precisely

Assure DQ

Operator's Guide

precisely

Trademarks

Infogix, the Infogix logo, ACR, ACR/Detail, ACR/Summary, ACR/Workbench, ACR/Connector, Infogix Assure, Infogix Insight, ACR/Instream, ACR/File, Infogix ER, Infogix Perceive, Data3Sixty, and Data360 are registered trademarks of Precisely. Data3Sixty Analyze, Data3Sixty Govern, Data3Sixty DQ+, Data360 Analyze, Data360 Govern and Data360 DQ+ are trademarks of Precisely. Any other trademarks or registered trademarks are the property of their respective owners.

1700 District Ave Ste 300
Burlington MA 01803-5231
USA

www.precisely.com

Contents

Chapter 4 Running Retention and Archiving

Chapter 5 Extracting Control Data

Chapter 6 Execute Script Actions

Chapter 7 Scheduling

Chapter 8 Troubleshooting

1

Introduction

This guide explains how to operate Assure DQ® in a production environment. The guide describes common system operation tasks, including scheduling and running controls, handling messages, managing data retention and archiving, and troubleshooting processing errors.

About This Guide

This guide has the following chapters

- Chapter 1, "Introduction"—System overview and processing.
- Chapter 2, "Running Controls"—Commands and syntax for running control points, halting processing, and running standalone processing.
- Chapter 3, "Message Handling"—Custom JMS message handling and message queue error handling.
- Chapter 4, "Running Retention and Archiving"—Running data retention and archiving from the console, and generating reports accessible to users outside Assure DQ.
- Chapter 5, "Extracting Control Data"— Running control data extraction from the console.
- Chapter 6, "Execute Script Actions"—Executing external scripts.
- Chapter 7, "Scheduling"—Managing the Assure DQ internal scheduler to execute control points and system maintenance processes.
- Chapter 8, "Troubleshooting"—Resolving common processing issues.

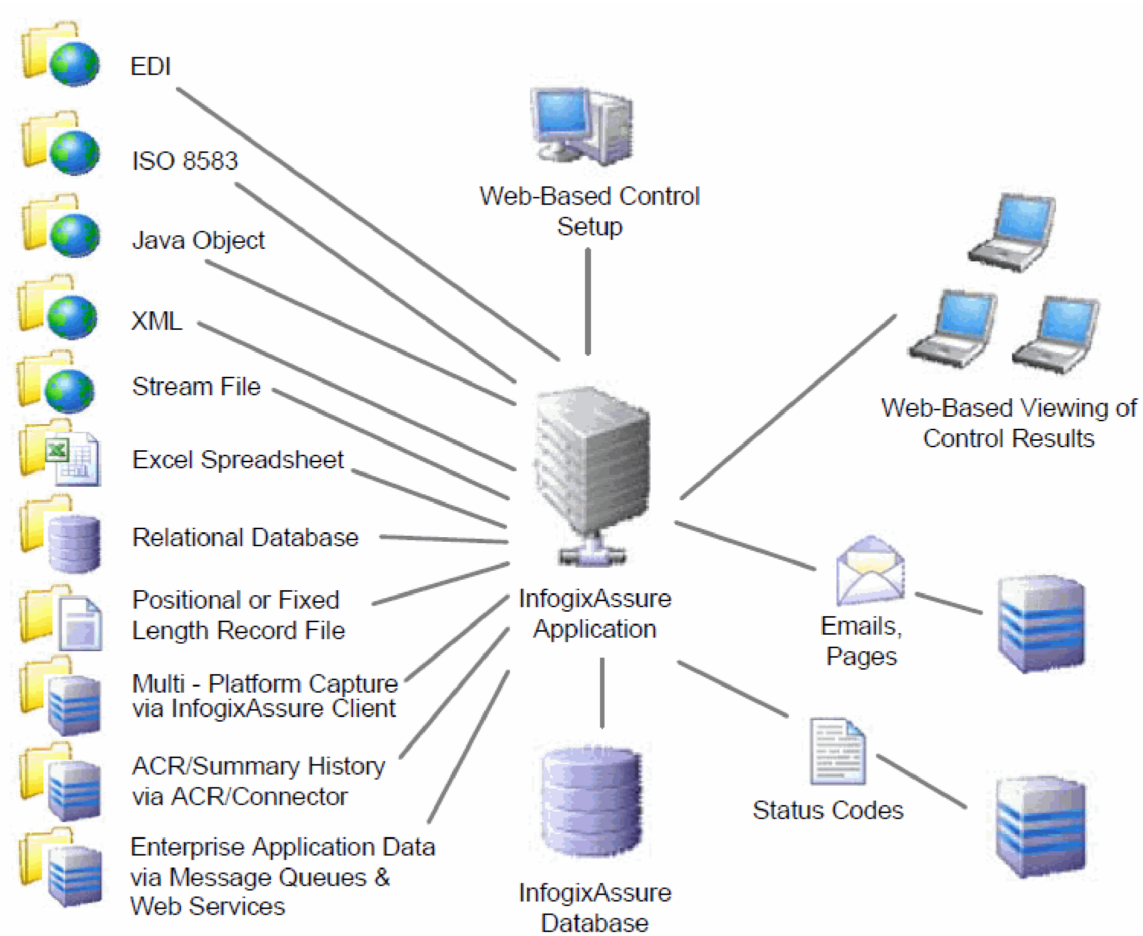
Assure DQ System Overview

Assure DQ is a J2EE™ application deployed under a JBoss® or WebSphere® application server on the Windows®, UNIX®, SuSE Linux®, or Red Hat Enterprise Linux platforms. Assure DQ clients can capture data for control processing on Windows, UNIX, Linux, and z/OS® platforms.

Supported databases are Oracle® and SQL Server®. Supported enterprise directories are IBM Tivoli® Directory Server, Microsoft® ADAM (Active Directory Application Mode), Novell® eDirectory, OpenDS, Oracle Sun

ONE™ Directory Server, and Windows Server® 2003 Active Directory. The system also offers internal (embedded) security and user information directories as an optional alternative to enterprise directories.

Assure DQ System Overview Diagram



Assure DQ Application

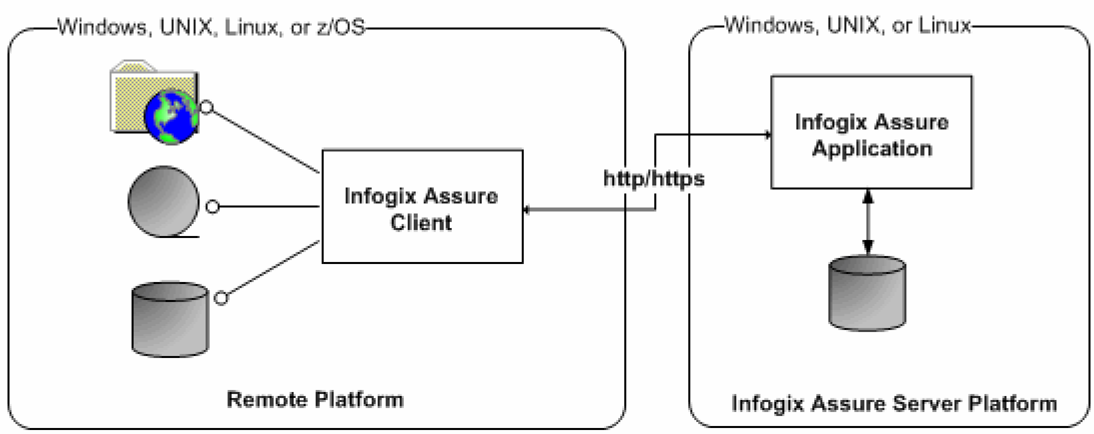
The Assure DQ application performs the following functions:

- Executes control processing when a request is received
- Captures data for control processing from various input sources
- Retrieves prior run data from the database, as needed
- Writes control processing results to the database

- Passes return codes indicating control outcomes to the operating system so that external applications can halt or launch additional processing
- Performs actions based on control outcomes, such as sending formatted email alerts to external email applications
- Retrieves and formats data from the database for display to users
- Provides control setup and security services

Assure DQ Client

You can use an Assure DQ client to initiate control point execution and data capture on a platform remote to the Assure DQ server application.



The Assure DQ client is useful in the following scenarios that involve initiating control processing or capturing data on a platform remote to the server:

- A process on another platform must trigger the execution of a control point.
- You want to extract and transfer a small amount of data from large files on a remote platform, rather than transfer the entire file to the server.
- You need to validate the integrity of a file on a remote platform before or after transfer from another platform.

The matrix below can help you decide whether to use the client or server to capture data.

Control type	File size	Data to transfer	Data transfer volume	Recommended data capture technique
Summary level controls	Large file	■ File level hash ■ File level field totals ■ Record counts	Small	Client
Detail level controls	Large file or a high volume of individual transactions	■ A small percentage of fields extracted from many transactions ■ A small percentage of complete records or transactions ■ Totals, tallies, or counts	Medium	Client or server, depending on network traffic
Detail level controls	Large file or a high volume of individual transactions	■ A large percentage of fields extracted from many transactions ■ A large percentage of complete records or transactions	Large	Server

Data capture is always more efficient on the server. However, file transfers between platforms may take too long or result in data loss or corruption. The goal of the controls designer should be to minimize the amount of data transferred to mitigate these risks. For example, using the client to select and extract a small percentage of fields or transactions may eliminate the need to transfer entire files or large volumes of transactions.

Client Processing

Control point processing and data capture are initiated on the remote platform by a job scheduler, console command, an Assure DQ control point action, or an external process. See "Executing Control Points" on page 17 for detailed operating instructions. The client and server communicate via http or https.

The following processing occurs when the client executes on a remote platform, relative to the Assure DQ server:

1. An Assure DQ control point action, an external application, or a job scheduler initiates client execution on the remote platform.
2. The client performs initial license checking and control point startup communication with the server.
3. The client retrieves the data capture rules for the control point.
4. The client performs capture processing for all the control point sources on its platform, and also performs any translations specified by the extraction rules.
5. The client combines related control data and sends it to the server.
6. The server applies the rules specified for the current control point, which may include merging the control data with data from other control points or from history, and validating the data.
7. The client gets the highest return code for the control run from the server and passes it back to the console or to the program or scheduler that initiated client execution.
8. The client or server perform any triggered control point actions.

