IBM WebSphere Partner Gateway V6.2

Archiver

This presentation covers the new archiver and restore feature that is part of WebSphere® Partner Gateway V6.2.

IBM

Agenda

- Archiver overview

- Archiver configuration and administration

- Archiver reports and events

- Archiver restore overview

- Restored documents search

- Performance and recommendations

- Trouble shooting

- Summary and references

Archiver

© 2009 IBM Corporation

2

The goal for this presentation is to provide an overview of the enhanced archiver feature. The presentation also provides an overview of configuration and administration of the archiver and restore features. It also details the recommendations on getting a better performance from the archiver, some trouble shooting tips in case of any failures, and the functionality.

Section

Archiver overview

Archiver

The next section covers an overview of the archiver functionality.

Archiver overview

- WebSphere Partner Gateway archiver is a tool for archiving and purging data

- Data is archived/purged from
  - Common file system
  - Application database

- Single task for archive and purge actions
  - No manual intervention of running separate scripts

- GUI for archiver task configuration and reports

WebSphere Partner Gateway is a business-to-business gateway product, which enables enterprises and their trading partners to exchange business documents. The intermediate documents during processing are stored in the common file system. WebSphere Partner Gateway also stores the document activity and associated message details in several activity and detail tables. Some of the tables are common across, while some are specific to the packaging and protocols used by the documents.

WebSphere Partner Gateway also retains the non-repudiation data in a separate table and the associated documents in a specific folder in the common file system. Over a period of time, as the data grows in the common file system and the database, it is advisable to archive and purge old data. The WebSphere Partner Gateway archiver is a GUI tool and a single scheduled task that provides the mechanism to do just this.

Archiver overview

- New enterprise application
  - ▸ Deployed on the bcgserver cluster
- Uses scheduled based tasks for archiving
- Improved performance and reporting capabilities
- Easy configuration
  - ▸ All configuration changes done using the WebSphere Partner Gateway console

5

The archiver tool is a new enterprise application that is deployed on the WebSphere Partner Gateway server cluster; hence, the BPE and the document manager. Archiver is also now a part of the router.

In a multi-machine setup, the archiver itself will run on one of the servers. If a particular server is not running, the archiver will run on any available server.

Archiver uses the scheduler provided by WebSphere Application Server, which in itself is quite robust. This also provides the ability to re-run the archiver task when the server is restarted.

The new archiver functionality is better in terms of performance and reporting capabilities. It is a single task that does archiving and purging as compared to the earlier archiver feature, wherein you had to manually run different scripts on different machines to perform archiving and purging.

The configuration in itself is very easy and can be done from the WebSphere Partner Gateway console. All changes are effective immediately and no server restart is required.
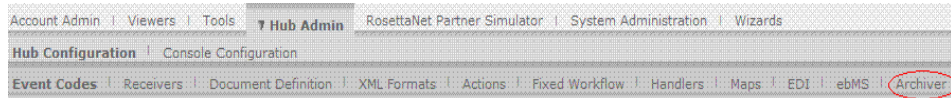
IBM

Section

Archiver configuration and administration

The next section covers the configuration and administration of the archiver.

The archiver tab is available on the WebSphere Partner Gateway console.

There are sub-tabs for the archiver configuration and archiver report as shown.

Archiver configuration and administration

- Archiver configuration
  - Four sections
    - General configuration
    - Purging
    - Purging schedule
    - Archive criteria
  - General configuration

Archiver Configuration

Task Name: ARCHIVER

Description: Archive and Purge Data From the System

Last Execution Time: 9/27/08 10:07:49 AM

Next Execution Time: 9/28/08 10:03:00 AM

Status: Scheduled [Run Now] [Disable]

The archiver configuration is segregated into sections such as general configuration, purging, schedule, and criteria.

The general configuration has the name, description, last execution time, next execution time, and the status fields. A user can also enable or disable the archiver or run the archiver on demand.

Archiver configuration and administration

- Purging and schedule
  - ▸ Data retention period (number of days data to be online)
  - ▸ Location where to store the archived data if data is required to be archived
  - ▸ When to run the archiver
    - Daily
    - Weekly
    - Custom

Purging

Data Retention Period: 15 days

Copy: ☑

File Destination: /tmp/dataBackup

Database Destination: /tmp/dbBackup

Purging Schedule

Daily Schedule : ◉   Weekly Schedule : ○   Custom Schedule: ○

Hour 10PM ▼   Minutes 30 ▼

The purging section has the configuration, which determines the parameters for the archiver runtime.

