

Gentran Integration Suite™

Archiving and Purging

Version 4.3

Sterling Commerce
An IBM Company

© Copyright 2007 Sterling Commerce, Inc. All rights reserved.
Additional copyright information is located on the Gentran Integration Suite Documentation Library:
<http://www.sterlingcommerce.com/Documentation/GIS43/HomePage.htm>

Contents

Archiving and Purging Activities	4
Archiving Business Process Data	4
Archiving Considerations	6
Configuring Archive Schedules	6
Archiving Business Process Data Manually	9
Configuring and Viewing the Archive Log	9
Archive Reporting	10
Working with Archived Business Process Data	10
Restoring Archived Business Process Data	10
Viewing Restored Business Process Data	10
Removing Restored Business Process Data	11
PurgeAll Utility	11
Running the PurgeAll Utility	12
Archiving and Purging Troubleshooting	13
Index	14

Archiving and Purging Activities

Archiving your business process data not only protects your critical data but also conserves database disk space and (when using documents on disk) file system disk space. This improves the efficiency of Gentran Integration Suite.

Note: Archive/purge is not the single point of cleanup within Gentran Integration Suite. Archive and purge mainly clean up business process data. Other services have special requirements or considerations, which are discussed in the documentation for these services and not in the archive/purge documentation.

For example, to purge mailbox messages, you must use the Mailbox Delete service, which deletes messages in one or more mailboxes. Since archive and purge have no knowledge of the data or the the data relationships on which they operate, archive does not try to understand what a mailbox is or what is in mailbox data.

Archiving Business Process Data

Gentran Integration Suite uses a set of components to move business process data that has reached its specified lifespan out of the live database. These components include:

- ◆ Index Business Process service
- ◆ Backup Business Process service
- ◆ Purge service

Data that has been removed using the Backup Business Process component can later be restored to the same version of Gentran Integration Suite using the Restore Business Process component.

Note: The level of detail written to the database as a business process runs is affected by the persistence level, which you specify with the Business Process Manager or in the `noapp.properties` file. Decreasing the persistence level increases the business process performance at the cost of full-tracking for each step of the business process.

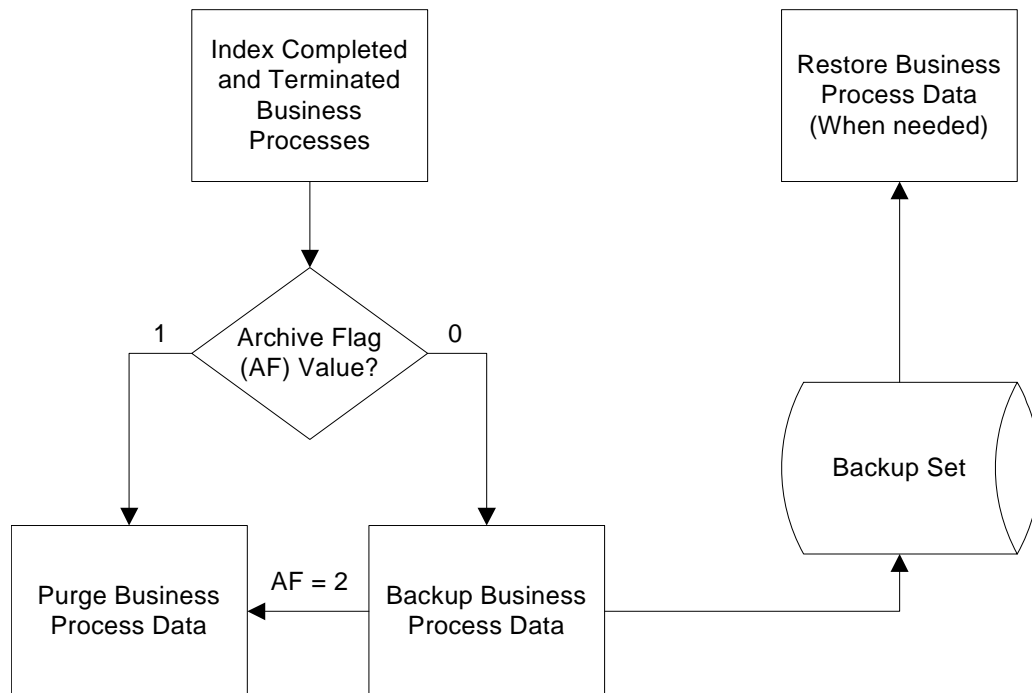
Business processes are eligible for archiving as soon as their archive flag is set by the Index Business Process service. You do not have to wait until a business process exceeds its lifespan before archiving it. After the archive flag is set, the business process is then archived the next time the Backup Business Process service is run either by schedule (at 2 A.M. each morning) or manually. The data is still available in the system because it cannot be purged (with the Purge service) until it has exceeded its lifespan.

