.

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**.

Validating a Job

You can check that a job will run successfully by validating it. Jobs should be validated before running them for the first time, or after making any significant changes to job parameters.

When a 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, UniVerse, or ODBC stages that use the local data source are created, if they do not already exist.

To validate a job:

1. Select the job you want to validate in the Job Status view.
2. Choose **Job ► Validate...** . The Job Run Options dialog box appears. See “Setting Job Options” on page 3-1.
3. Fill in the job parameters and check warning and row limits for the job, as appropriate.
4. Click **Validate**. Click **OK** to acknowledge the message. The job is validated and the job’s status is updated to Running.

Note: It may take some time for the job status to be updated, depending on the load on the server and the refresh rate for the client.

Once validation is complete, the job’s status updates to display 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 Chapter 6, “The Job Log File.”

If you want to monitor the validation in progress, you can use a Monitor window. For more information, see Chapter 5, “Monitoring Jobs.”

Running a Job

You can run a job in two ways:

- Immediately.
- By scheduling it to run at a later time or date. See “Job Scheduling” on page 3-5 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.

To run a job immediately:

1. Select the job in the Job Status view.
2. Do one of the following:
 - Choose **Job ► Run Now...** .
 - Click the **Run** icon on the toolbar.

The Job Run Options dialog box appears. See “Setting Job Options” on page 3-1.

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 may take some time for the job status to be updated, depending on the load on the server and the refresh rate for the client.

Stopping a Job

To stop a job that is currently running:

1. Select it in the Job Status view.
2. Do one of the following:
 - Choose **Job ► Stop**.
 - Click the **Stop** icon on the toolbar.

The job is stopped, regardless of the stage currently being processed, and the job’s status is updated to Stopped.

Note: It may take some time for the job status to be updated, depending on the load on the server and the refresh rate for the client.

Resetting a Job

If a job is stopped or aborted, it is difficult to determine whether all the required data was written to the target data tables. 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, return your target files to the state they were in before the job was run.

Note: You can only reinstate sequential files and hashed files to a prerun state.

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

To reset a job:

1. Select the job you want to reset in the Job Status view.
2. Choose **Job ► Reset**. 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 may take some time for the job status to be updated, depending on the load on the server and the refresh rate for the client.

Job Scheduling

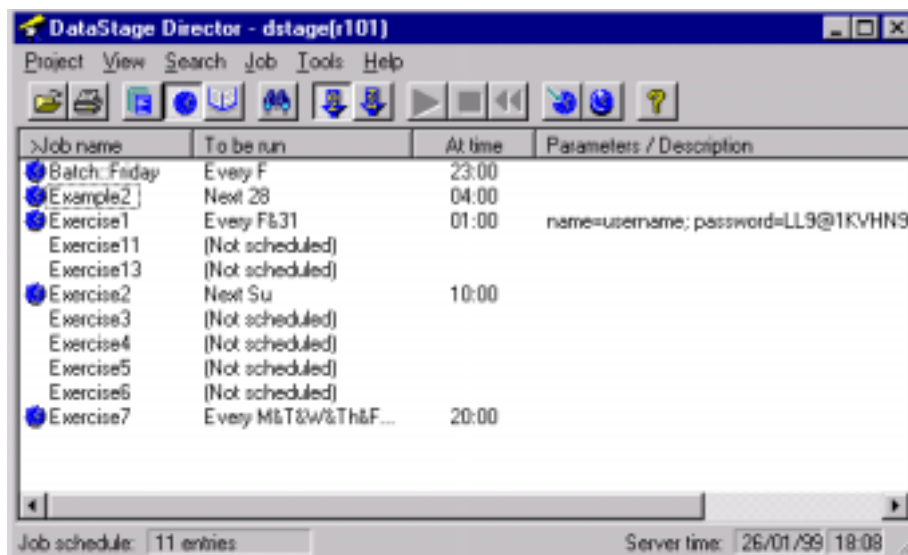
You can schedule a job to run in a number of ways:

- 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, using different job parameters if necessary. The scheduled jobs are displayed in the Job Schedule view.

Job Schedule View

The Job Schedule view displays details of all scheduled and unscheduled jobs and batches. To display the job schedule, choose **View ► Schedule** or click the **Schedule** icon on the toolbar.



The **Job name** column lists the jobs. By default this is all the jobs that are available, but you can filter the view to display specific categories of job. See “Filtering the Job Status or Job Schedule View” on page 2-10. The icon on the left indicates whether the jobs are 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 <i>n</i>	<i>n</i> 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 <i>x</i>	<i>x</i> 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.

To be run...	Means this...
Every <i>n</i> & <i>x</i>	<i>n</i> is a date and <i>x</i> 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 <i>n</i>	<i>n</i> is a date (as above). For example, Next 28 means the job is run on the next 28th of the month.
Next <i>x</i>	<i>x</i> is a day of the week. For example, Next W means the job is run the next Wednesday in the month.
Next <i>n</i> & <i>x</i>	<i>n</i> is a number and <i>x</i> 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.

