IBM InfoSphere DataStage
Version 11 Release 3

Troubleshooting Guide

IBM

IBM InfoSphere DataStage
Version 11 Release 3

Troubleshooting Guide

IBM

Contents

Troubleshooting InfoSphere DataStage

These topics contain troubleshooting information for IBM® InfoSphere® DataStage®. Some of the information is also useful for QualityStage users.

Troubleshooting problems when starting an InfoSphere DataStage and QualityStage client

If you cannot start your client and connect to the services tier (domain), the problem might be due to an invalid host name or invalid port, a startup script error, or incorrect user credentials.

The following table gives a list of the possible error messages, and tells you where to look for solutions.

Table 1. Error messages when starting an InfoSphere DataStage and QualityStage client

Error message	Related topic
Failed to authenticate the current user against the selected Domain: Server [servername] not found.	"Failure to connect to services tier: invalid host name"
Failed to authenticate the current user against the selected Domain: Could not connect to server [*servername*] on port [*portnumber*].	"Failure to connect to services tier: invalid port" on page 2
Failed to authenticate the current user against the selected Domain: Invalid user name (*username*) or password.	"Cannot authenticate user" on page 4

Failure to connect to services tier: invalid host name

You can verify whether the IBM WebSphere® Application Server has started, and test whether the host name that you are using is valid.

Symptoms
When you attempt to start one of the InfoSphere DataStage and QualityStage® clients, the following message is displayed:

Failed to authenticate the current user against the selected
Domain: Server [servername] not found.

Causes
You might be specifying an incorrect name for the computer that is hosting the services tier.

Diagnosing the problem
You can check whether the application server is running by attempting to connect to the application server using an Internet browser.

Connect to the application server by using a web browser:

1. Open a web browser.
2. Type the application server address in the form: http//*isserver*:*portnumber*/ibm/iis/console, where *isserver* is the name of the computer where the services tier is installed (or its IP address), and *portnumber* is the port for connecting to the services tier (by default, 9080).

If the application server has started, the login screen is displayed; otherwise an error message is displayed.

You can test whether you specified the correct name for the *isserver* by attempting to ping the computer that is hosting the services tier.

Resolving the problem

You can try to fix the problem by specifying a fully qualified path name for the computer that is hosting the application server. For example, instead of `isserver:9080`, you might have to type `isserver.mycompany.com:9080`.

If the application server is not running, attempt to start the service.

To start the application server where the services tier is installed on a Microsoft Windows computer, click **Start > All Programs > IBM WebSphere > Application Server v6 > Profiles > default > Start the server**

To start the application server where the services tier is installed on a UNIX or Linux computer, you must have root authority. To start the application server, do the following steps:

1. From a terminal window, change to the *WASInstDir*/ASBServer/bin/ directory. *WASInstDir* is the installation directory for the application server. The default installation directory is /opt/IBM/InformationServer/.

2. Run the following command:

   `./MetadataServer.sh start`

Failure to connect to services tier: invalid port

You can verify that the IBM WebSphere application server has started, test whether the port number is valid, and confirm that you specified the correct port number.

Symptoms

When you attempt to start one of the InfoSphere DataStage and QualityStage clients, the following message is displayed:

```
Failed to authenticate
the current user against the selected Domain:  Could not connect to
server [servername] on port [portnumber].
```

Causes

The port number is incorrect or is unavailable.

Diagnosing the problem

You can check whether the application server is running by attempting to connect to the application server by using an Internet browser.

Connect to the application server by using a web browser:

1. Open a web browser.
2. Type the application server address in the form: `http//isserver:portnumber/ibm/iis/console`, where *isserver* is the name of the computer where the services tier is installed (or its IP address), and *portnumber* is the port for connecting to the services tier (by default, 9080).

If the application server has started, the login screen is displayed; otherwise an error message is displayed.

Test whether the port is accessible from the client computer by typing at the command line:

`telnet hostname port`

If you get an error message, then the port is inactive. If you get no response, then the port is active.

You can also test which ports are listening on the server computer by typing the following command:

`netstat -a`

Look for an entry in the form: *isserver:port_number*

You can check whether you are specifying the correct port number in the WebSphere Administrative Console. To look up the port number:

1. From the start menu, select **IBM WebSphere** > **Application Server v6** > **Profiles** > **default** > **Administrative console** to start the WebSphere Administrative Console.
2. Log in using the websphere user name and password that was specified when IBM InfoSphere Information Server was installed.
3. In the left pane, select **Servers** > **Application servers**
4. Click the server1 link.
5. Select **Communications** > **Ports**.
6. Look for the port number for WC_defaulthost. This is the port number you should use when connecting to the application server.

Resolving the problem

If the application server is not running, attempt to start the service.

To start the application server where the services tier is installed on a Microsoft Windows computer, click **Start > All Programs > IBM WebSphere > Application Server v6 > Profiles > default > Start the server**

To start the application server where the services tier is installed on a UNIX or Linux computer, you must have root authority. To start the application server, do the following steps:

1. From a terminal window, change to the *WASInstDir*/ASBServer/bin/ directory. *WASInstDir* is the installation directory for the application server. The default installation directory is /opt/IBM/InformationServer/.
2. Run the following command:

   ```
   ./MetadataServer.sh start
   ```

You can also check whether there is a firewall between the client and the server. If there is a firewall, temporarily disable it to verify that all inbound and outbound ports are open.

IBM WebSphere Application Server fails to start: AIX and Linux

If the WebSphere application server has not started, you can try to fix the problem by changing one of the startup scripts.

Symptoms

The application server fails to start after system is restarted. No messages are generated in the application server logs.

Causes

The Metadata server startup script fails to finish. You must issue the **nohup** command for the Metadata server startup script.

Environment

IBM AIX® or Linux systems.

Diagnosing the problem

Check to ensure that WebSphere Application Server is running.

Connect to the application server by using a web browser:

1. Open a web browser.
2. Type the application server address in the form: http//*isserver*:*portnumber*/ibm/iis/console, where *isserver* is the name of the computer where the services tier is installed (or its IP address), and *portnumber* is the port for connecting to the services tier (by default, 9080).

If the application server has started, the login screen is displayed; otherwise an error message is displayed.

Resolving the problem

1. Run the following command to locate the WebSphere Application Server startup scripts on your computer:

```
cd /etc
find . -name "*" -print | xargs grep -i InformationServer
```

This command might return multiple files with various prefixes in the name. Some files might be links to other files and could reflect the change you made in the original file without needing to edit each file that was found. If you have multiple instances of WebSphere Application Server installed, unique files might exist for each WebSphere Application Server instance. You only have to modify the files that reference the instances of WebSphere Application Server that you have configured to start as non-root.

2. Identify the files that you need to modify. Typically, you must modify the following files:

Operating system	Files
AIX	/etc/rc#.d/S99ISFServer
	The number symbol (#) can have the value of 0 through 6. For example:
	/etc/rc0.d/S99ISFServer /etc/rc2.d/S99ISFServer /etc/rc5.d/S99ISFServer
Linux	/etc/init.d/ISFServer

3. In each file, change the following content.

   Locate the following text, where *IS_install_path* is the directory where you installed InfoSphere Information Server. The default installation path is /opt/IBM/InformationServer:

```
"IS_install_path/ASBServer/bin/MetadataServer.sh"
"$@"
```

   Change the text to match this example:

```
nohup "IS_install_path/ASBServer/bin/MetadataServer.sh"
"$@"
```

4. Save the modified files and restart your system.

Cannot authenticate user

You can verify whether the failure to start the client is due to a problem with the user credentials.

Symptoms

When you attempt to start one of the InfoSphere DataStage and QualityStage clients, the following message is displayed:

```
Failed to authenticate
the current user against the selected Domain:  Invalid user name (username)
or password.
```

Causes

There are several possible causes of this problem.

- The user name is invalid.
- The password is invalid or has expired.
- The user has no suite user role.
- Credential mapping is required, but has not been defined for this user.

- The user has no DataStage role or has the incorrect DataStage role.

Diagnosing the problem

You can diagnose the likely cause of your problem by identifying where the problem occurs:

- InfoSphere Information Server authentication is performed when you retrieve a list of available projects in the Attach to Project window. If you get an error when retrieving the project list, then the user name or password are invalid, or the user has no suite user role, or credential mapping is required, but has not been performed.
- DataStage role checking is performed when you attach to the project. If you can retrieve the list of projects but you cannot attach to a project, then the problem lies with the DataStage role for this user.

Resolving the problem

Information on resolving issues concerned with user roles is available in IBM InfoSphere DataStage and QualityStage Administrator Client Guide.

Information on creating users when configuring InfoSphere Information Server is available in IBM InfoSphere Information Server Planning, Installation, and Configuration Guide.

Information on credential mapping is available in IBM InfoSphere Information Server Administration Guide.

Troubleshooting scheduled jobs

You can schedule jobs to run when the system is less busy.

You schedule jobs from the Job Schedule view in the InfoSphere DataStage and QualityStage Director client window.

InfoSphere DataStage does not have its own separate scheduling program. Instead, whenever an InfoSphere DataStage user schedules a job, the underlying operating system controls the job. If scheduled jobs do not run correctly, the problem is usually with the operating system configuration on the engine.

Resolving scheduling problems on Windows engine tier hosts

On Microsoft Windows engine tier hosts, job scheduling is carried out by the Schedule service.

If your scheduled job did not run, there are a number of steps that you can take to identify the cause.

Viewing the schedule log

You can view the schedule log to diagnose problems with job scheduling.

Symptoms

Scheduled jobs do not run when expected.

Environment

This advice applies to the Windows environment.

Diagnosing the problem

The schedule log is a text file named dsr_sched.log. It is located in the project directory (by default: `c:\IBM\InformationServer\Server\Projects`). This file records any problems that occurred before control was transferred from the scheduler to InfoSphere DataStage. (After that point, messages are written to the appropriate job log file.) The schedule log contains a message if, for example, the server password that you specified has expired.

Testing user name and password

If you specified a user name and password to run all scheduled jobs in a project, you can test the user name and password.

Symptoms

Scheduled jobs do not run when expected.

Causes

The user ID used to run the schedule service has invalid user name or password details.

Environment

This advice applies to the Windows environment.

Diagnosing the problem

If the test works correctly but scheduled jobs still do not run, check that the user name you specified has permission to read and write the project directory.

If the test fails, there might be a problem with the user rights for the user name you specified. In which case, check the user rights.

To test the user name and password:
1. Open the Administrator client and attach to the engine that you are scheduling jobs for.
2. Click the **Projects** tab and select a project from the list.
3. Click **Properties**.
4. Click the **Schedule** tab.
5. Enter the user name and password to test.
6. Click **Test**.
7. Wait for the user name and password to be verified (this might take some time).

Checking user rights

If the Windows Schedule service on the engine tier host does not run under the default user name, try this procedure to ensure that the Schedule service has the correct user rights:

Symptoms

Scheduled jobs do not run when expected.

Causes

The user running the schedule service does not have sufficient user rights.

Environment

This advice applies to the Windows environment.

Resolving the problem

To check user rights and allocate new rights, if required:
1. From the Windows engine tier host, select **Start** > **Programs** > **Administrative Tools** > **Local Security Policy.**
2. Open the Local Policies folder and select **User Rights Assignment**.
3. In the list on the right of the Local Security Settings window, double-click **Act as part of the operating system**.
4. In the Properties window, check if the user name for the schedule server is included in the list.
5. If the user name is missing, click **Add User or Group**, add the user name in **Enter the object name to select** field and click **OK**.

6. In the list on the right of the Local Security Settings window, double-click **Replace a process level token**.
7. In the Properties window, check if the user name for the schedule server is included in the list.
8. If the user name is missing, click **Add User or Group**, add the user name in **Enter the object name to select** field and click **OK**.
9. In the list on the right of the Local Security Settings window, double-click **Increase scheduling priority**.
10. In the Properties window, check if the user name for the schedule server is included in the list.
11. If the user name is missing, click **Add User or Group**, add the user name in **Enter the object name to select** field, and click **OK**.

Resolving problems using the scheduler on non-English language systems
You might to need to localize the names of the days of the week if you schedule jobs on a non-English language system.

Symptoms
Scheduled jobs do not run when expected.

Causes
The AT command, which performs the Windows scheduling, only accepts day names in the local language.

Environment
This advice applies to the Windows environment.

Resolving the problem
If you run IBM InfoSphere DataStage on a system with a language other than English, and encounter problems when scheduling jobs to run on specific days of the week, you can try localizing the days of the week for each project.

To localize the names of the days:
1. Navigate to the project directory for your first project, which is located on the engine. By default the project directory is the folder `C:\IBM\InformationServer\Server\Projects`.
2. Edit the file DSParams by using a text editor such as Notepad.
3. Add the localized days of the week to the end of the file. The following example shows what you might add for a French system:
```
[SCHEDULER]
MONDAY=L
TUESDAY=M
WEDNESDAY=ME
THURSDAY=J
FRIDAY=V
SATURDAY=S
SUNDAY=D
```

   You might have to experiment with which day names the local AT command will accept. If in doubt, enter the full name (for example, LUNDI, MARDI, and so on).
4. Repeat these steps for each of your projects.

You might find that you get an error message equivalent to 'There are no entries in the list' when you use the scheduler on a non-English language system. This message is output by the AT command and passed on by the Director client. To prevent the Director client from passing on the message:
1. Identify a unique part of the message that the AT command is outputting (for example, 'est vide' in French).
2. For each project, add the following line to its DSParams file:

```
NO ENTRIES=est vide
```

The AT command usually accepts other keywords besides days of the week in English. If your system does not accept other keywords, you can add localized versions of the additional keywords NEXT, EVERY, and DELETE to your projects by doing the following tasks:

1. Edit the DSParams file for each project.
2. For each keyword, add a line of the form:

   ```
   KEYWORD=localized_keyword
   ```

   For example:

   ```
   NEXT=Proxima
   ```

Resolving scheduling problems on UNIX and Linux servers

On UNIX servers, the scheduling of IBM InfoSphere DataStage jobs is handled by the at and cron commands.

If your scheduled job did not run, there are a number of steps that you can take to identify the cause.

Viewing scheduled jobs

On UNIX servers, you can view only jobs that you scheduled yourself.

Symptoms

Administrator cannot see all the jobs that the users have scheduled.

Environment

This advice applies to the UNIX environment.

Diagnosing the problem

For an IBM InfoSphere DataStage administrator, the problem means that it is not possible to get a quick overall view of all the InfoSphere DataStage jobs that are scheduled to run over a particular period. The only way to find out which jobs are scheduled is to examine the files in the cron directory for each user ID. The naming and location of these files varies from system to system. For more information, see the reference page for the cron command on your system.

Dealing with scheduled jobs not running

If a scheduled job does not run, check that the user who scheduled the job has permission to use the cron command.

Symptoms

Scheduled job does not run when expected.

Environment

This advice applies to the UNIX environment.

Resolving the problem

To check user permissions, examine the cron.allow and cron.deny files which contain lists of users who can and cannot use the command. The location of these files varies from system to system. For more information, see the reference page for the cron command on your system.

Scheduled jobs do not run on an AIX server

If scheduled jobs are not running on your IBM AIX server, check your file permissions.

Symptoms

Scheduled jobs do not run.

Environment
This advice applies to AIX servers.

Resolving the problem
To schedule jobs on an AIX server, change the permissions of `/usr/ spool/cron/atjobs` from 770 to 775 (rwxrwxr-x).

Resolving job termination problems

If you experience delays in the termination of an IBM InfoSphere DataStage job when it is run, clear the &PH& directory.

Symptoms
Jobs take too long to terminate.

Causes
Each InfoSphere DataStage project directory contains a &PH& directory. The &PH& directory contains information about active stages that is used for diagnostic purposes. The &PH& directory is added to every time a job is run, and needs to be cleared periodically.

Resolving the problem
To clear the directory:

1. Ensure that there are no jobs running anywhere on the system.
2. Open the Administrator client, go to the Projects page, select the project whose file you want to clear, and click **Command**.
3. In the Command Interface window, type the following command:

   `CLEAR.FILE &PH&`
4. Click **Execute** to run the command and clear the file.

Resolving problems with database stages on 64-bit systems

If you are running jobs in a 64–bit bit environment, you must ensure that any database clients that are required by connectivity stages match the installed version of InfoSphere DataStage.

Symptoms
Failure of the stage with symptoms such as a memory fault and corresponding core dump.

Causes
If you are running a 64–bit version of InfoSphere DataStage, you must ensure any database clients you use are also 64–bit. If you are running a 32–bit version of InfoSphere DataStage, you must ensure any database clients you use are also 32–bit. For example, Oracle database is available with both 32- and 64-bit clients. You must use the 32-bit client with 32-bit InfoSphere DataStage, and the 64-bit client with 64-bit InfoSphere DataStage.

Environment
Applies to 64–bit UNIX, Linux, or Windows environments.

Resolving ODBC connection problems on UNIX and Linux systems

IBM InfoSphere DataStage relies on third-party ODBC drivers to connect to ODBC data sources. There are various steps to take to diagnose and fix problems.

Testing ODBC driver connectivity

You can test the whether your ODBC drivers can successfully connect to your data sources.

Symptoms

If a job fails to connect to a data source using an ODBC connection, test the connection outside the job to see if the ODBC connection is the source of the problem.

Environment

The procedure applies to ODBC connections in a UNIX environment.

Diagnosing the problem

To test the connectivity of your ODBC connections:

1. Change directory to $DSHOME and set up the IBM InfoSphere DataStage environment by running dsenv:

   `. ./dsenv`

2. Start the engine shell:

   `./bin/dssh`

3. In the engine shell, log to the project:

   `LOGTO project_name`

4. Get a list of available DSNs by typing:

   `DS_CONNECT`

5. Test the required connection by typing:

   `DS_CONNECT DSN`

   Where *DSN* specifies the connection that you want to test.

6. Enter the user name and password to connect to the required data source.

7. After you have connected to the data source, enter `.Q` to close the connection.

Checking the shared library environment

Connection errors can be caused by incorrect environment settings.

Symptoms

Cannot connect to database using ODBC connection.

Environment

This problem occurs when using ODBC connections in a UNIX environment.

Diagnosing the problem

If you see a message similar to the following message:

`ld.so.1: uvsh: fatal: libxxxx: can't open file: errno=2`

check that the ODBC driver shared library has been added to the environment variable used to locate shared libraries

Resolving the problem

When the ODBC access is configured for DataStage, entries specifying the environment are added to the file $DSHOME/dsenv. Check the dsenv file to ensure that your environment is configured correctly.

The name of the shared library environment variable that you need to check depends upon the type of UNIX system. The required entry depends upon the type of database that you are attempting to connect to. The environment variables for the UNIX platforms are in the following table. Consult your database documentation for the location of the shared libraries.

Table 2. Library path environment variables

Platform	Environment variable
Solaris	LD_LIBRARY_PATH
HP-UX	SHLIB_PATH
HP-UX Itanium	LD_LIBRARY_PATH
AIX	LIBPATH
Linux	LD_LIBRARY_PATH

Checking symbolic links

If your ODBC connection does not work, check your symbolic links.

Symptoms

Cannot connect to database using ODBC connection.

Causes

If you have moved shared libraries to a new directory or have installed a new ODBC driver manager, you might have broken symbolic links that the engine uses to access the shared libraries for the source database.

Environment

This problem occurs when using ODBC connections in a UNIX environment.

Resolving the problem

To reset the symbolic links to a new directory, run this command at the UNIX prompt:

```
# $DSHOME/bin/dspackinst relink.uvlibs pathname
```

$DSHOME is the home directory of the server engine. *pathname* is the full path name of the directory that contains the shared libraries.

To reset links for a new ODBC driver manager:

1. Install the ODBC driver manager according to the vendor's instructions.
2. Determine where the ODBC shared library libodbc.xx resides. For example, the library for the Intersolv driver is in $ODBCHOME/ dlls, and the library for the Visigenics driver is in $ODBCHOME/ libs.
3. Close all InfoSphere DataStage clients.
4. Run the `relink.uvlibs` command as described above.
5. Restart the InfoSphere DataStage clients.

Resolving configuration problems on UNIX systems

There are various problems that you might encounter when running IBM InfoSphere DataStage on a UNIX system.

Running out of file units

If you receive notification that jobs fail because they run out of file units you can increase the allocation of file units.

Symptoms

Jobs fail because they run out of file units.

Environment

This advice applies to UNIX systems.

Resolving the problem

The engine uses the parameter MFILES and the kernel parameter NOFILES to determine the number of open files allowed. The number of open files allowed is NOFILES - MFILES. If you encounter problems and run out of file units, you can decrease the value of MFILES in the server engine file uvconfig or increase the value of NOFILES for your operating system. The uvconfig file resides in the DSEngine directory.

If you change the value of MFILES, you must stop and restart the engine as follows:

1. To stop the engine:

   `$DSHOME/bin/uv -admin -stop`

2. To upgrade configuration information for the engine:

   `$DSHOME/bin/uv -admin -regen`

3. To start the engine:

   `$DSHOME/bin/uv -admin -start`

Ensure that you allow at least thirty seconds between executing stop and start commands.

Running out of memory on AIX computers

You can tune IBM InfoSphere DataStage to increase the configured memory on IBM AIX systems.

Symptoms

Jobs with large memory requirements cause `unable to locate memory` errors.

Environment

This advice applies to AIX systems.

Resolving the problem

Increase the memory allocation by doing the following steps:

1. Edit the $DSHOME/.uvconfig file in the engine directory. Make the following changes:
   * Change DMEMOFF to 0x90000000
   * Change PMEMOFF to 0xa0000000

2. Ensure that there are no active InfoSphere DataStage users on the system, then shut down the engine:

   `$DSHOME/bin/uv -admin -stop`

3. To upgrade configuration information for the engine, run the command:

   `$DSHOME/bin/uv -admin -regen`

4. Add the following line to the dsenv file (in the $DSHOME directory):

   `LDR_CNTRL=MAXDATA=0x30000000;export LDR_CNTRL`

5. Run the dsenv command to apply the new environment settings.

6. Restart the engine:

   `$DSHOME/bin/uv -admin -start`

Troubleshooting Designer client errors

Use the information provided by the Designer client to handle errors.

Handling exceptions in the Designer client

The Designer client collects the relevant information for error reporting when an exception occurs.

Exception handling modes

The Designer client handles exceptions in one of these modes:

Automatic error report mode
> Automatic error report mode is used when an exception is unexpected for the current operation.

Optional error report mode
> Optional error report mode is used when an exception might be unexpected for the current operation. Optional error report mode is also used when it is not possible to determine if an exception is unexpected.

No error report mode
> No error report mode is used when an exception is expected, or when a message is just a warning.

Automatic error report mode

Automatic error report mode is used when an exception is unexpected for the current operation.

The Designer client creates an error report containing all the information relating to the error. The error report is a `.zip` file called `ds_errorreport_YYMMDDHHmm.zip`, where YYMMDDHHmm is the date and time of the error.

The Designer client displays an Automatic error report message:
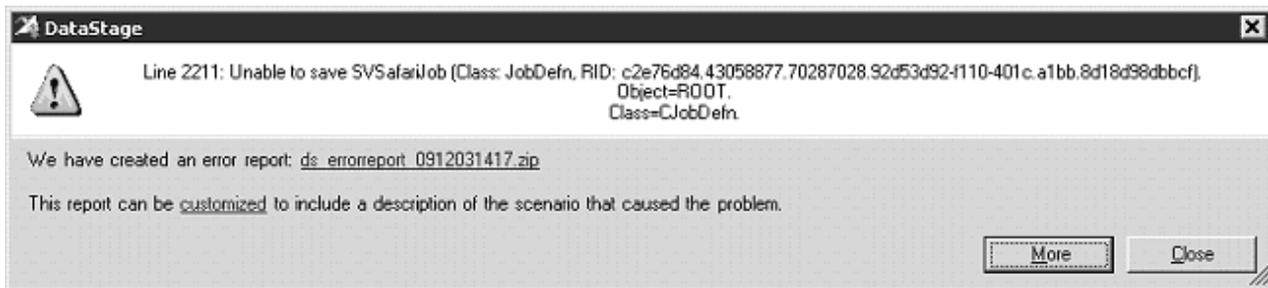
You can do the following actions on the Automatic error report message:

* Click **ds_errorreport_YYMMDDHHmm.zip** to view the directory containing the error reports using the Windows File Explorer.
* Click **customized** to open the Customize Report window where you can add a description of the scenario that caused the problem.
* Click **More** to display details of the exception and the client machine.

The `ds_errorreport_YYMMDDHHmm.zip` file contains the following information:
* the original error message
* the stack trace and exception details
* the client machine details
* the Client Version.xml file
* the associated dstage_wrapper_trace_NN.log file
* an optional user-defined description, entered on the Customize Report window

Optional error report mode

Optional error report mode is used when an exception might be unexpected for the current operation. Optional error report mode is also used when it is not possible to determine if an exception is unexpected.

The Designer client displays an Optional error report message:

```
DataStage                                                                    [X]

  /!\      Failed to authenticate the current user against the selected Services Tier

                      Invalid user name (admin) or password.

  If you wish to generate an error report for this problem, click here

                                              [ More ]   [ Close ]
```

You can do the following actions on the Optional error report message:
- Click **here** to create an error report for the exception. The Customize Report window opens, where you can add a description of the scenario that caused the problem.
- Click **More** to display details of the exception and the client machine.

No error report mode

No error report mode is used when an exception is expected, or when a message is just a warning.

The Designer client displays an No error report message:

```
DataStage                                                                    [X]

  /!\      Cannot open import file C:\temp\Copy of SVSafarioJogb.dsx (File not found)

                              [ Close ]
```
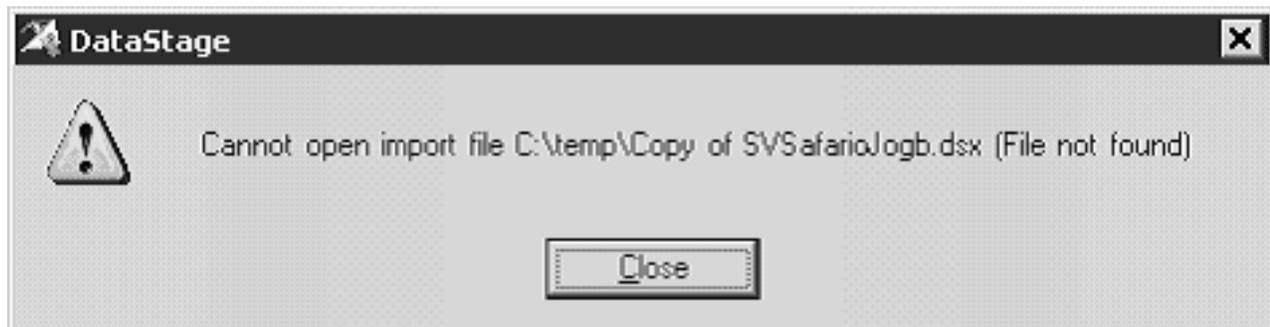
Viewing log files and error reports

View the log files and error reports that are created by the Designer client.

Viewing the current log file

Click **Help** > **Support** > **View Current Log** to view the contents of the current log file. The application that is registered to `.log` files in used.

Viewing the directory that contains the log files

Click **Help** > **Support** > **Open Log/Trace Folder** to view the directory containing the log files using the Windows File Explorer.

Viewing the directory that contains the error reports

Click **Help** > **Support** > **Open Error Reports Folder** to view the directory containing the error reports using the Windows File Explorer.

Troubleshooting a failure to submit jobs when you run a column analysis

Symptoms
You receive a "Failed to submit Job" error when you run Column Analysis.

Resolving the problem
If you require Workload Management (WLM) to run in your environment, first stop it manually with the \IBM\InformationServer\Server\DSWLM\stopwlm.cmd command. This command forces WLM to shut down if it is running. Then start WLM manually with the IBM\InformationServer\Server\DSWLM\startwlm.cmd command. WLM starts up with correct character set, which resolves the error.