To install the client on a remote platform, see the Assure DQ Client Installation Guide. The client is available for Windows, UNIX, Linux, and z/OS.

Assure DQ Database

During each control run, Assure DQ stores a history of controls applied, data captured, and control results in its database. This includes the following information:

- Data sources processed
- Data captured for control processing
- Control results
- Actions taken based on control results

Authorized users can sort, search, and view this information with the Web-based Assure DQ interface.

Web Service

The Assure DQ server includes a Web service that enables a calling client program to pass data to the server, request Assure DQ processing, and receive processing results. The client program can point to a file to process or pass individual transactions to Assure DQ. The client can also simply request control point execution or results. The Web service supports any client that can implement the SOAP protocol.

Invoking Control Processing

Assure DQ control processing can be triggered in the following ways:

- As a scheduled activity
- By a request from the user interface
- By a request from an external application
- By a request from a client program to the Assure DQ Web service
- By a request from the command prompt

Extracting Data

Assure DQ can capture data for control processing from one or more of the following sources:

- Relational databases
- IMS databases
- Stream files, such as delimited files
- Files containing fixed length records, or variable length records containing fixed length, positional fields
- EDI data (if licensed)
- ISO 8583 files (if licensed)
- XML data
- Microsoft Excel® spreadsheets
- Data in a Java class file in byte code or from the serialized object
- Data that resides on an OLAP server in a multi-dimensional format database
- Data that resides at a Web service endpoint
- Individual transactions passed to the Assure DQ Web service

Duplicate Source Control

Assure DQ can check to see if an input to a control point duplicates any that was processed during a prior run of that control point. It can also verify that none of the inputs for the current processing run are duplicates. If a duplicate is identified, Assure DQ can skip capturing data, generate an alert, and/or halt processing.

Contacting Customer Support

If you need assistance, contact Infogix Customer Support.

Support Phone: +1.630.505.1890

Support Email: support@infogix.com

Support Website: http://support.infogix.com

Fax Number: +1.630.505.1883

Visit our Website: www.infogix.com

2

Running Controls

You can use any of the following methods to trigger Assure DQ control processing:

- Execute the Run Control Point program from a job scheduler or the command line on a system where the Assure DQ server or client runs. See “Executing Control Points,” below.
- If you need to run many control points concurrently and J2EE/JVM memory constraints are an issue, execute the Standalone Run Control Point program from an external job scheduler or the command line on the system where the Assure DQ server runs. See “Running Controls in Standalone Mode” on page 22.
- In the user interface, set control points to sequentially trigger execution of additional control points.
- Set up an external program to pass an API request for control processing to Assure DQ. See the *Assure DQ Application Programming Interface Guide*.

Executing Control Points

Execute the Run Control Point (runcp) program to invoke control point processing for Assure DQ control entities or data quality applications. The runcp program is in <install_folder>/Infogix/server-util /<config_name>/IA.

You can install Assure DQ clients on remote platforms if you need to invoke control point processing remotely. Assure DQ clients are available for the Windows, UNIX, and z/OS platforms. The runcp program is in the Infogix Assure client installation bin directory.

To retrieve data for control entities from a remote platform, you can install the Assure DQ client on the remote platform and execute runcp with a -local parameter. The Assure DQ client will capture the required data on the remote platform and pass it to the server for control processing. The remote data capture capability is not supported for Assure DQ data quality applications.

To see syntax information for the program, go to the directory where the program is located and enter the name of the program followed by -?. On Windows, enter a command similar to the following example: runcp.bat -?

User IDs with the remote executor role can initiate control processing from the console, an external job scheduler, or API. Execute runcp.bat (for Windows) or runcp.sh (for UNIX or Linux), followed by the appropriate parameters from the table below.

Parameter	Description
-S *datasourcename=url*	Overrides the URL or path specified in the layout or control point definitions for one or more data sources. To specify more than one override, enter a space between each override pair. Enclose the source name or path in quotation marks if it contains spaces. Optional.
-c *configname*	If you invoke the run control point program from a remote platform, specify which client configuration to use to connect to the appropriate Assure DQ server instance. If you deployed only one client configuration (named "default"), this parameter is optional.
-ctlrptlimit *nnn*	For this control point run, overrides the system and control point settings for the maximum number of Control Reports to output. Optional.
-cycle *YYYYMMDD*	Overrides the cycle number defined for this control entity or application. Optional.
-ds	Disables duplicate source checking for the control run. This may be necessary in situations such as reruns. Duplicate source checking is optional and applies to control entities only. See “Duplicate Source Control” on page 15.
-entity *control entity name* or -application *application name*	Specifies the name of the control entity or application that contains the control point to execute. Enclose the name in quotation marks if it contains spaces. Case-sensitive.
-local	The Assure DQ client does not support local data capture for data quality control points. Specify the -local parameter to capture input data from a source local to the Assure DQ client, such as an input file located on the machine where the client runs. Do not use this parameter to capture data from an input source on the server or when running runcp_standalone. Local mode supports JDBC database connections, but not JNDI connections. Optional.
-point *control point name*	Specifies the name of the control point to execute. Enclose the name in quotation marks if it contains spaces. Case-sensitive.

Parameter	Description
-proxypassword	Specifies the password to use to connect to an http proxy server. Optional.
-proxyusername	Specifies the username to use to connect to an http proxy server. Optional.
-rerun	Specifies that the run is a rerun, meaning that the run has the same cycle and run number as a prior run. Optional.
-resultslimit *nnn*	For this control point run, overrides the system and control point settings for the maximum number of results to output. Optional.
-rtfield *name=value*	Specifies runtime field name/value pairs. Optional.
-rtfile *properties file*	Specifies the full path and name of a properties file that contains runtime field values. Optional.
-runno *nnn*	Overrides the run number defined for this control entity or application. Optional.
-showversion	Displays the release and version number of Assure DQ. Optional.
-username	Specifies the user ID that has remote executor permission to run the control point. User IDs with this role can execute control processing from the console, an external job scheduler, or API. Contact the Assure DQ security administrator to request remote executor permissions for one or more user IDs.
-password	Specifies the password for the above user. The username and password can be passed on the command line or preconfigured at installation.
-SP *control point name=N*	This split source parameter may improve performance when processing large input sources. Use this parameter to specify how many logical splits *N* of an input source to create and process concurrently. You must also specify SOURCE_DRIVEN for -cc concurrency mode, below. Optional. Applicable only for files that contain fixed length or delimited records. Not supported for data quality control points.
-cc *concurrency mode*	To split a single input into multiple streams for concurrent processing, specify SOURCE_DRIVEN and use the -SP parameter, above, to specify the number of streams *N*. Not supported for data quality control points. Optional.

Syntax Examples

Suppose you want to execute a control point on the server. The control point is “Control Point 1” within ControlEntity1. The input data sources for the control run are on the server in their expected locations, so no -S overrides are necessary. On Windows, enter a command similar to the following example:

```
<install_folder>\Infogix\server-util\<config_name>\IA>
runcp.bat -entity ControlEntity1 -point "Control Point 1"
```

Notice that “Control Point 1” is enclosed in quotation marks because it contains spaces. To run an assessment or measurement control point belonging to Application1, specify -application Application1 instead of -entity ControlEntity1.

Suppose you want to execute a control point using the client, because the data sources for the control run are on a different system than the Assure DQ server. The names of the data sources are Source1 and Source2. The source files are file1.csv and file2.csv in /home/user1. On the client, enter a command similar to the following example:

```
/home/user1>./runcp.sh -entity ControlEntity1
-point "Control Point 1" -S Source1=file1.csv
-S Source2=file2.csv -local
```

If you have multiple client configurations, enter a command similar to the following example:

```
runcp.sh –c <configname> -entity ControlEntity1 ...
```

where <configname> is the name of your configuration, such as test.

This enables the client to connect to the Assure DQ server instance of your choice, such as a test instance or a development instance. For instructions on setting up multiple configurations, see the *Assure DQ Client Guide*.

Halting a Control Run

You can stop control processing from the console or the Assure DQ user interface. The control point issues a fatal error and the highest return code set during control processing prior to the halt. If no return code has been set, the control point issues a return code of 16.

The command only stops control points that are currently running. Future control runs are not affected.

To halt a control run from the user interface, log in to the Assure DQ user interface as an administrator and knowledge worker. Select Results > By Control Point and choose the control point to stop. Click the Stop button in the bottom pane.

To stop processing from the console, execute the Stop Control Point command from <install_folder>/Infogix/server-util/<config_name>/IA. Execute signal_cp_stop.bat (for Windows) or signal_cp_stop.sh (for UNIX or Linux), followed by the appropriate parameters from the table below.

Parameter	Description
-resultCatalogID *nnnn*	A unique ID generated by the system for control runs. Pass this parameter to indicate which control run to stop. To determine the catalog ID, log in to the Assure DQ user interface as a knowledge worker. Choose Results > By Control Point and select the control point to stop. Go to the Summary tab for the control point and make a note of the catalog ID.
-waitTime	Specifies the maximum time in seconds to wait to confirm that the control point stopped. Optional. If you specify a wait time, the system will issue a confirmation message to the console when the control point stops. Specify a longer wait time if the response time for your system is slow. If you do not specify a wait time, you will see a console message indicating the control point stop is pending. However, to confirm that the control point stopped, you must log in to the Assure DQ user interface as a knowledge worker. Choose Results > By Control Point and select the control point run. Go to the Results tab and select and view the last result listed.
-username	Specifies the username to connect to an http server. Optional.
-proxyusername	Specifies the username to connect to an http proxy server. Optional.
-password	Specifies the password to connect to an http server. Optional.
-proxypassword	Specifies the password to connect to an http proxy server. Optional.

Syntax Example

Suppose you want to halt a control point run on the server, with a catalog ID of 9999 and a wait time of 60 seconds. Enter a command similar to one of the following examples:

Windows:

```
signal_cp_stop.bat -resultCatalogID 9999 -waitTime 60
```

UNIX/Linux:

```
./signal_cp_stop.sh -resultCatalogID 9999 -waitTime 60
```

Running Controls in Standalone Mode

If J2EE/JVM memory constraints are an issue in your production environment, consider running in standalone mode. Standalone programs are available for resource-intensive processes such as running control points and applying retention processing. The standalone programs run in their own JVM on the machine where the Assure DQ server runs, so the efficiency of the Assure DQ server is not adversely affected.