The Restore Business Process service restores archived business process data from physical media to a Gentran Integration Suite restored data location where it can be searched and viewed.

After running a business process, when a business process terminates, and before archiving or purging business process data, the business process must first be *indexed*. That is, the Index Business Process service runs a process that looks for archiving parameters defined in the business process. When the process finds these parameters, the process creates summary information and writes the information to a table in the database.

For information written to the table, the Backup Business Process service runs, according to how you have configured your archiving settings. The Backup Business Process service retains the business process and its related data to a file system as Java™ serialized objects for later backup to tape, some other long-term storage medium, or purging. The Purge service then runs and removes any data from the database or file system, or both, specified in archive settings not to retain.

The following diagram shows the archiving process.



The Archive Flag is initially set by the Index Business Process service to 0 or 1. A value of 0 causes the data to be backed up. A value of 1 causes the data to be purged without being backed up. A data backup sets the flag to 2, which causes data to be purged.

Note: You cannot archive active business process instances (for example, halting, halted or interrupted). Only data for completed or terminated business processes that have been indexed can be archived.

Note: You cannot delete a workflow definition of a business process (using the Source Manager of the Business Process Manager) while an instance of that process is being either indexed, archived or purged. If you try to delete such a definition, an error message appears.

In the same manner that you schedule business processes, you can schedule the intervals at which you want Gentran Integration Suite to index, archive, and purge the contents of the database.

You can also define the lifespan, or the length of time, in days and hours, to retain the data in Gentran Integration Suite live tables. This is the length of time until the business process instance expires.

Archiving Considerations

You can configure Gentran Integration Suite to automatically archive business process data as often as you need, or you can archive business process data manually. For more information about changing the default settings in your archive schedule, see *Configuring Archive Schedules* on page 6. For more information about running archiving manually, see *Archiving Business Process Data Manually* on page 9.

Caution: You should not create your own service that backs up data (either manually or on a schedule). Also you should not create your own service and then manually archive data at the same time that an archive process is automatically running (from a schedule in the Archive Manager). This could corrupt archived information because each process could be trying to archive the same data. Use the lock feature in the Archive Manager to prevent simultaneous manual and automatic archiving when using standard backup tools.

When archiving your business process data, consider that:

- ◆ You should indicate default values that apply to all business process instances that you run.
- ◆ How often you archive should depend on the volume of business processes that you run. Archiving can be a database-intensive operation. If possible, avoid archiving during peak times.
- ◆ Business level information should be archived for audit purposes, including:
 - ◆ Documents that support non-repudiation for AS1 and AS2 protocols
 - ◆ Correlations that provide searching ability
- ◆ For any data that you may potentially want to view in the immediate future, maintaining the data in live tables provides faster retrieval.
- ◆ You do not need a separate JVM to back up data, which allows you to allocate more resources to Gentran Integration Suite.
- ◆ Gentran Integration Suite can archive streamed-in documents that are much larger than the amount of available memory, because documents are not re-assembled in memory prior to writing them out to disk. Instead, the backup process reads and writes raw DATA_TABLE pages.
- ◆ In a clustered installation, run the backup process on only one node at a time.

Configuring Archive Schedules

When configuring default archive schedule settings, you can define where Gentran Integration Suite should archive data for business process instances. After you configure and save the archive schedule, the archive process is automatic and runs based on the parameters you provided.

You can still archive manually; however, you should consider the scheduled archive before running the archiving activities manually.

For more information about running archiving manually, see *Archiving Business Process Data Manually* on page 9.