The data retention period is the number of days for which data needs to be retained in the system. All data beyond the retention period can be purged.

It also has the option if the non repudiation data needs to be "archived." In this case, the non repudiation data from the table is exported to the directory specified in the "Database Destination" field, and the non repudiation documents are copied to the directory specified the "File Destination" field.

In the schedule section, you can configure the archiver task run time. One of the three available schedule kinds can be used.

The daily schedule; where the time of day when the archiver must run is specified.

You can also have weekly schedule; where the time and days of week when the archiver must run is specified.

If you want to run the archiver task as required, then create a custom schedule, wherein you can also specify the days of month when the archiver task must run.

Archiver configuration and administration

- Partner, package, protocol, document type
  - Sending partner, package , protocol, and document type
  - Used for archiving data from the LG_MSG_ARCHIVE table and non-rep folder

Archive criteria

Archive purged documents that meet any of the following sets of criteria :

Page 1 of 1   Total Rows:

| Beginning | Back 10 Pages | Back 1 Page | 1 | Forward 1 Page | Forward 10 Pages | E |

	Sending Partner	Package	Protocol	Document Type
	All	All	All	All

A criteria can be specified for the archiver configuration, which includes the sending partner name, package, protocol, and the document type.

If such criteria is enabled, it is used for exporting data from the non repudiation table and copying the non repudiation documents from the non-rep folder.

Note that the data will be purged from all the tables. The criteria is only for archiving of non-rep data from the table and file system.

Archiver configuration and administration

- Partner, package, protocol, document type
  - ▸ Multiple partner, package, protocol, and document type criteria can be specified

The default criteria can be changed from the pop-up window and multiple partner, package, protocol, and document type criteria can be specified.

Archiver configuration and administration

- Archiver task status
  - ▶ Scheduled – indicates the archiver task is scheduled and will run at the configured time
  - ▶ Running - indicates the archiver task is currently running
  - ▶ Disabled – indicates that the archiver schedule is disabled and the task will not run
  - ▶ Disabled_Failed – indicates that the system has disabled the archiver schedule as the previous run had errors

12

This slide describes the different possible status for the archiver. The archiver task will have exactly one of the given status at any point.

The scheduled status indicates that the archiver task is scheduled and will run at the configured time.

The running status indicates that the archiver task is currently running.

The disabled status indicates that the archiver schedule is disabled and the task will not run.

The disabled_failed status indicates that the system has disabled the archiver schedule as the previous run had errors.

IBM

Archiver configuration and administration

- Archiver task status
  - ▶ Stopped – indicates that the task was stopped by the user
  - ▶ Successful – indicates that the task execution was successful
  - ▶ Failed – indicates that the task execution failed
  - ▶ Intermediate states
    - Stopping – indicates that the current task execution is stopping as requested
    - Being scheduled to run now – indicates that the task will run now on demand
    - Disabling – indicates that the task is being disabled as requested

13

Archiver   © 2009 IBM Corporation

The stopped status indicates that the task was stopped by the user.

The successful status indicates that the task execution was successful.

The failed status indicates that the task execution failed.

Apart from the above mentioned status, there are some other intermediate states as well.

Stopping status is displayed when the current task execution is stopping as requested.

Being scheduled to run now status is displayed when the task is about to run on demand.

Disabling status is displayed when the task is being disabled as requested.

Archiver configuration and administration

- Changing the archiver configuration
  - Configuration is updated by the server
  - Intermediate state can be seen
  - Click the tab to see the updated configuration

  Task Name: ARCHIVER
  Description: Archive and Purge Data From the System
  Last Execution Time: 9/27/08 10:07:49 AM
  Next Execution Time: 9/29/08 10:30:00 PM
  Status: Disabling [Enable]
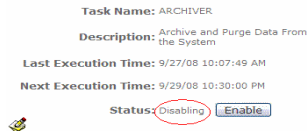
  - If copy is de-selected, on a save it will prompt:

  The page at http://9.184.251.51:58080 says:
  Non-repudiation data will not be backed up to the archive.It will be purged and will not be available for later use.Do you wish to proceed?
  [OK] [Cancel]

The archiver configuration can be updated from the "archiver configuration" tab. When the archiver configuration is changed, the server is notified of the change. You might see an intermediate state in the task field as the server updates the configuration. You need not restart the server after changing the configuration.