Configuring Standalone Operation

Before you can run controls in standalone mode, you must allocate separate memory for standalone processing. Your server machine must have the capacity to dedicate memory to standalone processing, while still allowing other processes to run normally.

Initially, 128 MB of memory is allocated for standalone processing. To increase memory, open <install_folder>/Infogix/server-util /<config_name>/IA/standalone-conf, and edit common.sh or common.bat. Increase the CLIENT_JAVA_OPTIONS memory setting in 128 MB increments until standalone control runs execute successfully. For example, increase the setting from -Xmx128m to -Xmx256m, then to -Xmx384m, and so on.

If the Assure DQ server is deployed in a cluster, an additional configuration change is necessary to run standalone. Open <install_folder>/Infogix/server-util/config.name/IA/standalone-conf/oscache.properties and set cache.is_clustered to false.

Executing Standalone Control Point Processing

The runcp_standalone program invokes control point processing using its own JVM on the machine where the Assure DQ server runs. Control runs that process large volumes of data run more efficiently in standalone mode, because they are not waiting for resources in use by other processes. Standalone control point processing is not supported for data quality control points.

User IDs with the remote executor role can initiate standalone control processing from the console or an external job scheduler. The program is in <install_folder>/Infogix/server-util/<config_name>/IA.

Execute runcp_standalone.bat (for Windows) or runcp_standalone.sh (for UNIX or Linux). The parameters are identical to those for runcp (see page 18).

Standalone mode supports JDBC database connections, but not JNDI connections. If your control point must use a JNDI connection to a data source or translation table, use the (non-standalone) runcp program on the server instead. The runcp program supports both JNDI and JDBC connections on the server.

Halting a Standalone Control Run

You can stop a standalone control run from the console, but not from the Assure DQ user interface. To halt a control run from the console, use a cancel command, such as Ctrl-C.

Close Control Point Catalogs

Sometimes control points may not complete and close the result catalog for the execution. This can occur when the application server is stopped while control points are running, as well as some failure scenarios for control points. The closeCatalogs utility can be used to clean up result catalogs that were left in an open state, which appears to users of the system as results catalogs with a "-1" return code. The closeCatalogs utility is in <install_folder>/Infogix/server-util /<config_name>/IA.

Execute closeCatalogs.bat (for Windows) or closeCatalogs.sh (for UNIX or Linux), followed by the appropriate parameters from the table below. You can close result catalogs for all that have been open for an amount of time, or close particular result catalogs based upon catalog IDs seen in the user interface.

Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters.

Apply Retention Parameters	Description
-catalogs	A comma separated list of result catalog IDs to close. Optional.
-olderThan	As an alternative to specific catalogs, specify a number of hours to indicate that all catalogs that have been open at least this long are to be closed.
-dryRun	When specified, does not close catalogs, and only lists those catalogs that would be closed.
-password	Specify the password to use to connect to an http server. Optional.
-proxypassword	Specify the password to use to connect to an http proxy server. Optional.
-proxyusername	Specify the username to use to connect to an http proxy server. Optional.
-username	Specify the username to use to connect to an http server. Optional.

Syntax Examples

To close any catalogs that have been open for more than 20 hours:

Windows:

```
closeCatalogs -olderThan 20
```

UNIX/Linux:

```
./closeCatalogs.sh -olderThan 20
```

To close specific result catalogs, enter a command similar to one of the following examples:

Windows:

```
closeCatalogs -catalogs 123456,123457
```

UNIX/Linux:

```
./closeCatalogs.sh -catalogs 123456,123457
```

3

Message Handling

This chapter provides information on Assure DQ message handling. The chapter has the following sections:

- Custom JMS Message Handling
- Message Queue Error Handling

Custom JMS Message Handling

Assure DQ provides a custom handler that can be enabled and configured to detect multipart mime messages that require extra extraction handling. When a mime message is detected, the system has a default, configurable behavior that determines which part of the mime to extract and make available as input to the control point.

By default, the handler extracts the first message part. You can change the default behavior for all control points, and override the default behavior for individual control points. The handler can select an individual message part to extract, based on a numeric index value (0, 1, 2, etc., or -1 for last part) or a content-type prefix (text/xml or application/octet-stream). The handler can also extract all parts of the message and send them to the control point as API data.

Enabling the Custom Message Handler

To enable and configure the custom handler, complete the following steps:

1. Open <install_folder>/igx-data/<config_name>/IA/config/api.properties.
2. Set the ENABLE_CUSTOM_MESSAGE_HANDLERS property to true.
3. Set the CUSTOM_MESSAGE_HANDLER_1 property to the following value, with no line wrapping:

   ```
   com.unitechsys.acr.component.controlsassure.ent.
   impl.MimeJMSAttachmentExtractor
   ```
4. Open the mime_jmsextractor.properties file, which is also in the config directory.

5. To edit the default behavior of the handler, change the DEFAULT_EXTRACT property to the appropriate value, below:
   - * (asterisk) extracts all message parts
   - 0 extracts the first message part
   - 1 extracts the second message part
   - 2 extracts the third part (and so on, up to 9, which extracts the tenth part)
   - -1 extracts the last part
   - text/xml extracts the first message part with text/xml content
   - application/octet-stream extracts the first message part with application/octet-stream content
6. To override the default extract setting for one or more individual control points, uncomment and edit the entity.control.point.name.EXTRACT property in the mime_jmsextractor.properties file.

   As shown in the properties file, you can override the extract behavior with either a message part setting (0,1,-1,*) or a content type.

   ```
   # Sample control-specific MIME body-part selectors
   # entity_name.point_name.EXTRACT=0
   # entity_name.point_name.EXTRACT=text/xml
   ```
7. Run the update-config script located in <install_folder> /Infogix/bin.

   Allow five minutes for your changes to take effect.

Creating Message Layouts

You must also create an Assure DQ layout for each message part you plan to extract. In addition, if you plan to extract all parts of a message, you must create a Java Object layout that references the layouts of its individual message parts.

To create a layout for an individual message part, complete the following steps:

1. In the user interface, select Rules > Layouts.
2. Click New Layout and create the appropriate layout type based on the content of the message part you want to extract.

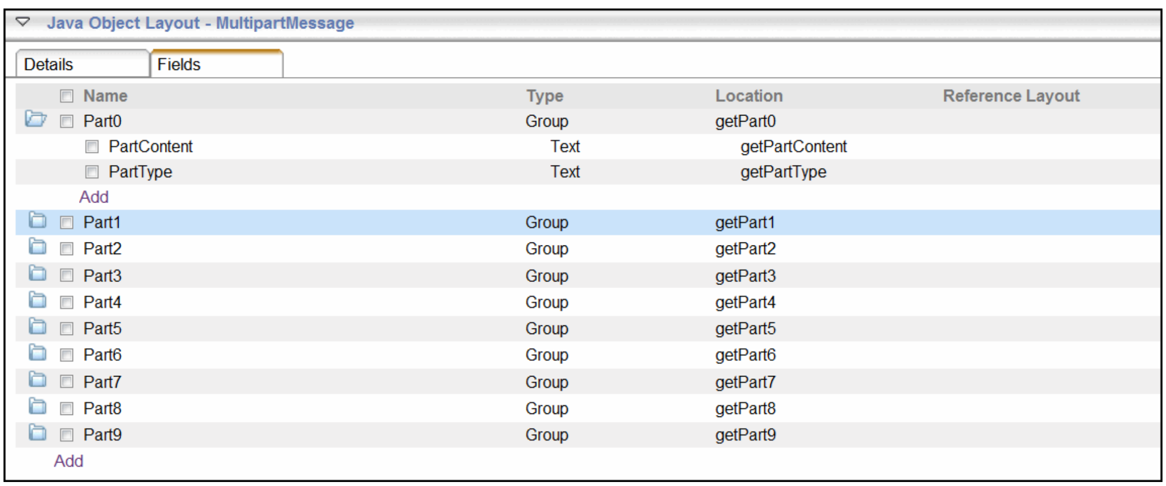
   For example, create a fixed layout for a message part containing COBOL data, or an XML layout for a message part containing XML data.

If you plan to extract all message parts, follow the procedure above to create layouts for each message part that contains data of interest. Then, follow the procedure below to create a Java Object layout for the complete message. The Java Object layout will reference the layout for each individual message part you plan to use.

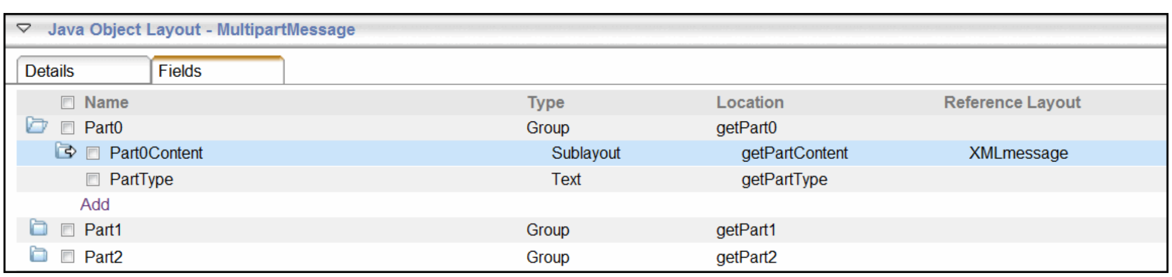
1. In the user interface, select Rules > Layouts > Java Object.
2. On the Java Object Layout Details tab, enter the following string in the Class Name field, without a line break:

   ```
   com.unitechsys.acr.component.controlsassure.ent.
   impl.MIMEJavaObject
   ```
3. On the Fields tab, click Generate > From Class to have the system generate a ten part message layout for you.

   The layout has ten groups, numbered 0-9. Each group has a Part Content and Part Type field.