If you do not require WLM, you can shut it down to eliminate the error. Edit the configuration file at \IBM\InformationServer\Server\DSODB\DSODBConfig.cfg, and change the WLMON setting to WLMON=0. Then issue the net stop DSEngine net start DSEngine command to restart the DataStage Engine without starting WLM, which resolves the error.

Troubleshooting login failures

When you encounter client login failures, you must analyze the stack trace to determine the reason why the login failed.

Client and server login failures

Login failures can happen either on the client side, or the server side.

Analyze the stack trace to determine the causes and location of the login failure.

Client side login failure
Some errors indicate failures to authenticate with selected servers.

Symptoms
You received the following error: Failed to authenticate current user against selected Domain Could not connect to server [rmanikon-2] on port [9080]

Causes
You specified an invalid server name or port.

Diagnosing the problem
The error indicates that your client cannot connect to the Information Server Services Tier (domain) server. Click the **More** option to get a stack trace for the error.

```
javax.security.auth.login.LoginException: Could not connect to server
[RMANIKON-2] on port [9081].
 at com.ascential.acs.security.auth.client.AuthenticationService.getLoginException
(AuthenticationService.java:991)
 at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.java:370)
Caused by: com.ascential.acs.registration.client.RegistrationContextManagerException:
 Caught an unexpected exception.
 at com.ascential.acs.registration.client.RegistrationContextManager.setContext
(RegistrationContextManager.java:76)
 at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.java:364)
```

```
Caused by: com.ascential.acs.registration.client.RegistrationHelperException:
 Caught an unexpected exception.
    at com.ascential.acs.registration.client.RegistrationHelper.getBindingProperties
(RegistrationHelper.java:672)
 at com.ascential.acs.registration.client.RegistrationHelper.getBindingConfigProperties
(RegistrationHelper.java:566)
 at com.ascential.acs.registration.client.RegistrationContextManager.setContext
(RegistrationContextManager.java:173)
 at com.ascential.acs.registration.client.RegistrationContextManager.setContext
(RegistrationContextManager.java:73)
    ... 1 more
 Caused by: java.net.ConnectException: Connection refused: connect
    at java.net.PlainSocketImpl.socketConnect(Native Method)
 at java.net.PlainSocketImpl.doConnect(PlainSocketImpl.java:391)
 at java.net.PlainSocketImpl.connectToAddress(PlainSocketImpl.java:252)
 at java.net.PlainSocketImpl.connect(PlainSocketImpl.java:239)
```

There are four important items to note in the stack trace. The stack trace contains no text that states `Trace from Server`, so the issue is on the client side. Look at the first highlighted message in the stack trace example. It is giving the host name and port number. The second highlighted message indicates that the error happens during the `RegistrationHelper` call. The last highlighted message indicates that the root cause is a socket connection error.

Resolving the problem

- Verify server name in DataStage login screen with a ping to the DataStage server address.
- Verify that the port number is correct
  - View theregistered-servers.xml file in the following location: `InformationServer/ASBNode/eclipse/plugins/com.ibm.isf.client_configuration_<is_version>/registered-servers.xml`
  - Find the following element: `<asb-server name="RMANIKON-2" host-name="RMANIKON-2" port="9080" is-primary-server="true" />`
  - Client has invalid entry in hosts file for DataStage server
- Check `Windows\System32\drivers\etc\hosts` for `ds_ip_address host_short_name host_long_name`

Several scenarios can cause an error where the client cannot connect to the server. The first issue can be an invalid server name or port. Verify that the server name is entered correctly in the DataStage login screen. Ping the DataStage server name from the client system to be sure that it is a valid host name.

Next, be sure that the Information Server port number you are using in the DataStage login screen is correct. Open the `registered-servers.xml` file to check the port number. Look for the string `<asb-server>` and check the port number in the `asb-server` tag.

A failure to connect to the server can also be caused by an incorrect entry in the client hosts file. Either the IP address for the DataStage server in the host file is incorrect, or the client does not have a valid entry for the server. Correct the entry in the hosts file for the DataStage server.

Server callback failure with localhost

You can determine the cause of a server callback failure by analyzing information in the stack trace.

Symptoms

Login fails, and the stack trace includes a `Trace from server` message.

Causes

WebSphere is not upgraded to the correct Java™ level.

Diagnosing the problem

Analyze the stack trace to determine the cause of this problem.

The `Trace from server` message indicates that the error is a server callback failure. The `initial and forwarded IOR inaccessible` message and client IP address are also included in the stack trace. The last message indicates that the exception is thrown from `com.ibm.rmi.iiop.CDRInputStream.read_value`. The stack trace includes `host= localhost`.

```
Trace from server: 1198777258 at host issun2 >>
org.omg.CORBA.MARSHAL: Unable to read value from underlying bridge :
 initial and forwarded IOR inaccessible :
Forwarded IOR failed with: java.net.ConnectException: Connection refused:
host=localHost ,port=33507
Initial IOR failed with: java.net.ConnectException: Connection refused:
host=localHost,port=33507 vmcid: IBM minor code: 89A completed: No
at  com.ibm.rmi.iiop.CDRInputStream.read_value (CDRInputStream.java:1993)
at
com.ascential.xmeta.shared.repository.core._EJSRemoteStatefulSandboxRemoteStatefulService
_4baa4bb1_Tie.executeQuery__CORBA_W StringValue__
CORBA_WStringValue__com_ascential_xmeta_crud_InternalQueryOptions__com_ascential
_xmeta_crud_InternalQueryCompil eOptions__java_util_Map(Unknown Source)
at com.ascential.xmeta.shared.repository.core.
_EJSRemoteStatefulSandboxRemoteStatefulService_4baa4bb1_Tie._invoke(Unknown Source)
at com.ibm.CORBA.iiop.ServerDelegate.dispatchInvokeHandler(ServerDelegate.java:614)
at com.ibm.CORBA.iiop.ServerDelegate.dispatch(ServerDelegate.java:467)
at com.ibm.rmi.iiop.ORB.process(ORB.java:439)
at com.ibm.CORBA.iiop.ORB.process(ORB.java:1761)
at com.ibm.rmi.iiop.Connection.respondTo(Connection.java:2376)
at com.ibm.rmi.iiop.Connection.doWork(Connection.java:2221)
at com.ibm.rmi.iiop.WorkUnitImpl.doWork(WorkUnitImpl.java:65)
at com.ibm.ejs.oa.pool.PooledThread.run(ThreadPool.java:118)
at com.ibm.ws.util.ThreadPool$Worker.run(ThreadPool.java:1475)
```

Resolving the problem
To resolve this issue, upgrade WebSphere Java to Java SDK 1.4.2 SR10 or later.

Server callback ORB failure
You can determine the cause of a server callback failure by analyzing information in the stack trace.

Symptoms
You receive the following unique message: `Read beyond end of data. No fragments available.`

Causes
This failure is caused by a known ORB defect.

Diagnosing the problem
The `read beyond end of data. No fragments available` message in the stack trace indicates that this problem likely caused by a known ORB defect.

```
Trace from server: 1198777258 at host green.bocaraton.ibm.com >>
org.omg.CORBA.MARSHAL: Unable to read value from underlying bridge :
 No available data: Request 18:
 read beyond end of data. No fragments available.
vmcid: IBM minor code: 89A completed: No
at com.ibm.rmi.iiop.CDRInputStream.read_value(CDRInputStream.java:1993)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.login(_EJSRemot
eStatelessAuthenticationService_e0d03809_Tie.java:146)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_Tie._invoke(_EJSRe
moteStatelessAuthenticationService_e0d03809_Tie.java:92)
at com.ibm.CORBA.iiop.ServerDelegate.dispatchInvokeHandler(ServerDelegate.java:614)
at com.ibm.CORBA.iiop.ServerDelegate.dispatch
(ServerDelegate.java:467)
at com.ibm.rmi.iiop.ORB.process(ORB.java:439)
at com.ibm.CORBA.iiop.ORB.process(ORB.java:1761)
at com.ibm.rmi.iiop.Connection.respondTo(Connection.java:2376)
at com.ibm.rmi.iiop.Connection.doWork(Connection.java:2221)
```

```
at com.ibm.rmi.iiop.WorkUnitImpl.doWork(WorkUnitImpl.java:65)
at com.ibm.ejs.oa.pool.PooledThread.run(ThreadPool.java:118)
at com.ibm.ws.util.ThreadPool$Worker.run(ThreadPool.java:1475)
```

Resolving the problem

Upgrade IBM WebSphere Application Server with interim fix: PK76826.

Bad hosts file on the server

You can determine the cause of a server-side login failure by analyzing the information in the stack trace.

Symptoms

You receive a message that indicates `An error occurred trying to instantiate an object of the entry point client`.

Causes

The hosts file on the server is formatted incorrectly.

Diagnosing the problem

Analyze the stack trace. The stack trace from the client might have an error that is caused by a bad hosts file on the server. The com.ascential.acs.security.auth.client.AuthenticationService.doLogin message indicates that the error is caused at the login time by the authentication service. The `com.ibm.ws.naming.jndicos.CNContextImpl.doLookup(CNContextImpl.java:1784` message indicates that the error occurred during the server lookup process.

```
com.ascential.asb.util.invocation.EntryPointInstantiationException:
An error occurred trying to instantiate an object of the entry point client
 implementation "com.ascential.acs.security.auth.ejb.EJBAuthenticationService"
at com.ascential.asb.util.invocation.ejb.EJBServiceCreator.
<init>(EJBServiceCreator.java:125)
at com.ascential.acs.security.auth.ejb.EJBAuthenticationService.
<init>(EJBAuthenticationService.java:39)
at com.ascential.acs.security.auth.JAASAuthenticationService.
getAuthService(JAASAuthenticationService.java:401)
at com.ascential.acs.security.auth.JAASAuthenticationService.
loginImpl(JAASAuthenticationService.java:381)
at com.ascential.acs.security.auth.JAASAuthenticationService.
login(JAASAuthenticationService.java:160)
at com.ascential.acs.security.auth.client.AuthenticationService
.doLoginImpl(AuthenticationService.java:879)
at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.java:365) Caused by: javax.naming.NamingException:
Error during resolve [Root exception is org.omg.CORBA.INTERNAL:
 initial and forwarded IOR inaccessible vmcid:
 IBM minor code: 58C completed: No]
 at com.ibm.ws.naming.jndicos.CNContextImpl.doLookup(CNContextImpl.java:1784 )
at com.ibm.ws.naming.jndicos.CNContextImpl.doLookup(CNContextImpl.java:1707)
at com.ibm.ws.naming.jndicos.CNContextImpl.lookupExt(CNContextImpl.java:1412)
at com.ibm.ws.naming.jndicos.CNContextImpl.lookup(CNContextImpl.java:1290)
at com.ibm.ws.naming.util.WsnInitCtx.lookup(WsnInitCtx.java:145)
at javax.naming.InitialContext.lookup(InitialContext.java:363)
at com.ascential.asb.util.invocation.ejb.EJBServiceCreator.
<init>(EJBServiceCreator.java:120)
```

Your /etc/hosts entry might use a single-line format, which can cause problems. The single-line format is similar to the following entry:

```
127.0.0.1 localhost.localdomain localhost machine_long_hostname machine_short_hostname
```

Resolving the problem

Separate the /etc/hosts entry to use the following double-line format:

```
127.0.0.1 localhost.localdomain localhost
<real ip address> machine_long_hostname machine_short_hostname
```

Server-rich client login failures

Server-rich clients sometimes have authentication failures that occur during login.

Analyze the stack trace to determine the causes and location of the login failure.

Server is not reachable

You are unable to connect to a server and receive a `Could not connect to server` message in the stack trace.

Symptoms

- The stack trace is from client side, and does not include the `Trace from server` message.
- The stack trace indicates that a socket connection error occurred.
- The host name and port number are included in error message.
- The exception comes from `RegistrationHelper` call.

```
Failed to authenticate the current user against the selected Domain:
Could not connect to server [purple1.bocaraton.ibm.com] on port [9080].
javax.security.auth.login.LoginException: Could not connect to server
 [purple1.bocaraton.ibm.com]on port [9080].
at com.ascential.acs.security.auth.client.AuthenticationService.getLoginException
(AuthenticationService.java:965)
at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.java:358)
Caused by: com.ascential.acs.registration.client.RegistrationContextManagerException:
Caught an unexpected exception.
at com.ascential.acs.registration.client.RegistrationContextManager.
setContext(RegistrationContextManager.java:67)
at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.java:352)
Caused by: com.ascential.acs.registration.client.RegistrationHelperException: Caught an
unexpected exception.
at com.ascential.acs.registration.client.RegistrationHelper.getBindingProperties
(RegistrationHelper.java:567)
at com.ascential.acs.registration.client.RegistrationHelper.getBindingConfigProperties
(RegistrationHelper.java:534)
at com.ascential.acs.registration.client.RegistrationContextManager.setContext
(RegistrationContextManager.java:167)
at com.ascential.acs.registration.client.RegistrationContextManager.setContext
(RegistrationContextManager.java:65)
... 1 more
Caused by: java.net.ConnectException: Connection refused: connect
at java.net.PlainSocketImpl.socketConnect(Native Method)
at java.net.PlainSocketImpl.doConnect(PlainSocketImpl.java:336)
at java.net.PlainSocketImpl.connectToAddress(PlainSocketImpl.java:201)
at java.net.PlainSocketImpl.connect(PlainSocketImpl.java:188).........
```

Causes

- Client has invalid entry in host file
- Server listening port might be blocked by a firewall
- Server is down

Resolving the problem

- Update the host file on client system so that the server host name can be resolved from client.
- Make sure the WebSphere TCP/IP ports are opened by the firewall.
- Make sure that the WebSphere Application Server is running.

Server callback failure

You are unable to connect to a server and receive a `itial and forwarded IOR inaccessible` message in the stack trace.

Symptoms

- Stack trace is from server (noted in trace as "Trace from server")
- Error happened during "initial and forwarded IOR"
- Exception comes from "com.ibm.rmi.iiop.CDRInputStream.read_value "
- The client IP is listed

```
Failed to authenticate the current user against the selected Domain:
CORBA MARSHAL 0x4942f89a No; nested exception is: org.omg.CORBA.MARSHAL:
 Trace from server: 1198777258 at host PURPLE1 >>
org.omg.CORBA.MARSHAL: Unable to read value from underlying bridge
 : initial and forwarded IOR inaccessible:Forwarded IOR failed with:
java.net.SocketException: Operation timed out:
connect:could be due to invalid address:host=10.38.86.83,
port=3953Initial IOR failed with:
java.net.SocketException: Operation timed out: connect:could be due to invalid
address:host=10.38.86.83,port=3953 vmcid: IBM minor code: 89A completed: No
at com.ibm.rmi.iiop.CDRInputStream.read_value(CDRInputStream.java:1993)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.
login(_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.java:146)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.
_invoke(_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.java:92)
at com.ibm.CORBA.iiop.ServerDelegate.dispatchInvokeHandler(ServerDelegate.java:614)
at com.ibm.CORBA.iiop.ServerDelegate.dispatch(ServerDelegate.java:467)
at com.ibm.rmi.iiop.ORB.process(ORB.java:439)
at com.ibm.CORBA.iiop.ORB.process(ORB.java:1761)
at com.ibm.rmi.iiop.Connection.respondTo(Connection.java:2376)
at com.ibm.rmi.iiop.Connection.doWork(Connection.java:2221)
at com.ibm.rmi.iiop.WorkUnitImpl.doWork(WorkUnitImpl.java:65)
at com.ibm.ejs.oa.pool.PooledThread.run(ThreadPool.java:118)
at com.ibm.ws.util.ThreadPool$Worker.run(ThreadPool.java:1475)
<< END server: 1198777258 at host PURPLE1
vmcid: IBM minor code: 89A completed: No
javax.security.auth.login.LoginException: CORBA MARSHAL 0x4942f89a No;
 nested exception is: org.omg.CORBA.MARSHAL:
```

Causes

- The client IP address is listed in the stack trace, and is not reachable from the server
- The client port is blocked

Resolving the problem

Upgrade toIBM InfoSphere Information Server Version 8.1 refresh 1 or later.

Server callback failure with localhost

You are unable to connect to a server and receive a `Connection refused:host=localHost` message in the stack trace.

Symptoms

The stack trace includes "host=localHost".

```
Trace from server: 1198777258 at host issun2 >>
org.omg.CORBA.MARSHAL: Unable to read value from underlying bridge
 : initial and forwarded IOR inaccessible:
Forwarded IOR failed with: java.net.ConnectException: Connection
 refused:host=localHost,port=33507
Initial IOR failed with: java.net.ConnectException: Connection
 refused:host=localHost,port=33507 vmcid:
IBM minor code: 89A completed: No
at com.ibm.rmi.iiop.CDRInputStream.read_value(CDRInputStream.java:1993)
at com.ascential.xmeta.shared.repository.core.
_EJSRemoteStatefulSandboxRemoteStatefulService_
4baa4bb1_Tie.executeQuery__CORBA_WStringValue
```

```
 __CORBA_WStringValue__com_ascential_
xmeta_crud_InternalQueryOptions__com_ascential
_xmeta_crud_InternalQueryCompileOptions__
java_util_Map(Unknown Source)
at com.ascential.xmeta.shared.repository.core.
_EJSRemoteStatefulSandboxRemoteStatefulService_
4baa4bb1_Tie._invoke(Unknown Source)
at com.ibm.CORBA.iiop.ServerDelegate.dispatchInvokeHandler
(ServerDelegate.java:614)
at com.ibm.CORBA.iiop.ServerDelegate.dispatch
(ServerDelegate.java:467)
at com.ibm.rmi.iiop.ORB.process(ORB.java:439)
at com.ibm.CORBA.iiop.ORB.process(ORB.java:1761)
at com.ibm.rmi.iiop.Connection.respondTo(Connection.java:2376)
at com.ibm.rmi.iiop.Connection.doWork(Connection.java:2221)
at com.ibm.rmi.iiop.WorkUnitImpl.doWork(WorkUnitImpl.java:65)
at com.ibm.ejs.oa.pool.PooledThread.run(ThreadPool.java:118)
at com.ibm.ws.util.ThreadPool$Worker.run(ThreadPool.java:1475) <<END>>
```

Causes
The IBM WebSphere Application Server SDK is outdated.

Resolving the problem
Install the WebSphere Java SDK 1.4.2 SR10 or later.

Server callback failure with a WebSphere Application Server SDK that is outdated
You are unable to connect to a server and receive a Read beyond end of data. No fragments available message in the stack trace.

Symptoms
The stack trace includes the following message. Read beyond end of data. No fragments available

```
Trace from server: 1198777258 at host green.bocaraton.ibm.com >>
org.omg.CORBA.MARSHAL: Unable to read value from underlying bridge : No available data:
Request 18:read beyond end of data. No fragments available. vmcid: IBM minor code: 89A
completed: No
at com.ibm.rmi.iiop.CDRInputStream.read_value(CDRInputStream.java:1993)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_
Tie.login(_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.java:146)
at com.ascential.acs.security.auth.server.
_EJSRemoteStatelessAuthenticationService_e0d03809_
Tie._invoke(_EJSRemoteStatelessAuthenticationService_e0d03809_Tie.java:92)
at com.ibm.CORBA.iiop.ServerDelegate.dispatchInvokeHandler(ServerDelegate.java:614)
at com.ibm.CORBA.iiop.ServerDelegate.dispatch(ServerDelegate.java:467)
at com.ibm.rmi.iiop.ORB.process(ORB.java:439)
at com.ibm.CORBA.iiop.ORB.process(ORB.java:1761)
at com.ibm.rmi.iiop.Connection.respondTo(Connection.java:2376)
at com.ibm.rmi.iiop.Connection.doWork(Connection.java:2221)
at com.ibm.rmi.iiop.WorkUnitImpl.doWork(WorkUnitImpl.java:65)
at com.ibm.ejs.oa.pool.PooledThread.run(ThreadPool.java:118)
at com.ibm.ws.util.ThreadPool$Worker.run(ThreadPool.java:1475)
<< END server: 1198777258 at host green.bocaraton.ibm.com
```

Causes
The ORB data is fragmented, which is a known issue.

Resolving the problem
- Upgrade to Version 8.1 Refresh 1 or later.
- Install IBM WebSphere Application Server interim fix PK77267

Invalid access information from server

You are unable to connect to a server and receive a `Read beyond end of data. No fragments available` message in the stack trace.

Symptoms

- The stack trace is from client side because it has no `Trace from server` message.
- The exception comes from `com.ibm.ws.naming.jndicos.CNContextImpl.doLookup`, which is the second level on stack trace.
- The exception comes from `com.ascential.asb.util.invocation.ejb.EJBServiceCreator`, which is the third level on stack trace.
- The first method of the stack trace is `com.ascential.acs.security.auth.client.AuthenticationService.doLogin`.

```
com.ascential.asb.util.invocation.EntryPointInstantiationException:
An error occurred trying to instantiate an object of the entry point client
 implementation "com.ascential.acs.security.auth.ejb.EJBAuthenticationService"
at com.ascential.asb.util.invocation.ejb.EJBServiceCreator.<init>
(EJBServiceCreator.java:125)
at com.ascential.acs.security.auth.ejb.EJBAuthenticationService.<init>
(EJBAuthenticationService.java:39)
at com.ascential.acs.security.auth.JAASAuthenticationService.getAuthService
(JAASAuthentication
Service.java:401)
at com.ascential.acs.security.auth.JAASAuthenticationService.loginImpl
(JAASAuthenticationService.
java:381)
at com.ascential.acs.security.auth.JAASAuthenticationService.login
(JAASAuthenticationService.
java:160)
at com.ascential.acs.security.auth.client.AuthenticationService.doLoginImpl
(AuthenticationService.
java:879)
at com.ascential.acs.security.auth.client.AuthenticationService.doLogin
(AuthenticationService.
java:365)
Caused by: javax.naming.NamingException: Error during resolve [Root exception is
org.omg.CORBA.INTERNAL: initial and forwarded IOR inaccessible vmcid: IBM minor code:
58C completed: No]
at com.ibm.ws.naming.jndicos.CNContextImpl.doLookup(CNContextImpl.java:1784)
at com.ibm.ws.naming.jndicos.CNContextImpl.doLookup(CNContextImpl.java:1707)
at com.ibm.ws.naming.jndicos.CNContextImpl.lookupExt(CNContextImpl.java:1412)
at com.ibm.ws.naming.jndicos.CNContextImpl.lookup(CNContextImpl.java:1290)
at com.ibm.ws.naming.util.WsnInitCtx.lookup(WsnInitCtx.java:145)
at javax.naming.InitialContext.lookup(InitialContext.java:363)
at com.ascential.asb.util.invocation.ejb.EJBServiceCreator.
<init>(EJBServiceCreator.java:120)
... 6 more
Caused by: org.omg.CORBA.INTERNAL: initial and forwarded IOR inaccessible vmcid: IBM
minor code: 58C completed: No
at com.ibm.rmi.corba.ClientDelegate.createRequest(ClientDelegate.java:1213)
at com.ibm.CORBA.iiop.ClientDelegate.createRequest(ClientDelegate.java:1320)........
```

Causes

- Some Linux computers automatically configure the host file with the following entry: `127.0.0.1 localhost.localdomain local host machinelonghostname machineshorthostname`
- The server has more than one IP address, and one IP address is not accessible from the client

Resolving the problem

Ensure that the host name of each endpoint in WebSphere Application Server is resolved as a client-accessible IP address.

The WebSphere Application Server endpoint configuration can be found from WebSphere administrator console: Servers -> Application servers -> server1 -> Ports. The server name that is specified on the endpoint must be resolved as a client-accessible IP address. IP address 127.0.0.1 or IP address 192.168.x.x are normally not accessible.

Troubleshooting job design issues

Some problems in InfoSphere DataStage can be corrected with changes to the job design.

IBM InfoSphere DataStage Error: Job xxx is being accessed by another user

A job can be accessed only by one user at a time.

Symptoms

You are unable to view a job, and receive the following error message.

`Error: Job xxx is being accessed by another user`

Causes

The job that you are trying to view is accessed by another user.

Resolving the problem

Perform the following actions:

1. Check the active sessions with the web console. If the job is listed, disconnect the session. In the web console, click the Administration tab. In the Navigation pane, select Session Management > Active Sessions. The Active Sessions pane shows the users that are currently connected to the server.
2. Check for universe locks and clear them.
   a. Start the DataStage Director
   b. Go to Select Job > Cleanup Resources. If the Cleanup Resources option is disabled, skip to step i, where you start the DataStage Administrator client.
   c. In the Processes pane, click Show All
   d. In the Locks pane, click Show All
   e. In the Locks pane, scroll to the job name in the **Item ID** field
   f. Note the PID/User # associated with the job
   g. Click the PID # in the upper pane (Processes)
   h. Click Show by process (Locks pane)
   i. Start the DataStage Administrator
   j. In the Projects tab, highlight the job
   k. Click Properties
   l. Check Enable job administration in Director
   m. Click OK
   n. Click close
   o. Exit DataStage Director and relaunch
   p. Repeat steps C through I.
   q. Log in to the server as the dsadm user
   r. cd to the DSEngine directory
   s. Enter ./dsenv to source the dsenv file
   t. Enter ./bin/uvsh to get into DataStage prompt
   u. At ">" DataStage engine prompt, enter LOGTO project name
   v. Run LIST.READU EVERY to list all the locks

w. Check active record locks under "Item Id" column for job name or RT_CONFIG# or RT_LOG# (# matches the job description number

x. Write down the Inode numbers and user numbers that are associated with these locks

y. Enter LOGTO UV. If the LOGTO command is disabled, enter the following command:

```
CHDIR path_to_the_DSEngine_folder
```

The UNLOCK command lives in the UV account.

z. Enter UNLOCK INODE inode# USER user# ALL

aa. You can use Q to get out of DataStage engine

3. Use the `xmetaAdmin cleanupAbandonedLocks` utility to clear any abandoned locks. The `xmetaAdmin cleanupAbandonedLocks` utility deletes session locks from the Information Server repository that were left over from some usage of an Information Server suite application such as DataStage. Log in to the services tier as the root user on UNIX or Linux or as an Administrator user on Windows.

On UNIX or Linux:

```
cd /opt/IBM/InformationServer/ASBServer/bin
```

```
./xmetaAdmin.sh cleanupAbandonedLocks -dbfile ../conf/database.properties
```

On Windows:

```
cd c:\IBM\InformationServer\ASBServer\bin
```

```
xmetaAdmin.bat cleanupAbandonedLocks -dbfile ..\conf\database.properties
```

DataStage Parameter Set - Parameter Set locked by non-existent user

You are unable to edit the parameter set.

Symptoms

The user receives the following error message.

```
Unable to edit the Item Properties The Parameter Set <parametersetname>
 is locked by user <username>.  The user was no longer connected to DataStage.
No session found in the web console for the user.
```

Causes

The user was no longer connected to DataStage. No session found in the web console for the user.

Resolving the problem

To clear an abandoned lock, run the `xmetaAdmin cleanupAbandonedLocks` utility in the `./IBM/InformationServer/ASBServer/bin` directory. Log in to the services tier as the root user on UNIX or Linux or as an Administrator user on Windows.

On UNIX or Linux:

```
cd /opt/IBM/InformationServer/ASBServer/bin
```

```
./xmetaAdmin.sh cleanupAbandonedLocks -dbfile ../conf/database.properties
```

On Windows:

```
cd c:\IBM\InformationServer\ASBServer\bin
```

```
xmetaAdmin.bat cleanupAbandonedLocks -dbfile ..\conf\database.properties
```

Cannot get exclusive access to the log for a job.

If you cannot get exclusive access to a log for a job, you are unable to delete the job.

Symptoms

The user deletes a job from the DataStage Designer, and receives the following error:

```
Unable to delete the item(s). Delete object for '\<path>\<jobname>' failed.
 Cannot get exclusive access to log for job <jobname>
```

Causes

A lock remains on the RT_LOG file for the job.

Resolving the problem

Make sure that the job is not running. If the job is running, stop the job if possible. From DataStage Director, select the job. Then click the **Job** menu and select **Reset**. Try to delete the job again from the Designer.