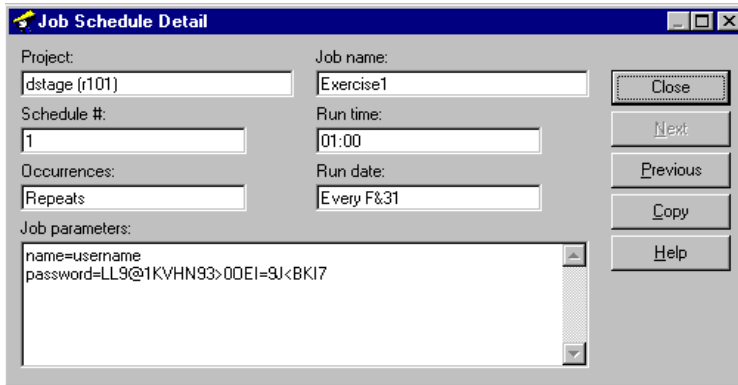
Viewing Details of a Job Schedule

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" on page 4-1.

If you choose a job, the Job Schedule Detail window appears:



This window 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 DataStage server.
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 <i>parameter name=value</i> .

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.

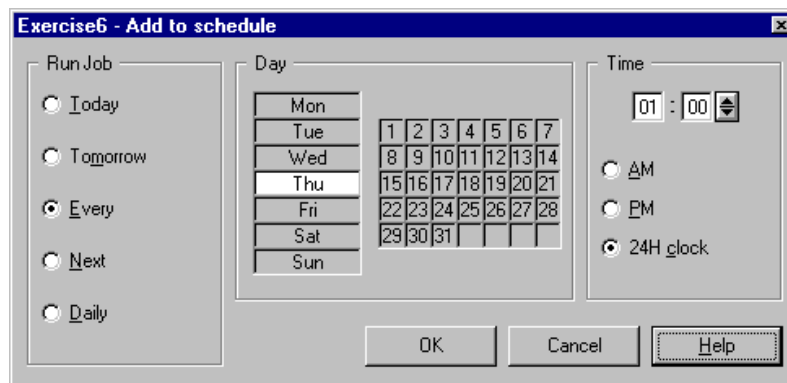
Scheduling a Job

To schedule a job:

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

Note: You cannot schedule a job with a status of Not compiled.

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** icon on the toolbar.



3. 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.

4. 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 server operating system, and you may not receive a warning of the invalid date. Refer to your server documentation for further information.

5. 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.

6. Click **OK**. The Add to schedule dialog box closes and the Job Run Options dialog box appears.
7. Fill in the job parameter fields and check warning and row limits, as appropriate.
8. Click **Schedule**. The job is scheduled to run and is added to the Job Schedule view.

Unschedulering a Job

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

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 shortcut menu.

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. To reschedule a job:

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 shortcut menu.
- Click the **Reschedule** icon 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 as follows:

1. Select the job in the Job Status view.
2. Choose **Job ► Delete**. A message confirms that you want to delete the chosen job.
3. Click **Yes** to delete the job. A message confirms the job has been deleted.
4. Click **OK**. The job 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 the job you deleted is part of a batch, edit the batch to remove the deleted job to prevent the batch from failing. See Chapter 4, “Job Batches.”

Job Administration

From the DataStage Administration client, the administrator can enable job administration options that let you clean up the resources of a job that has hung or aborted. These options 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 options:

- Cleanup Resources
- Clear Status File

Cleaning Up Job Resources

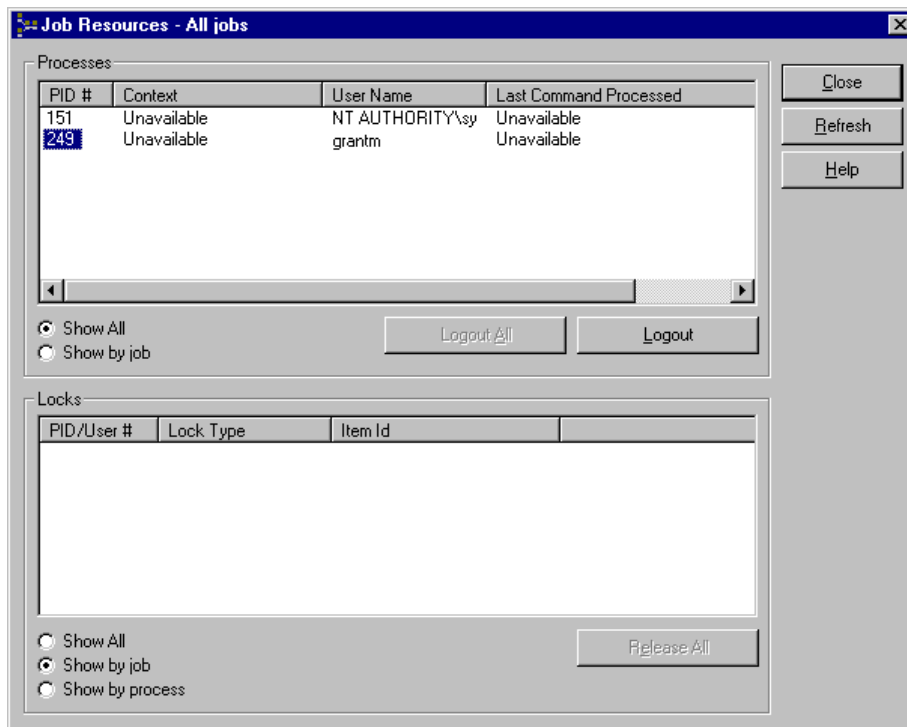
The **Cleanup Resources** option lets you:

- View and end job processes
- View and release the associated locks

To start this option, 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 may be reused during a job run. In that case the context field may 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 DataStage Director. 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 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

To end job processes:

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:

1. Log out of all DataStage clients.
2. Try to end the process using the Windows NT Task Manager or kill the process in UNIX.
3. Stop and restart the UniVerse service.
4. Reset the job from the DataStage Director (see “Resetting a Job” on page 3-4).