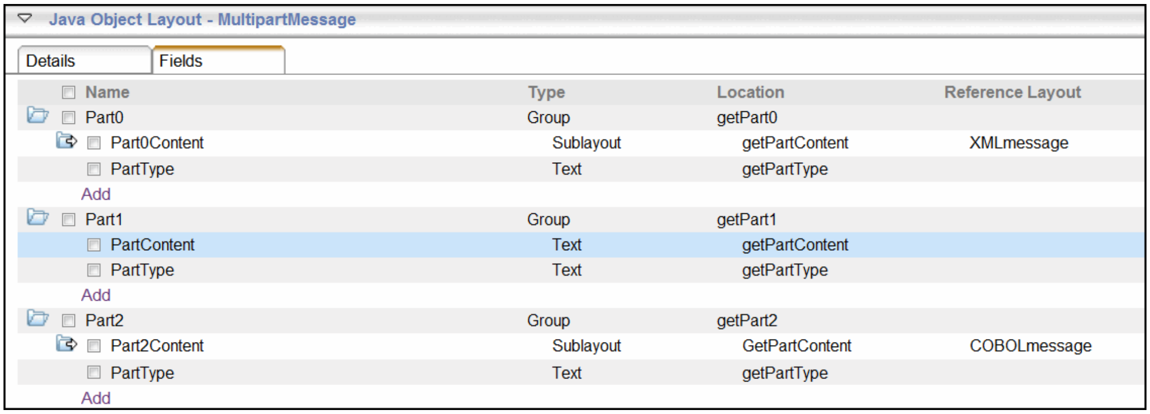


4. For each message part that contains data of interest, remove the Part Content text field. Then, add a Sublayout that references the layout for that message part.



5. Be sure to specify getPartContent as the Location for the sublayout.

6. It is unnecessary to add sublayouts for message parts that contain data you do not plan to use.
   For example, the user did not bother to create a sublayout for message part 1 in the example below.



Message Queue Error Handling

This section describes Assure DQ API and JMS message error handling. If an error occurs during message processing, Assure DQ places the message on an error queue to facilitate error recovery. The following potential errors may cause message processing to fail:

- A message is not well formed according to the schema
- Required fields are missing from an API request
- The message is trying to invoke a control point that is not defined in the current system or is disabled
- The Assure DQ internal database is offline

An error queue utility is available to clear the error queue or re-queue messages from the error queue to the input queue. If you have your own error handling processes, you can use the utility to clear the error queue on a regularly scheduled basis. You can manage the error queue from the user interface or the command line.

Managing the Error Queue From the User Interface

To manage the error queue from the user interface, complete the following steps:

1. Log in to the user interface as an Administrator or System Operator.
2. Select System > More > Error Queue Utility.
3. To clear all messages from the error queue, click Clear All.
4. To send all messages on the error queue back to the system input queue, click Requeue All.
5. To clear individual messages, select the messages and click Clear.
6. To requeue individual messages, select them and click Requeue.
7. To save the content of an individual message, click the Get Message Content icon to the right of the message. 
   You can save the message content to a file to use for testing.
8. To filter the messages to display, clear, or requeue, open the filter and specify a date in the Before Date field and click Apply.
9. To clear the filtered messages from the queue, click Clear Filtered.
10. To requeue the filtered messages to the input queue, click Requeue Filtered.

Managing the Error Queue From the Command Line

You can use the command line error queue utility to clear the Assure DQ error queue or requeue messages from the error queue to the input queue. The Error Queue utility is in <infogix_client_install_folder>/Infogix /server-util/<config_name>/IA.

To see syntax information for the utility, go to the directory described above and enter the name of the utility followed by -?. On Windows, enter a command similar to the following example: APIErrorQueueUtility -?

You must have administrator authority to run the command line error queue utility. Navigate to the directory where the utility is located and enter the name of the utility at the command prompt, followed by the appropriate parameters from the table below.

API Error Queue Utility Parameters	Description
-before	Selects messages prior to this date and time to clear from the queue or resubmit to the input queue. Specify dates (and optionally, times in hours, minutes, and seconds) in standard ISO 8601 datetime format, for example, 2010-07-04T10:00:00. See http://www.w3.org/TR/NOTE-datetime. The date portion of the parameter is required. The time portion is optional and defaults to 00:00:00.
-clear	Clears all messages from the error queue, or messages prior to a date and time specified with the -before parameter, above.
-requeue	Requeues messages from the error queue to the input queue. Requeues all messages, or messages prior to a date and time specified with the -before parameter, above. Optional.
-username	Specifies the username to use to connect to an http server. Optional.
-password	Specifies the password to use to connect to an http server. Optional.
-proxyusername	Specifies the username to use to connect to an http proxy server. Optional.
-proxypassword	Specifies the password to use to connect to an http proxy server. Optional.
-?	Describes the command line arguments.

To requeue messages prior to July 4, 2010 at 10 am, enter a command similar to the following example:

Windows:

```
APIErrorQueueUtility -before 2010-07-04T10:00:00 -requeue
-username user_id -password user_password
```

UNIX/Linux:

```
./APIErrorQueueUtility.sh -before 2010-07-04T10:00:00
-requeue -username user_id -password user_password
```

4

Running Retention and Archiving

Archiving and retention are scheduled system processes that normally run on a daily basis. However, from time to time, special retention or archiving runs may be necessary to clean up backlogs or take advantage of low production loads. This chapter describes how to run data retention and archiving from the console. The chapter also describes the generate reports utility, which you can use at any time to make specific reports and results accessible to users outside Assure DQ.

This chapter includes instructions on how to edit the system schedules to ensure archiving and retention occur during off peak processing times. The chapter also explains how to run retention or archiving in standalone mode, as an alternative to the scheduled system process. Standalone processing runs in its own JVM, so production processing is not adversely affected.

The chapter also includes tips on resolving retention and archiving problems. Commands are available to halt scheduled archival or retention processing if they interfere with control point execution. The chapter also describes what to do if the system cannot complete retention processing due to an archiving failure.

This chapter has the following sections:

- Running Retention From the Console
- Running Archiving From the Console
- Generating Reports From the Console
- Scheduling Archival and Retention Processing
- Running in Standalone Mode
- Resolving Retention and Archiving Problems

Running Retention From the Console

You can initiate retention processing from the command line to clean up backlogs or take advantage of low production loads. If the system administrator recently changed data retention policies to shorter timeframes, you may need to initiate a special retention run to purge high volumes of expired data and reports from the system.

User IDs with the superuser or administrator role can run retention processing from the console. Execute the program from <install_folder>/Infogix /server-util/<config_name>/IA.

Execute applyretention.bat (for Windows) or applyretention.sh (for UNIX or Linux), followed by the appropriate parameters from the table below. You can apply retention for all control entities and applications, or selectively roll off data and reports for a single control entity or application.

Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters.

Apply Retention Parameters	Description
-control	To run retention processing for a single control entity or application, specify the control entity or application name. Optional.
-password	Specify the password to use to connect to an http server. Optional.
-proxypassword	Specify the password to use to connect to an http proxy server. Optional.
-proxyusername	Specify the username to use to connect to an http proxy server. Optional.
-username	Specify the username to use to connect to an http server. Optional.

Syntax Examples

To run retention processing for all control entities and applications, enter a command similar to one of the following examples:

Windows:

```
applyretention
```

UNIX/Linux:

```
./applyretention.sh
```

To run retention processing for a single control entity named *ControlEntityABC*, enter a command similar to one of the following examples:

Windows:

```
applyretention -control ControlEntityABC
```

UNIX/Linux:

```
./applyretention.sh -control ControlEntityABC
```

Running Archiving From the Console

You can initiate archival processing from the command line or your own job scheduler in order to clean up existing backlogs of results and reports during off peak production times. This is advisable after upgrading from any release earlier than release 7.1, because the automatic archiving feature was not available prior to release 7.1. When archiving is enabled, results and reports must be successfully archived before retention processing can remove them from the system.

Before you run archiving from the command line, complete the steps in the Data Retention and Archiving section of the *System Administrator's Guide* to enable archiving and to specify the archive home directory.

User IDs with the superuser or administrator role can run archival processing from the console. To run archiving, go to <install_folder>/Infogix/server-util/<config_name>/IA.

Execute archiveReports.bat (for Windows) or archiveReports.sh (for UNIX or Linux), followed by the appropriate parameters from the table below. Note that if you specified the username and password in the application server properties file, you do not have to specify them as parameters.

archiveReports Parameters	Description
-username	Specifies the username to use to connect to an http server. Optional.
-proxyusername	Specifies the username to use to connect to an http proxy server. Optional.
-password	Specifies the password to use to connect to an http server. Optional.
-proxypassword	Specifies the password to use to connect to an http proxy server. Optional.

Syntax Examples

To run archival processing, enter a command similar to one of the following examples:

Windows:

```
archiveReports
```

UNIX/Linux:

```
./archiveReports.sh
```

Scheduling Archival and Retention Processing

The Infogix Archival Schedule runs archival processing daily at 21:00. Edit this schedule to ensure archival processing occurs during off peak processing times prior to retention processing. When possible, daily archival processing is desirable to prevent archiving from taking too long on any given day. Daily archiving limits the volume of control data and reports to archive to those generated within the previous 24 hours.

By default, the Infogix Apply Retention Policy Schedule applies retention processing daily at 00:00. This is a single, shared schedule that applies the retention policies to all controls. However, the rules writer can select an option on the control entity or application Storage tab to apply retention to an individual entity or application on a different, separate schedule. Consider using this feature if a control generates a lot of results or reports that you want to purge on a different, separate schedule or at a time when fewer resources are required for production processing.

Complete the steps below to edit the system schedules and any separate schedules you create, to ensure archival and retention processing occur during off peak processing times:

1. Log in to the Assure DQ user interface as an administrator or integrity rule administrator.
2. Select Rules > Calendars.

   The system displays the Manage Calendars window. The top pane lists all calendars.
3. Select the schedule you want to review and edit, such as the Infogix Archival Schedule or the Infogix Apply Retention Policy Schedule.
4. Click Edit Events.
5. Select the default event for the calendar.

   The bottom pane displays the settings for this event.

6. You can modify the following settings:
   - The name of the event
   - The start date and time
   - The end time, if this is not a recurring event
   - Whether the event occurs monthly, weekly, or on a specific working day of the month
   - The days of the week when the event occurs, if applicable
   - The working day of the month when the event occurs, if applicable
   - The number of hours, minutes, and seconds before the event reoccurs
7. Apply each change, click OK, and then save your changes.