If the delete still fails, manually delete the lock from uvsh with the following instructions where '$' is the OS prompt, and '>' is the uvsh prompt. That is, you do not type them.

1. Change the directory to **$DSHOME**, which is `./IBM/InformationServer/Server/DSEngine` by default. Source the dsenv file with the following command: `$ . ./dsenv`
2. Start uvsh with the `$ bin/uvsh` command. In uvsh, the shift key is not needed to produce uppercase character, but it is used to produce lowercase characters. The role of the shift key is reversed from normal. As a result, copy-and-paste behaves strangely and is not useful.
3. At the UV prompt, log on to the project. The project name is case-sensitive. Use the following command: `LOGTO projectname`
4. Find out what the job description number is with the `LIST DS_JOBS jobname` command.
5. List the active locks with the `LIST.READU EVERY` command. Check the active record locks under the Item Id column for the job name. You can also use `RT_CONFIG#`or `RT_LOG#` where # is the job description number.
6. Write down the inode number and user number for the lock that is not a valid lock.
7. Log in to the UV account where the `UNLOCK` command lives.
8. Run the following command: `UNLOCK INODE inode# USER user# ALL`
9. Type Q to logout of uvsh.
10. Delete the job from the Designer.

Troubleshooting problems when you create InfoSphere DataStage projects

You can encounter failures during the installation of IBM InfoSphere Information Server with the Administrator client or the `dsadmin` command.

Symptoms

Creating a project fails during the installation of InfoSphere Information Server, when you use the Administrator client, or when you use the dsadmin command.

Causes

- Incorrectly configured repository database
- Leftover metadata in repository database from a previously failed project creation
- Unable to create log file on the InfoSphere DataStage Server
- Incorrectly configured locale on the InfoSphere DataStage Server
- Incorrectly configured locale on the InfoSphere DataStage Server
- Failed to load JVM into theInfoSphere DataStage Server process (dsapi_slave)
- Firewall configuration
- Trusted authentication between InfoSphere DataStage Server system and the Domain system failed
- InfoSphere DataStage was not installed on the Domain system
- Locale regional settings that are customized on the Client system

- A full disk, partition, or user quota that is reached on InfoSphere DataStage Server system
- Project creation fails at `Initializing demo files...` within the Administrator client
- Stack Execution Disable (SED) is enabled on AIX
- Unable to increase the table space for the XMETA
- Error updating secondary indexes

Diagnosing the problem

The `dstage_wrapper_trace_N.log` indicates where the problem lies. These log files might require customer support, but some errors can be interpreted to and diagnosed.

Locate the error message. The error message location is based on how the project creation was started.

Installation
> Search in the installation logs at `/opt/IBM/InformationServer/logs/` for RUN BP DSR_QUICKADD.B. The error message is a few lines that follow this text.

Administrator client
> The error message is displayed in a message box.

dsadmin command
> The error message is displayed on the console.

After you locate the error message, you can find more detailed error information in the log file on the InfoSphere DataStage server system. You can find that log at the following file path: `/home/_Credential_Mapped_Username_/ds_logs/dstage_wrapper_trace_N.log`

- *N* is a number 1 - 20 for the log files for the last 20 InfoSphere DataStage sessions
- *_Credential_Mapped_Username_* is the InfoSphere DataStage Server user name that credentials are mapped to in web Administration console.

These logs go to *${user.home}*/`ds_logs` where *${user.home}* is the home directory of the credential mapped user or the user that runs the installation for projects that are created during installation. On Windows systems, the user home directory is `C:\Documents and Settings\_Credential_Mapped_UserName_`

Sometimes tracing the dsrpcd and child processes can produce useful information from which IBM support can narrow down the causes of a project creation problem. Operating systems differ in how to trace processes but a few examples are shown here:

```
AIX: "truss -o /tmp/dsrpcd_truss.out -fp <dsrpcdPID&gt;"RedHat:
"strace -fp <dsrpcdPID&gt; &gt; /tmp/dsrpcd_strace.out"
```

These commands attach to the running dsrpcd process and record all of the system calls that are made by that process and its child process during subsequent client/server sessions. For example, a call to create a project from the Administrator client or the dsadmin command line is recorded.

To produce extra diagnostic information for the JVM initialization after all of its libraries are successfully loaded, you can enable JVM startup tracing. Add the following lines to `/opt/IBM/InformationServer/Server/DSEngine/dsenv`:

```
XMOG_TRACE_LEVEL=TraceVerbose
XMOG_TRACE_FILE=/tmp/xmogtrace.txt
export XMOG_TRACE_LEVEL
export XMOG_TRACE_FILE
```

On Windows, these tracing options can be set as System Environment Variables by using the System Control panel. Remember to restart the InfoSphere DataStage Server engine processes after you add these variables, and to remove these environment variables after they are no longer needed.

Enable repository database tracing. To enable tracing of the code that populates the repository database, follow these steps:

1. Create a file on the InfoSphere DataStage Server system in /opt/IBM/InformationServer/ASBNode/ conf/ called NewRepos.debug.properties. The file name is case-sensitive.

2. In the file, add the following three lines:

```
log4j.logger.com.ascential.dstage=DEBUG
log4j.logger.com.ibm.datastage=DEBUG
NewRepos.spy.trace=true
```

The dstage_wrapper_trace_N.log contains extra tracing information the next time that you create a project. Ensure that you delete the NewRepos.debug.properties file when finished. In addition, spy trace files, such as dstage_wrapper_spy_N.log, are produced in the same directory as the log files. These files contain a detailed record of low-level method calls and can grow large.

Running project creation manually. The project creation code runs in the context of a dsapi_slave process that does not have any console output. Locate the full RUN BP DSR_QUICKADD.B command line from the domain installation log files on /opt/IBM/InformationServer/logs/.

Use the following commands to run the project creation code so that you can view the console output:

Linux and UNIX

1. `cd /opt/IBM/InformationServer/Server/DSEngine`
2. `./dsenv`
3. `bin/uvsh`
4. `RUN BP DSR_QUICKADD.B <<i>arguments from log file</i>>`
5. `QUIT`

Windows

1. `cd /opt/IBM/InformationServer/Server/DSEngine`
2. `bin/uvsh`
3. `RUN BP DSR_QUICKADD.B <arguments from log file> <newProjectName> C:\IBM\InformationServer\ Server\Projects\<newProjectName> CREATE`
4. `QUIT`

Resolving the problem

Incorrectly configured repository database.

> 8.0.x message: `Error creating DR elements, Error was -1", "Invalid node name: %1`

> 8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was Unique constraint violation.`

> These types of errors usually occur because the repository database returns an error when you attempt to make an update. The dstage_wrapper_trace_N.log file might contain more specific details about the exact database error.

> There might also be a database log, depending on what type of database the repository is running in, which contains more information. DB2® has the db2diag tool, which can be run to find out the exact reason why an update failed. The typical failures are running out of disk space, and memory configuration problems.

> For repository database errors it is important to confirm that the database was created with the scripts supplied on the installation media. These scripts configure important database parameters that might cause project creation problems.

It is also important that the database was created with the correct character set. The database creation script documentation on the installation media typically specifies UTF16/32. If a different character set was used, some of the metadata that is stored can become corrupted or might cause unexpected primary key violations. If the wrong character set was used, the product must be reinstalled.

For errors at this level the WebSphere Application Server logs might contain more information. The files `SystemOut.log` and `SystemErr.log` can be found in the following directory: `...WebSphere/AppServer/profiles/_profile_name_/logs/server1/`

Leftover metadata in repository database from a previously failed project creation

8.0.x message: `Error creating DR elements, Error was -1`

This problem occurs only on 8.0.x systems and can be identified by looking in the `dstage_wrapper_trace_N.log` file for a `unique constraint violation` error. This can occur when a project creation failed and did not remove all of its metadata from the repository. Even though the project cannot be seen in InfoSphere DataStage, attempting to create a project of the same name results in this error.

To work around this problem, you can create a project with a different name. Alternatively, IBM support can provide a tool and instructions for how to remove the leftover data from the repository.

Unable to create log file on the DataStage Server

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was log4j:ERROR setFile(null,true) call failed.`

Just before the metadata repository is populated with the default project contents, a log file is created on the InfoSphere DataStage Server system in `/home/_Credential_Mapped_Username_/ds_logs/`. If this log file cannot be created, the project creation fails. On Windows computers, the user home directory is at `C:\Documents and Settings\_Credential_Mapped_UserName_`

The usual reasons why this log was not created are either because the user has no home directory at all or they do not have appropriate permissions on it.

Incorrectly configured locale on the InfoSphere DataStage Server

8.0.x message: "Error creating DR elements, Error was -1"

8.1 and later message: "DSR.ADMIN: Error creating DR elements, Error was Unmatched quotation marks"

This problem is ultimately caused by bad locale configuration on the DataStage Server system. This problem manifests itself because the host name command is run during project creation. Instead of returning the correct host name, the command returns an error that the process `couldn't set locale correctly`.

Failed to load JVM into the InfoSphere DataStage Server process (dsapi_slave)

8.0.x message: `The connection is broken (81002)`

The JVM (Java virtual machine) can fail to load for several reasons. If it does fail to load, the dsapi_slave process is stopped, which results in broken connection errors on the client such as error 81002.

A core file might be produced which can be used to determine what caused the process to be stopped.

Possible causes of this problem are:

- The LIBPATH (or equivalent) is too long and caused a buffer overflow. This can be confirmed by using the Administrator client to run the `env` command with the **Command** button. If the

contents of LIBPATH are duplicated, then it is probable that dsenv are sourced twice. The dsenv does not need to be sourced when it starts the InfoSphere DataStage Server engine processes with the `uv -admin -start` command.

- Incompatible or missed patches on the Client, Server, and Domain systems. By looking in the `version.xml` file of each system you can confirm what patches are installed. Ensure that patches are installed on all appropriate systems.
- Environment variables such as LDR_CNTRL were added or modified in the `IBM/InformationServer/Server/DSEngine/dsenv` file. Generally speaking LDR_CNTRL settings in dsenv must not be modified unless otherwise directed by IBM.
- Incompatible operating system kernel parameters.

Firewall configuration

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was com.ascential.xmeta.exception.ServiceException`

The InfoSphere DataStage Server system must communicate with the domain system, and certain ports must be open between these systems if they are on separate servers.

This problem can be confirmed by looking in the `dstage_wrapper_N.log` file for the following error: `Connection refused:host=<hostname>,port=2809`. Ensure that the firewall is correctly configured and use telnet <hostname> <port> from the InfoSphere DataStage Server system to confirm that the port is accessible.

The necessary firewall configuration can be found in the installation guide.

Trusted authentication between InfoSphere DataStage Server system and the Domain system failed

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was Mapping failed to copy attributes: MetaTable -> DSTableDefinition (EObject: null, MetaTable)`

The InfoSphere DataStage Server system authenticates with the Domain system by a process called trusted authentication. This process uses a secure certificate exchange rather than explicit user name and password authentication. If the process fails, the project creation fails. Trusted authentication failure is identified by multiple exceptions in the InfoSphere DataStage Server ds_logs that says "Null session".

This can fail for a number of reasons:

- If the InfoSphere DataStage Server is installed onto a Windows system at `C:\IBM\InformationServer`. If you install the clients into a different directory, the certificate exchange might fail. If the certificate exchange fails, the project creation fails. See technote #1409412 and APAR JR34441 for more information.
- The number of trusted sessions reaches a maximum limit, so a new session cannot be started. This is identified by an entry in the WebSphere logs that says the limit is exceeded. If so, restarting WepSphere Application Server clears everything so that new sessions can be created and project creation can succeed.

DataStage was not installed on the Domain system

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was Mapping failed to copy attributes: MetaTable -> DSTableDefinition (EObject: null, MetaTable)`

When you install the Domain and InfoSphere DataStage Server onto different physical servers, the installation of DataStage Server fails to create projects that are specified in the installer if InfoSphere DataStage is not installed onto the Domain. These errors can be found in the installation logs.

Furthermore, attempting to create projects that use the Administrator client or command line fails. In both these cases, the exceptions state that `The package with URI "http:///1.1/ DataStageX.ecore" is not registered.`

InfoSphere DataStage can be added to the Domain system by rerunning the installer, selecting InfoSphere DataStage and clearing the other components.

Locale or regional settings that are customized on the Client system

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `Invalid Node Name %1`

If the regional language settings are modified to use a customized short date format (for example ddd dd/MM/yyyy) it can cause the DataStage Administrator client to send the wrong date information to the DataStage Server, causing project creation to fail. A patch for this issue is available under APAR JR34770.

Full disk, partition, or user quota that is reached on InfoSphere DataStage Server system

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, error was log4j: ERROR failed to flush writer.`

The project creation operation creates a log file on the InfoSphere DataStage Server system, called dstage_wrapper_trace_N.log, in the path indicated at the beginning of this document. The log creation fails when the disk partition is full or the user to which credentials are mapped to reaches their disk quota. Free up space as necessary and try the operation again.

Project creation fails at `Initializing demo files...` within the Administrator client

8.5 message: `Errors were detected during project creation that might render project <name> unstable. Caused by: DSR.ADMIN: Error creating DR elements, Error was <date timestamp> java.utils.prefs.FileSystemPreferences$2 run.`

This error states that there was a problem with being able to write Java preference data.

One of the following items causes these problems:
- SE (Security Enhanced) Linux is enabled.
- The user ID that is trying to create the project does not have a local home directory to write to

If SELinux is enabled, disable it.

To determine if SELinux is installed and in enforcing mode, you can do one of the following actions:
- Check the /etc/sysconfig/selinux file
- Run the `sestatus` command
- Check the /var/log/messages file for SELinux notices (Notice format might differ between RHEL 4 and RHEL 5.)

To disable SELinux, you can do one of the following actions:
- Set it in permissive mode and run the `setenforce 0` command as a superuser
- Modify /etc/sysconfig/selinux and restart the system

If there is no home directory for the user ID, create a local home directory with write permissions. (766) and have the group as part of the local dstage group.

Stack Execution Disable (SED) is enabled on AIX

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `Unable to confirm the JVM can be loaded into the DataStage server process 'DSR_CREATE.PROJECT.B TestJVM' failed`

If Stack Execution Disable (SED) is enabled in AIX, the JIT compiler fails when you try to run code it generated in the process data area. This error occurs with all of the InfoSphere DataStage executable items that include embedded JVMs. The solution to this problem is to turn off the SED at the system level and restart the system. To turn off the SED, use the `sedmgr -m off` command.

Unable to increase the table space for the XMETA

8.0.x message: `Error creating DR elements, Error was -1`

8.1 and later message: `DSR.ADMIN: Error creating DR elements, Error was unable to save`

The `SystemOut.log` might show the following error:
```
Oracle¨ORA-01653: unable to extend table
XMETAUSER.LOGGING_LOGGINGEVENT1466CB5F by 128 in tablespace XMETA
```

If DB2 is used for the XMETA look in the `<db2instance_home>/sqllib/db2dump/db2diag.log` file for errors.

To resolve this problem, increase the table space, and try the operation again. It might be necessary to manually delete any partially created project, which can be done by following the material here: http://www-01.ibm.com/support/docview.wss?uid=swg27021312

Error updating secondary indexes

Error message: `DSR.ADMIN: Error updating secondary indexes. Status code = -135 DSJE_ADDPROJECTFAILED`

A known cause for the error `error updating secondary indices` is one or more missing `I_*` directories in the `/opt/IBM/InformationServer/Server/Template` directory. If there is another InfoSphere DataStage engine installation with the same version and patch level available, it is possible copy the `Template` directory from the working engine and use it to replace the `Template` directory on the broken engine. However, backup the existing Template directory first. If a template directory is taken from a working engine of a different patch level, some of the patches on the broken engine might be rendered ineffective.

Troubleshooting job failures

Jobs can fail to start for several reasons, but you can determine the cause and resolve the problem.

Low system resource issues

Some job failures are caused when the host system runs out of resources.

Resources unavailable to create a InfoSphere DataStage parallel job

InfoSphere DataStage is unable to create parallel jobs when system resources are limited.

Symptoms

You receive the following message
```
DataStage parallel job fails with fork() failed, Resource temporarily unavailable
```

Causes

This error occurs when the operating system is unable to create all of the processes that are needed for the job at run time. Unfortunately, the exact reason for the failure is not available. This problem occurs on UNIX and Linux operating systems for the following reasons:

- The maximum process limit is reached
- The kernel or the maximum open file limit is reached
- The swap space allocation or pre-allocation is exceeded

Resolving the problem

The exact methods for identifying and modifying process and file limits vary between the different UNIX and Linux operating systems. The system administrator for the operating system must be engaged to assist with this.

If you are unable to adjust these limits, or if the limits are at maximum for your environment and the error still occurs, then the job run time is too large for this environment and you must scale back the job run time. This might reduce performance, but allow the job to complete. To reduce the number of processes, use the following methods:

- Reduce the number of logical nodes in the APT_CONFIG_FILE
- Ensure that APT_DISABLE_COMBINATION is not set

You can view the current value of the *maxuproc* variable on AIX with the following command: `lsattr -E -l sys0 | grep maxuproc`

A reasonable setting for environments that are running large jobs would be MAXUPROC = 1000. To optimize this value, you can monitor the number of processes over time that are running daily, and then set the value appropriately. Here is some sample shell script code that you can use to monitor the number of processes that belongs to the 'dsadm' user. The script loops 365 times and take a measurement every 5 seconds.

```
#!/bin/sh
COUNTER=360
rm dsadm_count.txt

until [ $COUNTER -lt 0 ]; do
let COUNTER-=1
sleep 5
date >>dsadm_count.txt
ps -ef | grep dsadm |wc -l >> DSADM_uproc_values.txt
done
```

There are special considerations for the Windows operating system. Read the related technotes for tuning Windows environments for large jobs. Tuning Windows Environments

InfoSphere DataStage error "The section leader on xxx died"

This error occurs when resources are temporarily unavailable and the conductor process times out.

Symptoms

TheInfoSphere DataStage job fails with the following message:

`main_program: The section leader on xxx died.`

Causes

This `section leader xxx died` error is related to temporary resource non-availability. The conductor process is timing out because it did not receive an acknowledgment from the player process that it started successfully.

Diagnosing the problem

Review the system logs to determine if any system issues occurred when the InfoSphere DataStage job ran.

Resolving the problem

You can avoid this problem by setting the *APT_PM_NODE_TIMEOUT* environment variable to a higher value. The *APT_PM_NODE_TIMEOUT* environment variable controls the number of seconds that the conductor waits for a section leader to start and load a score before you determine that something failed. Setting this environment variable to a higher value allows the conductor to wait longer before it times out, and might help to avoid this error in the future. The default for starting a section leader process is 30. The default for loading a score is 120.

Set the *APT_PM_NODE_TIMEOUT* to 300. This resolution might resolve the error message but might not resolve availability of resources on the system. See this note from the 'Parallel Job Advanced Developers Guide' on this environment variable: *APT_PM_NODE_TIMEOUT*

The need for long timeouts in the job startup process shows that the engine tier hardware is approaching overload. It is better to run fewer concurrent jobs order to keep startup times low.

Error setting up the internal communications (fifo RT_SCTEMP/jobName.fifo) file
Symptoms
InfoSphere DataStage jobs fail with the following error message:

```
Message: Error setting up internal communications (fifo RT_SCTEMP/jobName.fifo)
```

DataStage is unable to create, delete, read, or write a temporary fifo file for a job to the RT_SCTEMP directory within the project that owns the job.

Causes
The error occurs for the following reasons:
- InfoSphere DataStage cannot process the file because they are locked.
- InfoSphere DataStage cannot process the file because of inadequate file permissions or other file system problems.
- Virus Scan or backup program interferes with writing fifo files to temp or scratch directories, which is described in the following technote: https://www-304.ibm.com/support/docview.wss?uid=swg21445893&wv=1

Resolving the problem
If the failure is caused by locks, then the error must have more text at the end or in subsequent message that states the lock status, for example:

```
Error setting up internal communications (fifo RT_SCTEMP/MyTestJob.fifo
 LOCKED STATUS () -1); file is locked.
```

In that situation, read the following technote for instructions on how to clear locks for a job: https://www-304.ibm.com/support/docview.wss?uid=swg21438482

If the issue is not caused by locks, then review the following checklist to resolve other common causes for this error:
- Check the file limits at job run time, especially if all jobs run under a common user ID such as DSADM. You can check the limits that are used at job runtime even if you cannot run jobs, by running the command through the Administrator client. Log in to the Administrator client, select the failing project, click the **COMMAND** button, and then enter the following command: `sh ulimit -a` If the command returns a number that is less than 2048, consider increasing it. On busy systems the limit must be higher. In this situation, you can add command to set the limit to `$DSHOME/dsenv` script. For example, `ulimit -n 10240` increases the limit to 10240. After you change the limit, you must stop and restart InfoSphere DataStage. Perform the test again to ensure that new limit is in effect.
- Check available space on the volume that contains the `RT_SCTEMP` directory. If the project that contains failing jobs is named "MyProject", then the path to RT_SCTEMP is similar to the following path: `/opt/IBM/InformationServer/Server/Projects/MyProject/RT_SCTEMP`
- Check permissions for the RT_SCTEMP directory and the files inside it. You can determine which user ID is running jobs from the event messages in the job log. Ensure that the user ID has permission to read and write to the directory and the files within it either, The ID must have permission through the ownerid, group membership, or public permissions. A quick test to confirm if permissions are the problem is to set directory permissions temporarily to 777 so that all users can write to it.
- Confirm that DSADM or the user ID that runs the failing jobs can create a file in this directory using the following steps:
  - Log in to the operating system as the user ID that runs the failing jobs.

- Change the directory to `InformationServer/Server/Projects/projectname/RT_SCTEMP` directory.
- Enter the following command: `touch test.fifo`
- If the above command fails, then the user ID is unable to create a file at that location and that issue must be resolved beforeInfoSphere DataStage jobs can run correctly.

If this issue is not caused by locks, then the InfoSphere DataStage error occurs when it is unable to correctly create, read, write, or delete the temporary fifo files. If the tests do not isolate the cause of file system i/o problem, then it might be necessary to run a system trace of the dsapi process that starts the failing jobs. Use the system trace to determine which operating system processes that are failing.

InfoSphere DataStage job fails with a `Could not map table file` message

Jobs start to fail when memory is fragmented, or the amount of data that is used in the Lookups exceeds its limit.

Symptoms
Information Server or InfoSphere DataStage job fails with the following message:

`Could not map table file`

DataStage fails when it loads lookup data into memory or creates a lookup file.

Causes
Other applications might be running concurrently with resources that are no longer available to InfoSphere DataStage. Available memory might be reconfigured from creating or moving lpars. The **
LDR_CNTRL environment setting on AIX might limit the *ulimit -d (data)* setting even if you have hard limit set higher.

InfoSphere DataStage is limited to the amount of memory that can be allocated for a Lookup.

A single Lookup stage in designer has multiple Lookup inputs. These stages analyze the corresponding number of Lookup operators in the generated osh script. When the operator compatibility is optimal, each Lookup operator has one physical process for each partition that is defined by the configuration file. Each physical process can address only up to 2 GB of memory because it is a 32-bit application. The Windows version of the InfoSphere DataStage Parallel Engine is available only with 32-bit pointers.

Each lookup requires contiguous memory allocation. Each process is limited to the *ulimit* setting of the InfoSphere DataStage Environment that can be limited by LDR_CNTRL on AIX. Each lookup data set uses the entire partitioning method by default. With the entire partitioning method, one memory segment is used and shared across all partitions for a physical server. The method is defined by the *fastname* option in the configuration file.

For MPP environments, each server has an individual copy of a memory segment. If you use a method other than entire or auto partitioning, each partition uses its own copy of data in memory and only up to 2 GB or *ulimit -d (data)***. This method is the most restraining. All lookup data is processed to a file in scratch and then loaded to a mmap structure in memory. The mmap function is a C++ function. Allocation of this structure requires contiguous memory, and happens before any source data is processed for the lookup.

Resolving the problem
You must design the job to be more scalable. By default the Lookup stage uses Entire partitioning for the Lookup data. This ensures that no matter what partition the source data is in, all the Lookup data is available to be matched.

Use hash partitioning. Create a hash on the lookup keys for both the source and lookup data. As the amount of lookup data increases, add more nodes to the configuration file to distribute data across more processes and more memory segments.

Parallel startup failed for job runs on multiple nodes across multiple servers

Parallel jobs can fail from configuration errors.

Symptoms

A parallel InfoSphere DataStage job with configuration file that is set up to run multiple nodes on a single server fails with the following error:

```
Message: main_program: **** Parallel startup failed ****
```

Causes

The full text for this `parallel startup failed` error provides some additional information about possible causes of the problem. The problem is often caused by one of the following configuration errors:

This problem is caused by configuration errors

- The Orchestrate® installation directory is not properly mounted on all nodes.
- The rsh permissions are not set correctly with `/etc/hosts.equiv` or `.rhosts`.
- The job runs from a directory that is not mounted on all nodes

The messages in the server log that precede the `startup failed` message contain more information about the cause of the failure.

For the situation where a site is attempting to run multiple nodes on multiple servers, the statement above is correct. More information about setting up ssh or rsh and parallel processing can be found in the following topics:

- Configuring remote and secure shells
- Configuring a parallel processing environment

In the case where all nodes are running on a single server, the "Parallel startup failed" message is usually an indication that the fastname defined in the configuration file does not match the name output by the `hostname` command on the server. In a typical node configuration file, the server name where each node runs is indicated by the fastname in the `/opt/IBM/InformationServer/Server/Configurations/default.apt`

```
{
node "node1"
{fastname "server1"
pools ""
resource disk "/opt/resource/node1/Datasets" {pools ""}
resource scratchdisk "/opt/resource/node1/Scratch" {pools ""}
}
node "node2"
{
fastname "server1"
pools ""
resource disk "/opt/resource/node2/Datasets" {pools ""}
resource scratchdisk "/opt/resource/node2/Scratch" {pools ""}
}
}
```

Diagnosing the problem

Log in to the InfoSphere DataStage server and at the operating system command prompt, enter `hostname` command.

If the host name output EXACTLY matches the fastname that is defined for local nodes, then the job runs correctly on that server. However, if the `hostname` command outputs the host name in a different format (such as with domain name appended) then the names that are defined for fastname are considered remote nodes and a failed attempt is made to access the node through rsh or ssh.

Using the example above, if the output of the `hostname` command was *server1.mydomain.com*, then before the `Parallel startup failed` error in job log you likely see the following error: `Message: main_program: Accept timed out retries = 4 server1: Connection refused`

This problem occurs even if your `/etc/hosts` file maps *server1* and *server1.mydomain.com* to the same address since it is not the inability to resolve either address that causes this issue, but rather that the fastname in node configuration file does not exactly match the system host name or the value of *APT_PM_CONDUCTOR_NODE*.

Resolving the problem
You can resolve the problem with the following actions:
- Change fastname for nodes in configuration file to exactly match the output of `hostname` command.
- Set APT_PM_CONDUCTOR_NODE to the same value as fastname. This must be defined either in every project or every job.

Do not change the host name of server to match fastname. InfoSphere DataStage stores some information that is based on the current host name. If you change the host name after installation of InfoSphere DataStage, then you must contact the support team for more instructions on how to allow InfoSphere DataStage to work correctly with the new host name.

InfoSphere DataStage Parallel Job fails with reading connection message error
InfoSphere DataStage parallel job fails with a reading connection message error and no data is processed.