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” on page 3-15).

Releasing Locks

To release locks:

1. From the Job Resources dialog box, choose the range of locks to list either by clicking the **Show by job** option button in the **Locks** area or by selecting a process in the **Processes** area, then clicking 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.)

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. You should therefore use this option with great care, and only when you believe a job has hung or aborted. In particular, before you clear a status file you should:

- Try to reset the job (see “Resetting a Job” on page 3-4).
- Ensure that all the job's processes have ended (see “Ending Job Processes” on page 3-14).

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.

4

Job Batches

This chapter describes how to create and run DataStage 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

What Is a Job Batch?

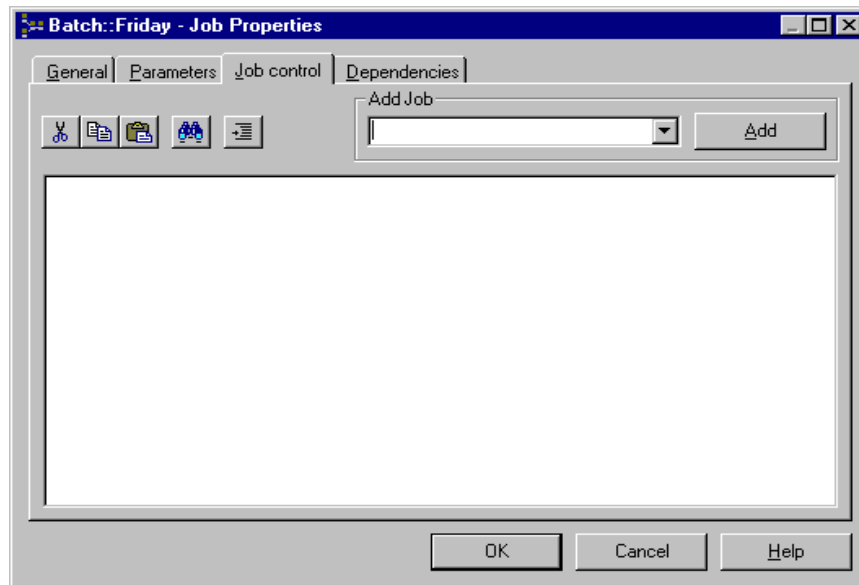
A job batch is a group of jobs or separate instances of the same job (with different job parameters) that you want to run sequentially. DataStage treats a batch as though it were a single job. If any job fails to complete successfully, the batch run stops. You can follow the progress of jobs within a 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

To create a job batch:

1. Choose **Tools** ► **Batch** ► **New...** . The New Batch dialog box appears.

2. Enter a name for the new batch in the **Batch name** field and click **OK**. The Job Properties dialog box appears:



3. Select the jobs to add to the batch from the **Add Job** list on the **Job control** page. This list displays all the 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.
4. 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.
5. Optionally, enter a more detailed description of the batch in the **Full job description** field.
6. Click the **Parameters** tab. The **Parameters** page appears at the front of the Job Properties dialog box.
7. 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.
8. 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 package the batch job for deployment on another system using the DataStage Manager. Information on dependencies and packaging jobs is in *DataStage Developer's Guide* and DataStage Developer's Help.

Running a Job Batch

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:

1. Select the batch in the Job Status view.
2. Do one of the following:
 - Choose **Job ► Run Now...**
 - Click the **Run** icon on the toolbar.

The Job Run Options dialog box appears. See “Setting Job Options” on page 3-1.

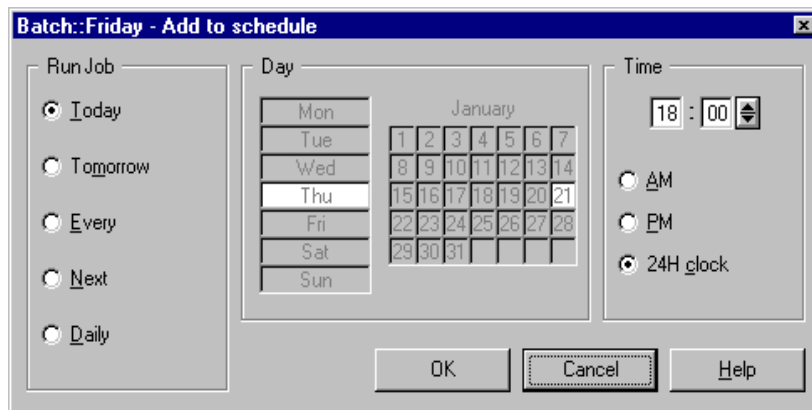
3. Fill in the job parameters and check warning and row limits for the batch, as appropriate.
4. Optionally, click **Validate** to validate the job.
5. Click **Run**. The batch is started and its status is updated to Running.

Note: It may take some time for the job status to be updated, depending on the load on the server and the refresh rate for the client.

Scheduling a Job Batch

To schedule a batch:

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 shortcut menu.
 - Click the **Schedule** icon 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 server operating system, and you may not receive a warning of the invalid date. Refer to your server 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. 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 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. To reschedule a batch:

1. Select the batch in the Job Schedule view and do one of the following:
 - Choose **Job ► Reschedule...**
 - Choose **Reschedule...** from the shortcut menu.
 - Click the **Reschedule** icon 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 NT Schedule service is running on the server (see *DataStage Administrator's Guide*).
- You belong to a Windows NT 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.