Generating Reports From the Console

The generateReports utility generates reports and results in the format of your choice and places them in a secure directory on your file system or network. Anyone who can access this directory does not need to log in to Assure DQ to view or print reports or results.

By default, results and most reports are generated in pdf format. If you prefer, you can select text format for result catalogs, and text or csv format for most types of reports. Match set audit reports are always generated in text format, regardless of the output format specified. Result catalogs are available in pdf or text format, but not csv. Failed value reports are always generated in csv format, regardless of the output format specified. if you select the output format which is not available, the default format, specific to the reports or results, is generated.

You can generate specific result catalogs, or result catalogs and reports for all control points that executed during a specific timespan. If you prefer, you can generate specific reports for a specific control entity, application, or control point.

By default, the generateReports utility generates reports on the Assure DQ server in the “report-outputs” folder within the “igx-data” location (see the –outputpath parameter description in the table below for the exact location). You can override this location using the –outputpath parameter and specify an absolute path on the Assure DQ server that the server has write access for. Alternately, you can generate the reports where the generateReports utility is invoked from using the –local parameter. The generateReports utility is also available in the Assure DQ Client installation, allowing you to generate reports on systems remote from the Assure DQ server. When you use the –local parameter, the system generates the reports in the current working directory where the utility is invoked by default. You can use the –outputpath parameter with the –local parameter, in which case the reports are generated in the relative (or absolute) path specified from the current working directory where the utility is invoked.

The User IDs with the superuser or administrator role can generate reports from the console. The generateReports utility is in <install_folder> /Infogix/server-util/<config_name>/IA.

Enter generateReports.bat (for Windows) or generateReports.sh (for UNIX or Linux), followed by the appropriate parameters from the table below. Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters.

generateReports Parameters	Description
-catalogs	Use this option if you want to output pdfs or text files of specific result catalogs only. If you prefer to output result catalogs and reports, use the startdate option instead. Specify the catalog ID(s) of the result catalogs you want. Multiple catalog IDs may be specified, separated by commas. To determine a catalog ID, log in to the user interface and select Results > By Control Point. Select a result of interest. The catalog ID appears on the Summary tab for the result.

generateReports Parameters	Description
-startdate *datetime* -enddate *datetime*	Use this option to generate result catalogs, results, and reports from control runs created on or between specified dates. To generate result catalogs only, specify the catalogs option (above) instead. To generate results and reports from a single date, enter the same date for both parameters. Specify dates (and optionally, times in hours, minutes, and seconds) in standard ISO 8601 datetime format. See http://www.w3.org/TR/NOTE-datetime. The date portion of the startdate is required. The time portion is optional and defaults to 00:00:00. If you run the utility using a job scheduler, specify a relative date based on how frequently you run the utility. For example, startdate –P7D would generate results and reports for the last seven days, including today. When the utility generates results and reports, it does not delete them from the system. If you run the utility again with the same absolute startdate, the reports and results from that date forward will be generated again (unless retention processing has deleted them). The enddate is optional and defaults to the current date and time.
-excluderesults	If you specify the startdate option, result catalogs, results, and reports are output. If you prefer to output reports only, specify -excluderesults. Optional.

generateReports Parameters	Description
-reportTypes	By default, all report types are produced if you specify the startdate option. To generate specific reports only, specify the numbers corresponding to the report types to produce, separated by commas. 0 = Control Report 1 = User Report 2 = Reconciliation Match Report 3 = Match Set Audit Report 4 = Recap Report 5 = Data Element Control Report 6 = Data Element Recap Report 7 = Failed Value Report Optional.
-control	Use in conjunction with the startdate option if you want to output reports and results for a single control entity or application only. Specify the control entity or application name. Optionally, to further narrow the scope of reports to generate, specify a control point, below. Optional.
-cp	Use in conjunction with the startdate option and the control option, above, if you want to output reports and results for a single control point only. Specify the control point name. Optional.
-csv	By default, reports are output in pdf format. If you prefer comma separated variable length files, specify -csv. This option does not apply to result catalogs and match set audit reports. The system generates result catalogs as pdfs if you specify csv format. Match set audit reports are always generated in text format, regardless of the output format specified. Optional.
-text	By default, results and reports are output in pdf format. If you prefer text files, specify -text. Match Set Audit reports are always generated in text format, and Failed Value reports are always generated in csv format, regardless of the output format specified. Optional.

generateReports Parameters	Description
-local	Use this option to output result catalogs, results, and reports to the remote platform. Optional.
-outputpath	Specify the full path of the directory on your file system where you want to output the result and report files. The default location is *infogix.install.home*\igx-data*config.name*\IA\data \report-outputs. Optional.
-username	Specify the username to use to connect to an http server. Optional.
-proxyusername	Specify the username to use to connect to an http proxy server. Optional.
-password	Specify the password to use to connect to an http server. Optional.
-proxypassword	Specify the password to use to connect to an http proxy server. Optional.

Syntax Examples

To output result catalogs 99999 and 1234 in pdf format, enter a command similar to one of the following examples:

Windows:

```
generateReports -catalogs 99999, 1234
```

UNIX/Linux:

```
./generateReports -catalogs 99999, 1234
```

To generate results and reports created on or between one day ago and the current date and time, use a command similar to one of the following examples:

Windows:

```
generateReports -startdate -P1D
```

UNIX/Linux:

```
./generateReports.sh -startdate -P1D
```

By default, reports and results are output in pdf format. However, Match Set Audit reports are always generated in text format, and Failed Value reports are always generated in csv format, regardless of the output format specified.

To generate results and reports created on or between March 1, 2011 at 8:30 am and March 31, 2011 at 10:00 pm and place them in text files, enter a command similar to the following example:

Windows:

```
generateReports -text -startdate 2011-03-01T08:30:00
-enddate 2011-03-31T20:00:00
```

UNIX/Linux:

```
./generateReports.sh -text -startdate 2011-03-01T08:30:00
-enddate 2011-03-31T20:00:00
```

To generate results and reports created in the last day by a control entity and place the results in a local folder where the utility is run from, enter a command simiar to the following example:

Windows:

```
generateReports -control "Some Name" -startdate P1D -local -
outputpath "controls/Some_Name"
```

UNIX/Linux:

```
./generateReports.sh -control "Some Name" -startdate PID -
local -outputpath "controls/Some_Name"
```

Running in Standalone Mode

Standalone programs are available for resource-intensive processes, such as control point execution and retention processing. If J2EE/JVM memory constraints are an issue in your production environment, consider running retention, archiving, and report generation in standalone mode. The standalone programs run in their own JVM on the machine where the Assure DQ server runs, so the efficiency of the Assure DQ server is not adversely affected.

Configuring Standalone Operation

Before you can run standalone programs, you must allocate separate memory for standalone processing. Your server machine must have the capacity to dedicate memory to standalone processing, while still allowing other processes to run normally.

Initially, 128 MB of memory is allocated for standalone processing. To increase memory, open <install_folder>/Infogix/server-util/<config_name>/IA/standalone-conf and edit common.sh or common.bat. Increase the CLIENT_JAVA_OPTIONS memory setting in 128 MB increments until standalone control runs execute successfully. For example, increase the setting from -Xmx128m to -Xmx256m, then to -Xmx384m, and so on.

If the Assure DQ server is deployed in a cluster, an additional configuration change is necessary to run standalone. Open <install_folder>/Infogix/server-util/<config_name>/IA/standalone-conf/oscache.properties and set cache.is_clustered to false.

Executing Standalone Retention Processing

The applyretention_standalone program is similar to the apply retention program described on page 33. Running retention processing in standalone mode has several advantages. You can manually initiate processing when production loads are low and sufficient time is available for processing to complete. In addition, if necessary, you can stop retention processing without shutting down the Assure DQ server.

Be sure to disable scheduled retention processing before you use the standalone utility. The scheduled process and the standalone process cannot run simultaneously. If an apply retention process starts while another apply retention process is executing, the system terminates the second process.

To disable scheduled retention processing, log in to the user interface and select Rules > Calendars. Edit the Infogix Apply Retention Policy Schedule. Remove any scheduled apply retention events from the schedule.

User IDs with the remote executor role can initiate standalone retention processing from the console or an external job scheduler. The program is in <install_folder>/Infogix/server-util/<config_name>/IA.

Execute applyretention_standalone.bat (for Windows) or applyretention_standalone.sh (for UNIX or Linux). The parameters are identical to those for applyretention (see page 34).

Executing Standalone Archival Processing

The archiveReports_standalone program is similar to the archive reports program described on page 35. Running archival processing in standalone mode has several advantages. You can manually initiate processing when production loads are low and sufficient time is available for processing to complete. In addition, if necessary, you can stop archival processing without shutting down the Assure DQ server.

Before you run standalone archival processing, complete the steps in the Data Retention and Archiving section of the *System Administrator's Guide* to enable archiving and to specify the archive home directory.

If an archive reports process starts while another archive reports process is executing, the system terminates the second process. If you would like to disable scheduled archival processing, log in to the user interface and select Rules > Calendars. Edit the Infogix Archival Schedule. Remove any scheduled archive reports and results events from the schedule.

User IDs with the remote executor role can initiate standalone archival processing from the console or an external job scheduler. The program is in <install_folder>/Infogix/server-util/<config_name>/IA.

Execute archiveReports_standalone.bat (for Windows) or archiveReports_standalone.sh (for UNIX or Linux). The parameters are identical to those for archiveReports (see page 34).

Executing Standalone Report Generation

The generateReports_standalone program is similar to the generate reports program described on page 37. The standalone program runs in its own JVM, so production processing is not adversely affected.

User IDs with the remote executor role can initiate standalone report generation from the console or an external job scheduler. The program is in <install_folder>/Infogix/server-util/<config_name>/IA.

Execute generateReports_standalone.bat (for Windows) or generateReports_standalone.sh (for UNIX or Linux). The parameters are identical to those for generateReports (see page 38).

Resolving Retention and Archiving Problems

This section describes what to do if scheduled archival or retention processing interferes with control point execution. In addition, the section describes what to do if the system cannot complete retention processing due to an archiving failure.