To configure archive settings (either the default settings or your manual settings):

1. From the **Administration** menu, select **Operations > Archive Manager**.
2. In the Archive Configuration section, next to Configure Archive Settings, click **Go!**
3. Click **OK** to enable the lock for the Archive Manager and prevent others from updating archiving settings at the same time that you are updating them. The Configure page appears.
4. Type numeric values in the first pair of **Days** and **Hours** fields to indicate how long business process data should remain active in Gentran Integration Suite. This is the lifespan of the process.
5. Type numeric values in the second pair of **Days** and **Hours** fields to indicate the additional time that trackable business process data (such as certificates that contain security information and documents that support non-repudiation) should remain active in Gentran Integration Suite.

This trackable data includes archive/purge-related tables in a business process. These tables are retained longer than tables in the same process that are deleted using the lifespan of the process that was specified in the first pair of **Days** and **Hours** fields. These tables of trackable data help with data searches.

For example, if you enter two days at step 4 and one day at step 5, then trackable business process data is active for three days (2 + 1).

6. In the **Backup Directory** field, type the complete path to the directory where you want to store the archived data.
7. Indicate whether to archive or purge expired business processes by selecting either **Archived** or **Purged**.
8. Click **Next**. The Business Process Definitions page appears. The left-hand window is an alphabetical list of business processes that do not have the lifespan set as the system default. The right-hand window is initially automatically populated with business processes that have the lifespan set as the system default.
9. To apply your settings to a business process, move it to the right-hand window, using the arrow keys. To not apply your settings to a business process, move it to the left-hand window.

To filter the unselected list on the left-hand side, use the **by Business Process Name** field.

Note: The lifespan setting from the Configure page applies to all business processes that are set with a lifespan of system default. Any processes in the right-hand window will be set to the system default lifespan, and will pull the lifespan setting at indexing that is set on the Configure page. If you change the removal method (archive or purge), business processes that are not moved to the right-hand window retain their previous removal method.

10. Click **Next**. Click **OK** at the prompt about the updating of business processes. The Schedule Type page appears.
11. Indicate whether your schedule uses a 24 hour clock display (that is, hour numbers from 1 to 24, rather than cycling through 12 hours twice).

12. Choose how to run the archive/purge process by completing one of the following procedures:
 - ◆ To archive or purge business process data every several hours or minutes, select **Run based on timer**.
 - ◆ To archive or purge the business process data on a daily basis, select **Run daily**.
 - ◆ To archive or purge business process data on specific days of a week, select **Run based on day(s) of the week**.
 - ◆ To archive or purge business process data on specific days of a month, select **Run based on day(s) of the month**.
13. Click **Next**. The Schedule Settings page appears. The fields on this page depend on the schedule selection that you just made on the Schedule Type page.
14. Choose how to run the archive/purge process by completing one of the following procedures.

For all schedule types except **Run based on timer**, you also can schedule activity on a time interval (that is, at regularly spaced times during a specified time period). To do this, select **Check here to select time interval** and specify when to start and end the time period, and the interval (that is, how often the process runs during the time period).

 - ◆ If you selected **Run based on timer** on the Schedule Type page, enter numeric data in the **Hour(s)** and **Min(s)** to show how often the process should run.
 - ◆ If you selected **Run daily** on the Schedule Type page, enter when you want to run the process each day. Click **add** to specify an entry. You can enter more than one item. Click **delete** to remove an entry.
 - ◆ If you selected **Run based on day(s) of the week** on the Schedule Type page, enter when you want to run the process each day. For each time that you pick, also enter the day when you want to run the process. Click **add** to specify an entry. You can enter more than one item. Click **delete** to remove an entry.
 - ◆ If you selected **Run based on day(s) of the month** on the Schedule Type page, enter the time when you want to run the process each month. For each time that you pick, also enter the day of the month when you want to run the process. Enter **LDOM** to choose the last day of the month. Click **add** to specify an entry. You can enter more than one item. Click **delete** to remove an entry.
15. To have the scheduled job run when Gentran Integration Suite is started (as well as on the specified schedule), select **At startup**. This option is not available if you selected **Run based on day(s) of the month** on the Schedule Type page.
16. Click **Next**. The Schedule Exclusions and/or the Date Exclusions page appears, depending on your run selection on the Schedule Type page. When you are finished with the exclusion page(s), the Confirm page appears.
17. To indicate exclusion dates that archiving should not occur, complete the following steps:
 - a. Enter the time(s) or date(s) when not to run archive.
 - b. Click **add** to specify the exclusion time(s) or date(s). Click **delete** to remove time(s) or date(s).
 - c. Click **Next**.
18. Check your settings on the Confirm page and click **Finish** to update your archiving or purging schedule in Gentran Integration Suite.