Symptoms
The InfoSphere DataStage job log contains the following unrecoverable error:

```
Item #: 13
Event ID: 1960
Timestamp: 2011-09-01 06:30:44
Type: Fatal
User Name: dsadm
Message Id: IIS-DSEE-TFPM-00154
Message: main_program: APT_PMConnectionRecord::start:
 Reading connection message returned 28, expected 40, Error 0

Item #: 14
Event ID: 1961
Timestamp: 2011-09-01 06:30:44
Type: Fatal
User Name: dsadm
Message Id: IIS-DSEE-TFPM-00356
Message: main_program: **** Parallel startup failed ****
This is usually due to a configuration error, such as
not having the Orchestrate install directory properly
mounted on all nodes, rsh permissions not correctly
set (via /etc/hosts.equiv or .rhosts), or running from
a directory that is not mounted on all nodes. Look for
error messages in the preceding output.
```

Causes
The framework that is used by InfoSphere DataStage to start all the parallel processes uses TCP/IP connections during the startup phase, even on single-host configurations. The processes are listening for specific responses on these ports, to coordinate the startup. This error means that one of the processes received an unexpected response and stopped.

These ports are used for low-level coordination between the specific InfoSphere DataStage processes, not for user requests, so there is very little error handling capability. When this error occurs it is an indicator that some process other than InfoSphere DataStage connected to one or more of the ports and put invalid data there. The InfoSphere DataStage process that receives this unauthorized connection has no other alternative except to print the error and exit. The default port range used by InfoSphere DataStage for

this communication is 10,000 - 11,000 and 11,000 and up (there is no upper bound, but it reasonably will not be more than a few thousand.) These are not common port ranges for other applications to use. When this problem occurs it usually means that applications for scanning ports, monitoring networks, or detecting intrusions might be causing the problem.

Resolving the problem
Ensure that no other applications are using the ports in the 10,000 - 15,000 range. If the problem persists, change the default start point for these port ranges to something else. You can use the operating system utility netstat -an to see which ports are in use. Select starting port ranges where netstat shows that nothing is in use.

```
APT_PM_STARTUP_PORT=50000
APT_PLAYER_CONNECTION_PORT=51000
```

You can set these environment variables at the job or project level for testing. When you find suitable values and the problem does not recur, set and export these variables in the `/opt/IBM/InformationServer/dsenv` file so that they take effect for all projects.

InfoSphere DataStage Parallel job failed to start because of a process fork failure in Solaris.

Symptoms
A parallel job stops with the following message: `Fatal Error: Unable to start ORCHESTRATE process on node node1`
`(sun01): APT_PMPlayer::APT_PMPlayer: fork() failed, Not enough space`

Causes
This error indicates that the `fork()` system call failed and returned an ENOMEM error. There is not enough swap space to support the virtual memory that is required by the call. The ENOMEM error for `fork()` can occur in other Operating Systems like AIX, Linux, or HP-UX, but it is seen more frequently in Solaris because Solaris requires much more virtual memory when you use the `fork()` command because it does not have a memory overcommit feature.

Linux, AIX, and HP-UX operating systems have a memory overcommit or lazy swap allocation feature. In a memory over commit mode, `malloc()` does not reserve swap space and always returns a none NULL pointer, regardless of whether there is enough virtual memory on the system to support it or not. The swap space must be made available only when the memory is referenced.

In contrast, under the Solaris OS, when the application calls `malloc()` and internally starts `sbrk(2)` to get more memory from the system, the kernel goes through its free memory lists and finds the requested amount of virtual memory. If it finds the requested amount of virtual memory, the kernel returns a pointer to that memory and reserves the swap space for it such that no other process can use it until the owner releases it. If the requested amount of virtual memory is not found, `malloc()` fails with an ENOMEM error and returns a NULL pointer.

For a large-memory process in Solaris, the `fork()` system call can fail because an inadequate amount of virtual memory because the `fork()` call requires twice the amount of the parent memory. This can happen even when the `fork()` call is immediately followed by an `exec()` call that would release most of that extra memory.

Resolving the problem
You can reduce the swap space that is required for InfoSphere DataStage with the following methods:
- Increase virtual memory by adding swap device. Consult your Solaris administrator for assistance. Adding more physical memory also increases virtual memory.
- Decrease the number of UNIX processes for InfoSphere DataStage jobs. You can decrease the number of UNIX processes for a parallel job by using an APT configuration file of fewer nodes. Unsetting the *APT_DISABLE_COMBINATION* variable also helps.

- Don't use the /tmp directory for temporary files. In Solaris, /tmp is allocated from virtual memory and requires swap space. You can reduce /tmp usage by adjusting the InfoSphere DataStage environment variables *TMP*, *TMPDIR*, *APT_FIFO_DIRECTORY*, *APT_PM_SCORE_DIR*, and the *uvconfig* parameter *UVTEMP* to specify another location for these files. Specify that this new location is on a locally mounted disk. The maximum size for /tmp can be configured at mount time. Consult your Solaris administrator for assistance.

InfoSphere DataStage Jobs fail to start or run poorly when temporary directories are large

InfoSphere DataStage jobs write multiple files to temporary directories, which are not automatically cleaned.

Symptoms

When the number of temporary files grows large, InfoSphere DataStage jobs have slower performance and can hang. Sites that run InfoSphere DataStage for a year or more and did not clean up the temporary directories can contain 100,000 or more files.

Causes

Normally temporary files are cleaned up when a job ends. However, terminated jobs can leave behind files.

Diagnosing the problem

Check these temporary directories for old InfoSphere DataStage files, which can be removed, or many files in general since other programs also write to these directories. You can check the job log of any InfoSphere DataStage job to determine the value of the *TEMP*, *TMPDIR*, and *TMP* variables at job run time. On UNIX systems, also check location /tmp. InfoSphere DataStage writes non-temporary files to these locations, which are not automatically cleaned up. The directories are pointed to by the following environment variables:

- TEMP
- TMPDIR
- TMP
- /tmp on UNIX

You can check the job log of any InfoSphere DataStage job to determine the value of these variables at job run time.

Resolving the problem

Perform periodic cleanup of old InfoSphere DataStage files in these directories to prevent performance issues that relate to the number of files in the directories.

You can remove InfoSphere DataStage files from these directories that are older than longest job run time or older than the last InfoSphere DataStage engine restart. Unless you have long running jobs or "always-on" web service jobs, it is safe to remove files more than two days old from the following directories:

- The &PH& directory within each project directory
- The directory that is pointed to by *UVTEMP* variable in the uvconfig file in the `DSEngine` directory
- The directory that is pointed to by *UVSPOOL* variable in the uvconfig file in the `DSEngine` directory on UNIX systems. On Windows, this variable might not exist in uvconfig.

When you remove files from the directory that is pointed to by *UVTEMP* variable in the uvconfig file in the `DSEngine`, the primary files to clean up are the capture files that capture the system output of external commands that are called from InfoSphere DataStage jobs. For example: `capture0001360098aa1312969887`.

On UNIX operating systems, the *UVTEMP* setting points to /tmp directory. If the *UVTEMP* variable is undefined on a windows operating system, it defaults to the *UVTEMP* directory under DSEngine. /opt/IBM/InformationServer/Server/DSEngine/UVTEMP

Jobs with join stages unexpectedly terminated by UNIX Signal 11 SIGSEGV

InfoSphere DataStage jobs with a join stage terminate with the following error "Unexpected terminated by UNIX Signal 11 (SIGSEGV)"

Symptoms

InfoSphere DataStage job with a join stage terminates with the following error `Unexpected terminated by UNIX Signal 11 (SIGSEGV)`

Causes

If the size of the record is larger than the default setting of 20 MB, the sort that is inserted for the join fails.

Resolving the problem

Calculate the size of your largest record in bytes. Double that size and set the environment variable *APT_TSORT_STRESS_BLOCKSIZE* to that value in bytes. If this environment variable does not exist, create it first. The *APT_TSORT_STRESS_BLOCKSIZE* variable sets the size of the shared memory block that is used to pass data between the writer, sorter, and merger processes within the sort operation.

The default is set so that each sort in a job attempts to map 20 MB each for these in-memory functions. In the case when a single record is larger than 20 mb, this step is necessary to increase the size of the block. You must set the *APT_TSORT_STRESS_BLOCKSIZE* environment variable at the job level if possible. The *APT_TSORT_STRESS_BLOCKSIZE* environment variable can also be set at the project level by defining a project level default value.

Disk space issues

If the system runs out of available disk space, it can cause several problems for InfoSphere DataStage jobs.

Errors when the system writes the OshExecuter.osh record to a file

You can encounter an `Error writing record OshExecuter.sh to file` message when you try to import a job into your project.

Symptoms

When you import a job into a project, you receive the following error: `Error writing record OshExecuter.sh to file RT_SCxxx - error number is 40019` The job is corrupted.

Causes

The following items commonly cause job corruption:
- Full disk space in the /tmp *UVTEMP* directory or the InfoSphere DataStage project directory.
- A 32-bit hash file becomes larger than 2 GB
- System power outages
- System failures
- Rebooting the server while a job is running
- Virus checker/scanner running while job is running
- Backup running while job is running
- Failures on the system

Diagnosing the problem

To determine if the job is corrupted, run the following process:

1. Log on to the InfoSphere DataStage Administrator client
2. In the **projects** tab, select the corrupted project.
3. Click **Command**
4. Issue the following command: `LIST <FILENAME> F1 DET.SUP`

If the job is not corrupted, you can see a number that indicates the number of records in the file. If the file is corrupted, you receive an error message.

Resolving the problem
If you confirm that the file RT_SC file is corrupted, you can re-create the job.

Unable to view data or run a Parallel job with a Sequential File stage
In a InfoSphere DataStage Parallel job with a Sequential File stage, cannot view data or run the job.

Symptoms
The following error message is generated: `File archive: Trouble creating file`

For a parallel job with a Sequential File stage plug-in, viewing the data results in the following error:

`IIS-DSEE-TFAR-00015 00:10:13 <main_program> File archive: Trouble creating file "/tmp/...."`

Run time results in the following errors:
* Message Id: IIS-DSEE-TFAR-00015 `Message: main_program: File archive: Trouble creating file "/tmp/...."`
* Message Id: IIS-DSEE-TFPX-00002 `Message: main_program: Fatal Error: Null archive.`

Causes
The program is searching for relative path called "tmp" that is not present. This occurs on Windows installations when the project is not on the same drive as the engine.

Resolving the problem
Create a directory that is called `temp` at the root of the drive where the InfoSphere DataStage project is located. For example, if the InfoSphere DataStage projects are on the `D:` drive, create the following directory: `D:\tmp` If the directory exists, check the remaining disk space on your drives to ensure that limited disk space is not the cause of the problem.

Detecting corrupted files in InfoSphere DataStage
If the InfoSphere DataStage project or temp directory runs out of disk space, some of the buffered write requests might not be completed and the project might have hash files that are corrupted.

About this task

Run the `uvbackup` process and redirect the backup to null devices. The `uvbackup` process produces an output file that can be used to identify the corrupted files.

Procedure
1. Source your `dsenv` file in *$DSHOME*. `./dsenv`
2. Go to your project directory `../InformationServer/Project/<project name>`
3. List all files and direct them to a file with the `ls > myfiles.txt` command. This file is used to list of files for the uvbackup.
4. Run the `uvbackup` and redirect output to null with the following command: `$DSHOME/bin/uvbackup -V -f -cmdfil myfiles.txt -s uvbackupout.txt -t /dev/null  2>&1 > testing123.txt"`
5. Search for warnings with the `grep "WARNING:" uvbackupout.txt` command.

Results

The `uvbackupout.txt` file helps to identify corrupted files in the project.

Example

Here is an example what you might see in the `uvbackupout.txt` file:

```
WARNING: Unable to open file 'RT_STATUS3' for reading. File not saved!
```

The `uvbackup` verifies the integrity of the files and does not back up any files that are corrupted.

InfoSphere DataStage PX job errors with no available output files

Symptoms
You receive the following error message:

```
Error Message: Sort_17,2: write failed: Output file full, and no more
    output files [sort/merger.C:1393]
Message: Sort_17,2: Fatal Error: Tsort merger aborting:  mergeOneRecord(
) punted [sort/merger.C:1214]
Message: Sort_17,2: Fatal Error: Pipe read failed: short read [sort/m
  erger.C:1740]
```

Causes
This error indicates that the job is running out of scratch, temporary, or swap space.

Diagnosing the problem
Checking the amount of space available without the job that is running can give a false indication that there is enough space.

Open the configuration file that is being used by the job. This can be found by checking the APT_CONFIG_FILE parameter of the job. You can find the configuration file in ../IBM/ InformationServer/Server/Configurations/*.apt. Check the location of the scratch space that is identified for each node. While the job is running, issue the following command: `df -k` This command displays the amount of space that is allocated and being used on the system. Make sure that there is space available for `/tmp`, `/swap`, and the scratch area. Also make sure that the location that is identified in the configuration file for scratch space also has enough space.

Resolving the problem
Increase the disk space for `/tmp`, `/swap`, and `/scratch`. You can create an extra scratch area in your configuration file. For instance, you might add a second scratch disk area with the following command:

```
node "node1"
{
fastname "gcp5bs03"
pools ""
resource disk "/opt/local/InformationServer/Server/Datasets" {pools ""}
resource scratchdisk "/opt/local/InformationServer/Server/Scratch" {pools ""}
resource scratchdisk "/etc/svc/volatile" {pools ""}
}
```

Temporary lookuptable files that use up disk space
When a InfoSphere DataStage job with a lookup stage fails, files with the `lookuptable` prefix are left in the resource directories.

Symptoms
The resource directories and common space contain files with names that are similar to the following file name: `lookuptable.20091210.513biba`

Causes

When a job fails, it leaves the temporary files for postmortem review in the resource directories. Temporary files are left in scratch, but lookup files are created in the resource directories. Lookup file sets are not removed. A lookup file set is similar to the following file set: /opt/IBM/InformationServer/ Server/Datasets/export.dsadm.abcdefg.P000000_F0000 A lookup file has a structure that is similar to the following file: /opt/IBM/InformationServer/Server/Datasets/lookuptable.20091210.513biba

Diagnosing the problem
Look for files with file names similar to the following structure: `lookuptable.yyyymmdd.nnnnnnn` You can find these files on the disk when no jobs are running.

Resolving the problem
When no jobs are running, remove all of the files that have the `lookuptable` prefix on the file name. These files get re-created with every new run of the job and are never reused. If the job runs successfully, then only the lookuptable file that is created during that job run is removed. Create a maintenance script that cleans up the `lookuptable` files regularly when jobs are not running.

Disk lookup issues
Disk lookup issues are often caused by a lack of disk space.

A parallel job with lookup fails with a `File too large` message
Sometimes the lookup table is too large to fit in available memory, which can cause a parallel job to fail.

Symptoms
An InfoSphere DataStage parallel job that contains a lookup failure with an error that is similar to the following error: `Lookup_107,0: Error writing table file "/d01/Ascential/DataStage/Datasets/ lookuptable.20100217.abcde": File too large`

Causes
The lookup table is too large to fit in available memory.

Resolving the problem
Create a hash partition on the lookup data and add more nodes to the parallel configuration file to break the lookup data into smaller pieces.

InfoSphere DataStage jobs fail to map table files
Some jobs fail to map table files even though the disk volumes are not full.

Symptoms
InfoSphere DataStage jobs fail with the following error message: `Message:: stage_lkup_FS,0: Could not map table file "/Projects/Datasets/lookuptable.20110720.oja0s0c (size 29056 bytes)": Invalid argument.` This error occurs even though the file is small and the disk volumes are not full.

Causes
Normally when a InfoSphere DataStage job fails with the logged error message `Could not map table file` the message ends with `not enough space`. In that situation, the issue is either insufficient disk space, or table too large to map into a single process and can be resolved by the steps that are described in the following topic: "InfoSphere DataStage job fails with a `Could not map table file` message" on page 34.

Information Server or InfoSphere DataStage job fails with `Could not map table file` For the special case where the error message ends with "invalid argument" this is typically caused by an I/O error unrelated to disk full. The most common cause of the error is that one of the directories where InfoSphere DataStage is writing is mapped to a volume that was mounted with the CIO option.

Diagnosing the problem
In the job log, locate the error message that contains `could not map table file`. That file contains the name of directory we were writing to. Additionally, the next message in file might be a second error that

points to a different directory. For each of these directories and any other resource and scratch disks that are identified in the node configuration file that is used by this job, use the df command to confirm the volume that contains the directory. For example: df /Projects/Datasets Then issue the mount command to list the mounted volumes and the options that they are mounted with.

Check to see if the volume with failing I/Os was mounted with the "cio" option. If the volume was mounted with the "cio" option, it causes problems for InfoSphere DataStage jobs and severe problems for jobs that use the lookup stages or other stage types that rely on memory map files.

The reason why the "cio" option causes problems is explained in the AIX mount command reference. http://publib.boulder.ibm.com/infocenter/pseries/v5r3/index.jsp?topic=/com.ibm.aix.cmds/doc/aixcmds3/mount.htm The explanation for the cio option states why this option causes problems for programs that use the mmap() function and other related functions.

Resolving the problem
Stop the InfoSphere DataStage engine. Then instruct the system administrator to remount that volume without the cio option. Use the default rw mode to remount the volume, and restart InfoSphere DataStage. This action resolves the invalid argument error that occurs when you map table files.

Data processing failures

Several problems can cause data processing failures.

Issues with NULL values in input columns in Parallel Engine Transformer stage

InfoSphere Information Server 8.5 changed the way that the parallel engine transformer stage handles columns that contain null values.

Symptoms
Jobs that process nulls in transformer stages show different behavior when you upgrade from InfoSphere Information Server 8.1 to 8.5, even when legacy null handling is enabled.

Causes
In InfoSphere Information Server version 8.1 and prior versions, the job design had to explicitly handle null column values in the Transformer stage. If the Parallel Engine encountered null column values outside of specific contexts, the entire row that contains the null was dropped, or sent to a reject link if the Transformer stage had a reject link.

Note: NULL is the SQL value NULL, and not an empty string or the character with the value of 0x00.

Explicit null handling caused the Transformer stage code to be too complex, and allowed inconsistent behavior. In InfoSphere Information Server version 8.5, the default behaviors were changed and explicit null handling was no longer required. It was recognized that some customers would want to retain the original null-handling behavior so an environment variable, APT_TRANSFORM_COMPILE_OLD_NULL_HANDLING, was introduced. When you define the environment variable, it preserves compatibility with the behavior before InfoSphere Information Server 8.5.

After version 8.5 released, differences in the default null handling behavior of version 8.5 and problems with the implementation of the InfoSphere Information Server 8.1 compatibility mode were discovered. There are issues with null handling in InfoSphere Information Server 8.5 when you enable compatibility with an earlier version. There are issues with null handling in InfoSphere Information Server 8.1, and earlier versions. Most of these issues were caused by the lack of a clear explanation about how null values must be handled in the Transformer stage.

Resolving the problem

InfoSphere Information Server 8.1 NULL handling where explicit handling is required.

If you use input columns in an output column expression, a NULL value in that input column causes the row to be dropped, or rejected if a reject link is defined. This applies for the following conditions:

- An input column is used in an output column derivation expression like ?DSLink4.col1 + 1?.
- An input column is used in an output column constraint.
- An input column is used in a stage variable derivation.

It does not apply where an output column is mapped directly from an input column, with a straight assignment expression.

In InfoSphere Information Server 8.1, there are three legal ways to explicitly handle NULL values in input columns in the Transformer stage.

Conversion of NULL values:

Nulls can be converted to other values with any of the following functions:

- NullToEmpty()
- NullToZero()
- NullToValue()

See the following example:

```
DSLink3.OrderCount + 1 --> If DSLink3.OrderCount is NULL,
the whole record containing DSLINK3.OrderCount
 will be dropped or sent to the reject link
```

This expression can be changed to the following command:

```
NullToZero(DSLink3.OrderCount) + 1
--> If DSLink3.OrderCount is NULL,
 the target field will be assigned the integer value 1.
```

Check for NULL value.

To test if a value is NULL in a logical expression, use one of the following functions.

- IsNotNull()
- IsNull()

For example:

```
DSLink3.OrderCount + 1
--> If DSLink3.OrderCount is NULL,
record will be dropped or rejected.
This expression can be changed to:
If(IsNotNULL(DSLink3.OrderCount)
Then DSLink3.OrderCount + 1 Else 1
--> If DSLink3.OrderCount is NULL,
the target field will be the integer 1 .
```

Each nullable column in an expression must be properly NULL checked or the NULL value must be converted to a concrete value.

IF-ELSE operations on NULL values.

Handling NULL values in IF-ELSE conditions can be complex. Consider the following examples to get familiar with using NULL checks in IF-ELSE statements.

Example 1: Simple IF ELSE statement.

```
If (DSLink1.Col1 > 0) Then xxx Else yyy
In InfoSphere Information Server 8.5 code will be
 generated to drop records in case DSLink1.Col1 is NULL.
```

This statement must be written as the following statements:

```
If (IsNotNull(DSLink1.Col1) and (DSLink1.Col1 > 0))
 Then xxx Else yyy
or
If (IsNull(DSLink1.Col1) or (DSLink1.Col1 > 0))
 Then xxx Else yyy
based on the business requirement on how the NULL
 case should be treated
```

The statement can also be written as the following statements:

```
If (NullToZero(DSLink1.Col1) > 0) Then xxx Else yyy
or
If (NullToZero(DSLink1.Col1) >= 0) Then xxx Else yyy
again based on the business requirement.
```

```
NUllToValue() or NullToEmpty() also can be used instead of
 NullToZero(), based on the requirement.
```

```
NULL conversion functions should
 not be used for checking a NULL
Example: If ((NullToZero(DSLink1.Col1) = 0) or
(DSLink1.Col1 > 0)) Then xxx Else yyy
```

```
NullToZero(DSLink1.Col1) = 0 is not considered a NULL check,
 code will be generated to drop records
 in case DSLink1.Col1 is NULL.
```

Example 2: IF ELSE statement with multiple conditions.
```
If ((DSLink1.Col1 = 5) or (DSLink1.Col2 = 8))
Then xxx Else yyy --> Col2 is a nullable field
```

This statement cannot be written as the following statements:

```
If (DSLink1.Col1 = 5 or IsNotNull(DSLink1.Col2)
and (DSLink1.Col2 = 8)) Then xxx Else yyy
This is not the proper way of checking.
If(condition1 or condition2 and condition3) will
be treated as If((condition1 or condition2) and condition3))
```

The statement must be written as the following statement:

```
If (DSLink1.Col1 = 5 or (IsNotNull(DSLink1.Col2)
and (DSLink1.Col2 = 8))) Then xxx Else yyy
```

The condition that contains the nullable column must be properly pre-"AND"ed with an IsNotNull() check, or pre-"OR"ed with an IsNull() check. The order of the conditions must be clearly specified with parentheses.

Example 3: IF ELSE statement in which the nullable field is used multiple times.
```
If ((DSLink1.Col1 = 3) or (DSLink1.Col1 = 5))
 Then xxx Else yyy
Records will be dropped in case Col1 is NULL.
```

This statement cannot be written as the following statement:

```
If ((IsNotNull(DSLink1.Col1) and (DSLink1.Col1 = 3))
or (DSLink1.Col1 = 5)) Then xxx Else yyy
```

This statement must be written as the following statement:

```
If ((IsNotNull(DSLink1.Col1) and (DSLink1.Col1 = 3))
or (IsNotNull(DSLink1.Col1) and (DSLink1.Col1 = 5))
 Then xxx Else yyy
```

Each instance of nullable field must be pre-"AND"ed or pre-"OR"ed with NULL check.

Example 4: Using two nullable columns in a condition.
```
If (DSLink1.Col1 = DSLink1.Col1) Then xxx Else yyy
```

Both the columns must be NULL checked or NULL conversion functions must be used on both the columns.
```
If (IsNotNull(DSLink1.Col1) and (IsNotNull(DSLink1.Col2)
and (DSLink1.Col1 = DSLink1.Col1))) Then xxx Else yyy
```

InfoSphere Information Server 8.5 NULL handling where explicit handling is not required.

In InfoSphere Information Server 8.5, NULL value in that input column does NOT cause the row to be dropped nor is it be sent to reject link.

A NULL value in that input column is handled by the Transformer stage, following specific logic. The job designer can skip explicit NULL handling. More information on explicit null handling is available in the following technote:http://www-01.ibm.com/support/docview.wss?uid=swg21514921

The environment variable APT_TRANSFORM_COMPILE_OLD_NULL_HANDLING can be used, in case the designer wants to have 8.1 behavior in 8.5. Enabling old NULL handling can be done at the three following stages:

- Setting APT_TRANSFORM_COMPILE_OLD_NULL_HANDLING at project level
- Setting APT_TRANSFORM_COMPILE_OLD_NULL_HANDLING at job level
- Checking the **Legacy null processing** option in the InfoSphere DataStage Designer for individual Transformer stages in a job

IBM discovered a previously undocumented difference in behavior of InfoSphere Information Server 8.1 and InfoSphere Information Server 8.5 with old NULL handling enabled. InfoSphere Information Server 8.5 allowed the three "NullToxxxx()" functions to be used as Null tests.

The following is an example of the IF-ELSE condition:
```
If ((NullToZero(DSLink1.Col1) = 0) or (DSLink1.Col1 > 0)) Then xxx Else yyy
```

The `(NullToZero(DSLink1.Col1) = 0)` section is considered as a NULL check and records were not dropped or sent to the reject link because of inconsistency in the code. In InfoSphere Information Server 8.5 this code inconsistency was eliminated and only `IsNull()` and `IsNotNull()` can be used as Null checks. InfoSphere Information Server 8.1 Jobs that used `NullToZer0()`, `NulltoValue()`, or `NulltoEmpty()` for null checking must be changed to use `IsNull()` or `IsNotNull()`.

Note: The `!IsNull()` function can be used instead of the `IsNotNull()` function. The `!IsNotNull()` function can be used instead of the `IsNull()` function for NULL checking.

InfoSphere DataStage jobs have transformer compile errors after you upgrade to version 8.1 or later

If you upgrade from

Symptoms

Your InfoSphere DataStage 7.0 jobs that have a transformer stage where a stage variable is set to null compile successfully. However, the same job fails on InfoSphere DataStage v8.1 with the following compile error:
```
<transform> Error when checking composite operator:
Setting null to this non-nullable field: StageVar0_myStageVariable
```

Resolving the problem

Stage variables that are created in the transformer stage do not have a nullable option. InfoSphere DataStage considers stage variables to be non-nullable objects. The `SetNull()` function for stage variables is not supported. Information Server 8.1 traps this condition and reports it as an error at job compile

time. InfoSphere DataStage 7.5 did not trap this condition at compile time. The job compiled successfully, but it had the potential to fail at run time when it handled null data. The capability to support null stage variables was added in InfoSphere Information Server 8.5. This capability is not available for InfoSphere Information Server 8.1.

For InfoSphere Information Server 8.1, if a transformer stage must assign a nullable field to a stage variable, assign an empty string with two single quotation marks instead of a null. For example: If IsNull(InLink.FieldName) then '' Else InLink.FieldName

For InfoSphere Information Server 8.5 and later, the transformer stage properties dialog now has new options on the Stage -> General page: Legacy null processing Abort on unhandled null

When this is enabled, the legacy option handles nulls as with prior releases. When disabled, null values can be output. Newly created InfoSphere Information Server 8.5 jobs have the legacy option that is disabled by default. Jobs that are imported or migrated from older releases might have this option enabled by default. More information about these options is available from the **Help** button for the Transformer properties dialog in InfoSphere Information Server 8.5.

Converting a nullable source warning message in InfoSphere DataStage

Symptoms
The job log contains the following warning message:

```
When checking operator: When binding output interface field "<field name>"
to field "<field name>": Converting a nullable source to a non-nullable result;
a fatal runtime error could occur;
use the modify operator to specify a value to which the null should be converted.
```

Causes
The warning occurs when the field comes in with Nullable=Yes and the output gets set to the value of that field, and the output field is Nullable=No. Setting a constraint in the database does not allow the incoming value to be Null, but it does not eliminate the warning

Resolving the problem
Ensure that both fields are set to either Nullable=Yes or Nullable=No. Note that the warning refers to the job design, and does not refer to the actual data or tables. Because the table definitions are generated from the tables themselves, the source table might have a column that is nullable and the same column in the target table is not nullable. This results in the respective column definitions when the tables are imported.

You can also eliminate the warning by changing the source and target tables to both be nullable or not nullable and regenerating the table definitions.