Note that you should not refresh the screen to check the updated status. You must click the "archiver configuration" tab to view the changed status.

If the copy option is de-selected, no data is archived. Data will only be purged from the system. The user is alerted by the prompt when a save is attempted with the copy option de-selected.

Another administrative task that can be performed from the archiver configuration screen is stopping the running task from the screen.

You will be prompted to confirm if you want to stop the running task.

Archiver configuration and administration

- Running a scheduled task on demand
  - ▸ A scheduled task can be run anytime from the console
  - ▸ Intermediate state "being scheduled to run now" is displayed
  - ▸ The task subsequently changes to running state
  - ▸ The last execution time reflects that the task ran before the schedule
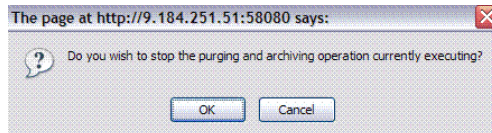
Archiver Configuration

Task Name: ARCHIVER

Description: Archive and Purge Data From the System

Last Execution Time: 9/27/08 12:54:09 PM

Next Execution Time: 9/29/08 10:30:00 PM

Status: Being Scheduled to Run Now

Archiver Configuration

Task Name: ARCHIVER

Description: Archive and Purge Data From the System

Last Execution Time: 9/27/08 12:58:55 PM

Next Execution Time: 9/29/08 10:30:00 PM

Status: Running [Stop]

A scheduled archiver task can be run outside of a schedule as well. This "on demand" execution is done by clicking "run now."

The task will subsequently start running and the last execution time is updated to reflect the same.

Section

Archiver reports and events

Archiver

The next section covers the archiver reports that are available. It also covers the different events that are generated during the archiver runtime.

IBM

Archiver reports and events

- Archiver report for the data archived and purged by a task

- Paginated report list is available

- Shows start time, end time, and the status

- Detailed report available for each task that ran

- Three new events are available for the archiver

Archiver © 2009 IBM Corporation

Each task that runs generates a report. A paginated set of report listing is available on the "archiver report" page. This list shows the start time, end time, and status for each listing. Further, three new events have been introduced specifically for different archiver runtime statuses.

This slide displays the archiver task listing. Detailed report is also available for each task that ran, which can be navigated to from this page.

Archiver reports and events

- Archiver detail report page
  - ▸ Start time, end time and status
  - ▸ Grouped into
    - Message archive table
    - Activity logging tables
    - State engine tables
    - Summary tables
  - ▸ Shows the purge start and purge end time for each table
  - ▸ Shows the number of rows that were purged for each table
  - ▸ Shows the settings that were used for this archiver task

20

The "archiver report details" page provides details on the archiver task activity. It displays the start time, end time, and the status of the task.

With regard to purging, the purged tables are grouped under the sections message archive table, activity logging table, state engine table, and summary table. Each section has details of the tables from which the data was purged. It shows the number of rows that were purged and the time taken for purging the table. The details also show the archiver criteria that were used while running the archiver task.

Archiver reports and events

- Archiver report details page

This slide displays a snapshot of the archiver report details page.

Archiver reports and events

- Three new events are available for the archiver
  - ▶ When the archiver task runs successfully
    - BCG700001 archiver task successful – informational event
  - ▶ When the archiver task runs with failures
    - BCG700002 archiver task error – error event
  - ▶ When the archiver task is stopped by a user
    - BCG700003 archiver task stopped by user – warning event
- All the three events can be made alert-able and alerts can be configured for each

22

Three new events have been introduced for the archiver.

The first event is BCG700001 archiver task successful, which is an informational event.

The second event is BCG700002 archiver task error, which is an error event.

The third event is BCG700003 archiver task stopped by user, which is a warning event.

It is possible to make the events as alert-able and generate alerts for any of them.

Section

Archiver restore overview

The next section covers the restore feature of the archiver.