Halting Retention Processing

If *scheduled* retention processing interferes with control point execution, use the signal_retention_stop command to halt processing. The system rolls back the transaction it is currently processing and logs a message that the operator stopped retention processing. Retention processing resumes at the next scheduled time.

To stop retention processing running in standalone mode or initiated from the command line or your job scheduler, simply kill the process.

User IDs with the superuser or administrator role can stop retention processing from the console. Execute the command from <install_folder>/Infogix/server-util/<config_name>/IA. Enter signal_retention_stop.bat (for Windows) or signal_retention_stop.sh (for UNIX or Linux), followed by the appropriate user name and password. Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters.

Parameter	Description
-username	Specifies the username to connect to an http server. Optional.
-proxyusername	Specifies the username to connect to an http proxy server. Optional.
-password	Specifies the password to connect to an http server. Optional.
-proxypassword	Specifies the password to connect to an http proxy server. Optional.

Syntax Examples

To halt retention processing from the command line, enter a command similar to one of the following examples:

Windows:

```
signal_retention_stop.bat
```

UNIX/Linux:

```
./signal_retention_stop.sh
```

Halting Archival Processing

If *scheduled* archiving interferes with control point execution, use the signal_archive_stop command to halt processing. The system rolls back the transaction it is currently processing and logs a message that the operator stopped archival processing. Archival processing resumes at the next scheduled time.

To stop archival processing initiated from the command line or your job scheduler, simply kill the process.

User IDs with the superuser or administrator role can stop archival processing from the console. Execute the command from <install_folder>/Infogix/server-util/<config_name>/IA. Enter signal_archive_stop.bat (for Windows) or signal_archive_stop.sh (for UNIX or Linux), followed by the appropriate user name and password. Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters.

Parameter	Description
-username	Specifies the username to connect to an http server. Optional.
-proxyusername	Specifies the username to connect to an http proxy server. Optional.
-password	Specifies the password to connect to an http server. Optional.
-proxypassword	Specifies the password to connect to an http proxy server. Optional.

Syntax Examples

To halt archival processing from the command line, enter a command similar to one of the following examples:

Windows:

```
signal_archive_stop.bat
```

UNIX/Linux:

```
./signal_archive_stop.sh
```

Skipping Archiving

A message displays in the log if the system cannot complete retention processing due to an archiving failure. The message indicates the report ID or catalog ID of the problematic report or result catalog.

Results and reports must be successfully archived before retention processing can remove them from the system. When necessary, you can use the skipArchive utility to override this safeguard. The utility applies a “skip archiving” flag to unwanted or problematic results or reports, so retention processing can remove them from the system on a timely basis.

The skipArchive utility is in <install_folder>/Infogix/server-util/<config_name>/IA. To see syntax information for the utility, go to the directory where the utility is located and enter the name of the utility followed by -?. On Windows, enter a command similar to the following example: archivereports.bat -?

To skip archiving of reports and results, execute skipArchive.bat (for Windows) or skipArchive.sh (for UNIX or Linux), followed by the appropriate parameters from the table below.

skipArchive Parameters	Description
-username	Specifies the username to use to connect to an http server. Optional.
-proxyusername	Specifies the username to use to connect to an http proxy server. Optional.
-password	Specifies the password to use to connect to an http server. Optional.
-proxypassword	Specifies the password to use to connect to an http proxy server. Optional.
-catalogID	To skip archiving of a problem result catalog, specify the catalog ID to skip. The catalog ID displays in the system log.
-reportID	To skip archiving of a problem report, specify the report ID to skip. The report ID displays in the system log.
-startdate *datetime* -enddate *datetime*	Use these parameters to skip archiving of all result catalogs, results, and reports from control runs created on or between specified dates. To skip results and reports from a single date, enter the same date for both parameters. If you specify a catalog ID or report ID, above, do not specify a start or end date. Specify dates (and optionally, times in hours, minutes, and seconds) in standard ISO 8601 datetime format. See http://www.w3.org/TR/NOTE-datetime. The date portion of the startdate is required. The time portion is optional and defaults to 00:00:00. The enddate is optional and defaults to the current date and time.

Syntax Examples

Skip a Result Catalog

To skip archiving of results and reports for catalog 9999, enter a command similar to the following example:

Windows:

```
skipArchive -catalogID 9999
```

UNIX/Linux:

```
./skipArchive -catalogID 9999
```

Skip a Report

To skip archiving of an individual report with report ID 99999, enter a command similar to the following example:

Windows:

```
skipArchive -reportID 99999
```

UNIX/Linux:

```
./skipArchive -reportID 99999
```

Skip All Results and Reports for a Time Period

To skip archiving of results and reports created on or between one day ago and the current date and time, use a command similar to the following example:

Windows:

```
skipArchive -startdate -P1D
```

UNIX/Linux:

```
./skipArchive.sh -startdate -P1D
```

To skip archiving of results and reports created on or between March 1, 2008 at 8:30 am and March 31, 2008 at 10:00 pm, enter a command similar to the following example:

Windows:

```
skipArchive -startdate 2008-03-01T08:30:00 -enddate 2008-03-
31T20:00:00
```

UNIX/Linux:

```
./skipArchive -startdate 2008-03-01T08:30:00 -enddate 2008-
03-31T20:00:00
```

5

Extracting Control Data

The control data of a control entity can be exported via the command line utility, with control over the delimited output structure (header or not, delimiter, quotation), with simple time or catalog-based selection criteria, if the installation is licensed for the export utility feature. This chapter describes how to run extraction processing from the console and extract the control data either to a server location or transmit to a client location.

Extracting Data From the Console

User IDs with the superuser, administrator, or remote executor role can run export processing from the console. Execute the program from <install_folder>/Infogix /server-util/<config_name>/IA. The program can also be run from the Assure DQ client or Assure DQ managed client.

Execute exportControlData.bat (for Windows) or exportControlData.sh (for UNIX or Linux), followed by the appropriate parameters from the table below.

Note that if the username and password are specified in the application server properties file, you do not have to specify them as parameters. For options with date and time values, specify dates (and optionally, times in hours, minutes, and seconds) in standard ISO 8601 datetime format. See http://www.w3.org/TR/NOTE-datetime.

exportControlData Parameters	Description
-password	Specifies the password to use to connect to an http server. Optional.
-proxypassword	Specifies the password to use to connect to an http proxy server. Optional.
-proxyusername	Specifies the username to use to connect to an http proxy server. Optional.
-username	Specifies the username to use to connect to an http server. Optional.
-entity	Specifies the control entity name.

exportControlData Parameters	Description
-all	Specifies that all control data be selected. Only one of -all, -catalog, -from, -fromday or -fromCycle may be specified. Optional.
-catalog	Specifies the result catalog ID. Use this option only with Match Data control entities to export all the data corresponding to the single Summary ID instance that was captured by a single execution of a match data capturing control point. Only one of -all, -catalog, -from, -fromday or -fromCycle may be specified. Optional.
-ext	Specifies a file extension to use, without the period. If omitted, defaults to csv (if fielddelimiter is specified as comma), tab (if fielddelimiter is specified as tab), or cdt. Optional.
-fromday	Specifies a relative duration period for the start of control data extraction with a default of the time-of-day to the start of the day. For example, -fromday "-P7D" would extract data with a starting range of midnight 7 days ago. Only one of -all, -catalog, -from, -fromday or -fromCycle may be specified. Optional.
-from *datetime*	Specifies a date or period for the start of control data extraction.The date portion of the option is required. The time portion is optional and defaults to 00:00:00. You can also specify a relative duration period. For example, specifying -from "-P1D" would extract control data of previous day at the current time of day. Only one of -all, -catalog, -from, -fromday or -fromCycle may be specified.Optional.
-fromCycle cycle.run	Specifies a cycle and run number for the start of control data selection. The run number portion may be omitted. Only one of -all, -catalog, -from, -fromday or -fromCycle may be specified. Optional.
-toCycle cycle.run	Specifies a cycle and run number for the end of control data extraction. This option can only be used in conjunction with the -fromCycle option. If -toCycle is omitted, it defaults to be the same as -fromCycle. Optional.

exportControlData Parameters	Description
-to *datetime*	Specifies a date or period for the end of control data extraction. Use in conjunction with either the -from or -fromday option. This option defaults to current date and time if not specified. This option can only be specified if -from or -fromday is also specified. Optional.
-includeHeader	Includes a header row of field names in the output file. Optional.
-localdest	Transmits the file produced to the remote platform. Use in conjunction with the -destdir option. Optional.
-destdir *path*	Specifies a destination directory where the produced file is saved; If the -localdest option is also specified, the specified location of the directory must exist on the remote platform where the utility is executed.
-noquote	Bypasses enclosing text and date values in double quotes. Optional.
-fielddelimiter *delimiter*	Specifies a delimiter to separate the fields of a record in the produced file. Common delimiters such as a comma or a tab can be used. You can also define a custom delimiter. After -fielddelimiter, the system will handle the word "comma" or "tab" typed out. For anything else, the characters typed are the delimiter. If the option is not specified, comma is the default delimiter. Optional
-recordelimiter *delimiter*	Specifies a delimiter to mark the end of a record. Standard delimiters such as a newline (\n), carriage return (\r), tab (\t), or form feed (\f) can be used. You can also define a custom delimiter.The default delimiter is a newline. Optional.
-charset *charset*	Specifies the codepage to indicate the character set of the produced file.Optional.
-dateFmt *format*	Specifies the date value format to parse as a JAVA SimpleDateformat.The default date format is "yyyy-MM-dd HH:mm:ss.SSS". Optional.
-dateToUTC	Shifts the timestamp values to UTC time. Optional.