Problems when you run multiple instances of a job from a job sequence, or from a script that uses dsjob.
A number of related problems can occur when you are running multiple instances of a job from job sequence, or a script that sequences jobs with the dsjob command.

Symptoms
* Multiple job instances run from a sequence or a script, and the sequence reports that status=99 for one or more of the job instances.
* Multiple job instances run from a sequence or a script, and the job instances take a long time to start and to finish.
* More than 25 job instances run from a sequence or a script, and the sequence reports that status=99 for one or more of the job instances.
* The system does not have enough resources because of a heavy work load, and the sequence reports a code=-99 error for a parallel job.

- On Intel RedHat and Suse systems, jobs can hang despite having successfully run the underlying OSH code.
- Some jobs run with missing parameters, or parameters that are erroneously set to default values.

Resolving the problem

For release 8.0.x /8.1:

- If your system uses RedHat, Suse, or Intel, install JR30015v5.
- If your system does not use RedHat, Suse, or Intel, install JR30015v6 on the server.
- If auto-purge is enabled, and more than 25 instances of a job are run simultaneously, then the JR30015v3 client patch must be installed on all client systems.
- Recompile all parallel jobs after you install the patch or fix pack.

For release 8.1:

- Install fix pack 1.
- Recompile all parallel jobs after you install the patch or fix pack.

The fix introduces the following optional capabilities:

Environment variable: DSWaitResetStartup

> When multiple instances of a job are run from a sequence, and one or more of the job instances are set to reset, the sequence might report that `status=99`. This can occur because the controlling sequence did not give the job instances enough time to reset before polling its status. The startup time for a job reset must be increased. The environment variable DSWaitResetStartup can be used for this purpose. The maximum value that can be set for *DSWaitResetStartup* is equal to the value of *DSWaitStartup*, which is 60. For example, if a value of 120 is required for DSWaitResetStartup, you must ensure that DSWaitStartup is also set to a minimum of 120.

Environment variable: DS_NO_INSTANCE_PURGING

> If the system is under extreme load, it might be necessary to use the DS_NO_INSTANCE_PURGING environment variable if Status=99 errors still occur when you run many multi-instance jobs and auto-purge is enabled. This environment variable must be set to 1. This stops the auto-purge from deleting the status records for the job instance, allowing the controlling job to read its status when system resource becomes available. (In other situations, you might want clean logs with no persistent instance entries, so the default behavior is to purge instance entries.)

Client change, and environment variable DSJobStartedMax

> The number of recorded instance identifiers increased to 100 from a value of 25. The increase prevents status records from being purged when more than 25 instances run simultaneously. If you use N-instance auto-purging and run more than 25 simultaneous instances, then the N-Instance auto-purge limit must be set to more than 25. The limit is set in the Director or Administrator clients. If you must run more than 100 Instances simultaneously, then the environment variable *DSJobStartedMax* must be set to the required value. The maximum value is 9999. The APAR number for this issue is APAR JR30015.

DataStage timeout variables

You can correct several errors by setting timeout variables correctly.

An InfoSphere DataStage job completes successfully but dsjob returns `Error 81002 waiting for job`

Symptoms

When you run `dsjob` from UNIX, the status of success or failure is returned correctly for most jobs. However, long-running jobs return the `Error 81002 waiting for job` error even though the job runs

successfully. Specifically, if the job runs longer than the timeout set in InfoSphere DataStage Administrator -> General tab -> Inactivity timeout, the 81002 error is returned. If the job completes before the timeout, the correct status is returned.

Resolving the problem

Increase the timeout in the InfoSphere DataStage Administrator to accommodate the longest-running job. However, increasing the timeout also increases the time that inactive client connections remain on the system before they are cleaned up.

If the timeout increase in InfoSphere DataStage Administrator does not work, create a script that starts dsjoband uses a loop to query dsjob periodically to avoid the timeout.

Parallel jobs on remote nodes fail with a broken pipe in IBM InfoSphere Information Server

Symptoms

You run a parallel job in a cluster environment with a remote node and receive the following message in the job log:

```
main_program: Fatal Error: Service table transmission failed for node2
 (<node name>-svc:Broken pipe.
```

Resolving the problem

Increase the value of APT_PM_CONDUCTOR_TIMEOUT to a larger value. The default value is 60 seconds. Double the value to 120 seconds.

If an increase to the value of APT_PM_CONDUCTOR_TIMEOUT does not solve the problem, there might be a problem with starting the resource tracker. You can use the resource tracker to gather system statistics from the system where jobs are running. These statistics have no effect on the actual job that is running, and are unrelated to the information that is captured in the InfoSphere DataStage and QualityStage job monitor or job log. If you are not using the resource tracker functions, it can easily be turned off to avoid this problem. To turn off the resource tracker, add the following APT_DISABLE_TRACKER_STARTUP environment variable at the project level, and set the default value to 1.

A parallel job fails intermittently on systems with a heavy workload

An InfoSphere DataStage parallel job can fail intermittently on systems with a heavy workload.

Symptoms

A parallel job fails with one of the following error messages:
- ds_ipcopen() - call to OpenFileMapping() failed - The system cannot find the file specified
- ds_ipcput() - timeout waiting for mutex

Causes

In systems with a heavy load, timeouts can cause the system to fail.

Resolving the problem

Increase the active link performance *DSIPC_OPEN_TIMEOUT* timeout variable, which is

To increase these values, use the following steps:
1. Open the Administrator client.
2. Open the project properties for your project.
3. Open the Tunables page and increase the *Timeout* value under Active-to-active link performance. If the timeout field is not enabled, you must select both **Enable row buffer** and **Inter process**.
4. Open the General page and click Environment.

5. Open the **Parallel** > **Operator-specific** branch.

6. Increase the value of the *DSIPC_OPEN_TIMEOUT* environment variable.

If an increase to the timeout does not help, or if the problem always affects your job, contact IBM support for further advice.

Parallel job ends with an APT_PMsyncWithSectionLeaders: Non-zero status 4 error.

Symptoms
The InfoSphere DataStage Parallel job log shows the following information:

```
APT_PMsyncWithSectionLeaders: non-zero status 4 from APT_PMpollUntilZero
broadcastStepIR: step timed out sending 66,135,142-bytescore;
status = 4  Error during score broadcast or reload. Score size is 66,135,142 bytes
```

Causes
The error usually indicates that a resource issue is the cause of the problem.

Resolving the problem
The environment variable APT_PM_NODE_TIMEOUT controls the number of seconds that the conductor waits for a section leader to start and load a score before you decide that something failed. The default for starting a section leader process is 30 seconds. The default for loading a score is 120 seconds. Set the following environment variable at the project level: APT_PM_NODE_TIMEOUT=300

You can increase the value of this environment variable to 600 if 300 does not resolve the problem. If the APT_PM_NODE_TIMEOUT environment variable does not correct the issue, monitor processor, disk space, memory, and swap when the job is running. Check with your network administrator to see if the nodes are on a SAN or NFS mount.

Troubleshooting specific stages

Each stage has unique troubleshooting scenarios.

Troubleshooting errors when you load the connector library

When you use the InfoSphere DataStage and QualityStage Designer to test connections, view data, or generate or validate SQL, you might get an error that relates to loading connector libraries.

Symptoms
The InfoSphere DataStage jobs that include connectors run successfully, even if the error is displayed.

Causes
The library path environment variable is not set. The ASB Agent process fails to find libraries that the connector library requires and issues an error.

Resolving the problem
Add the location of the directory that contains the client libraries to the library path environment variable, and then restart the ASB Agent.

The name of the library path environment variable depends on your operating system.

Operating system	Library path environment variable
Microsoft Windows	`PATH`
HP-UX	`SHLIB_PATH`
IBM AIX	`LIBPATH`

Operating system	Library path environment variable
Other supported Linux, UNIX, and HP-IA operating systems	`LD_LIBRARY_PATH`

Troubleshooting the Big Data File stage

You can troubleshoot errors that are specific to the Big Data File stage.

Potential issues with sharing libhdfs by NFS mount

When the Hadoop Distributed File System (HDFS) NameNode and the InfoSphere DataStage parallel engine are on different systems, you must ensure that the Big Data File stage can still access various HDFS `.jar` files, the HDFS configuration directory, and the libhdfs library. One method for providing access to these HDFS components is to NFS mount directories from the HDFS server on the parallel engine server.

Symptoms

Some HDFS distributions install the libhdfs library in a system directory (for example `/usr/lib64`). Such directories contain other libraries that are important to system operation. Therefore, NFS mounting this directory from a remote system can cause system-wide issues when the library versions for the remote and local systems are mismatched. It can be especially problematic if the remote directory is mounted over the existing local `/usr/lib64` directory.

If the remote `/usr/lib64` directory is mounted over the local `/usr/lib64` directory, even small library version mismatches between the systems can cause system-wide program failures when the programs try to load the shared libraries that are needed from the `/usr/lib64` directory. Programs such as vi, ls, rpm, shutdown, and others can fail with `Failed to load shared library` errors in this configuration. Mounting over the local `/usr/lib64` directory is not recommended.

Even if a new mount point is created on the server where the parallel engine is installed (for example, `/usr/lib64_hdfs`), and the new directory is added only to the *LD_LIBRARY_PATH* environment variable that is used by InfoSphere DataStage (in the `dsenv` file), version mismatches between the systems can still cause jobs to fail. In this configuration, the failures are limited to InfoSphere DataStage and do not affect other system programs.

Resolving the problem

If the HDFS distribution that is used provides the libhdfs library in a separately installable package, install the package on the server where the parallel engine is installed rather than trying to access the libhdfs library from the remote system. When the library is installed locally on the server where the parallel engine is installed, mounting the `/usr/lib64` directory from the HDFS system is no longer required.

If the libhdfs library must be accessed remotely, complete the following steps:

1. Ensure that both systems are of the same distribution, release, and patch level (for example, SUSE Linux Enterprise 11 Service Pack 1 is installed on both computers, not SUSE Linux Enterprise 11 Service Pack 1 on one computer and Red Hat Enterprise Linux 6 Update 3 on the other computer).
2. Create a mount point directory on the computer where the parallel engine is installed (for example `/usr/lib64_hdfs`), and NFS mount over that new mount point.
3. Add the new mount point directory to only the *LD_LIBRARY_PATH* environment variable in the `dsenv` file. Do not add the new mount point directory to the system-wide *LD_LIBRARY_PATH* environment variable. The goal is to confine any possible version mismatch side effects to InfoSphere DataStage.

Troubleshooting the DB2 Connector Stage

Several common errors are specific to the DB2 Connector stage.

The stage cannot obtain error messages from the server

Sometimes you are unable to obtain error messages when you view data or run a job.

Symptoms

You receive the error `Could not obtain the error message from server (function=SQLAllocHandle(SQL_HANDLE_ENV))` when you try to view data or run a job.

Resolving the problem

Verify that the instance property specifies a valid DB2 instance. If the Instance property is blank, verify that the *DB2INSTANCE* environment variable specifies a valid DB2 instance. More information on known issues with DB2 instances is available in the release notes.

Error loading connector library when you try to view data
Symptoms

When you try to view data, you receive the following error: `Error loading connector library`. The error message mentions symbols that are not exported in *libfcl*.

Resolving the problem

The *LIBPATH/LD_LIBRARY_PATH/PATH* environment variable that is used by the ASBAgent does not have directories listed in the correct order. The fcl library that is located in the ASBNode directory must be first in the *LIBPATH/LD_LIBRARY_PATH/PATH* environment variable that is used by the ASBAgent. Change your environment variable so that the ASBNode fcl library is found first. For example, the path must be similar to the following path: `LIBPATH=/opt/IBM/InformationServer/ASBNode/lib/cpp:....[add'l directories]`

Action String parameter is not valid or too long

You can encounter errors with the Action String parameter when the write mode of a job is set to "Bulk load".

Symptoms

When you run a job with a write mode that is set to "Bulk load", you receive the following error: `SQL3009N The Action String parameter is not valid or too long`.

Resolving the problem

This error occurs because of a bug in DB2 9.5 FP2 and is documented in the following technote: http://www-01.ibm.com/support/docview.wss?rs=14&uid=swg21322938. You must install DB2 9.5 FP3. After you install DB2 9.5 FP3, you must rebind your packages by following the directions in the following technote: http://www-01.ibm.com/support/docview.wss?uid=swg21449630

SQL0443N Routine error when you run a job

You can resolve some errors by rebinding packages.

Symptoms

You receive the following error message: `SQL0443N Routine "SYSIBM.SQLCOLUMNS" (specific name "COLUMNS") has returned an error SQLSTATE with diagnostic text "SYSIBM:CLI:-805". SQLSTATE=38553` when you run a job.

Resolving the problem

Rebind your packages. Follow the directions in the following technote: http://www-01.ibm.com/support/docview.wss?uid=swg21449630

Potential data loss and corruption from the connector

Symptoms

When you run a InfoSphere DataStage job, you see warnings similar to the following warning:

```
DB2: [IIS-CONN-DAAPI-000396] Writing the WCHAR column COL2 into a CHAR database
column can cause data loss or corruption due to character set conversions.
```

Causes

WCHAR in the message refers to a wide character type, which is equivalent to NChar types.

The length of WCHAR column COL2 cannot be validated because the database column is CHAR and character set conversion is involved. Inadequate column lengths can lead to data truncation or unexpected errors.

This problem can happen for two reasons:
- The column type in the job is specified as NChar or NVarChar
- The column type in the job is specified as Char or VarChar with the extended attribute 'Unicode' set.

The column stores a Unicode string, but the CHAR database column stores data in a non-Unicode format. The **Unicode** field can have characters that cannot be converted into the database code page that is used for the CHAR column.

Even if all the characters can be converted, the connector cannot determine if the size of the database CHAR column is adequate to store every possible combination of Unicode characters that are contained in the job's Unicode column. The converted characters might occupy more than 1 byte.

Resolving the problem

Make sure that you use the same column type in both the job and the database. You must use type Char with type VarChar, and NChar with NVarChar. If the columns in the job are loaded from a saved table definition in DataStage, the columns automatically apply the extended attribute Unicode set. This makes Char and VarChar columns equivalent to NChar and NVarChar columns, which can also cause the warning message. Make sure that you manually remove the extended attribute.

Errors for jobs that access a remote DB2 instance

Symptoms

When you run a job that accesses a remote DB2 node and you specified an instance in the **Instance** property, you receive the following error: `An error occurred while getting the DB2 instance. Verify that a valid DB2 instance is specified in the Instance property. The method sqlegins returned reason code 0, SQLCODE -1,390.`

Resolving the problem

The instance that you specified in the **Instance** property might be invalid on the nodes that are defined in your configuration file. This problem often occurs when the connector is set to run on the DB2 nodes themselves. You can work around this issue by unsetting the *DB2INSTANCE* environment variable according to the following technote: http://www-01.ibm.com/support/docview.wss?uid=swg21381234

Jobs that fail are slow to report connection errors

Sometimes a job that uses the DB2 Connector can fail, and the failure is not immediately reported.

Symptoms

Jobs that fail are slow to report connection errors.

Causes

If the DB2 server fails when the DB2 Connector attempts to connect, the DB2 Connector might take hours to report the failure if the job hangs. This problem occurs when the DB2 server fails during the DB2 Connector's call to the `SQLConnect()` function.

Resolving the problem

This timeout is a DB2 feature. The timeout can be adjusted by setting the *ConnectTimeout* CLI/ODBC configuration keyword in the `db2cli.ini`, or by setting the *DB2TCP_CLIENT_CONTIMEOUT* registry

variable on the DB2 client. The default timeout is to wait forever. The timeout can be set through the database alias. For information on the *ConnectTimeout* variable, see the following topic:
http://pic.dhe.ibm.com/infocenter/db2luw/v9r7/topic/com.ibm.db2.luw.apdv.cli.doc/doc/r0021533.html

User-defined functions fail when the DB2 Connector queries the DB2 environment
Some user-defined functions can fail when they do not have privileges to access information about the DB2 environment

Symptoms
When you run a job with the DB2 Connector, you receive the following error: `"SQL0430N User defined function SYSPROC.ENV_GET_SYS_INFO" (specific name "ENV_GET_SYS_INFO") has abnormally terminated. SQLSTATE=38503 (CC_DB2Connection::queryServerHostName, file CC_DB2Connection.cpp, line 3,248)`

Causes
The DB2 Connector runs the following queries to obtain information about the DB2 environment:
- select HOST_NAME from SYSIBMADM.ENV_SYS_INFO
- select IS_INST_PARTITIONABLE from SYSIBMADM.ENV_INST_INFO

These queries require the following items:
- SELECT or CONTROL privilege on the ENV_SYS_INFO administrative view and EXECUTE privilege on the ENV_GET_SYS_INFO table function
- SELECT or CONTROL privilege on the ENV_INST_INFO administrative view and EXECUTE privilege on the ENV_GET_INST_INFO table function

Resolving the problem
Add these privileges to the user that is specified in the DB2 Connector's *User name* property, and run the job again.

The DB2 Connector cannot find any available nodes in the APT configuration file
When the DB2 connector cannot find available notes in the APT configuration file, you must add the DB2 node pool in that APT file.

Symptoms
When you run a job, you receive the following error: `The connector could not find any available nodes in the APT configuration file. This usually occurs when a node pool constraint is specified for a connector stage, but it is not defined in the APT configuration file (CC_DB2Configuration::validateEnvironment, file CC_DB2NodeNegotiation.cpp, line 787`

Causes
This error occurs when a node pool constraint is specified for a DB2 Connector stage, but the node pool is missing from the APT configuration file. This problem occurs after you migrate a job that uses the IBM DB2/UDB Enterprise stage to use the DB2 Connector stage with the Connectivity Migration tool.

The IBM DB2/UDB Enterprise stage requires a node pool that is called *DB2* is defined. When jobs are migrated, the DB2 Connector stage in those jobs has a *DB2* node pool constraint specified. Edit your APT file and remove the "DB2" node pool before you run your migrated jobs, you encounter this error.

Resolving the problem
Add the *DB2* node pool back into the APT file, and then edit the node pool constraints on the DB2 Connector stages.

Errors when you set the active DB2 instance value
An error occurred setting the active DB2 instance to value `NewInstance`. The active DB2 instance is already set to value `CurrentInstance`.

Symptoms

The error message indicates that the DB2 connector cannot set the active DB2 instance to the specified value.

Causes

This error happens because each process that uses the DB2 connector can be associated with only one DB2 instance. The error occurs when two or more DB2 connector stages are in a job, and a different instance is specified in each stage. The error can also occur when the **Instance** property is changed when you view data or import metadata. The DB2 connector can test connections, view data, and import metadata from only one DB2 instance. A single continuously running ASBAgent process handles all of the activity that happens at design time. The first DB2 instance that is used to test a connection, view data, or import metadata, is the instance that the connector uses for any design-time actions.

Resolving the problem

To use a different DB2 instance, complete one of the following actions:

* Restart the ASBAgent service on Microsoft® Windows or the daemon on UNIX.
* Enable Connector Access Service multiple instance support.

To enable the Connector Access Service multiple instance support, complete the following steps:

1. Stop the ASBAgent and ensure that the `ASBAgent.exe` process is shut down.
2. Specify the free socket port and the port number range by adding the following lines to `ASBNode/conf/cas.properties`. The port number range depends on your environment.

   ```
   multipleprocess.enable=1
   multipleprocess.port.range.min=30000
   multipleprocess.port.range.max=30010
   ```

3. Restart the ASBAgent service.

Timeout error when the connector waits for the LOAD utility to begin reading data

The FTP server timed out waiting for the LOAD utility to begin reading data. A timeout can occur when the FTP server is configured with a short FIFOOPENTIME or when the LOAD utility starts slowly.

Symptoms

You receive the following error while the connector waits for the LOAD utility to begin reading data:

```
Transfer to dataset dataset failed with error: 550 DELE fails:
dataset does not exist.
...
125-Waiting for read process to open /tmp/dataset
125 Transfer request aborted450 timer expired waiting for read process to
open /tmp/dataset
450 timer expired waiting for read process to open /tmp/dataset
```

The *dataset* variable is the name of the data set that is generated from the data set names that are specified in the DB2 connector.

Resolving the problem

To resolve the error, complete one of the following steps:

* Increase the FIFOOPENTIME setting by using the **Usage > Bulk load to DB2 on z/OS > Transfer > Transfer command** property.
* Set the **Transfer command** property to **quote site fifoopentime=**n, where n is the number of seconds the FTP server waits for the LOAD utility to start.

You can find more information about the FIFOOPENTIME statement in the following topic: http://publib.boulder.ibm.com/infocenter/zos/v1r12/index.jsp?topic=%2Fcom.ibm.zos.r12.halz001 %2Ffifoopentime.htm

Hierarchical Data stage

You can troubleshoot errors that are specific to the Hierarchical Data stage.

Running out of memory

Symptoms
When you run a job that includes a Hierarchical Data stage with the Regroup, Sort, or HJoin steps, these steps run out of memory.

Causes
The HJoin, Regroup, and Sort steps perform non streaming operations that require complete data in memory. So, these steps must cache the whole data set before the step can begin to process data.

Diagnosing the problem
Most of the steps in the Hierarchical Data stage process input data in real time, so they do not need to cache data. Memory usage in these cases depends on the size of the input data. Selecting the **Disk-based** option in the **Optimization type** field of the HJoin step does not resolve the issue. In the same way, clearing the **Data set fits into memory; optimize sort** check box does not resolve the memory issue in the Sort step.

Resolving the problem
Complete one or more of the following tasks:
- Sort data before it is used as input for the Regroup step. If the Regroup step is not the first step in the stage, ensure that data is not reordered before it is used as input for the Regroup step. When you sort the data records, use the Sort stage instead of using the Sort step of the Hierarchical Data stage.
- Use the Wave Generator stage to divide the data into waves, and then run the job for each wave of the data. This approach works only for parallel jobs.
- Break a job run with a large data set into multiple job runs. Each run processes a subset of the large data set.

Troubleshooting the Java Integration stage

Several common errors are specific to the Java Integration stage.

Errors in the stage editor
When using the Java Integration stage, you might encounter certain GUI errors.

Symptoms
- When you click **Configure** on the **Java Integration stage property editor**, you might encounter the following warning message:

  ```
  Failed to instantiate the resource wrapper class
  ```
- When you click **Select** on the stage property **Usage>User class** property, you might encounter the following warning message:

  ```
  Failed to send the request to the handler: The agent at HOST123:31531
  is not available.
  ```

Resolving the problem
Verify that the ASB agent is working and restart the ASB agent if necessary.

The ASB agent log file can help you understand the errors that occur in the stage editor. The file is named `asb-agent-xx.out` and is in the `ASBNode/logs` folder if `com.ibm.iis.cas.level=ALL` is added into the `ASBNode/conf/asbagent-logging.properties` file.

Runtime errors

After you configure and compile your Java Integration stage job, you might encounter certain run time errors.

Stage configuration issues:

When you run a job that includes a Java Integration stage, you get an error that is related to the configuration of the stage.

Symptoms

When a user class `userClass` is not found in the classpath that is specified at the **Java Integration stage property editor**, the following error might occur:

```
com.ascential.e2.common.CC_Exception: java.lang.ClassNotFoundException:
userClass
 at java.net.URLClassLoader.findClass(URLClassLoader.java:588)
 at java.lang.ClassLoader.loadClassHelper(ClassLoader.java:743)
 at java.lang.ClassLoader.loadClass(ClassLoader.java:711)
 at java.lang.ClassLoader.loadClass(ClassLoader.java:690)
 at com.ibm.is.cc.javastage.connector.CC_JavaConnection.connect
(CC_JavaConnection.java:155)
 at com.ibm.is.cc.javastage.connector.CC_JavaConnection.connect
(CC_JavaConnection.java: 250)
```

If a user class fails the connection test at the initialization time, the following error might occur for a class that is not inherited from `com.ibm.is.cc.javastage.api.Processor` class or `com.ascentialsoftware.jds.Stage` class.

```
Connection test failed (PXBridgeOp::negotiate, file pxbridge.c, line 1,684)
```

Resolving the problem

Add the user class file that is inherited from the Processor or Stage class, or a `.jar` file that contains the class to a directory that the parallel engine can access. In the Java Integration stage, set the **Usage - Java - Classpath** property to the path name that contains the file.

To confirm that the correct class path was specified, click **Select** on the **Usage - User class** property and select the user class.

JVM related issues:

You might encounter an error related to a JVM creation when running a Java Integration stage.

Symptoms

The conductor runs as a single process and creates a single JVM. Options for the JVM are specified in one of Java Integration stages in the job created. When you run a job which has multiple Java Integration stages, you might encounter the following error:

```
JVMDUMP039I Processing dump event
"systhrow", detail "java/lang/OutOfMemoryError"
at 2013/08/15 11:46:11
com.ibm.tools.attach.enable please wait.
JVMDUMP032I JVM requested System dump using
'C:\IBM\InformationServer\Server\Projects\dstage1\core.20130815.
114611.3932.0001.dmp'
in response to an event
```

This error might occur if maximum heap size is enough for one Java Integration stage but not all of the Java Integration stages.

Resolving the problem

1. Verify which JVM options are required.

2. Set an environment variable **CC_MSG_LEVEL** to 2.

3. Compile and run the job.

   You can see the JVM options in the following section of the job log:

```
CC_JNICommon.cpp:(634)
The JVM is started with these options =
-Dcom.ibm.tools.attach.enable=no
-Dcom.ibm.is.cc.options=noisfjars
-Xmx256m -Djava.class.path=
C:\IBM\InformationServer\Server\DSEngine/../DSComponents/bin/ccapi.jar;.;
C:\IBM\SQLLIB\java\db2java.zip;C:\IBM\SQLLIB\java\db2jcc.jar;
C:\IBM\SQLLIB\java\sqlj.zip;C:\IBM\SQLLIB\java\db2jcc_license_cu.jar;
C:\IBM\SQLLIB\bin;C:\IBM\SQLLIB\java\common.jar
```

4. You need to verify how much of the JVM is required and set the environment variable **CC_JVM_OVERRIDE_OPTIONS** with the JVM options after changing the value of the maximum heap size with a JVM option -Xmx.

```
$CC_JVM_OVERRIDE_OPTIONS =
-Dcom.ibm.tools.attach.enable=no
-Dcom.ibm.is.cc.options=noisfjars
-Xmx512m -Djava.class.path=
C:\IBM\InformationServer\Server\DSEngine/../DSComponents/bin/ccapi.jar;.;
C:\IBM\SQLLIB\java\db2java.zip;C:\IBM\SQLLIB\java\db2jcc.jar;
C:\IBM\SQLLIB\java\sqlj.zip;C:\IBM\SQLLIB\java\db2jcc_license_cu.jar;
C:\IBM\SQLLIB\bin;C:\IBM\SQLLIB\java\common.jar
```

Java Pack related issues:

You might encounter a fatal error when running a Java Integration stage migrated from an existing Java Pack job.

Symptoms

1. You might encounter `java.lang.ClassNotFoundException` when running a Java Integration stage job migrated with the Connector Migration Tool although the original Java Pack job did not throw any error.

```
Java_Client: java.lang.ClassNotFoundException:
myClass.MyTester at
java.net.URLClassLoader.findClass(URLClassLoader.java:588)
at java.lang.ClassLoader.loadClass(ClassLoader.java:743)
at java.lang.ClassLoader.loadClass(ClassLoader.java:711)
at java.lang.ClassLoader.loadClass(ClassLoader.java:690)
at com.ibm.is.cc.javastage.connector.CC_JavaConnection.connect
(CC_JavaConnection.java:155)
at (com.ibm.is.cc.javastage.connector.CC_JavaConnection::connect
(CC_JavaConnection.java: 250)
```

2. You might encounter the following fatal error when running a Java Integration stage job using Java Pack API:

```
JavaPackTransformer: java.lang.IllegalAccessError:
Class com/ibm/is/cc/javastage/connector/CC_JavaPackProcessor
illegally accessing "package private" member of class
com/ascentialsoftware/jds/Stage
at com.ibm.is.cc.javastage.connector.CC_JavaPackProcessor.<init>
(CC_JavaPackProcessor.java: 121)
at com.ibm.is.cc.javastage.connector.CC_JavaConnection.connect
(CC_JavaConnection.java: 167)
```

Resolving the problem

The first error might occur because in Java Pack, the $DSHOME directory is considered the root directory of the relative path specified at the stage property **User's Classpath** on the **Java Pack property editor**. In the Java Integration stage $DSHOME directory might not be the root directory. The subject exception occurs if the user class files were only placed at that directory relative to $DSHOME.