Note: Do not edit a job batch while it is running.

To edit a job batch:

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 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

To copy a job batch:

1. In the Job Status view, select the batch and choose **Tools ► Batch ► Save As...**. The Save As New Batch dialog box appears.
2. Enter the name for the new copy of the batch.
3. Click **OK** to copy the batch.

Deleting a Job Batch

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

1. Select the batch 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.

5

Monitoring Jobs

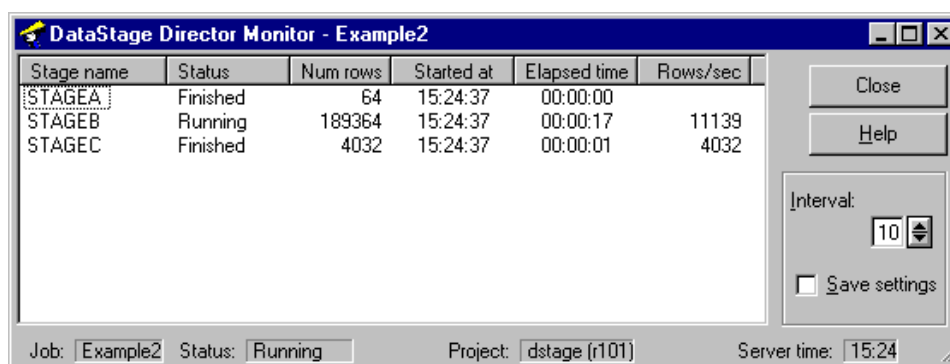
This chapter describes how to monitor running jobs using the **Monitor** option. Processing details that occur at each active stage of the job are displayed in a Monitor window. These details include:

- 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

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.



By default, this window displays summary information about the active stages in a job. The columns in the display are described in the following table:

This column...	Contains this information...																
Stage name	The names of the active stages in the job. These are the stages that perform the processing (for example, Transformer stages). Stages that represent data sources or data warehouses are not displayed in this window.																
Status	<p>The status of each active stage. The possible states are:</p> <table> <tr> <th>State</th><th>Meaning</th></tr> <tr> <td>Aborted</td><td>The processing terminated abnormally at this stage.</td></tr> <tr> <td>Finished</td><td>All data has been processed by the stage.</td></tr> <tr> <td>Ready</td><td>The stage is ready to process the data.</td></tr> <tr> <td>Running</td><td>Data is being processed by the stage.</td></tr> <tr> <td>Starting</td><td>The processing is starting.</td></tr> <tr> <td>Stopped</td><td>The processing was stopped manually at this stage.</td></tr> <tr> <td>Waiting</td><td>The stage is waiting to start processing.</td></tr> </table>	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.
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.																
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 server.																
Elapsed time	The elapsed time since processing started.																
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 DataStage server, and the current time on the server.

Setting the Server Update Interval

The Monitor window display is updated with new information from the server 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.

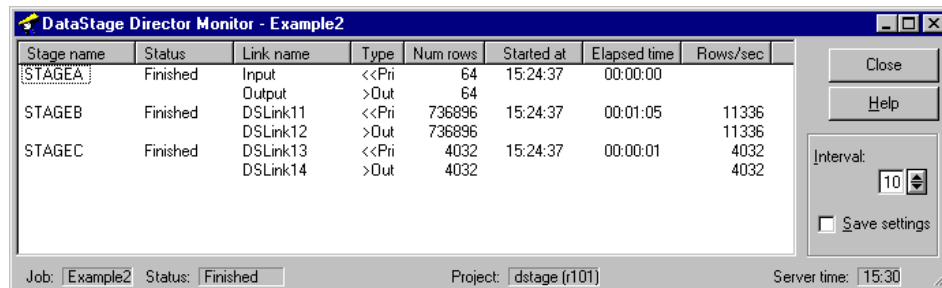
Saving Monitor Window Settings

You can save the Monitor window size and position (and the Interval setting) by selecting the **Save settings** check box. These settings are used the next time you open a Monitor window for any job.

You can also save the window settings immediately by choosing **Save window settings** from the Monitor shortcut menu. If you want to clear any saved settings and return to the default values, choose **Clear window settings** from the shortcut menu.

Showing Links

As well as the default Monitor window view, you can display link information for each active stage in a job. To show the links, choose **Show links** from the Monitor shortcut menu. The updated Monitor window displays the link information:



The screenshot shows the 'DataStage Director Monitor - Example2' window. It contains a table with columns: Stage name, Status, Link name, Type, Num rows, Started at, Elapsed time, and Rows/sec. The table lists three stages: STAGEA, STAGEB, and STAGEC, all with a status of 'Finished'. STAGEA has an input link and an output link. STAGEB and STAGEC have multiple links. To the right of the table are buttons for 'Close', 'Help', and 'Interval' (set to 10). At the bottom, there is a 'Save settings' checkbox. The status bar at the bottom shows 'Job: Example2', 'Status: Finished', 'Project: dstage (r101)', and 'Server time: 15:30'.

Stage name	Status	Link name	Type	Num rows	Started at	Elapsed time	Rows/sec
STAGEA	Finished	Input	<<Pri	64	15:24:37	00:00:00	
		Output	>Out	64			
STAGEB	Finished	DSLink11	<<Pri	736896	15:24:37	00:01:05	11336
		DSLink12	>Out	736896			11336
STAGEC	Finished	DSLink13	<<Pri	4032	15:24:37	00:00:01	4032
		DSLink14	>Out	4032			4032

Each link to or from an active stage is represented by one row in the Monitor window. Primary input links are listed first, followed by other (reference) inputs, then outputs and rejects.