Syntax Examples

To extract control data of a Match Data control entity named *ControlEntityABC* with the catalog ID *123* to a directory named *dir*, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat -entity ControlEntityABC" -destdir
"C:/some/server/dir" -catalog 123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -destdir
"/opt/some/server/dir" -catalog 123
```

To place the output file in a local folder where the utility is run from, enter a command similar to the following example:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -destdir
"C:/some/local/dir" -localdest -catalog 123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -destdir
"/opt/some/local/dir" -localdest -catalog 123
```

To include a header row of field names in the produced file, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -
includeHeader -destdir "C:/some/server/dir" -catalog 123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -
includeHeader -destdir "/opt/some/server/dir" -catalog 123
```

To specify a tab as the field delimiter in the produced file, enter a command similar to the following example:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -
fielddelimiter tab -destdir "C:/some/server/dir" -catalog
123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -
fielddelimiter tab -destdir "/opt/some/server/dir" -catalog
123
```

To end each record with a double new line in the produced file, enter a command similar to the following example:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -
recorddelimiter "\r\n\r\n" -destdir "C:/some/server/dir" -
catalog 123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -
recorddelimiter "\r\n\r\n" -destdir "/opt/some/server/dir" -
catalog 123
```

To avoid enclosing text and date values in double quotes, use a command similar to one of the following examples:

Windows

```
exportControlData.bat -entity "ControlEntityABC" -noquote –
destdir "C:/some/server/dir" -catalog 123
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -noquote –
destdir "/opt/some/server/dir" -catalog 123
```

To extract the control data created or modified since the start of Sept 11 till the current time, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat –entity "ControlEntityABC" –destdir
"C:/some/server/dir" –from “2017-09-11”
```

UNIX/Linux:

```
./exportControlData.sh -control "ControlEntityABC" -destdir
"/opt/some/server/dir" from “2017-09-11”
```

To extract the control data created or modified since noon on Sept 11 till the present time, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat –entity “ControlEntityABC” –destdir
"C:/some/server/dir" –from “2017-09-11T12:00:00”
```

UNIX/Linux:

```
./exportControlData.sh –entity “ControlEntityABC” –destdir
"/opt/some/server/dir" –from “2017-09-11T12:00:00”
```

To extract the control data created or modified since noon on Sept 11 through the beginning of Sept 12(midnight), enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat –entity “ControlEntityABC” –destdir
"C:/some/server/dir" –from “2017-09-11T12:00:00” -to "2017-
09-12T00:00:00"
```

UNIX/Linux:

```
./exportControlData.sh –entity “ControlEntityABC” –destdir
"/opt/some/server/dir" –from “2017-09-11T12:00:00” -to
"2017-09-12T00:00:00"
```

To extract the control data for a particular cycle, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat –entity “ControlEntityABC” –destdir
"C:/some/server/dir" –fromCycle 20170911
```

UNIX/Linux:

```
./exportControlData.sh –entity “ControlEntityABC” –destdir
"/opt/some/server/dir" -fromCycle 20170911
```

To extract the control data for a range of cycle and run combinations, enter a command similar to one of the following examples:

Windows:

```
exportControlData.bat –entity “ControlEntityABC” –destdir
"C:/some/server/dir" -fromCycle 20170911.2 -toCycle
20190913.1
```

UNIX/Linux:

```
./exportControlData.sh –entity “ControlEntityABC” –destdir
"/opt/some/server/dir" –fromCycle 20170911.2 -toCycle
20190913.1
```

To shift the date values to UTC time, use a command similar to one of the following examples:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -destdir
"C:/some/server/dir" -dateToUTC –from “2017-09-11T12:00:00”
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -destdir
"/opt/some/server/dir" -dateToUTC –from “2017-09-
11T12:00:00”
```

To specify UTF-8 character set encoding, use a command similar to one of the following examples:

Windows:

```
exportControlData.bat -entity "ControlEntityABC" -charset
"UTF-8" -destdir "C:/some/server/dir" -dateToUTC –from
“2017-09-11T12:00:00”
```

UNIX/Linux:

```
./exportControlData.sh -entity "ControlEntityABC" -charset
"UTF-8" -destdir "/opt/some/server/dir" -dateToUTC –from
“2017-09-11T12:00:00”
```

6

Execute Script Actions

This chapter describes Execute Script actions. Control points and validations can trigger the execution of external scripts coded by your organization. You can use these scripts to perform additional processing based on a validation result or control point completion. You can also create your own scripts to invoke external applications or to execute utilities shipped with Assure DQ. Place any scripts you create in the execute scripts home directory configured at installation time.

Any type of control point can trigger script actions, except data quality control points. You can limit the number of script actions executed by individual Scan and Multiple Capture control points, or accept the default limit of 100 actions.

This chapter has the following topics:

Execute Script Actions

The system can execute one or more script actions when a control point executes or a validation produces a pass, fail, or complete outcome. Optionally, rules writers can specify arguments and parameters to pass to a script. Parameters can include file paths or data values available from control processing.

Most organizations find numerous uses for script actions. The following examples describe a few potential uses:

- Use the Assure DQ file monitoring feature to detect new input files as they arrive for processing. When the system detects that files have arrived, run a script to invoke your application processing and Assure DQ control processing of the files.
- Invoke an external application function of your choice when a file or individual instance of data fails validation.
- Have Assure DQ control points run application processes in environments where an external scheduler is not available.

Assure DQ Script Example

This example demonstrates how to set up a script action in the user interface. The example executes a custom script, myruncp, that in turn executes the Run Control Point script that ships with Assure DQ. See page 17 for more details on the Run Control Point script.

To configure an existing validation to trigger a script, complete the following steps:

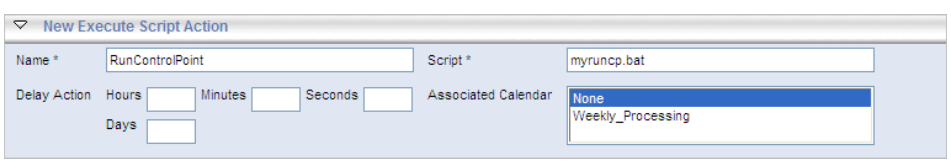
1. Create your custom script.

   In this example, the operator creates "myruncp.bat," which invokes the Assure DQ script named runcp.bat. The custom script, myruncp, passes three parameters to runcp. These parameters are the control entity and control point name and the data source.

2. Copy your script into the script repository created at installation time.

   To configure or change the location of the repository post-installation, modify the EXECUTE_SCRIPTS_HOME property in the IA.properties file. For a description of the property, see the *Infogix Properties Guide*. Run the update-config script in <install_folder>/Infogix/bin to apply your changes.

3. Log in to the Assure DQ user interface.
4. Select Rules > Controls.
5. Select the control entity and control point where you want to add the script action.
6. On the Validate tab, click the Validation Rule Set where you want to add the script action.
7. On the Actions tab, click New Action.
8. In the bottom pane, select Execute Script as the Action Type and click Apply.
9. Enter a name for the script action and the name of the script, such as myruncp.bat.



10. Optionally, specify a time delay for executing the script.

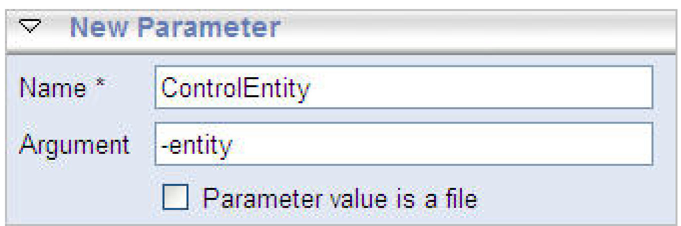
    Enter the number of hours, minutes, seconds or days to wait before executing the script. If you assign the script action to a validation outcome, the script will execute after the specified delay, after the validation outcome occurs.

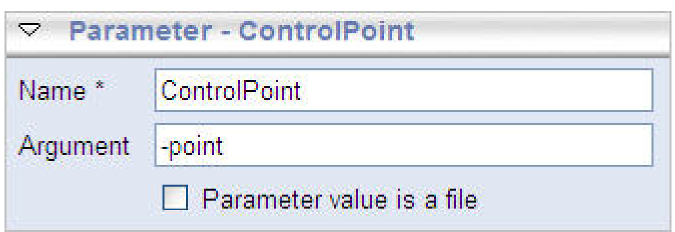
11. Optionally, if you specified a time delay, select a duration calendar from the Associated Calendar list to control when the script can execute.

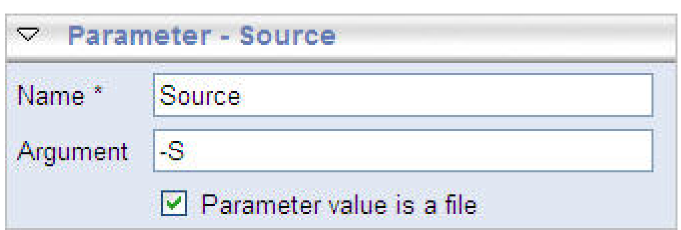
    The script will execute at the next permissible time, according to the duration calendar. The system will calculate the delay only during time spans specified on the duration calendar.

    Duration calendars specify days, dates, or time spans when processing can or cannot occur. For instructions on how to set up duration calendars, see "Creating Duration Calendars" on page 61.

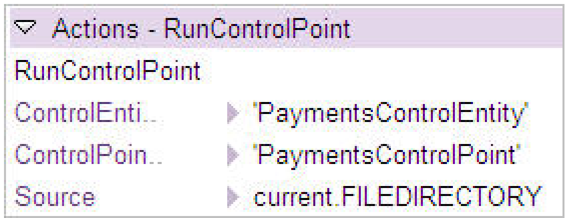
12. Click New Parameter and enter a name and the argument for each parameter to pass to the script.







13. Use the **Add Field, Add Function, Add Operation,** and **Add Literal** buttons, as needed, to assign a value to each parameter.



14. Assign the script action to When Passed, When Failed, or When Complete validation outcomes on the Validate tab.

Execute Action Limits

By default, Scan and Multiple Capture control points can trigger a maximum of 100 execute actions. Rules writers can override this limit for individual control points, as needed. After the limit is met, no more actions execute, but control point execution continues. Any actions that are bypassed after the limit is met appear in the server log and in the control point results displayed in the user interface.

To change an execute action limit, complete the following steps:

1. Log in to the Assure DQ user interface.
2. Select Rules > Controls.
3. Select the control entity and control point where you want to adjust the limit.
4. On the Details tab, enter the limit you prefer in the Execute Action Limit field.

   Valid limits are 0 to 1000.
5. Click Save.

7

Scheduling

Assure DQ has an internal (embedded) job scheduler that executes control points and system utilities. The scheduler executes jobs based on internal schedule calendars. Log in to the Assure DQ user interface if you need to temporarily suspend all scheduled jobs (page 59) or change the run status of individual jobs (page 58).