To correct the first error, you need to specify the absolute path at the stage property **Usage - Java - Classpath** on the **Java Integration stage property editor** or copy the class files at the directory relative to the project directory $DSHOME/../Projects/<project_name>.

The second error might occur because Java Integration stage runtime wrongly would invoke a Java Pack method not for Java Integration stage. You might specify a jar file named tr4j.jar for Java Pack into classpath.

To correct the second error, you need to remove tr4j.jar which is a jar file for Java Pack from the following:
* CLASSPATH environment specified at the Operating System, the DataStage project properties and the job properties.
* **Usage - Java - Classpaths** on the Java Integration stage property editor.
* **-classpath** or **-cp** option specified at **Usage - Java - VM option** on the **Java Integration stage property editor**.

Stage capabilities related issues:

You might encounter errors and warning related to stage capabilities when running a Java Integration stage job.

Symptoms
1. You might encounter the following fatal warning when running a Java Integration stage job:

   A wave maker cannot be emitted.
2. You might encounter the following fatal error when running a Java Integration stage job:

   Stage design is incompatible with Java code.
   Stage has 0 output links but Java code expects min.=1 max.=1.

Resolving the problem
The first error warning might occur because in your user class, you might have tried to emit a wave marker into an output link by calling OutputLink.writeWaveMaker(), but, it was not allowed since it set the capability to false in getCapabilities().

To fix the first issue, you need to modify your Java code to solve a conflict of the capability.

The second error might occur because in your user class, you expected the Java Integration stage must have 1 output link, but the Java Integration stage did not have any output link on your job design.

To fix the second issue, you need to modify your job design based on the capabilities defined in your Java code.

Column mapping related issues:

You might encounter a fatal error and warning when running a Java Integration stage job using JavaBeans.

Symptoms
1. You might encounter the following fatal error when running a Java Integration stage job using JavaBeans:

   Com.ascential.e2.common.CC_Exception: Java bean could not be set.
   Class name: userClass
   The mapping from columns to bean properties was not defined.
   Launch the Column mapping editor to define the column mappings for the bean.
   Bean class name: userInputBean
   at com.ibm.is.cc.javastage.connector.CC_JavaLinkImpl.setBeanInfo

```
(CC_JavaLinkImpl.java: 146)
at com.ibm.is.cc.javastage.connector.CC_JavaInputLinkImpl.setBeanClass
(CC_JavaInputLinkImpl.java: 128)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.updateLinkWithBeanInfo
(CC_JavaAdapter.java: 208)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.preRunNode
(CC_JavaAdapter.java: 421)
```

2. You might encounter the following warning when running a Java Integration stage job using JavaBeans:

```
com.ascential.e2.common.CC_Exception: Java bean could not be set.
Class name: userOutputBean
Type mismatch occurred. Column name: aaa, Column type class: java.sql.Date,
Bean property name: bbb, Bean property type class: java.lang.String,
Bean class name: userOutputBean
at com.ibm.is.cc.javastage.connector.CC_JavaLinkImpl.setBeanInfo
(CC_JavaLinkImpl.java: 146)
at com.ibm.is.cc.javastage.connector.CC_JavaOutputLinkImpl.setBeanClass
(CC_JavaOutputLinkImpl.java: 278)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.updateLinkWithBeanInfo
(CC_JavaAdapter.java: 218)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.preRunNode
(CC_JavaAdapter.java: 421)
```

Resolving the problem

This fatal error might occur because you specified `userInputBean` for an input link by coding `getBeanForInput()` in your user class *userClass*, but you did not set column mapping for that link.

The warning might occur because you might have specified `userOutputBean` for an output link and also configured that link, but, you might have selected a wrong DataStage column aaa of which column type is `java.sql.Date` and map it into the Java bean property bbb of which type is `java.lang.String`.

To fix the issues, perform the following actions:

1. Verify the link property **Usage - Column mapping**.
2. Click **Configure** to update the link property.
3. You can see the current setting on **Column Mapping Editor**. Modify the mapping by selecting the required setting from the pull down list of each DataStage column.
4. If you want to see all the JavaBean properties, click **Browse Objects**. You can see and select them by clicking check boxes on the **Select Bean Properties** dialog.
5. If you want to save new selections, click **OK** on the **Select Bean Properties** dialog and click **Finish** on the **Column Mapping Editor**.

Link related issues:

You might encounter a fatal error when running a Java Integration stage job with multiple input links or reject links.

Symptoms

1. You might encounter the following fatal error when running a Java Integration stage job with multiple input links:

```
[Input link 0] com.ascential.e2.common.CC_Exception:
User code returned the bean class "YourInput" for this link,
but this is not matched to "MyInput" specified in the design.
Launch the Column mapping editor to reconfigure the column mappings.
at com.ibm.is.cc.javastage.connector.CC_JavaInputLinkImpl.setBeanClass
(CC_JavaInputLinkImpl.java: 123)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.updateLinkWithBeanInfo
(CC_JavaAdapter.java: 208)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.preRunNode
(CC_JavaAdapter.java: 421)
```

2. You might encounter the following fatal error when running a Java Integration stage job with reject links:

```
com.ascential.e2.common.CC_Exception: Reject link is not configured.
Configure reject link and save it.
at com.ibm.is.cc.javastage.connector.
CC_JavaRecordDataSetConsumer.setRejectManager
(CC_JavaRecordDataSetConsumer.java: 182)
at com.ibm.is.cc.javastage.connector.CC_JavaAdapter.getRejectDataSetProducer
(CC_JavaAdapter.java: 349)
```

Resolving the problem

The first error might occur because you might have changed the link order after configuring the column mapping.

To fix the error, click the **Configure** button on the **Java Integration stage property editor** and configure the column mapping again. You always need to verify the link order on the **Link Ordering** tab on the **Java Integration stage property editor** if your job has multiple input and output links to prevent any inconsistencies between your job design and Java code.

To fix the second error, open the **Reject** tab and confirm the right input row was selected at the stage property **Reject - From Link** on the **Java Integration stage property editor**.

Troubleshooting the JDBC Connector stage

Several common errors are specific to the JDBC Connector stage.

The configuration file cannot be accessed

The `isjdbc.config` file is configured, but an error indicates that the configuration file cannot be accessed.

Symptoms

An error indicates that the configuration file cannot be accessed.

Resolving the problem

Complete the following steps:

- Ensure that the configuration file is saved in the `DSEngine` directory of the InfoSphere Information Server engine tier installation.
  - On Windows operating systems, the default location for the `DSEngine` directory is `C:\IBM\InformationServer\Server\DSEngine`.
  - On Linux and UNIX operating systems, the default location is `/opt/IBM/InformationServer/Server/DSEngine`.
- Ensure that the name of the file is `isjdbc.config`, with all lowercase characters and no leading and trailing white space characters.
- Ensure that the file is saved on the engine tier computer.
- Ensure that when the engine tier spans multiple hosts, the file is accessible from the same location on all of the hosts. You can make a copy of the file on the same `DSEngine` location on each of the hosts, or use a shared `DSEngine` network location that is accessible by all of the hosts.
- Ensure that the file has the read access for the InfoSphere DataStage user and the local Administrator user (Windows) or system root user (Linux and UNIX). On Linux and UNIX, check the permissions of the `.odbc.ini` file that is stored at the same location and set the same permissions on the `isjdbc.config` file.
- Ensure that the content of the configuration file is in the correct format.

  For more information about the format, see Setting up the driver configuration file topic.

Suitable driver not found when you use the JDBC Connector stage

Errors occur when the driver class path is not correctly specified in the JDBC configuration file.

Symptoms

An error occurs indicating that no suitable driver is found. The connector failed to locate a suitable driver for the specified URL value.

Resolving the problem

Ensure that the information about the driver class path and the name of the Driver class is included in the isjdbc.config configuration file as follows:

- The *CLASSPATH* entry in the configuration file must include the fully qualified class path of the driver. If the driver consists of more than one JAR file or depends on more JAR files, all the JAR file locations must be included.

  If the driver is implemented in one or more JAR files, ensure that the specified JAR files are present in the system. Also, ensure that the JAR files have the read and run permissions that are enabled for the InfoSphere DataStage user and the local Administrator user on Windows or system root user on Linux and UNIX.

- If you are not sure about the JDBC version of your driver, check for the presence of the file META-INF\services\java.sql.Driver in the driver JAR archive. If this file is present and contains a single value, which indicates the Driver JDBC class name, you do not need to specify this class name in the *CLASS_NAMES* entry.

  The *CLASS_NAMES* entry in the configuration file must include the fully qualified name of the Java class in the driver that implements the java.sql.Driver JDBC interface. This information is only required if the driver is based on the JDBC API version before version 4.0. For the name of this class, see your driver documentation.

The connector failed to detect and report all schema mismatches

The connector is configured to report mismatches between the column definitions on the link and the column definitions in the JDBC data source. However, the connector failed to detect and report all of the mismatches.

Symptoms

The connector tries to determine the JDBC data types of the columns in the data source that correspond to the columns on the link. If the connector determines that the data types are incompatible, it reports the schema mismatch warning.

The driver, the data source, and the usage scenario for the stage might prevent checks for data type mismatches from happening.

Causes

The connector checks for a schema mismatch on a best-effort basis. It might not always be able to detect all the discrepancies.

Resolving the problem

This feature must be used only as an aid for configuring the column definitions on the link. Schema mismatch checking happens during the job setup phase. Irrespective of the schema mismatch checking, the connector checks for more schema mismatches during the job processing phase. Extra checking happens when the connector provides values to the JDBC driver, or receives values from the JDBC driver. If then the connector detects data truncation or character set conversion issues, the connector reports the error and the job fails.

CREATE TABLE failed to run successfully

The connector is configured to create the target table automatically, but the CREATE TABLE that was generated failed to run successfully.

Symptoms

Depending on the column definitions on the input link, the driver, and the data source, the connector might not be able to generate a valid CREATE TABLE statement.

Causes

To generate the CREATE TABLE statement, the connector first maps each column definition on the input link from the InfoSphere DataStage data type to the closest matching JDBC data type. Then, the connector queries the driver to determine the data source native type that corresponds to that JDBC data type.

Depending on the input link data type, the connector might try several different JDBC data types. However in some cases the driver does not report a native data type for any of the JDBC data types that the connector presented. In those cases the connector uses the first JDBC data type in the CREATE TABLE statement. This in turn can cause the statement to fail to run. The connector writes the complete generated CREATE TABLE statement text in the job log.

Resolving the problem

To correct the error, complete one or more of the following steps:

- Use different data types for the input link columns. Ensure that the input link column definitions are valid for the actual data values that are represented by those columns when the job runs.
- Enter the CREATE TABLE statement manually. You can start with the generated CREATE TABLE statement by copying it from the job log to the **Create table statement** property in the stage editor.

  For more information, see the documentation for the data source where you want to create the table. Locate the native data types to use for the required columns in the statement, and then specify those data types in the statement directly.

  Make sure that the native data types that you specify for the table columns are valid for the actual data values that are written to those columns when the job runs.

For more information about mapping, see the following topic: ../com.ibm.swg.im.iis.conn.jdbc.usage.doc/topics/jdbc_mapping_datatype_to_nativedatatypes.dita

The job fails and reports data truncation or character set conversion errors

When the connector reads values from or write values to a character-based column in the data source by using a character-based column on the link, the job fails and reports data truncation or character set conversion errors, even if the Length property of the column on the link matches the length of the column in the data source.

Symptoms

The job fails and reports data truncation or character set conversion errors.

Causes

For the Char, VarChar, LongVarChar, NChar, NVarChar, and LongNVarChar columns on the link, the connector exchanges values with the JDBC driver as double-byte Java Unicode values.

For the NChar, NVarChar and LongNVarChar columns, and, the Char, VarChar, and LongVarChar columns for which the **Extended** property is set to Unicode, the connector exchanges values with the InfoSphere DataStage framework as double-byte Unicode values.

For the Char, VarChar and LongVarChar columns for which the **Extended** property is not set to Unicode, the connector exchanges values with the InfoSphere DataStage framework based on the character set encoding specified for the stage in the **Character set for non-Unicode columns** property. When this property is set to value **Default**, the default system locale character set encoding is used. When this property is set to value **Custom**, the character set encoding to use is specified in the **Character set name** property.

The Length property for columns on the link is measured in double-byte units for NChar, NVarChar, and LongNVarChar columns and Char, VarChar, and LongVarChar columns for which **Extended** property is set to Unicode and is measured in bytes for Char, VarChar, and LongVarChar columns for which **Extended** property is not set to Unicode. These differences might lead in incompatible column definitions even when the same Length is specified for the column on the link and the column in the data source.

For example, a column in the data source can be defined as VARCHAR(10) column and be able to store 10 double-byte Unicode values. Assume that the corresponding column on the link is defined as VarChar(10) column and **Extended** property is not set to Unicode. When the connector fetches the value from this column in the data source, it fetches it as a Java Unicode string that consists of 10 double-byte Unicode characters.

Assume that the connector must convert this value to UTF-8 encoding, either because **Character set for non-Unicode columns** property was set to Default and UTF-8 is the character set encoding in the current system locale, or because **Character set for non-Unicode columns** property was set to value **Custom** and UTF-8 character set encoding was specified in the **Character set name** property. Assume further that some of the double-byte characters in the fetched value result in 2, 3 or 4-byte UTF-8 characters. The converted value in that case can require anywhere between 10 - 40 bytes of storage. If it requires more than 10 bytes of storage, the connector detects that the Length of the link column is not sufficient and reports an error.

Resolving the problem
To correct the error as mentioned in the previous example, complete one of the following actions:
* Change the data type of the column definition on the link to NVarChar(10).
* Set the **Extended** property for the column to Unicode.
* Increase the Length in the column definition to 40 to accommodate all the values.

For more information, see Data type mappings and Character set for non-Unicode columns property.

The job fails and unused columns were detected
When the connector detects a column on the link that is not used by the statements that the connector runs for the records on that link, the connector reports an error for the detected unmatched column.

Symptoms
The connector reports an error for the unmatched column on the link and the job fails.

Resolving the problem
The resolution of this error depends on whether the connector is configured to read data from a JDBC data source or write data to a JDBC data source

When the connector is configured to read rows from the data source, ensure that the SELECT statement that the connector uses to read the rows includes a column in its select list that corresponds to the column on the link. Alternatively, remove the unmatched column from the link.

The following table includes examples of scenarios and respective problems that you might encounter when the connector is configured to read data from a JDBC data source and the resolution for the problems.

SELECT statement	Output link columns	Problem	Resolutions

`SELECT C1,`
`C3 FROM TABLE1`
	C1, C2 and C3	The link column C2 is treated as an unmatched column and results in error	Complete one of the following steps:
			• Change the statement to
			  `SELECT C1, C2,`
			  `C3 FROM TABLE1`
			• Remove column C2 from the output link.

The SELECT statement is generated automatically and the source table TABLE1 contains columns C1, C3 and C4	C1, C2 and C3	The connector does not include the table column C4 in the generated statement and does not fetch data for it because it determines that a corresponding column is not defined on the output link. But the connector includes column C2 in the generated statement because it is present on the output link and that leads to an error because the table does not have the column with that same name.	Complete one of the following steps: • Add column C2 to the table. • Remove column C2 from the output link.

When the connector is configured to write records from the input link to the data source, each column on the input link must be referenced by one or more SQL statements that the stage uses to write records to the data source. If a column on the input link is not referenced the behavior depends on the value of the **Drop unmatched fields** property. When this property is set to **No**, the connector reports an error, and the job fails. When the property is set to **Yes**, the connector reports an informational message for each unused input link column to indicate that the column is ignored, and the job continues.

The following table includes examples of scenarios and respective problems that you might encounter when the connector is configured to write data to a JDBC data source and the resolution for the problems.

Write modes	Input link columns	Problem	Resolutions

The connector is configured to run in Insert write mode and to run the statement
`INSERT INTO`
`TABLE1(C1, C2, C3)`
`VALUES (`
`ORCHESTRATE.C1,`
`'job 1',`
`ORCHESTRATE.C3)`
	C1, C2 and C3	The connector determines that the input link column C2 is not referenced by the INSERT statement and acts based on the **Drop unmatched fields** property value.	Complete one of the following steps:
			• Change the INSERT statement to
			  `INSERT INTO`
			  `TABLE1(C1,`
			  `C2, C3)`
			  `VALUES`
			  `(ORCHESTRATE.C1,`
			  `ORCHESTRATE.C2,`
			  `ORCHESTRATE.C3)`
			• Remove column C2 from the input link.

Write modes	Input link columns	Problem	Resolutions
The connector is configured in run in Insert write mode and automatically generate the INSERT statement to use to insert records into the table TABLE1 with columns C1, C3 and C4	C1, C2 and C4	The connector does not include the column C4 in the generated INSERT statement because a corresponding column is not on the input link. Because column C2 is on the input link and does not have a matching column in the target table, the connector treats this input link column as an unmatched column.	Complete one of the following steps: • Add column C2 to the table. • Remove column C2 from the input link.

Troubleshooting the Greenplum Connector stage

Several common errors are specific to the Greenplum Connector stage.

Connection with the gpfdist program fails

The Greenplum parallel file distribution program, gpfdist, runs on each InfoSphere DataStage node, but an error indicates that the program cannot be accessed.

Symptoms

You get an error with the 110 error code, which indicates that network access cannot be reached between the gpfdist program and the Greenplum server.

Resolving the problem

At run time, the gpfdist program is accessed by the Greenplum server. To ensure that the Greenplum server has network access to the InfoSphere DataStage node where the gpfdist runs, complete one or more of the following tasks:

- Test connectivity from the Greenplum server to the InfoSphere DataStage node by using the **ping** or **wget** commands.
- Because the gpfdist program uses the http protocol, ensure that an http connection is open and accessible on the Greenplum server and the InfoSphere DataStage node. To test http connectivity, use a web browser or the **lynx** command.
- Ensure that the host name for the InfoSphere DataStage node as it appears in the InfoSphere DataStage parallel configuration file fastname entry can be resolved by the Greenplum server. You might need to add the IP address and host name of the InfoSphere Information Server engine tier to the network hosts file.

The connector was not able to start the gpfdist program

The Greenplum parallel file distribution program, gpfdist, runs on each InfoSphere DataStage node, but an error indicates that the program could not be started.

Symptoms

You get an error that indicates that the program could not be started. You might also get an operating system error message that indicates that the gpfdist executable file is not in the system path.

Causes

The gpfdist program is not in the system path on the InfoSphere Information Server engine tier.

Resolving the problem

Complete the following tasks:

- Update the following:

- – ▓Windows▓ The system environment if you are using Windows, or
- – ▓Linux▓ The `dsenv` file if you are using Linux.
- Verify if you have followed all the instructions included in the Greenplum parallel file distribution program (`gpfdist`) topic in the IBM InfoSphere Information Server Connectivity Guide for Accessing Greenplum Databases.
- Ensure that the operating system that the user associated with the DataStage has access to `gpfdist` and the permission to execute `gpfdist` . You can allow access and permissions to execute `gpfdist` on the directory `$GPHOME_LOADERS/bin`.

External table has more `gpfdist` instances than available primary segments

The Greenplum connector reports an error when the connector uses more processing nodes than the number of primary segments that are available in the Greenplum server.

Symptoms

An error from the driver reports that the external table has more `gpfdist` instances (or URLs) than the number of primary segments that are available.

Resolving the problem

The number of `gpfdist` instances invoked by the Greenplum connector is determined by the number of processing nodes used by the DataStage job. Each processing node invokes and connects to an instance of the `gpfdist` program that serves a location that is defined in the external table.

To resolve the error, identify the number of processing nodes for the connector and complete one of the following tasks:

- Decrease the number of nodes defined in the parallel configuration file (`$APT_CONFIG_FILE`), to a value less than or equal the number of the primary segments that are available.
- Create a node pool to define a subset of processing nodes for the Greenplum connector with a lesser number of processing nodes than available segments.
- Verify that the Greenplum segments are not occupied by other processes and that the Greenplum server is configured to allow sufficient number of segments to scan external tables. You can control the number of segments that will scan external tables using the Greenplum server configuration parameter `gp_external_max_segs`.

Data truncation with Greenplum geometric or multiple format data types

When the Greenplum connector uses the InfoSphere DataStage Varchar data type to read from Greenplum geometric or multiple format data types, some data might get truncated.

Symptoms

When using the InfoSphere DataStage Varchar data type to read from Greenplum geometric or multiple formats data types, the data that is sent to the output link is truncated. The connector does not report a warning or error.

Causes

Some Greenplum data types do not map directly to an InfoSphere DataStage type and use the Varchar data type on the column grid. If the length that is defined for the Varchar column is too short to contain all of the data for one of these Greenplum data types, data is truncated.

Resolving the problem

Increase the column length of the Varchar column on the link so that the column can contain all of the data in the record.

Jobs using key columns fail with ERROR: operator does not exist

The Greenplum server does not support using the equals operator (=) for at least one of the columns in the SQL statement.

Symptoms

An error is reported by the driver from the Greenplum server indicating that an operator does not exist for the type *<type>*. The error message might appear as follows:

```
ERROR: operator does not exist:<type>=<type> (Hint No operator matches the given name and
argument type(s). You may need to add explicit type casts.
```

Causes

When the connector generates the SELECT, DELETE, or UPDATE statements the columns specified in the **Key columns** property are used in the WHERE clause of the statements. The columns used in the WHERE clause must support the equals operator. This error occurs because one or all of the data types of the key columns (for example, xml, polygon) is not supported with the equal operator.

Resolving the problem

Perform one of the following steps:

* Use a User defined statement and cast the offending column of type *<type>* to a type that supports the equal operator, like text.
* Remove the offending column from the Key columns list.

Jobs fail with permission error

The Greenplum connector reports an error when it reads or writes data by using the `gpfdist` program and external tables.

Symptoms

An error from the driver indicates that you do not have the privilege that is required to create external tables.

```
ERROR: permission denied: no privilege to create a type gpfdist(s) external table
```

where, *type* is the type of external table that the connector is creating.

Causes

The job fails because the user or role that is provided to the connector does not have the privileges that are required to create external tables.

Resolving the problem

To resolve the error, you must grant the CREATEEXTTABLE privileges to the user or role that the connector uses to connect to the Greenplum database.

For example, `ALTER ROLE dsuser WITH CREATEEXTTABLE;`

Join stage

You can troubleshoot errors that are specific to the Join stage.

Join stage exits with a heap allocation error
Symptoms

The join stage exits with a heap allocation error.

Causes

The Join stage processes the data with the primary link or driving data sets first. It gets a row from this link and then retrieves all rows with matching values in the secondary link, which is also called a reference link.

These rows are temporarily stored in memory. If there are many rows in the secondary link that match the current record in the primary link, then a large amount of memory is allocated to hold the result. If there is not enough heap memory, the job fails.

Because of this, the amount of memory that is used by the Join stage depends heavily on the cardinality of the reference side. The lowest the cardinality of the reference side, the larger the amount of memory that it uses for each row.

Note: Cardinality is here the uniqueness of data values of column. A column with all unique values is said to have high cardinality, while a column with repeated values is said to have low cardinality.

Resolving the problem

If you are working with a symmetric join such as Inner Join or a Full Outer join and you are experiencing heap memory problems, you might be able to work around these problems. Make sure that the data set with the lowest cardinality is used as the primary link.

Changing this link order does not impact the outcome of the Join because Inner Joins and Full Outer Joins are symmetric. The sides are interchangeable and the order does not change the outcome of the Join.

If you are working with non-symmetric joins such as Left Outer and Right Outer, then changing the link order does have an impact in the output of the data and therefore you must not change the link order unless you understand the consequences of this change. For a more detailed explanation about the differences between the different types of joins that this stage can perform, refer to the "Parallel Job Developer Guide".

Join stages fail with a `UNIX Signal 11 (SIGSEGV)` message

A job with a join stage fails with an `Unexpected terminated by Unix Signal 11 (SIGSEGV)` message.

Symptoms

An InfoSphere DataStage job with a join stage fails with the following error: `Unexpected terminated by Unix Signal 11 (SIGSEGV)`

Causes

If the size of the record is larger than the default setting of 20 MB, the sort that is inserted for the join fails.

Resolving the problem

First, you must calculate the size of your largest record, in bytes. Then, double that size and set the *APT_TSORT_STRESS_BLOCKSIZE* environment variable to that value in bytes. If this environment variable does not exist, create it first.

The *APT_TSORT_STRESS_BLOCKSIZE* variable sets the size (in bytes) of the shared memory block that is used to pass data between the writer, sorter, and merger processes within the sort operation. The default is set so that each sort in a job tries by default to `mmap` 20 MB each for these in-memory functions. In the case when a single record is larger than 20 MB, this step is necessary to increase the size of the block.

The *APT_TSORT_STRESS_BLOCKSIZE* environment variable can be set at the job level. The *APT_TSORT_STRESS_BLOCKSIZE* environment variable can also be set at the project level by defining a project level default value.

Lookup Stage

You can troubleshoot several errors with the Lookup Stage.

InfoSphere DataStage Lookup Stage fails on Linux
Symptoms
The InfoSphere DataStage Lookup Stage fails on the Linux operating system.

Causes
Lookup stage creates files in the resource disk area that use the C++ `mmap` function. When those files are used on an NFS or shared mount, the `mmap` function might fail. This issue is a known issue on Linux. The problem is caused by the C++ libraries, and not InfoSphere DataStage.

Resolving the problem
To allow the lookups to create the files that are required on a Linux NFS mount without using the `mmap` function at file creation, set the *APT_LUTCREATE_NO_MMAP* environment variable to 1 at the project level. There is no requirement to recompile the jobs as this environment variable takes effect at job run time.

Parallel job with lookup fails with a File too large error
Symptoms
A InfoSphere DataStage parallel job that contains a lookup fails with the following error:

```
Lookup_107,0: Error writing table file "/d01/Ascential/DataStage/Datasets/lookuptable.20100217.abcde":
  File too large
```

Causes
The lookup table is too large to fit in available memory.

Resolving the problem
Hash partition the lookup data and add more nodes to the parallel configuration file to break the lookup data into smaller pieces.

Errors when jobs finalize or save a table
When a job attempts to write to a lookup data set, you might receive an error that the job failed to `finalizing / saving table`.

Symptoms
When an InfoSphere DataStage job attempts to write to a lookup data set, the following error occurs:
```
Event: test_lkup_by_key_fs,2: Error finalizing / saving table /home/data/datasets/
test_lkup_by_key.fs
```

Causes
This error is caused by one of the following issues:
- The user ID that runs the InfoSphere DataStage job does not have permission to write to the directory that is shown in the error. The user ID is identified in each event message in job log for the failing job.
- The volume that contains the output directory that is stated in error message does not have enough free space to write the file.
- An `out of memory` error precedes the data set write error in login InfoSphere DataStage 8.1. In InfoSphere DataStage 7.5.x, the only error is the space error. There is no additional memory-related error message even when that is the cause of failure when the job writes the lookup data set.
- Output of temporary data sets might occur to the directory specified by the *UVTEMP* setting in `uvconfig` file. Parallel jobs can also write output to the directory specified by *tmpdir* environment variable. Ensure that those directories have sufficient space for the file.

Resolving the problem
Check for conditions that can cause the job to fail.

1. Log in to server that uses the user ID that was running the job, then go to the failing directory and confirm you can create a new file. For UNIX operating systems, use the touch command to write a small test file. If that test fails, then you must adjust permissions of target directory to ensure that InfoSphere DataStage users can write to it.