Archiving Business Process Data Manually

After you have configured archive settings, you can archive completed or terminated business process instances at any time.

To archive business process instances:

1. From the **Administration** menu, select **Operations > Archive Manager**.
2. In the Archive Execution section, click **Go!** to run the archiving process.
3. When prompted, click **OK** to execute the archive process.

Configuring and Viewing the Archive Log

Each day, Gentran Integration Suite compiles an archive log of business process data archived during the day. The archive log provides details about business process instances, including errors encountered during archiving.

To configure the archive log settings:

1. From the **Administration** menu, select **Operations > System > Logs**.
2. In the Application Logs section for your specific Web application environment, click the **Archive** icon.
3. In the **Archive: Log Settings** page, complete the following steps as necessary:
 - a. Type a new location to store the archive log.
 - b. Select the maximum number of lines the log can contain before it rolls over to a new log.
 - c. Indicate whether to run the logging level.
 - d. Click **Save** to return to the Gentran Integration Suite interface.

The following table describes the two ways you can view the archive.log file:

Method	Procedure
Archive Manager page	To view the archive log using the Archive Manager: <ol style="list-style-type: none"> 1. From the Administration menu, select Operations > Archive Manager. 2. In the Archive Log section, next to View Archive Log, click Go!
System Logs page	To view the archive log using the system logs: <ol style="list-style-type: none"> 1. From the Administration menu, select Operations > System > Logs. 2. In the Application Logs section for your specific Web application environment, click the link archive.log to view the archive log.

Archive Reporting

The Archive Status report has two sections:

- ◆ The values of all of the configuration parameters, along with the results of the archive run and the name of the backup set directory.
- ◆ The results of the archive for each table, including the name of the table, the expected and actual row counts, and the elapsed time for each table. The tables are sorted by the actual table names. If no data is eligible for archiving, this second section does not contain any table information.

The status report does not show error and exception information, which is shown in the archive log.

Status reporting uses the standard Gentran Integration Suite application programming interfaces to capture stdout and stderr to report results. Reporting does not capture results from the Command Line Adapter 2.

Working with Archived Business Process Data

To view business process data that has been archived to an offline location, you must first restore that data within Gentran Integration Suite.

Restoring Archived Business Process Data

To restore archived business process data:

1. From the **Administration** menu, select **Operations > Archive Manager**.
2. In the Restore Manager section, next to Manage Restoring Business Processes, click **Go!**
3. In the **Restore Directory** field, type the complete path to the directory of the archived data that you want to restore.
4. Click **Go!** to restore the data in Gentran Integration Suite.

Viewing Restored Business Process Data

To view archived business process data restored in Gentran Integration Suite:

1. From the **Administration** menu, select **Business Processes > Advanced Search > Business Processes**.
2. Under Search Location, select **Restore Tables**.

3. On the Business Process Monitor Advanced Search (Restore) page, complete one of the following steps:

Click **Go!** to view all restored business process data.

or

Search for specific restored business process data. Indicate any combination of the following criteria and click **Go!**

- ◆ Process ID – Display data identified by the number assigned by Gentran Integration Suite when the business process ran.
- ◆ Business Processes – Display data for business processes with names that contain the specified character or string.
- ◆ System Business Processes – Display data for Gentran Integration Suite system business processes with names that contain the specified character or string.
- ◆ State – Display data for business processes with a final state (that is, All, Completed, and Terminated) that corresponds to the specified state.
- ◆ Status – Display data for business processes with a final status (that is, All, Success, Error) that corresponds to the specified status.
- ◆ Start date/time range – Display data for business processes run within the specified start dates and times.
- ◆ Deadline range – Display data for business processes run by a deadline, along with the status of the deadline (that is, All, Met, and Missed) and the start and end times of the deadline, that corresponds to the specified deadline.