The link information is displayed in two additional columns in the Monitor window. **Link name** is the name of the link in the job design. **Type** displays the type of link. The types available are:

primary input link	<<Pri
input link	<Ref
output link	>Out
output link for rejected rows	>Rej!

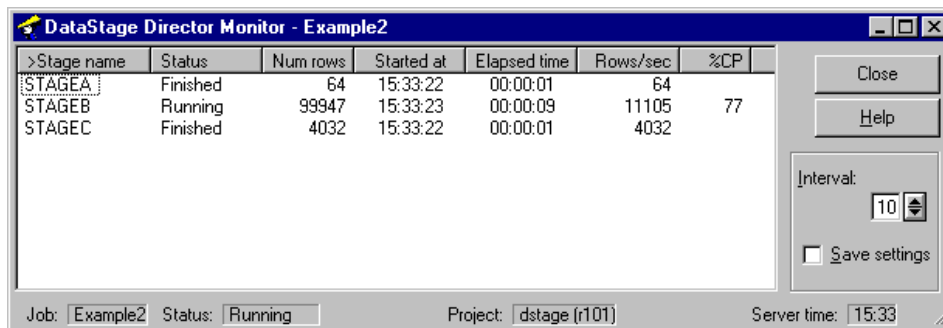
The properties of the stage as a whole (stage status and the start and elapsed times) are displayed in the row for the primary input link. The order used to display the rows is determined by the server. If you try to reorder the columns, a message box appears. Click **OK** to close the message box.

To disable the link information, choose **Show links** from the shortcut menu again.

Showing CPU Usage

While a stage 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 following screen shows the **%CP** column in the default Monitor window:



The screenshot shows the 'DataStage Director Monitor - Example2' window. It contains a table with columns: Stage name, Status, Num rows, Started at, Elapsed time, Rows/sec, and %CP. The table lists three stages: STAGEA (Finished), STAGEB (Running), and STAGEC (Finished). STAGEB is the primary input link and shows a %CP value of 77. The window also features a 'Close' button, a 'Help' button, an 'Interval' spinner set to 10, and a 'Save settings' checkbox. The status bar at the bottom shows 'Job: Example2', 'Status: Running', 'Project: dstage (r101)', and 'Server time: 15:33'.

>Stage name	Status	Num rows	Started at	Elapsed time	Rows/sec	%CP
STAGEA	Finished	64	15:33:22	00:00:01	64	
STAGEB	Running	99947	15:33:23	00:00:09	11105	77
STAGEC	Finished	4032	15:33:22	00:00:01	4032	

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.

Note: If you are viewing link information in a Monitor window, the percentage of computer processor time is displayed only in the row for the primary input link.

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

Viewing and Cleaning Up Job Resources

If the DataStage administrator has enabled job administration within the Director, you can view and clean up resources associated with the job being monitored. To do so, choose **Cleanup Resources** from the Monitor shortcut menu. The Job Resources window is displayed, from which you can:

- View and end processes
- View and release locks

For full information about these options, refer to “Cleaning Up Job Resources” on page 3-12.

Switching Between Monitor Windows

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

To bring a Monitor window to the front of the screen:

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 are displaying the default Monitor window or link information when you choose a stage.

The following screen shows the Stage Status window from a default Monitor window:

Project:	dstage (r101)	Row count:	75376	Close Copy Help
Job name:	Example2	Control:	0	
Stage:	STAGEB	Wave #:	1	
Status:	Running	User:	-175	
Started at:	1999-01-28 09:29:06	Retrieved:	1999-01-28 09:29:14	
Ended at:				

The fields in the window cannot be edited.

This field...	Displays this information...
Project	The name of the project and the DataStage server.
Job name	The name of the job.
Stage	The name of the stage.
Status	The status of the stage. This is one of the states described on page 5-2.
Started at	The date and time (on the server) the processing started at this stage.
Ended at	The date and time (on the server) 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 DataStage.
Wave #	An internal number assigned by DataStage.
User	The name or process number of the user who ran the job.
Retrieved	The date and time (on the server) the stage retrieved the data to process.

The following screen shows the Stage Status window if you are displaying link information:

The screenshot shows a window titled "Stage Status" with a standard Windows-style title bar (minimize, maximize, close buttons). The window contains several text input fields and buttons. The fields are arranged in two columns. The left column contains: "Project:" with value "dstage (r101)", "Job name:" with value "Example2", "Stage.Link:" with value "STAGEC.DSLink13", "Status:" with value "Finished", "Started at:" with value "1999-01-28 09:29:06", and "Ended at:" with value "1999-01-28 09:29:07". The right column contains: "Row count:" with value "4032", "Link type:" with value "Primary", "Wave #:" with value "1", "User:" with value ".229", and "Retrieved:" with value "1999-01-28 09:32:11". To the right of the "Row count" field is a "Close" button. Below the "Link type" field is a "Copy" button. Below the "Wave #" field is a "Help" button.