2. If directory permissions are set correctly, confirm that target volume has sufficient free space. Data sets can often be several gigabytes in size, so some analysis of the job might be necessary to estimate how much data is written to the data set.

3. If the target volume has sufficient free space, obtain a **full detail** job log for the failing job and look for more errors. One common error when you write large data sets is the following out of memory error: `Event: test_lkup_by_key,2: Could not map table file "/home/data/datasets/n3/lookuptable.20091216.xxhxipb (size 2956470664 bytes)": Cannot allocate memory`

This memory error typically occurs when you build large data sets because of the requirement to load the data set into memory first. The lookup stage uses memory mapped files. You must have enough system memory available to store the entire contents of the file AND enough disk space to shadow the file in memory.

In a 32-bit system, there is a 2 GB limit on size. Refer to the next troubleshooting tip for more information on dealing with the "could not map table file" error:

`DataStage job fails with "Could not map table file"`

InfoSphere DataStage job fails with `Could not map table file`
Symptoms
An InfoSphere DataStage job fails with the following error: `Could not map table file`. InfoSphere DataStage fails to load lookup data into memory or create lookup file.

Causes
InfoSphere DataStage is limited to the amount of memory that can be allocated for a Lookup.

Resolving the problem
The job must be more scalable. By default the Lookup stage uses Entire partitioning for the Lookup data. This ensures that no matter what partition the source data is in, all the Lookup data is available to be matched.

The following list provides the types of system constraints that can result in the memory allocation failure for a Lookup stage.

*   A single Lookup stage in the designer with multiple Lookup inputs parses out to the same corresponding number of Lookup operators in the generated osh script.
*   Each Lookup operator has it's own physical process for each partition that is defined by the configuration file, though it can combine with an upstream or downstream operator.
*   Each physical process can address only up to 2 GB of memory for 32-bit server systems such as Windows.
*   Each Lookup REQUIRES contiguous memory allocation.
*   Each process is limited to the *ulimit* setting of the InfoSphere DataStage environment, which can be limited by *LDR_CNTRL* on AIX)
*   Each Lookup data set uses Entire partitioning method by default.
*   With Entire or Auto partitioning, one memory segment is used and shared across all partitions for a physical server, which is defined by *fastname* in the configuration file. For MPP environments, each server gets its own copy.
*   When you use partitioning other than Entire or Auto, each partition uses its own copy of data in memory. Each partition can use up to 2 GB for 32-bit systems, or the *ulimit -d* data setting depending on which one is the most restraining

- All Lookup data for Lookups is processed to a file in scratch that is then loaded to a `mmap` (C++ function) structure in memory. This happens before ANY source data is processed for the Lookup.
- Use of hash partitioning. Hash on the lookup keys for both the source and lookup data.

As the amount of Lookup data increases, you can add more nodes to the configuration file to further distribute data across more processes and thus more memory segments.

Several factors can lead to a job that runs successfully then starts to fail. The server might have fragmented memory if it is not rebooted, or the amount of data that is used in the Lookups might grow to just over the limit. Other applications might be running concurrently, which uses resources that were previously available to InfoSphere DataStage. Available memory might be reconfigured by creating or moving lpars.

The *LDR_CNTRL* environment setting on AIX might limit the *ulimit -d* data setting even if you have hard limit set higher.

Lookup file does not span across multiple scratch resource disks

You can force lookup to a keyed partition to reduce the amount of data in each partition if you have more than one node.

Symptoms

The lookup file that is generated by a InfoSphere DataStage lookup does not span across multiple scratch resource disks that are defined per node when scratch fills up.

Causes

InfoSphere DataStage lookup files are memory mapped files, so there can be only one file per lookup process.

Resolving the problem

If the lookup file is too large, you can force the lookup to a keyed partition method like hash, which reduces the amount of data per partition if you have more than one node.

Note: The lookup stage uses Entire partitioning by default, which forces all records into a single memory segment. Adding nodes to the configuration file reduces the amount of data per lookup process.

Troubleshooting the Oracle Connector stage

You can troubleshoot errors that are specific to the Oracle Connector stage.

Oracle connector runs in sequential mode when a reject link has a constraint violation reject condition

When an Oracle Connector stage is configured to bulk load data and has a reject link where the **SQL error - constraint violation** reject condition is selected, the connector runs in sequential mode.

Symptoms

When you run the job that contains the Oracle Connector stage, you get the following message:

```
[IIS-CONN-ORA-003004] The connector was configured to load data in parallel but the reject
condition for checking constraints was selected for the reject link. This combination is not
supported. The connector will run in sequential mode.
```

Causes

When the following conditions are met, the Oracle connector must run in sequential mode:
- The connector is configured to write data in bulk load mode
- A reject link is defined for the stage
- The **SQL error - constraint violation** reject condition is specified for the reject link

Suppose that a stage is configured to reject rows that violate the constraints, to disable constraints before the load, and to enable them after the load. The following steps occur:

1. The connector disables the constraints before the load and then loads the data.
2. The ROWID values of the rows that violated the constraints are stored in the exceptions table on the Oracle database.
3. The connector sends the rows that failed the constraints to the reject link.
4. The connector deletes the rows that failed the constraints from the target table and enables the constraints.

All rows that violate the constraints are rejected. For example, suppose that two rows that have the same primary key value are loaded. Because this condition violates the primary key constraint, both rows are rejected.

In this scenario, the Oracle connector must run in sequential mode because of the way that parallel jobs that contain the Oracle connector work. A parallel job uses one conductor process and one or more player processes for each stage. When the Oracle connector uses player processes, the processes are independent of each other and cannot detect when other player processes start or end. Only the conductor process can detect when the player processes are complete, and rows cannot be rejected until all the player processes are complete. However, only player processes can access the reject link for a stage. As a result, the connector must run in sequential mode and use only one player process to load the data.

Resolving the problem

Complete one of the following tasks:

* If constraints are not defined for the target table, clear the **SQL error - constraint violation** reject condition for the reject link. The job can then run in parallel mode.

* Use the insert write mode instead of the bulk load write mode. When the connector uses the insert write mode, constraints remain enabled while the player processes insert data to the table. Constraint violations are reported immediately to the player processes, and the player processes can send rows that violate constraints to the reject link.

* Enforce constraints after data is bulk loaded to the target table. Instead of configuring the stage to reject rows that violate the constraints automatically, complete the following steps:

    1. Disable the constraints on the table.
    2. To bulk load data to the table, run the job.
    3. Process the exceptions table and enable the constraints manually, or specify a PL/SQL block in the **After SQL** property of the stage.

* On the Advanced page for the stage, set the **Execution mode** property to **Sequential**. The stage runs in sequential mode, and message IIS-CONN-ORA-003004 is not logged if the **SQL error - constraint violation** reject condition is selected on a reject link.

Sequential File Stage

You can troubleshoot errors that are specific to the Sequential File Stage.

Unable to view data or run a Parallel job with a Sequential File stage

In a InfoSphere DataStage Parallel job with a Sequential File stage, cannot view data or run the job.

Symptoms

The following error message is generated: `File archive: Trouble creating file`

For a parallel job with a Sequential File stage plug-in, viewing the data results in the following error:

`IIS-DSEE-TFAR-00015 00:10:13 <main_program> File archive: Trouble creating file "/tmp/...."`

Run time results in the following errors:

- Message Id: IIS-DSEE-TFAR-00015 Message: `main_program: File archive: Trouble creating file "/tmp/...."`
- Message Id: IIS-DSEE-TFPX-00002 Message: `main_program: Fatal Error: Null archive.`

Causes

The program is searching for relative path called "tmp" that is not present. This occurs on Windows installations when the project is not on the same drive as the engine.

Resolving the problem

Create a directory that is called `temp` at the root of the drive where the InfoSphere DataStage project is located. For example, if the InfoSphere DataStage projects are on the `D:` drive, create the following directory: `D:\tmp` If the directory exists, check the remaining disk space on your drives to ensure that limited disk space is not the cause of the problem.

Teradata Connector Stage

You can troubleshoot errors that are specific to the Teradata Connector Stage.

Error opening the message catalog errmsg.cat file

You must set the Teradata environment variables correctly to avoid errors with the `errmsg.cat` file.

Symptoms

You receive the following error: `file errmsg.cat can't be opened" from any of the Teradata Stages (Teradata plugins / TDEE / TDCC)`

Causes

The issue is caused by the incorrect settings of the Teradata *COPLIB* and *COPERR* environment variables.

Resolving the problem

The *COPLIB* environment variable must be set to point to the directory-path, which contains the `clispb.dat` file. The *COPERR* environment variable must be set to point to the directory-path, which contains the *errmsg.txt* and *errmsg.cat* files.

By default, the Teradata client install-program installs all the `clispb.dat`, `errmsg.txt`, and `errmsg.cat` files into /opt/teradata/client/liband creates links in /usr/lib. On HP-Itanium, the files are copied to /opt/teradata/client/lib64with links in /usr/lib/hpux64.

Setting ENV-VARs: COPERR and COPLIB to the correct folder-path resolves the problem.

Unable to load the Teradata connector library when you view data

You can correct failures to connect to the Teradata connector library by modifying the *dsenv* variable.

Symptoms

You are unable to load the Teradata connector library when you view data with the connector. The runtime functions properly, but the view data and test connection are not working.

Causes

This problem is caused when the ASB agent that runs on the engine system is unable to find the connector library or the Teradata libraries. The agent is started with the /local/IBM/InformationServer/ASBNode/bin/NodeAgents script. This script sources the /local/IBM/InformationServer/ASBNode/bin/NodeAgents_env_DS.sh file for any DS-specific environment. This script is sourcing the /local/IBM/InformationServer/Server/DSEngine/dsenv. The script eventually starts the Agent.sh script to start the agent.

Resolving the problem

If you modify the *dsenv* variable, restart the ASB agent.

Blocking and deadlock while stream jobs are running in parallel

When you run write operations in parallel mode, row-hash collisions might occur.

Symptoms

You experience blocking or deadlocks while stream jobs are running in parallel.

Causes

When InfoSphere DataStage runs write operations in parallel mode through immediate-mode or through the stream-operator, row-hash collisions might occur. These collisions can cause blocking and deadlocks.

Any updates for a particular row must come from the same partition to avoid a deadlock. Each partition uses a separate connection to the database. If multiple updates for a row do not come from the same connection, it would cause blocking or deadlocks.

Resolving the problem

If you are using Auto partitioning, you must have the partition. Create a hash on the key that was used in the WHERE clause of the UPDATE or DELETE statement. For an INSERT statement, create a hash on the table's primary key.

Hash partitioning helps to avoid blocking and deadlocks, but it does not always resolve the problem. You might continue to encounter blocking and deadlocks if a key in partition 0 has the same row-hash as a key in partition 1.

You can also reduce blocking and deadlocks by committing often. If the job has the record count property that is set to 0, then the commit happens at the end of the job. You can set the record count to be equal to the array-size and ensure that the commits happen at shorter intervals.

You can also reduce both the array size and the commit interval to lower values. By committing often, the sessions release their write locks and allow other sessions to write.

Sort Stage

You can troubleshoot errors that are specific to the Sort Stage.

InfoSphere DataStage outputs a warning message about a partition Sort key "CO_ID" no longer exists in data set schema

You can correct warnings about a partition Sort key with an environment variable.

Symptoms

InfoSphere DataStage outputs a warning message about a partition key. `main_program: Sort key "CO_ID" no longer exists in data set schema. It is dropped from the inserted sort and merge collector. main_program: There are no sort keys in the data set schema. No parallel sort and merge operator is inserted.`

Resolving the problem

Add the *APT_NO_SORT_INSERTION* environment variable to job and set the value to 1.

Transformer Stage

You can troubleshoot errors that are specific to the Transformer Stage.

Checking composite operator errors

Compilers at higher or lower versions than the requested level might not function correctly.

Symptoms

InfoSphere DataStage jobs contain a transformer stage call to an external C compiler, and they compile correctly when the InfoSphere DataStage job runs. However, under some conditions the compile fails with many composite operator errors. The operator errors are similar to the following errors:

- `##E IIS-DSEE-TBLD-00076 15:20:18(000) <main_program> Error when checking composite operator: Subprocess command failed with exit status 256.`

- `##W IIS-DSEE-TFTM-00012 15:20:18(002) <transform> Error when checking composite operator: The number of reject datasets "0" iless than the number of input datasets "1".`

- `##W IIS-DSEE-TBLD-00000 15:20:18(007) <main_program> Error when checking composite operator: Output from subprocess: Error 8: "/usr/include/machine/sys/_types.h", line 65 # Invalid type specifier combination in declaration: "short double".`

Causes
If many errors occur when the job checks the composite operator, the compiler that is used with InfoSphere DataStage is incompatible, unsupported, or the compiler uses the incorrect compiler or linker options.

Resolving the problem
Each version of InfoSphere DataStage expects to run with a specific set of compilers at a known and tested maintenance level. Compilers at higher or lower versions than the requested level might not function correctly. The links below list the compilers that are supported by InfoSphere Information Server releases 8.0, 8.1, 8.5 and 8.7:

- InfoSphere Information Server 8.0 http://www-01.ibm.com/support/docview.wss?uid=swg27008923
- InfoSphere Information Server 8.1 http://www-01.ibm.com/support/docview.wss?uid=swg21315971
- InfoSphere Information Server 8.5 http://www-01.ibm.com/support/docview.wss?uid=swg27016382
- InfoSphere Information Server 8.7 http://www-01.ibm.com/support/docview.wss?uid=swg27021833

If the compile error mentions only a few errors when it checks the composite operator and mentions many errors for missing include files, review the following related question in the troubleshooting guide: https://www-304.ibm.com/support/docview.wss?uid=swg21469811

Another similar compile failure can occur on operating systems that support multiple processor types if the compiler is not the correct edition for that processor type. For example, the errors can occur when you compile a InfoSphere DataStage job on an HP-UX Itanium system and the `/opt/aCC` compiler is the PA-RISC edition of compiler instead of the Itanium edition of the compiler.

You can use the file command on the compiler file to check the processor type file in `/opt/aCC/bin/aCC`.

InfoSphere DataStage jobs with Transformer Stage fail to compile on AIX
Confirm that the required files for Transformer Stage compiles are included in the `vacpp.cmp.include` file set.

Symptoms
When you compile an InfoSphere DataStage job that contains the transformer stage on AIX, the compile fails with the following errors:

- `##W IIS-DSEE-TBLD-00000 17:52:00(010) <main_program> Error when checking composite operator: Output from subprocess: "/opt/IBM/InformationServer/Server/PXEngine/include/apt_components/ transformop/transformbasehdrs.h", line 41.10: 1540-0836 (S) The #include file <map> is not found`

- `"/opt/IBM/InformationServer/Server/PXEngine/include/apt_framework/operator.h", line 70.10: 1540-0836 (S) The #include file <vector> is not found`

- `"/opt/IBM/InformationServer/Server/PXEngine/include/apt_util/custreport.h", line 36.10: 1540-0836 (S) The #include file <string> is not found`

- "/opt/IBM/InformationServer/Server/PXEngine/include/apt_util/iostream_s.h", line 23.10: 1540-0836 (S) The #include file <iostream.h> is not found

Diagnosing the problem

The following errors indicate that required include files that are normally part of the `vacpp.cmp.include` file set were not installed:

- The #include file <map> is not found
- The #include file <vector> is not found
- The #include file <string> is not found
- The #include file <iostream.h> is not found

Resolving the problem

If you are using the vacpp or XL C compiler on AIX, issue the following command to confirm if the required file set was installed: `lslpp -f vacpp.cmp.include` If you receive the `lslpp: 0504-132 Fileset vacpp.cmp.include not installed` message, the AIX administrator must install the missing file set.

If the file set is already installed, then the `lslpp -f vacpp.cmp.include` command gives the location of the include files. Usually the include files are located at the following path: `/usr/vacpp/include/` Confirm that the missing include files that were listed in the compile errors do exist in that directory and have public read permission set.

Errors when you import a job with Transformer Stages into a different system

The *APT_TRANSFORM_OPERATOR_DEBUG* environment variable can cause errors when you import a job with Transformer Stages into a different system

Symptoms

When you import jobs with Transformer Stages into a different system, you receive the following error message:

```
'RT_BP123.O/V0S11_TESTJOB1_Transformer.C' is of unknown format
Error processing file RT_123.O/V0S11_TESTJOB1_Transformer.C. file not modified


Command 'CATALOG RT_BP123
V0S11_TESTJOB1_Transformer.C
V0S11_TESTJOB1_Transformer.C LOCAL FORCE error:
Program V0S11_TESTJOB1_Transformer.C was not compiled with a supported
 version of the BASIC compiler. It must be recompiled.
```

Causes

This issue is related to having the *APT_TRANSFORM_OPERATOR_DEBUG* environment variable set. When this environment variable is set, the file with the "C" code is maintained in the BP.O directory.

When you then export with Job Executables, the `.C` file is included in the export. When this binary section is included, you always get this error on import.

Resolving the problem

1. Remove the *APT_TRANSFORM_OPERATOR_DEBUG* environment variable on original system.
2. Recompile and re-export the job.
3. Import job into the production environment.

Transformer code cannot be loaded at run time when you use non-US ASCII characters in module names

NFS or OS does not recognize non-US ASCII characters in module names.

Symptoms

You receive an error that is similar to the following error:

```
"V195S1_EMA_Bad_If_logic_Québec"
(for class "APT_TransformOperatorImplV195S1_EMA_Bad_If_logic_Quebec9_u")
 failed to load: Could not load "V195S1_EMA_Bad_If_logic_Québec": Invalid argument.
```

Causes
Globalization settings allow non-US ASCII characters to be used in Transformer Stage Naming. However, the NFS or OS does not recognize non-US ASCII characters in module names.

Resolving the problem
Use only the character sets that are supported by the OS and NFS mounts so object modules are found and available on all compute nodes.

Replacing é with e in the transformer stage name resolves this issue.

Jobs with Transformer Stages that use remote nodes fail
The projects directory must exist and be accessible on the remote node.

Symptoms
InfoSphere DataStage parallel jobs with Transformer stages that use remote nodes fail with the following error:

```
Item #: 19
Event ID: 126
Timestamp: 2010-06-23 13:37:03
Type: Fatal
User Name: t2etl01
Message: trn: Failed to distribute the shared library
"/datastage/DataStage/Projects/ProjectName/RT_BP123.O/V10S0_xxxxxxxx_trn.o"to
node "nodeName". [transform/transform.C:1827]
```

Causes
This error is a result of the projects directory either not in existence or not accessible on the remote node.

Resolving the problem
You can resolve this problem with one of the following actions:

• Ensure that the Projects directory is NFS mounted to the remote node to avoid the need for copying of the transformer objects to the remote node. NOTE: This is the preferred method of resolution.

• If you chose not to have the Projects directory NFS-mounted to the remote node, create a Projects directory structure on the remote node and ensure that permissions allow access to all InfoSphere DataStage users. Then, ensure that the InfoSphere DataStage uses can run these commands. If you disallow rcp and require scp to be used, ensure that you created $APT_ORCHHOME/etc/remcp file to use scp instead of rcp. Transformer objects are copied to the remote node with either rcp or scp.

Director logs are not showing warning messages for all records dropped by the Transformer stage
You can set the maximum number of log reject messages to unlimited, which prevents records from being dropped.

Symptoms
The InfoSphere DataStage Director logs are not showing warning messages for all records dropped by the Transformer stage.

Causes
InfoSphere DataStage displays only 50 warning messages in the Director log per node when records are dropped by the Transformer stage. When more than 50 records are dropped by a Transformer stage, the InfoSphere DataStage Director log displays up to only 50 warning messages per node before it goes into silent mode with the following warning message: `Warning, all other rejected records will be silent.`

Resolving the problem

Increase or set the maximum number of reject messages to an unlimited value by altering the **Maximum log reject messages** property in the Transformer stage.

1. Right-click the Transformer Stage and select Properties
2. Open Stage properties
3. Go to Stage Tab -> General Tab
4. Set **Maximum log reject messages** to the required value, or set it to -1 for unlimited. If no value is specified, the maximum number of log reject messages defaults to 50 messages per node.

Unable to compile a transformer stage in a InfoSphere DataStage Parallel job

Symptoms

When you try to compile a transformer stage in a InfoSphere DataStage Parallel job, you receive the following error: `Error compiling Parallel Transformer.`

The following error messages are in the compiler report:

```
<transform> Error when checking composite operator:
 The number of reject datasets "0" is less than the number of input datasets "1".
<transform> Error when checking composite operator: Expected semi-colon;
```

Causes

This error might be caused by an unbreakable space that is entered into the column definition of the transformer. This is typically caused by copying and pasting from a Microsoft Word or Excel document. The InfoSphere DataStage Designer prohibits inserting a normal space, but does not check for "unbreakable spaces".

Diagnosing the problem

Copy the full column name from the InfoSphere DataStage Designer client to a hex viewer. If you see the hex character `A0`, you have an unbreakable space in your column name.

Resolving the problem

Remove the unbreakable spaces from the column names.

Job transformer compile errors with smallint and bigint

You might encounter errors if you import an InfoSphere DataStage job that contains smallint and bigint data types.

Symptoms

You receive an InfoSphere DataStage job transformer compile error after you import an InfoSphere DataStage job from a previous release to version 8.x, and the job contains smallint and bigint data types. Compiling a InfoSphere DataStage job that contains a transformer produces the following error:

```
Error when checking composite operator: Setting null to this non-nullable field:
```

Causes

InfoSphere DataStage 8.x is coded to maintain compatibility with an earlier version for how nulls are handled in the transformer stage. For previous releases that use a transformer stage and use smallint and bigint data types produced an error message.

This is caused by the transformer property that is called **Legacy null processing** that is automatically set when the job is imported to InfoSphere DataStage 8.x.

InfoSphere DataStage jobs that are created with the transformer stage and not imported from a previous release does not have the **Legacy null processing** option set.

Resolving the problem

1. Right-click on the transformer stage and click properties or double-click to access properties.
2. In the upper left-corner, click the stage properties **smart** button.
3. Remove the check mark in the "**Legacy null processing**" field.

InfoSphere DataStage Parallel framework changes that require job modifications

Changes in the InfoSphere DataStage parallel framework might require job modifications when jobs are upgraded from earlier releases.

IBM tries to avoid making code changes that require customers to modify their existing InfoSphere DataStage jobs. However, sometimes it is necessary to make such changes to introduce new features or to fix errors. This topic documents changes in InfoSphere DataStage releases that might require you to correct jobs that were created in earlier versions.

Partitioning and Sort Insertion

> Information Server releases affected: 8.0.1 Fix Pack 1 and higher, 8.1 GA and higher, 8.5 GA

> For any stage that requires data to be partitioned by hash and sorted, the parallel framework automatically inserts a hash partitioner and a sort on each input link to ensure that input data is partitioned and sorted properly. Before InfoSphere Information Server 8.0.1 Fix Pack 1, the parallel framework would not automatically insert the partitioner or sort if the **Preserve Partitioning** flag was set on the input link. Not repartitioning or resorting might result in unexpected results because input data might be partitioned and sorted with different keys from the keys that are specified by the stage.

> To avoid this problem, the parallel framework was changed in Information Server 8.0.1 Fix Pack 1 so that a hash partitioner and sort would automatically be inserted even in the presence of a (framework inserted) Preserve Partitioning flag, but not in the case of user-specified partitioning, as the latter takes higher precedence. However, the problem can still occur if the user-specified partitioning and sort keys do not match the keys that are required by the stage. Here are some example scenarios that might experience problems as a result of this change:

> - A join stage has A and B keys. The user explicitly specifies a hash partitioning method and inserts a sort stage on the producing side of the primary link. The hash key is "A", and the sort keys are "A" and "B". Input data of the reference link is partitioned or sorted upstream or in another job. The partitioning method on both the primary link and the reference link of join is set to Auto. When the parallel framework analyzes partitioning and sort requirements at job startup time, it inserts hash and tsort stages on the reference link with the same two keys as specified by join, and keeps what the user defined on the primary link. This can cause data to be distributed to wrong partitions.

> - A join stage has one key. The user explicitly specifies a hash partitioning method and inserts a sort stage upstream of the primary link. The hash and sort key is "A" with the case-sensitive property. Input data of the reference link is not pre-partitioned and pre-sorted. The partitioning method on both the primary link and the reference link of join is set to Auto. When the parallel framework analyzes partitioning and sort requirements at job startup time, it inserts a hash partitioner on both links and the hash key does not have the case-sensitive property. The framework also inserts a tsort on the reference link, but not on the primary link because data is already sorted. This can break the sort order of input data on the primary link. A sequential stage or a parallel stage that runs in a sequential mode produces this warning message if its producing stage is partitioned by hash: "Sequential operator cannot preserve the partitioning of the parallel data set on input port 0. "

> These issues can be worked around by setting the environment job parameters *APT_NO_PART_INSERTION* and *APT_NO_SORT_INSERTION* to true. Then modify the job to ensure that the partitioning and sorting requirements are met by explicit insertion.

Default Decimal Separator

Information Server releases affected: 8.0.1 Fix Pack 1 and higher, 8.1 Fix Pack 1 and higher, 8.5 GA

Before Information Server Version 8.0.1 Fix Pack 1, the default decimal separator that is specified with Job Properties->Defaults was not recognized by the APT_Decimal class in the parallel framework. This caused problems for the DB2 API stage where decimals with a comma decimal point are not processed correctly. This issue was fixed in release 8.0.1 Fix Pack 1, as APAR JR31597. The default decimal separator can be specified in a job parameter. However, if the job parameter does not contain any value, '#' is taken as the decimal separator. This can cause the following error if the actual decimal separator is not '#':

```
Fatal Error: APT_Decimal::assignFromString: invalid format for the source string.
```

If you encounter this problem after you upgrade, make sure the job parameter that represents the default decimal separator contains the actual decimal separator character that is used by input data. If you cannot change the job parameter is not an option, you can set the environment variable APT_FORCE_DECIMAL_SEPARATOR. The value of APT_FORCE_DECIMAL_SEPARATOR overrides the value that is set for the "Decimal separator" property. If more than one character is set for this environment variable, the decimal separator defaults to a dot character, '.'

Embedded Nulls in Unicode Strings

Information Server releases affected: 8.1 Fix Pack 1 and higher, 8.5 GA

Before Information Server 8.1 Fix Pack 1, nulls that are embedded in Unicode strings were not treated as data, but rather they were treated as string terminators. This caused data after the first null to be truncated. The issue was fixed in Fix Pack 1, as APAR JR33408 for Unicode strings that were converted to or from UTF-8 strings. As a result of this change, you might observe a change in job behavior where a bounded-length string is padded with trailing nulls. These extra nulls can change the comparison result of two string fields, generate duplicate records, or make data conversion fail depending on the job logic. To solve this problem, the job must be modified to set the *APT_STRING_PADCHAR* variable to 0x20 and call the Trim() function in the transformer stage if necessary.

Null Handling at column level

Information Server releases affected: 8.1 GA and higher, 8.5 GA

In parallel jobs, nullability is checked at runtime. It is possible for the user to set a column as nullable in the InfoSphere DataStage Designer, but at run time the column is mapped as non-nullable to match the actual database table. Before 8.1 GA the parallel framework issued a warning for this mismatch, but the job would potentially fail with a segmentation violation as a result. The warning was changed to an unrecoverable error in 8.1 GA as ECASE 124987 to prevent the job from failing with SIGSEGV. After this change, jobs that used to run with this warning present will now fail with an unrecoverable error. For an example, this problem is often seen in the lookup stage. To solve the problem, modify the job to make sure the nullability of each input field of the lookup stage matches the nullability of the same output field of the stage, which is upstream to the lookup.