Archiver restore overview

- Archived documents can be restored

- Data from the archived LG_MSG_ARCHIVE table is restored

- Non-repudiation folder data is restored

- Hub admin -> hub configuration - > Archiver -> Archiver restore

| Account Admin | Viewers | Tools | ▼ Hub Admin | RosettaNet Partner Simulator | System Administration | Wizards | Logout |

Hub Configuration | Console Configuration

Event Codes | Receivers | Document Definition | XML Formats | Actions | Fixed Workflow | Handlers | Maps | EDI | ebMS | **Archiver**

Archiver Configuration | Archiver Report | **Archiver Restore**

A console based restore feature is provided such that the archived data can be restored from the console.

The archived data from LG_MSG_ARCHIVE table and the non-repudiation documents that were copied to the destination folder is restored.

The restore feature is available under the "archiver restore" tab on the console as shown in the screen capture.

Archiver restore overview

- Table data is restored in an internal independent table

- Non rep documents are restored in the common file system under a different folder

- Data is restored between two dates

- Option of appending to the existing restored data

Archiver Restore

Welcome, Hub Adminis

・Search Restored Documents

Start Date	12/1/08 *	End Date 12/5/08 *
	Example: 12/31/08	Example: 12/31/08
NonRep File Archive Location:	/tmp/dataBackup *	
NonRep Database Archive Location:	/tmp/dbBackup *	
Append:	☐ **	
Status:		

[Restore] [Refresh]

The table data is restored in an internal table and not in the runtime LG_MSG_ARCHIVE. The non repudiation documents too are restored in a different folder in the common file system.

You can select a date range between which the data has to be restored. You also have the option of appending the data to the existing restored data as well.

Section

Restored documents search

26

The next section covers the search feature for the restored documents.

Restored documents search

- Restored documents can be searched

- Link is available on the restore page

- Documents can be searched based on date and sending partner, package and so on

Search Restored Documents

Start Date	12/1/08	* **End Date**	12/5/08 *
	Example: 12/31/08		Example: 12/31/08
Source Partner	All		
Package	All		
Protocol	All		
Document Type	All		
esults Per Page	10		

Search   Reset

27

Archiver   © 2009 IBM Corporation

Once the archived documents have been restored, you can search through these documents as well. The search feature is available on the archiver restore page itself.

You can search for a document by using the search criteria like date, partner, package, protocol, and document type.

Restored documents search

- The restored documents viewer shows the list of restored documents

- Sending partner and document information is shown

- Option of viewing the raw document is also available

- Decrypted document content is shown

Restored Documents Viewer   Welcome, Hub Ad

· Search Restored Docume

Page 1 of 21   Tol

| | | | 1 2 3 4 5 6 7 8 9 10 | Forward 1 Page | Forward 10 Pages |

View raw document	Sending Partner	Package	Protocol	Document Type	SubVolumeName	Certificate Name
	Internal	None (N/A)	customxml (1.0)	customtype (1.0)	20081204/0	
	Internal	None (N/A)	customxml (1.0)	customtype (1.0)	20081204/0	

28

Archiver   © 2009 IBM Corporation

The restored document viewer shows the list of the restored documents as seen in the image. You can view the raw document as well.

In case the document was encrypted, the decrypted document is shown in the document viewer.

Section

Performance and recommendations

The next section covers the performance aspects for the archiver.

Performance and recommendations

- The new archiver tool has a much improved performance

- Performance depends on database processor availability and disk speed

- Recommendation is to run the archiver when no document processing is happening or very low peak times

- Database statistics need to be current before the archiver runs. Recommendation is to run the "runstats" for the database before the archiver execution starts

- Consider using the archiver tuning parameters

Archiver

The new console based archiver tool has a much improved performance. The performance of the archiver depends on database and server processors availability and disk speeds. The archiver uses the database processor for purging data from the different tables and therefore you should run the archiver when no document processing is happening or very low document processing is happening.

The archiver can potentially delete a large number of rows from the different tables. It is very important that the database statistics are current before the archiver runs. This is an important step. You should make sure that "runstats" is run for the database before archiver execution starts.

Performance and recommendations

- The following tuning parameters are available for the archiver
  - bcg.archive.db.rowsPerBatch – batch of data used by temporary tables during purging. Larger batch size helps reducing the overheads of temporary tables
  - bcg.archive.db.rowsPerStatement – delete chunk size used for purging from the batch. Larger delete can cause slow purging
  - bcg.archive.db.parallelism– parallelism for deleting data from tables. Larger parallelism helps improve db server utilization. However too many threads can thrash the DB server. If the number of DB cores is 'x', then the recommended value is around 'x/2'
  - bcg.archive.activityParallelism - copy/purging parallelism. The maximum value is 4
  - bcg.archive.maxThreads – number of threads allocated to copy/delete for the common file system data
- The default values have been found optimal on an eight core database system in the lab. It is recommended to test with other values before arriving at the optimum

31

This slide gives the different tuning parameters that are available for the archiver.