Project:	dstage (r101)	Row count:	4032	Close
Job name:	Example2	Link type:	Primary	Copy
Stage.Link:	STAGEC.DSLink13	Wave #:	1	Help
Status:	Finished	User:	.229	
Started at:	1999-01-28 09:29:06	Retrieved:	1999-01-28 09:32:11	
Ended at:	1999-01-28 09:29:07			

The same fields described earlier are displayed in this window except:

- **Stage** is replaced by **Stage.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"

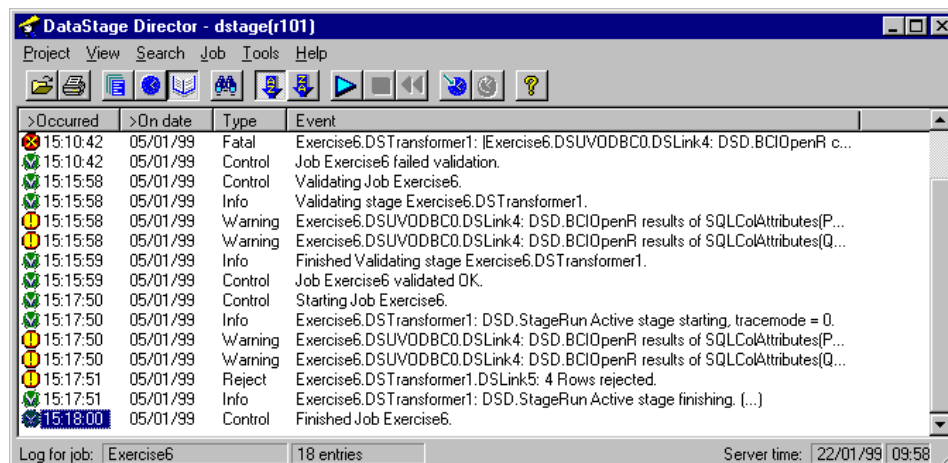
6

The Job Log File

This chapter describes the job log file and the Job Log view. The job log file is updated when you validate, run, or reset a job. You can use the log file to troubleshoot jobs that fail during validation, or that are aborted when they are run.

Job Log View

A job's log file is displayed in the Job Log view. To display the log file, select the job in the Job Status view and choose **View ► Log** or click the toolbar icon. The display is updated to list the job log file.



Each log file contains entries describing events that occurred during the last (or previous) runs of the job.

Entries are written to the log when:

- The job starts
- The job finishes
- A stage starts
- A stage finishes
- Rejected rows are output
- Errors are generated
- A scheduled job starts or a batch starts or finishes

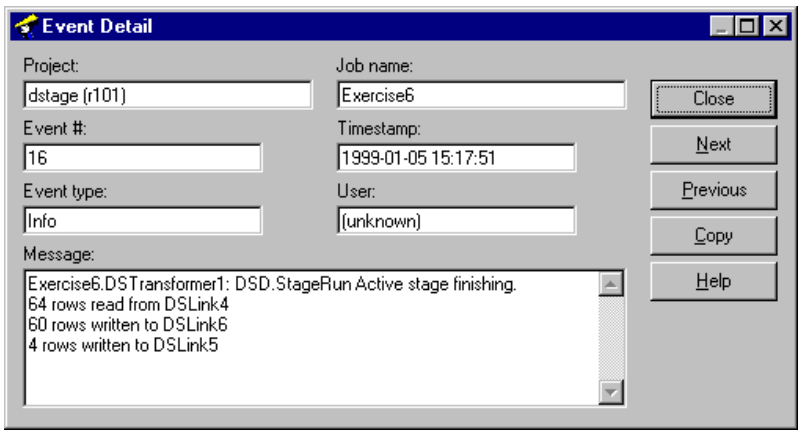
The four columns in the Job Log view are described in the following table:

This column...	Contains this information...	
Occurred	The time the event occurred.	
On date	The date the event occurred.	
Type	The type of event:	
	This event...	Occurs when...
	Info	Stages start and finish. No action required.
	Warning	An error occurred. You may need to investigate the cause of the warning, as this may 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.
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 one line. You can view the full message in the Event Detail window.	

The Event Detail Window

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

This field...	Contains this information...
Project	The name of the project and the DataStage server.
Event #	The ID assigned to the event during the job run.
Event type	The type of event. The event types are: <ul style="list-style-type: none">• Info• Warning• Fatal• Control• Reject• Reset For a description of each event type, see “Job Log View” on page 6-1.

This field...	Contains this information...
Job name	The name of the job.
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.

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.

Viewing Related Logs

To examine the log entries for a job that was run as part of a batch:

1. Switch to the log entries for the batch itself by selecting a log entry.
2. Do one of the following:
 - Choose **Tools ► Batch ► 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

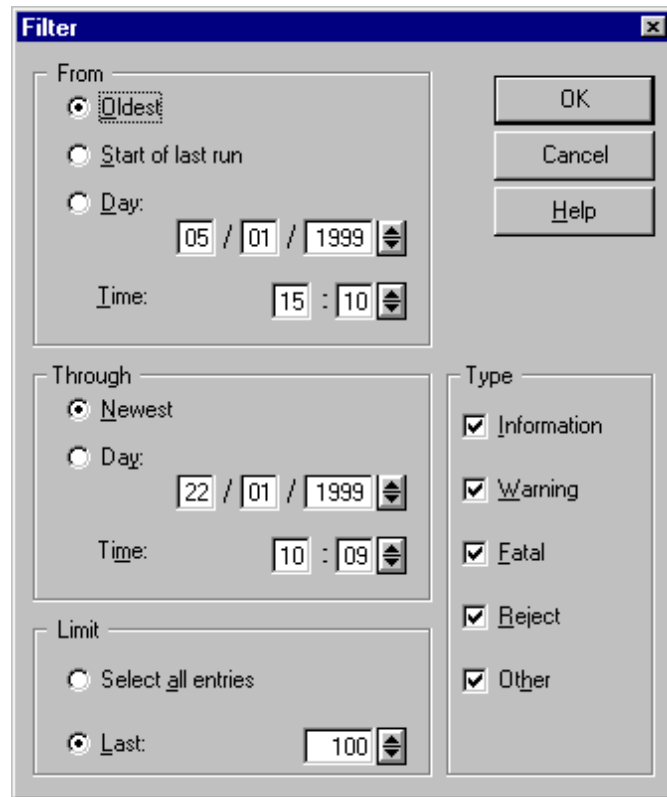
By default, the information displayed in the Job Log view includes all the events that have occurred for the chosen job, regardless of how many times the job has been run. On an established system, the display may include more information than you need.