Transformer Stage: Run time Column Propagation (RCP)

DataStage releases affected: 7.5 and higher

Information Server releases affected: 8.0 GA and higher, 8.1 GA and higher, 8.5 GA

When RCP is enabled at any InfoSphere DataStage 7.x release before 7.5, for an input field "A" which is mapped to an output field "B", both "A" and "B" are present in the output record. Starting with InfoSphere DataStage 7.5, apparently "A" is being renamed to "B", so that only "B" is in the output. To improve transform performance, a straight assignment like "B=A" is considered as renaming "A" to "B". Before the change, the straight assignment was considered as creating an

extra field by copying "A" to "B". With this change in place, the user now must explicitly specify both "A" and "B" in the output schema to prevent "A" from being renamed to "B" and to create a new field "B".

Transformer Stage: Decimal Assignment

Information Server releases affected: 8.0 GA and higher, 8.1 GA and higher, 8.5 GA

The parallel framework issued a warning if the target decimal had smaller precision and scale than the source decimal. The warning was changed to an error in Information Server 8.0 GA, and as a result the input record is dropped if a reject link is not present. This behavior change was necessary to detect the error earlier to avoid data corruption. The user must modify the job to make sure that the target decimal is large enough to hold the decimal value. Alternatively, the user can add a reject link to prevent records from being dropped.

Important: This change in behavior does not apply to any Linux operating system such as Redhat, Suse, or Linux on System z. The parallel framework does not enable exception handling on Linux operating systems, so the behavior remains the same as it was before 8.0 GA.

Transformer Stage: Data Conversion

Information Server releases affected: 8.0 GA and higher, 8.1 GA and higher, 8.5 GA

Before Information Server 8.0 GA, an invalid data conversion in the transformer would result in the following behavior:

- A warning message is issued to the InfoSphere DataStage job log
- A default value was assigned to the destination field according to its data type
- The record was written to the output link
- If a reject link was present, nothing was sent to the reject link

The behavior changed in the 8.0 GA release when a reject link is present. Instead of writing the record to the output link with a default value, it is written to the reject link instead. This might lead to data loss if the job is expecting those records to be passed through to the output. To get to the original behavior of passing the records through, the job must be modified to remove the reject link.

An environment variable was added along with this change to add the capability of stopping the job. To use this option, ensure that there is no reject link and then set the environment variable `APT_TRANSFORM_ABORT_ON_CONVERSION_ERROR=True`. The job now fails from an invalid data conversion scenario.

Surrogate Key Generator

Information Server releases affected: 8.0.1 Fix Pack1 and higher, 8.1 Fix Pack 1 and higher, 8.5 GA

The surrogate key stage reserves keys in blocks. Before Information Server 8.1 Fix Pack 1, if only one record was generated, the surrogate key generator would use values that begin with 6 and greater as available keys for incoming records. The surrogate key generator was changed in 8.1 Fix Pack 1, as APAR JR29667. With this change, InfoSphere DataStage considers values 1 - 4 and any value 6 and greater as available keys. This behavior change might cause the SCD stage to produce incorrect results in the database or generate the wrong surrogate keys for the new records of the dimension. If required, the job can be modified to revert to the old behavior (start generating keys from the highest key value last used) by setting option 'Generate Key From Last Highest Value' to Yes. This approach however might result in gaps in used keys. You must understand how the key file is initialized. Based on business logic, you must decide if it is necessary to modify the job.

Sequential File Format on Windows

Information Server releases affected: (Windows Platforms) 8.1 GA and higher, 8.5 GA

Prior to Information Server 8.1 GA, the default format for sequential files was Unix format which requires a newline character as the delimiter of a record. The default format for the Sequential File stage was changed to Windows format in the Information Server 8.1 GA release. Due to this change, data files previously created with UNIX format will not import properly. To solve this issue, set the environment variable APT_USE_CRLF=FALSE at the project level or within the system environment variables (requires a Windows reboot).

Troubleshooting for specific operating systems

Some issues are specific to the operating system on which InfoSphere DataStage is running.

Troubleshooting slow jobs that use data sets in cluster environments

InfoSphere DataStage jobs can sometimes become very slow over a period, but you can add environment variables to correct this problem.

Symptoms

InfoSphere DataStage Jobs that use data sets become very slow over a period.

Causes

Data sets use a `sync()` call, but according to solaris it must be making `fsync()` call.

Environment

This problem occurs only in cluster environments. The same jobs that are experiencing bottlenecks run normally in a non-cluster environment.

Diagnosing the problem

Check to see if your data sets are using a sync() call when they must use a `fsync()` call.

Resolving the problem

Add the *APT_DATASET_FLUSH_NOSYNC* and *APT_DATASET_FLUSH_NOFSYNC* environment variables to disable `sync()` and `fsync()` system calls.

Table 3. APAR information

APAR number	JR37466
Reported component name	WIS DATASTAGE
Reported component	ID5724Q36DS
Reported release	753
Status	CLOSED PER
PE	NoPE
HIPER	NoHIPER
Special Attention	NoSpecatt
Submitted date	2010-08-20
Closed date	2010-09-13
Last modified date	2011-05-27

Table 4. Fix information

Fixed component name	WIS DATASTAGE
Fixed component ID	5724Q36DS
R753 PSN	UP

Heap allocation errors with InfoSphere DataStage Parallel Jobs on AIX

You can correct heap allocation errors in several ways based on which version of InfoSphere DataStage you are running.

Symptoms

InfoSphere DataStage parallel jobs ends with the following error message:

```
APT_BadAlloc: Heap allocation failed.
```

Causes

AIX divides memory address space into segments. If the InfoSphere DataStage jobs must allocate more memory than exists in the number of available segments, the job ends with a heap allocation or failure to allocate memory.

Resolving the problem

Verify that *ulimit* is not restricting the memory allocation of the process. Check the *ulimit* from within InfoSphere DataStage itself because the value is not the same as the value in the interactive shell. To capture *ulimit* for all nodes, use the following process:

1. Create a new Parallel job.
2. Add an External Source stage (under File on the palette) connected to a peek stage (under Development/Debug on the palette).
3. Access the advanced properties of the External Source stage, and make sure its running in Parallel mode.
4. In the External source stage, enter `ulimit -a; ulimit -aH` without the quotations in the Source Program property and a column as VarChar with length of 255.
5. Use a configuration file that includes at least one node for each fast name (host) in your cluster or GRID.
6. Compile the job, run it, and look in the Director log.

The director contains the soft limits and the hard limits for each node in the configuration file. If the hard limit for data is too low, you must contact your AIX administrator to increase that value. This value can be set in the file `/etc/security/limits`. After you increase the hard limit settings, you can set the *ulimit* settings for the user in the ds.rc file that is located under $DSHOME/sample. You can add a line like `ulimit -d unlimited` at the beginning of the file, after the umask settings.

The ds.rc file is owned by root, and is writable only to root, so your system administrator must change the file permissions. For security reasons, do not change the owner or grant write permission to any non-root user. Important: do not set the number of file descriptors to unlimited with the `ulimit -n` command. That setting causes a problem with InfoSphere DataStage. Ensure that the value for this limit is set sufficiently high. A safe value is 100000 in nearly all situations

InfoSphere DataStage Version 7.5.x

> The InfoSphere DataStage Software is a 32-bit application for all 7.5.x releases, even when installed onto an AIX server with a 64-bit kernel. To obtain the maximum amount of process address space for your parallel job processes, set the *LDR_CNTRL* variable with a value of `MAXDATA=0x80000000@DSA` as the default value at the project level (for all jobs in a project) or within specific jobs.

> **Important:** Do not add *LDR_CNTRL* to your `dsenv` file. That setting might interfere with the memory model used by the Server Engine.

InfoSphere DataStage Version 8.0.x

> The InfoSphere DataStage Software is a 32-bit application for all 8.0.x releases, even when installed onto an AIX server with a 64-bit kernel. Starting with the Information Server 8.0 GA release, InfoSphere DataStage now starts Java components to integrate with the services tier. For

these Java components to function properly, the LDR_CNTRL=MAXDATA=0x60000000@USERREGS environment variable is added to the `dsenv` file. This variable must not be removed or modified to ensure the correct operation of the Java components.

For parallel jobs that require more than 1.5GB of memory per process, the *LDR_CNTRL* variable can be set to a larger value. This variable must be given a default value at the project level if you want it to take effect for all jobs in the project, or by leaving the project default value blank and assigning a value to specific jobs only. As stated previously, do not alter *LDR_CNTRL* within the `dsenv` file.

To obtain the maximum amount of process address space for your job processes, set the *LDR_CNTRL* variable with a value of MAXDATA=0x80000000@DSA in your job or as a project default.

InfoSphere DataStage Version 8.1.x

Starting with the 8.1 GA release, InfoSphere DataStage is now a 64-bit application and requires a 64-bit AIX kernel. The osh item is compiled with the MAXDATA=x80000000 property, so the amount of memory address space available to the parallel job process is limited to 2 GB in the default configuration. The improvement of being a 64-bit application allows for the allocation of more segments and a larger private memory address space. For situations where large amounts of heap memory are required for each process, set *LDR_CNTRL* with a value of MAXDATA=0x00000001000000000. This value allocates up to 64 Gb for private data for each process.

Set this large value at the job level rather than at the project level to avoid large consumption of memory by jobs where you did not intentionally want this behavior.

InfoSphere DataStage Version 8.5

InfoSphere DataStage is a 64-bit application and requires a 64-bit AIX kernel just like release 8.1. A significant improvement at this release, is that the MAXDATA parameter is removed from the executable. With this change, InfoSphere DataStage is now able to access all of the available memory address segments in the default configuration. Any jobs or projects that had *LDR_CNTRL* specified with the *MAXDATA* parameter must be modified to remove this parameter after you upgrade to 8.5 so that you are able to access all of the segments.

Important: The *LDR_CNTRL=USERREGS* environment variable must not be removed from the `dsenv` file. The variable is required for correct operation of Java components that are loaded by the InfoSphere DataStage processes. The *USERREGS* property does not affect the memory usage of InfoSphere DataStage jobs.

Node or host names cannot be found during parallel jobs on HP-UX Itanium systems

HP-UX Itanium systems have a default maximum lengths for node and host names, and truncated names can cause jobs to fail.

Symptoms

Parallel jobs on HP-UX Itanium systems fail with an error that indicates that a host or node name cannot be found. The node names or host names are truncated. The node or host cannot be found during job processing. For example, if your host name is *H1234567890*, you receive an error that is similar to the following output:

```
##E IIS-DSEE-TUTL-00096 00:54:15(000) <main_program> APT_PMaddrInfoFor():
  getaddrinfo(host = H1234567, port = ) failed with error 8 - host nor service provided,
  or not known: node H1234567 cannot be used
##F IIS-DSEE-TFCF-00014 00:54:15(001) <main_program> Fatal Error:
  The configuration does not contain the invoking node truncate.
```

Causes

By default, HP-UX Itanium node names are limited to 8 bytes, and host names are limited to 64 bytes. The job cannot find the host or node during processing if the name is truncated. The kernel **expanded_node_host_names** parameter is required for long host name support on HP-UX Itanium and is not enabled by default.

Diagnosing the problem

Run the `uname -a` command on the host or node system to access the system information. Read the host name in the system information to see if the host or node name is truncated to eight characters.

Resolving the problem

Modify the kernel parameters to allow for expanded node and host names.

1. Edit the `/etc/sysctl.conf` file and set `expanded_node_host_names=1`.

2. Restart the HP-UX Itanium system.

3. After the system restarts, view the system information again with the `uname -a` command to verify that the host or node name is no longer truncated.

Tuning engine parameters

You can tune engine parameters to improve performance and avoid troubleshooting problems in the future.

Using tunable parameters in the UVCONFIG file

The `UVCONFIG` file contains several parameters that you can configure to improve performance and avoid troubleshooting.

The most commonly used parameters in the `UVCONFIG` file are the following parameters:

MFILES

This parameter defines the size of the server engine (DSEngine) rotating file pool. This is a per process pool for files such as sequential files that are opened by the InfoSphere DataStage server run time. It does not include files that are opened directly at operating system level by the parallel engine (PXEngine running osh).

The server engine logically opens and closes files at the InfoSphere DataStage application level and physically close them at the OS level when the need arises.

Increase this value if InfoSphere DataStage jobs use numerous files. Generally, a value of around 250 is suitable. If the value is set too low, then performance issues might occur. The server engine makes more calls to open and close at the physical operating system level to map the logical pool to the physical pool.

Note: The operating system parameter of **nofiles** must be set higher than *MFILES*. The **nofiles** parameter must be at least 512 to allow the InfoSphere DataStage process to open up to 512 files.

On most UNIX systems, the `proc` file system can be used to monitor the file handles opened by a process. For example:

```
ps -ef|grep dsrpcd

root     23978    1  0 Jul08 ?
00:00:00 /opt/ds753/Ascential/DataStage/DSEngine/bin/accdsrpcd

ls -l /proc/23978/fd

lrwx------  1 root dstage 64 Sep 25 08:24 0 -> /dev/pts/1 (deleted)
l-wx------  1 root dstage 64 Sep 25 08:24 1 -> /dev/null
l-wx------  1 root dstage 64 Sep 25 08:24 2 -> /dev/null
lrwx------  1 root dstage 64 Sep 25 08:24 3 -> socket:[12928306]
```

The dsrpcd process (23978) has four files open in the example.

T30FILE

This parameter determines the maximum number of dynamic hash files that can be opened system-wide on the InfoSphere DataStage system. If this value is too low, expect to find an error message that is similar to `T30FILE table full`.

You can see the number of dynamic files in use if you run the following engine command from $DSHOME:

```
echo "`bin/smat -d|wc -l` - 3"|bc
```

Use this command to assist with tuning the **T30FILE** parameter. See the following technote: https://www-304.ibm.com/support/docview.wss?uid=swg21390117

Every running InfoSphere DataStage job requires at least three slots in this table. The three slots are **RT_CONFIG**, **RT_LOG**, and **RT_STATUS**. However, that multi-instance jobs share slots for these files. Although each job run instance creates a separate file handle, it increments a usage counter in the table if the file is already open to another instance.

On AIX, the *T30FILE* value must not be set higher than the system setting for *ulimit -n*.

GLTABSZ

This parameter defines the size of a row in the group lock table. Tune this value if the number of group locks in a slot is getting close to the value defined.

Use the LIST.READU EVERY command from the server engine shell to assist with monitoring this value. LIST.READU lists the active file and record locks; the EVERY keyword lists the active group locks in addition.

For example, with a Designer client and a Director client both logged in to a project named dstage0:

```
Active Group Locks:                                    Record Group Group Group
Device.... Inode..... Netnode Userno Lmode G-Address. Locks ...RD ...SH ...EX
838222719 2039334646       0   5620 62 IN       800     1    0     0     0

Active Record Locks:
Device.... Inode..... Netnode Userno Lmode    PID Item-ID....................
 838222719 2039334646       0  64332 62 RL   1204 dstage0&!DS.ADMIN!&
 838222719 2039334646       0  62412 62 RL   3124 dstage0&!DS.ADMIN!&
```

Device
    A number that identifies the logical partition of the disk where the file system is located.

Inode   A number that identifies the file that is being accessed.

Netnode
    A number that identifies the host from which the lock originated. 0 indicates a lock on the local system, which is usually the case for InfoSphere DataStage. If other than 0, then on UNIX it is the last part of the TCP/IP host number that is specified in the `/etc/hosts` file. On Windows it is either the last part of the TCP/IP host number or the LAN Manager node name, depending on the network transport that is used by the connection.

Userno
    The phantom process that set the lock.

Pid   A number that identifies the controlling process.

Item-ID
    The record ID of the locked record.

Lmode
    The number that is assigned to the lock, and a code that describes its use.

G-Address
> Logical disk address of group, or its offset in bytes from the start of the file, in hex.

Record Locks
> The number of locked records in the group.

Group RD
> Number of readers in the group.

Group SH
> Number of shared group locks.

Group EX
> Number of exclusive group locks.

When the report describes record locks, it contains the following *Lmode* codes:

FS, IX, CR
> Shared file locks.

FX, XU, XR
> Exclusive file locks.

When the report describes group locks, it contains the following *Lmode* codes:

EX   Exclusive lock.

SH   Shared lock.

RD   Read lock.

WR   Write lock.

IN   System information lock.

When the report describes record locks, it contains the following *Lmode* codes:

RL   Shared record lock.

RU   Update record lock.

RLTABSZ

> This parameter defines the size of a row in the record lock table. From a InfoSphere DataStage job point of view, this value affects the number of concurrent InfoSphere DataStage jobs that can be run, and the number of InfoSphere DataStage Clients that can connect.

> Use the `LIST.READU` command from the InfoSphere DataStage Engine shell to monitor the number of record locks in a slot. With one Director client logged in to a project named *dstage0*, and two instances of a job in that project that are running, the active record locks are similar to the following example:

```
Active Record Locks:
Device.... Inode..... Netnode Userno Lmode   Pid Item-ID.............
 838222719 2039334646       0  64332 62 RL  1204 dstage0&!DS.ADMIN!&
 838222719 2039334646       0  62128 62 RL  3408 dstage0&!DS.ADMIN!&
 838222719 2039334646       0  65252 62 RL   284 dstage0&!DS.ADMIN!&
 304877956  328255620       0  62128 62 RL  3408 RT_CONFIG456
 304877956  328255620       0  65252 62 RL   284 RT_CONFIG456
```

> In the report, `Item-ID=RT_CONFIG456` identifies that the running job is an instance of job number 456, whose compiled job file is locked while the instance is running so that, for example, it cannot be recompiled in that time. A job's number within its project can be seen in the Director job status view, the detail dialog, for a particular job.

> The unnamed column in-between *UserNo* and *Lmode* relates to a row number within the Record Lock table. Each row can hold **RLTABSZ** locks. In the example, three slots out of 75 are used for

row 62. Default value for **RLTABSZ** is 75. When the number of entries for a row gets close to the **RLTABSZ** value, it is time to consider retuning the system.

Jobs can fail to start, or generate -14 errors, if **RLTABSZ** is being reached.

InfoSphere DataStage Clients might see an error message similar to `DataStage Project locked by Administrator` when they attempt to connect. The error message can be misleading. The message means that a lock cannot be acquired because the lock table is full, and not because another user already has the lock.

MAXRLOCK

This parameter must always be set to the value of **RLTABSZ** − 1.

Each `DSD.RUN` process takes a record lock on a key name `<project>&!DS.ADMIN!&` of the `UV.ACCOUNT` file in $DSHOME, which you can see in the example. Each InfoSphere DataStage client connection (for example, Designer, Director, Administrator, dsjob command) takes this record lock as well. This is the mechanism by which InfoSphere DataStage determines whether operations such as project deletion are safe, operations cannot proceed while a project lock is held by any process.

MAXRLOCK must be set to accommodate the maximum number of jobs and sequences plus client connections that are used. And **RLTABSZ** must be set to **MAXRLOCK** + 1. Keep in mind that changing RLTABSZ greatly increases the amount of memory needed by the disk shared memory segment.

Using settings of 130/130/129 for **RLTABSZ**, **GLTABSZ**, and **MAXRLOCK** work successfully on most customer installations. Some customers use higher settings of 300/300/299 depending on their environment.

If sequencers or multi-instance jobs are used, start with the settings of 130/130/129, and increase to 300/300/299 if necessary.

Before InfoSphere DataStage v8.5 the following settings were pre-defined:
- **MFILES** = 150
- **T30FILE** = 200
- **GLTABSZ** = 75
- **RLTABSZ** = 75
- **MAXRLOCK** = 74 (75-1)

InfoSphere DataStage v8.5 has the following settings pre-defined:
- **MFILES** = 150
- **T30FILE** = 512
- **GLTABSZ** = 75
- **RLTABSZ** = 150
- **MAXRLOCK** = 149 (150-1)

These are the lowest suggested values to accommodate all system configurations, so tuning of these values is often necessary.

DMEMOFF, PMEMOFF, CMEMOFF, NMEMOFF

These are the shared memory address offset values for each of the four InfoSphere DataStage shared memory segments such as Disk, Printer, catalog, and NLS. Depending upon the system, **PMEMOFF**, **CMEMOFF**, and **NMEMOFF** must be increased to allow for a large disk shared memory to be used.

Where these values are set to 0x0 on AIX , the operating system takes care of managing these offsets. Otherwise, the value of **PMEMOFF** minus the value of **DMEMOFF** is the largest disk shared memory segment size. Additionally, on Solaris for example, these values are increased to allow for a greater heap size for the running InfoSphere DataStage job.

When you run the `shmtest` utility, great care must be taken with interpreting its output. The utility tests the availability of memory that it can allocate at the time it runs. Memory availability is affected both by the current **uvconfig** settings, by how much shared memory is already in use, and by other activity on the system at the time.

Enabling tracing for InfoSphere DataStage parallel jobs

You can enable tracing for InfoSphere DataStage parallel jobs to help you troubleshoot problems.

Procedure

1. Enable the following administrator project level parameters for the project or the job and set them to true
   - **APT_DUMP_SCORE**
   - **APT_PM_SHOWRSH**
   - **APT_PM_SHOW_PIDS**
   - **APT_RECORD_COUNTS**
   - **APT_SHOW_COMPONENT_CALLS**
   - **APT_STARTUP_STATUS**
   - **OSH_DUMP**
   - **OSH_ECHO**
   - **OSH_EXPLAIN APT_DISABLE_COMBINATION**
2. Add a user-defined environment variable that is called *DS_PXDEBUG* in the InfoSphere DataStage Administrator. The value must be undefined for the project. Leave that value blank or set it to 0 at the project level. Add this new environment variable to the job level and set the value to 1. The DS_PXDEBUG variable causes the job to report debugging information.

Results

Debug information is collected under a new project-level directory called `Debugging`. Subdirectories are created on a per-job basis, and are named after the job. Multi-instance jobs run with a non-empty invocation ID, and the directory is named with the job name and the invocation ID.

What to do next

Run the job. Send an export of the job with the detailed job log and the *project path/Debugging/Jobname* folder to support.

Contacting IBM

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

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

Table 5. IBM resources

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

Accessing the product documentation

Documentation is provided in a variety of formats: in the online IBM Knowledge Center, in an optional locally installed information center, and as PDF books. You can access the online or locally installed help directly from the product client interfaces.

IBM Knowledge Center is the best place to find the most up-to-date information for InfoSphere Information Server. IBM Knowledge Center contains help for most of the product interfaces, as well as complete documentation for all the product modules in the suite. You can open IBM Knowledge Center from the installed product or from a web browser.

Accessing IBM Knowledge Center

There are various ways to access the online documentation:
- Click the **Help** link in the upper right of the client interface.
- Press the F1 key. The F1 key typically opens the topic that describes the current context of the client interface.

  Note: The F1 key does not work in web clients.
- Type the address in a web browser, for example, when you are not logged in to the product.

  Enter the following address to access all versions of InfoSphere Information Server documentation:

  `http://www.ibm.com/support/knowledgecenter/SSZJPZ/`

  If you want to access a particular topic, specify the version number with the product identifier, the documentation plug-in name, and the topic path in the URL. For example, the URL for the 11.3 version of this topic is as follows. (The ⇒ symbol indicates a line continuation):

  ```
  http://www.ibm.com/support/knowledgecenter/SSZJPZ_11.3.0/⇒
  com.ibm.swg.im.iis.common.doc/common/accessingiidoc.html
  ```

  Tip:

  The knowledge center has a short URL as well:

  `http://ibm.biz/knowctr`

  To specify a short URL to a specific product page, version, or topic, use a hash character (#) between the short URL and the product identifier. For example, the short URL to all the InfoSphere Information Server documentation is the following URL:

  `http://ibm.biz/knowctr#SSZJPZ/`

  And, the short URL to the topic above to create a slightly shorter URL is the following URL (The ⇒ symbol indicates a line continuation):

  ```
  http://ibm.biz/knowctr#SSZJPZ_11.3.0/com.ibm.swg.im.iis.common.doc/⇒
  common/accessingiidoc.html
  ```

Changing help links to refer to locally installed documentation

IBM Knowledge Center contains the most up-to-date version of the documentation. However, you can install a local version of the documentation as an information center and configure your help links to point to it. A local information center is useful if your enterprise does not provide access to the internet.

Use the installation instructions that come with the information center installation package to install it on the computer of your choice. After you install and start the information center, you can use the `iisAdmin`

command on the services tier computer to change the documentation location that the product F1 and help links refer to. (The ⇒ symbol indicates a line continuation):

Windows

```
IS_install_path\ASBServer\bin\iisAdmin.bat -set -key ⇒
com.ibm.iis.infocenter.url -value http://<host>:<port>/help/topic/
```

AIX Linux

```
IS_install_path/ASBServer/bin/iisAdmin.sh -set -key ⇒
com.ibm.iis.infocenter.url -value http://<host>:<port>/help/topic/
```

Where <host> is the name of the computer where the information center is installed and <port> is the port number for the information center. The default port number is 8888. For example, on a computer named server1.example.com that uses the default port, the URL value would be http://server1.example.com:8888/help/topic/.

Obtaining PDF and hardcopy documentation

- The PDF file books are available online and can be accessed from this support document: https://www.ibm.com/support/docview.wss?uid=swg27008803&wv=1.
- You can also order IBM publications in hardcopy format online or through your local IBM representative. To order publications online, go to the IBM Publications Center at http://www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss.

Product accessibility

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

The IBM InfoSphere Information Server product modules and user interfaces are not fully accessible.

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

Accessible documentation

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

The documentation that is in the information center is also provided in PDF files, which are not fully accessible.

IBM and accessibility

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

Notices and trademarks

This information was developed for products and services offered in the U.S.A. This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

Notices

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

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

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

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

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

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

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

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

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

97

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

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

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

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

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

Privacy policy considerations

IBM Software products, including software as a service solutions, ("Software Offerings") may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering's use of cookies is set forth below.

Depending upon the configurations deployed, this Software Offering may use session or persistent cookies. If a product or component is not listed, that product or component does not use cookies.

Table 6. Use of cookies by InfoSphere Information Server products and components

Product module	Component or feature	Type of cookie that is used	Collect this data	Purpose of data	Disabling the cookies
Any (part of InfoSphere Information Server installation)	InfoSphere Information Server web console	• Session • Persistent	User name	• Session management • Authentication	Cannot be disabled
Any (part of InfoSphere Information Server installation)	InfoSphere Metadata Asset Manager	• Session • Persistent	No personally identifiable information	• Session management • Authentication • Enhanced user usability • Single sign-on configuration	Cannot be disabled
InfoSphere DataStage	Big Data File stage	• Session • Persistent	• User name • Digital signature • Session ID	• Session management • Authentication • Single sign-on configuration	Cannot be disabled
InfoSphere DataStage	XML stage	Session	Internal identifiers	• Session management • Authentication	Cannot be disabled
InfoSphere DataStage	IBM InfoSphere DataStage and QualityStage Operations Console	Session	No personally identifiable information	• Session management • Authentication	Cannot be disabled
InfoSphere Data Click	InfoSphere Information Server web console	• Session • Persistent	User name	• Session management • Authentication	Cannot be disabled
InfoSphere Data Quality Console		Session	No personally identifiable information	• Session management • Authentication • Single sign-on configuration	Cannot be disabled

Notices and trademarks   **99**

Table 6. Use of cookies by InfoSphere Information Server products and components (continued)

Product module	Component or feature	Type of cookie that is used	Collect this data	Purpose of data	Disabling the cookies
InfoSphere QualityStage Standardization Rules Designer	InfoSphere Information Server web console	• Session • Persistent	User name	• Session management • Authentication	Cannot be disabled
InfoSphere Information Governance Catalog		• Session • Persistent	• User name • Internal identifiers • State of the tree	• Session management • Authentication • Single sign-on configuration	Cannot be disabled
InfoSphere Information Analyzer	Data Rules stage in the InfoSphere DataStage and QualityStage Designer client	Session	Session ID	Session management	Cannot be disabled

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, see IBM's Privacy Policy at http://www.ibm.com/privacy and IBM's Online Privacy Statement at http://www.ibm.com/privacy/details the section entitled "Cookies, Web Beacons and Other Technologies" and the "IBM Software Products and Software-as-a-Service Privacy Statement" at http://www.ibm.com/software/info/product-privacy.

Trademarks

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

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

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

Intel and Itanium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

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

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

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

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

Index

IBM ®

Printed in USA