Removing Restored Business Process Data

To remove business process data restored from archive tables:

1. From the **Administration** menu, select **Operations > Archive Manager**.
2. In the Restore Manager section, next to Manage Restoring Business Processes, click **Go!**
3. In the Cleanup section, next to Clean restore area, click **Go!**
4. When prompted whether to clean the restore area, click **OK** to remove the business process data from Gentran Integration Suite.

PurgeAll Utility

The command line tool provides you with the ability to systematically clear out all business process data from the application database. The business process data cleared by this utility is limited to truncate safe tables. The following tables are considered truncate safe:

- ◆ EDI_COMPLIANCE_RPT
- ◆ EDI_DOCUMENT_STATE

- ◆ WF_INST_S
- ◆ WORKFLOW_CONTEXT
- ◆ WORKFLOW_LIFESPAN
- ◆ WORKFLOW_DATA
- ◆ EDIINTDOC
- ◆ MSGMDNDUP
- ◆ MSGMDNCORRELATION
- ◆ WEBX_MINED_DATA
- ◆ GENERIC_ACK
- ◆ CORRELATION_SET
- ◆ ACT_SESSION
- ◆ ACT_AUTHENTICATE
- ◆ ACT_AUTHORIZE
- ◆ ACT_XFER
- ◆ ACT_NON_XFER
- ◆ DATA_FLOW
- ◆ DMR_ROUTE
- ◆ DMI_ROUTE_FACT

Note: The standard purge process is responsible for clearing the remaining table data.

Running the PurgeAll Utility

To run the PurgeAll utility:

1. Stop the application.

Note: If the application is running, you will get a warning message and the PurgeAll command process will abort.

2. Do one of the following:

- ◆ For Unix, enter `PurgeAll.sh -t -l`
- ◆ For Windows, enter `PurgeAll.cmd -t -l`

Where:

- ◆ `t` - Truncate eligible database tables to remove all rows (this will free up space in your database tables)
- ◆ `l` - List the names of the database tables that are going to be truncated

Note: After the purge is completed, any remaining business process table data will be purged through the standard process upon restart of the application.

Archiving and Purging Troubleshooting

If you have addressed the issues described in *Archiving Considerations* on page 6 and the archiving or purging process did not appear to work properly, check for the following issues:

- ◆ From the **Administration** menu, select **Business Processes > Central Search** and search for the `Schedule_BackupService` business process. Verify that this business process completed successfully. `Schedule_BackupService` must run for data to be removed from the live database tables.
- ◆ From the **Administration** menu, select **Deployment > Schedules** and verify that `IndexBusinessProcessService` is scheduled to run. The default time for this service to run is every ten minutes.
- ◆ From the **Administration** menu, select **Business Processes > Central Search** and search for the `Schedule_IndexBusinessProcess` business process. Verify that this business process completed successfully. `Schedule_IndexBusinessProcess` must run to set the expiration dates for completed business processes.
- ◆ From the **Administration** menu, select **Business Processes > Advanced Search > Business Processes**. In the Search Location area, select Live Tables and perform the following actions:
 - a. Select ALL – Business Processes, System Business Processes, State, and Status.
 - b. Set the Start Date and End Date to include business processes that should have been archived according to the default value set in the Archive Manager configuration.
 - c. Click **Info**, under the Expires column, to verify that the date reflects a future runtime.
 - d. If **Info** does not display or the date reflects an expired date, the business process has not been configured for archiving. Therefore, you must terminate the business process manually, configure archive settings, and then run the business process again.
- ◆ Check the archive log for errors.
- ◆ From the **Administration** menu, select **Business Processes > Central Search** and search for the `Schedule_BackupService` business process. Verify that this business process completed successfully.

A

archive log 9
Archive Manager 9
archiving
 Archive Manager 6
 configuring default settings 6
 considerations 6
 log settings 9
 offline 10
 removing restored data 11
 reporting 10
 restoring data 10
 scheduling 6
 troubleshooting 13
 viewing data 10

B

Backup Business Process service 4

C

configuring archive schedules 6

D

deadline setting 11
deadline, business process 11

I

Index Business Process service 4

P

Purge service 4
PurgeAll utility 11
purging
 data 6
 scheduling 6

troubleshooting 13

R

reporting 10
Restore Business Process service 4

S

service
 Backup Business Process 4
 Index Business Process 4
 Purge 4
 Restore Business Process 4