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

To start the **Filter** option, 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.

To use the **Filter** option:

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\mm\yyyy` or `mm\dd\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.

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 by the DataStage administrator. See *DataStage Administrator's Guide*.
- 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 the Clear Log dialog box, select the job in the Job Status view, then choose **Job ► Clear Log...**

Clear Log

For job: Exercise6

Oldest entry: 05/01/99 15:10:41

☒ Immediate purge

☐ Auto-purge

Immediate

☒ Clear all entries

☐ Up to last run

☐ Before date:

15 / 01 / 1999

Auto-purge

☐ Enabled for this job

☐ Up to previous (job runs):

1

☒ Older than (days):

1

OK

Cancel

Help

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 may look different on your screen depending on your Windows settings.)

Purging Log Entries Immediately

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

Click **Auto-purge** to select the **Auto-purge** area. You set the purging criteria by clicking one of the following:

- **Up to previous n job runs.** Purges old log entries, leaving the last n entries in the file.
- **Over n days old.** Purges all log entries over n days old.

Specify n 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.

Glossary

This glossary contains terms that have a specific meaning in DataStage.

active stage	A stage in a job that carries out processing. Only active stages can be monitored from a Monitor window. See Chapter 5.
data source	A source for the data to be processed in a DataStage job, for example, a database, file, or device.
DataStage Administrator	The client program used for administering DataStage projects. See <i>DataStage Administrator's Guide</i> .
DataStage Designer	A graphical design program used to design and develop DataStage jobs. See <i>DataStage Developer's Guide</i> .
DataStage Director	A program used to run and monitor DataStage jobs.
DataStage Manager	A program used to view and manage the contents of the Repository. See <i>DataStage Developer's Guide</i> .
DataStage Server	The DataStage component that links the DataStage client programs to the server engine.
developer	The person designing and developing DataStage jobs.
job	A program that defines how to extract, cleanse, transform, integrate, and load data into a target database. Jobs are designed and packaged by the developer, using the DataStage Designer. Compiled jobs are run through the DataStage Director.
job batches	A group of jobs that are run together as a sequence. See Chapter 4.
job log	A file containing entries for events that occur during a job run. You view log entries from the Job Log view. See Chapter 6.

job parameter	A variable included in the job design, for example, a file name or a password. A suitable value must be entered for the variable when the job is validated, scheduled, or run.
link	A link joins stages of a job, and comprises either a flow of data, or a reference look-up. You can monitor links in a job using a Monitor window. See Chapter 5.
operator	The person running, scheduling, and monitoring DataStage jobs.
project	A collection of jobs and the components required to develop or run them. The level at which you enter DataStage from a client. DataStage projects must be licensed.
Repository	A server area where projects and jobs are stored. The Repository also holds definitions for data, stages, and so on. The Repository is accessed through the DataStage Manager. See <i>DataStage Developer's Guide</i> .
server engine (UniVerse)	The server software that holds, cleanses, and transforms the data during a DataStage job run.
stage	A component that represents a data source, a processing step, or a data mart in a DataStage job. See also active stage .

Index

A

- active stages 5-1
 - definition G1-1
- Add to schedule dialog box 4-4
- aggregating data 1-2
- alternative printer 2-16
- alternative project 2-22
- Attach to Project dialog box 2-2

B

- batches, *see* job batches

C

- changing the printer setup 2-16
- choosing an alternative printer 2-16
- cleaning up job resources 3-11, 5-5
- clearing
 - job log file 6-6
 - job status file 3-15
- client components of DataStage 1-2
- columns, sorting 2-14
- CPU usage, showing 5-4
- creating job batches 4-1

D

- data
 - aggregating 1-2
 - extracting 1-2
 - transforming 1-2
- data sources, definition G1-1
- data warehouses, overview 1-1

- DataStage
 - client components 1-2
 - introduction 1-2
 - projects 1-3
- DataStage Administrator,
 - definition G1-1
- DataStage Designer 1-2
 - definition G1-1
- DataStage Director
 - definition G1-1
 - exiting 2-23
 - starting 2-1
- DataStage Director options 2-16
 - display settings 2-19
 - Filter settings 2-18
 - limits 2-18, 2-23
 - Main window size and position 2-18
 - Options dialog box 2-17
 - priority 2-21
 - Refresh setting 2-17
 - save settings 2-18
 - tracing 2-20
- DataStage Director window 2-3
 - display area 2-6
 - menu bar 2-3
 - shortcut menus 2-8
 - status bar 2-5
 - toolbar 2-5
- DataStage Manager 1-2
 - definition G1-1
- DataStage Server 2-1
 - definition G1-1
- default display options 2-16