Assure DQ internal schedule calendars control execution of control points and the following system utilities:

- **Apply Retention**—Applies both the standard and alternate data retention policies specified for control entities in the user interface.
- **Delete Alerts**—Removes expired alerts or alerts older than a specified date or age from user inboxes.
- **Action Rule Processing**—Specifies when and how often to execute actions linked to passed/failed/completed validation outcomes. These actions consist of inbox alerts, email alerts, pager alerts, log actions, script actions, and execute actions.

You can exclude one-time or recurring system down times from control point execution schedules, instead of suspending scheduled jobs. See “Managing Calendars” on page 59 for more information on calendars.

Managing the Internal Scheduler

This procedure describes how to use the Assure DQ internal job scheduler to manage one-time and recurring jobs scheduled for future execution. To temporarily suspend all scheduled jobs, see “Suspending All Scheduled Jobs” on page 59.

Managing Scheduled Jobs

To manage scheduled jobs in Assure DQ, complete the following steps:

1. Log in to the Assure DQ user interface as a superuser or administrator.
2. Select System > Scheduler.

   The system displays the Manage Scheduler window. The top pane lists all control points and utilities (labeled *jobs*) scheduled to run one or more times in the future. For each job, the list also shows the next run time and the job's status. The list does not include jobs that ran in the past, but are not scheduled to run again. The list shows each job only once, even if the schedule calls for the job to execute more than once.
3. To delete a job scheduled to run only once, select the job from the Job list and click Remove.

   The Remove button is only enabled for one-time jobs.
4. To suspend a job or change the next run time, select the job you want to change from the Job list.

   In the bottom pane, drop down the Action list and select one of the following options for running or suspending the job:

   - **Enable Job**—Runs the job according to the schedule calendar. If a one-time job was scheduled to run while the entire scheduler was disabled, the job will run as soon as the entire scheduler is enabled again.
   - **Disable Job**—Suspends all future executions of the job until an operator chooses to enable or run it.
   - **Run Job At Next Scheduled Time**—Cancels the Next Run Time shown on the Job list and reschedules the job for the following run time. After that, the job will continue to run according to the schedule calendar.
   - **Run Job Immediately**—Runs the job when you click Save and continues to run the job after that according to the schedule calendar.
   - **Run Job At Specific Time**—Runs the job on the date and time you specify and, after that, according to the schedule calendar.

     Enter the next date and time to run the job. You can select the run date and time by clicking the calendar and clock icons adjacent to the Date and Time fields.

Suspending All Scheduled Jobs

To temporarily suspend all scheduled jobs in Assure DQ, complete the following steps:

1. Log in to the Assure DQ user interface as a Superuser or Administrator.
2. Select System > Scheduler.

   The system displays the Manage Scheduler window. The top pane lists all control points and utilities (labeled *jobs*) scheduled to run one or more times in the future.
3. To suspend all jobs on the list, click Disable Scheduler.
4. To delete a job scheduled to run only once, select the job from the Job list and click Remove.

   The Remove button is only enabled for one-time jobs.
5. To change the run time or enabled/disabled status of any job on the list while the scheduler is disabled, select the job you want to change and edit its properties in the bottom pane of the window.

   See "Managing Scheduled Jobs" on page 58 for more information on changing jobs.
6. To restart the scheduler, click Enable Scheduler.

   One-time jobs scheduled to run while the scheduler was disabled will run immediately in scheduled sequence, up to 10 jobs at a time. All other jobs will run at the next scheduled time.

Managing Calendars

Schedule calendars specify when to execute control points. To exclude system down times from a single schedule calendar, edit the calendar by adding an Exclude Duration Event. To exclude days or time spans from multiple schedule calendars, create a duration calendar (see page 61).

Editing Individual Schedule Calendars

To add an Exclude Duration Event to a schedule calendar, complete the following steps:

1. Log in to the Assure DQ user interface as an administrator, integrity rule administrator, or superuser.
2. Select Rules > Calendars.

   The system displays the Manage Calendars window. The top pane lists all existing calendars.
3. Select the calendar you want to change from the calendar list.
4. In the bottom pane, click Edit Events to exclude system down times from the calendar.

   The system displays the Events for Schedule Calendar window.
5. Click New Event and select Exclude Duration Event.

   An Exclude Duration Event can exclude days, dates, or time spans from the calendar.
6. Specify the following information for the date or time span you want to exclude from the schedule:

   Event Name—Choose a name for the event you are defining, such as System Maintenance.

   Date—To exclude a particular date, click the calendar icon and select the date of your choice.

   Start Time—Specify a time for the event to begin. Click the clock/calendar icon to select the time in hours, minutes, and seconds.

   End Time—Specify how long the event should continue. Click the clock/calendar icon to select the time in hours, minutes, and seconds. To span two days, create two events with the first event ending at 23:59:59 and the second event starting at 00:00:00 the next day.

   Recurring—To specify a reoccurring (vs. one time) event, choose Days or Monthly. Events that reoccur will begin on the date and start time you specify and then repeat on the same days in subsequent weeks or months.

   Days—To specify a Recurring Day event, select the days of the week to include or exclude.
7. Click Apply and then OK.

Creating Duration Calendars

If your organization uses schedule calendars to execute multiple control points, you can use duration calendars to exclude system down times from the schedules. Multiple schedule calendars can share the same duration calendar.

To create and apply duration calendars to schedule calendars, complete the following steps:

1. Log in to the Assure DQ user interface as an administrator, integrity rule administrator or superuser.
2. Select Rules > Calendars.

   The system displays the Manage Calendars window. The top pane lists all existing calendars.
3. Click the New Calendar button and select Duration Calendar.
4. Specify a name for the calendar.
5. Click Edit Events.

   The system displays the Events for Duration Calendar window.
6. Click New Event and select Duration Event.

   Duration Events can exclude days, dates, or time spans from the calendar.
7. Specify the following information for the event:

   Event Name—Choose a name for the event you are defining, such as System Maintenance.

   Exclude—Select this option to exclude the days, dates, or time spans you specify.

   Date—To exclude a particular date, click the calendar icon and select the date of your choice.

   Start Time—Specify a time for the event to begin. Click the clock/calendar icon to select the time in hours, minutes, and seconds.

   End Time—Specify how long the event should continue. Click the clock/calendar icon to select the time in hours, minutes, and seconds. To span two days, create two events with the first event ending at 23:59:59 and the second event starting at 00:00:00 the next day.

Recurring—To specify a reoccurring (vs. one time) event, choose Days or Monthly. Events that reoccur will begin on the date and start time you specify and then repeat on the same days in subsequent weeks or months.

Days—To specify a Recurring Day event, select the days of the week to include or exclude.

8. Click Apply and then OK.

To change a duration calendar, select it from the calendar list. Click Edit Events to change dates or time spans on the calendar.

Applying Duration Calendars

To exclude a duration calendar from a schedule calendar, complete the following steps:

1. Select Rules > Calendars.

   The system displays the Manage Calendars window. The top pane lists all existing calendars.
2. Select the schedule calendar where you want to exclude a duration calendar.
3. In the bottom pane, use the Exclude From Calendar list to select the duration calendar to exclude from the schedule calendar.
4. Click Save.
5. Repeat the steps above to exclude the same duration calendar from other schedule calendars, as needed.

8

Troubleshooting

This chapter provides information on resolving common problems that may occur during processing. The chapter has the following sections:

- System Logs
- Troubleshooting Processing Errors

System Logs

By default, the system writes informational log messages to igx.log and debugging messages to igx-alarms.log. By default, both files are in <install_folder>/Infogix/logs/<config_name>/IA.

If you have difficulty finding the logs, the following properties files indicate the path and name of the log subdirectories for the client and server:

- **Server**—Search for the LOG_FILE_DIR properties in the appserver.properties file, located in <install_folder>/Infogix /config/<config_name>.
- **Client**—Search for the CLIENT_LOG_FILE_DIR property in the client.properties file in <infogix_client_install_folder> /Infogix_Client/config/<config_name>.

Troubleshooting Processing Errors

To troubleshoot system processing errors in the user interface, complete the following steps:

1. Log in to the Assure DQ user interface.

   You must have knowledge worker authority to access control run results. You must also have View Results permission for the control entity you want to view.
2. Select Results > By Control Point.

   A list of control runs within the last 24 hours appears in the top pane of the window.
3. To list results for a different time period and/or a specific control entity/control point, click the Filter button at the top right of the window.

4. Click any control run on the list to see a summary of the results it produced and to review individual results or reports.

   The table that follows describes common system processing results that you may see in the user interface. Most of the system processing errors below set a return code of 16. If your rules writers have set up custom return codes, the highest return code set during processing will appear instead. At the start of control processing, the system sets the return code to -1. If the system shuts down or an undiagnosed error occurs during processing, you may see a return code of -1.

System Processing Messages

Result	Default Severity	Default Return Code	Description
Assure DQ Cannot Process Capture Source	Fatal	16	Indicates the name and expected path/URL of an unavailable data source. This error halts control processing. The system does not store data in the database or apply controls.
Assure DQ Cannot Read Record From Capture Source	Error	16	Provides detailed information on the problem encountered, and the name and path/URL of the data source.
Assure DQ Control Point Execution Action Limit Exceeded	Warn	0	Indicates Assure DQ bypassed control point execution actions that exceeded a specified action limit. Typically, control point executions are triggered by specific control outcomes. View the result details to see the action limit and the number of executions that were bypassed.
Assure DQ Control Point Processing Error	Fatal	16	Indicates processing failed for a particular control entity or application and control point. This error halts control processing. The system does not store data in the database or apply controls.

Result	Default Severity	Default Return Code	Description
Assure DQ Duplicate Data Source	Fatal	16	Indicates the name, path/URL, size, last modified date, and checksum of an input to the current run that duplicates an input processed previously. This error halts control processing. The system does not store data in the database or apply controls.
Assure DQ Failed Value Limit Exceeded	Warn	8	Indicates Assure DQ did not record individual failed values that exceeded a specified limit. View the result details to see the failed data limit and the number of failed values that were not recorded.
Assure DQ Missing History Data Input	Error	12	Indicates a particular history input to a control point was not found. Missing input data may lead to unreliable validation results.