- deleting
 - job batches 4-7
 - jobs 3-11
- dependencies, specifying 4-3
- Designer, *see* DataStage Designer
- developer 1-3
 - definition G1-1
- Developer's Edition 1-3
- Director, *see* DataStage Director
- displaying
 - Monitor window 5-1
 - Stage Status window 5-5
 - ToolTips 2-5
- documentation conventions viii

E

- editing job batches 4-6
- ending job processes 3-12, 5-5
- entering job parameters 3-1
- Event Detail window 6-3
- examples
 - filtering views 2-12
 - Job Log view 6-1
 - Job Schedule view 3-5
 - Job Status view 2-6
- exiting DataStage Director 2-23
- extracting data 1-2

F

- Filter dialog box 6-5
- Filter Jobs dialog box 2-11
- Filter option 2-10, 6-4
- filtering
 - examples 2-12
 - Job Schedule view 2-10
 - Job Status view 2-10
- Find dialog box 2-13
- Find option 2-13

H

- Help system, starting 2-4

I

- icons, hiding 2-4

J

- job administration 3-11
- job batches 4-1
 - copying 4-6
 - creating 4-1
 - definition G1-1
 - deleting 4-7
 - editing 4-6
 - related log entries 6-4
 - rescheduling 4-5
 - running 4-3
 - scheduling 4-3, 4-4
 - unscheduling 4-5
- job log file 6-1
 - purging entries 6-6
- Job Log view 2-4, 6-1
 - Event Detail window 6-3
 - filtering the display 6-4
 - shortcut menus 2-9
- job log, definition G1-1
- job parameters 3-1
 - definition G1-2
- job processes
 - ending 3-12, 5-5
 - viewing 3-12, 5-5
- Job Properties Job control page 4-2
- Job Resources dialog box 3-12
- job resources, cleaning up 3-11, 5-5
- Job Run Options dialog box 3-1
- Job Schedule Detail window 3-7
- Job Schedule view 2-3, 3-5
 - filtering the display 2-10
 - shortcut menu 2-9

- Job Status Detail dialog box 2-7
- job status file, clearing 3-15
- Job Status view 2-3, 2-6
 - filtering the display 2-10
 - shortcut menus 2-8
- jobs
 - definition G1-1
 - deleting 3-11
 - overview 1-3
 - rescheduling 3-10
 - resetting 3-4
 - running 3-3
 - scheduling 3-5
 - status 2-6
 - stopping 3-4
 - unscheduling 3-10
 - validating 3-2

L

- limiters
 - row 2-19
 - warning 2-19
- link information, showing 5-3
- links, definition G1-2
- locks
 - releasing 3-12, 5-5
 - viewing 3-12, 5-5
- log file, *see* job log file

M

- menus
 - pull-down 2-4
 - shortcut 2-8
- Monitor window
 - % CPU 5-4
 - cleaning up job resources 5-5
 - clearing settings 5-3
 - contents 5-2
 - displaying 5-1
 - link information 5-3
 - saving settings 5-3

- shortcut menu 2-10
- switching between 5-5

O

- operator 1-3
 - definition G1-2
- Operator's Edition 1-3
- options, DataStage Director 2-16
- overview
 - of data warehouses 1-1
 - of DataStage 1-2
 - of jobs 1-3
 - of projects 1-3

P

- parameters, *see* job parameters
- Print dialog box 2-14
- Print Setup dialog box 2-16
- printers, choosing alternative 2-16
- printing current view 2-14
- priority of Director process 2-21
- process priority 2-21
- projects
 - choosing alternative 2-22
 - definition G1-2
 - overview 1-3
- purging log file entries 6-6

R

- Refresh setting 2-17
- related logs 6-4
- releasing locks 3-12
- Repository, definition G1-2
- rescheduling
 - job batches 4-5
 - jobs 3-10
- resetting jobs 3-4
- row limiters 2-19

- running jobs 3-3
 - immediately 3-3
 - scheduling 3-3

S

- saving window settings 2-18
- scheduling
 - job batches 4-4
 - jobs 3-5
- server engine, definition G1-2
- setting default display options 2-16
- shortcut menus 2-8
 - in Monitor windows 2-10
 - in the Job Log view 2-9
 - in the Job Schedule view 2-9
 - in the Job Status view 2-8
- showing
 - CPU usage 5-4
 - link information 5-3
- sorting columns 2-14
- Stage Status window
 - contents 5-6
 - displaying 5-5
- stages, definition G1-2
- starting
 - DataStage Director 2-1
 - Help system 2-4
- status
 - jobs 2-6
 - stages 5-2
- status bar 2-5
- stopping jobs 3-4
- switching between Monitor windows 5-5

T

- times, comparing client and server 2-18
- toolbar 2-5
- ToolTips 2-5

Index-4

- trace level 2-20
- tracing options 2-20
- transforming data 1-2

U

- unscheduling
 - job batches 4-5
 - jobs 3-10
- using
 - Filter option 6-4
 - Find option 2-13
 - job parameters 3-1

V

- validating jobs 3-2
- viewing
 - event details 6-3
 - job processes 3-12, 5-5
 - job status details 2-7
 - jobs in another project 2-22
 - jobs on a different server 2-23
 - locks 3-12, 5-5
 - schedule details 3-7
- views
 - filtering 2-10
 - Job Log 2-4, 6-1
 - Job Schedule 2-3, 3-5
 - Job Status 2-3, 2-6

W

- warning limiter 2-19
- window settings, saving 2-18