The bcg.archive.db.rowsPerBatch parameter is the batch of data used by temporary tables during purging. Larger batch size helps in reducing the overheads of temporary tables.

The bcg.archive.db.rowsPerStatement parameter deletes chunk size used for purging from the batch. Larger delete can cause slow purging.

The bcg.archive.db.parallelism parameter is the parallelism for deleting data from tables. Larger parallelism helps improve db server utilization. However too many threads can thrash the DB Server. If the number of DB cores is 'x', then the recommended value is around 'x/2'.

The bcg.archive.activityParallelism parameter is the copy/purging parallelism. The maximum value is four.

The bcg.archive.maxThreads parameter denotes the number of threads allocated to copy/delete for the common file system data.

During tests, the default values have been found to be optimal for an eight core DB system. The recommendation is to test with other values before arriving at the optimal values.

Section

Troubleshooting

32

This section covers some trouble shooting tips if in case an issue is seen during running the archiver task.

IBM

Troubleshooting

- The archiver task can fail during runtime in different scenarios
  - ‣ The destination to copy non-rep files and table data is not available either because of non-existence, no permissions, no disk space, and so on
    - In these scenarios the archiver task will fail and the status of the configuration will show "Disabled_Failed"
    - This indicates that the archiver task has failed and all further schedules have been disabled
  - ‣ Correct the situation by checking the permissions, disk space for the destination folders

Archiver   © 2009 IBM Corporation

Ensure that the destination directory where the non-rep data is exported from the LG_MSG_ARCHIVE table is accessible by the database machine and has enough disk space. Also, for DB2®, the fenced user must have all the permissions, and for Oracle, the user must have permissions to write to this directory.

For the destination directory where the non-rep documents will be copied to, ensure that it is accessible by the server and has enough disk space. The user must have write permissions to be able to copy documents to this directory.

The archiver task will be marked as "disabled_failed" if it detects a condition where it fails permanently.

IBM

Troubleshooting

- Database failures
  - ▸ The archiver purging can create an SQLException with error code as -911 on DB2
  - ▸ This is not an unrecoverable exception, but affects the performance of the archiver task
  - ▸ Make sure that the database index statistics are current by running the "runstats" on the database

Archiver

34

© 2009 IBM Corporation

If -911 deadlock error is seen on DB2, ensure that database index statistics are current by running the "runstats."

Troubleshooting

- Server crash – single machine
  - ▸ On a single machine setup, if the server stops while the archiver task is running
    - The task does not complete and shows as running on the console
  - ▸ When the server is brought up again
    - The archiver task resumes. In this scenario, a new task is shown in the reports
    - If you stop the task while export of data is going on, there might be further errors when the task resumes. In this scenario, backup the existing exported files, stop the task and run it again manually

Archiver   © 2009 IBM Corporation

If the server stops while the archiver task is running, the task will not complete and the status is not updated. It will still be shown as running. The task will resume when the server is brought up again.

Troubleshooting

- Server crash – multi machine
  - One of the servers will run the archiver task
    - Server stops before the archiver task runs
      - Other available server will automatically run the task on the next schedule
    - Server stops while running the archiver task
      - Other available server will try to run the task
    - If the export of data is not complete at this time, there can be errors and the task does not complete
      - Stop this task from the WebSphere partner gateway console

In a clustered environment with more than one node, if the server stops before the archiver task has run, the other available server will automatically run the task. If the server stops when the task is running, the other available server will try to run the task.

Troubleshooting

- Log files
  - ▸ The archiver activities are logged in the bcg_server.log
    - The archiver configuration that is used is logged
      - – The archiver configuration uses the server's date and time when calculating the cutoff date
    - When the actual archiver task executes it logs
      - – "Archiver execution is starting"
      - – "Archiver execution is done. Time taken" *<seconds>*
      - – All archiver task execution is logged between the above two
    - The common file system activity and database activities are logged as well

The log files will record any errors that are encountered while the archiver task runs. The log files also record the time when the archiver task "started" and the time when the archiver task "ended." It also records the time taken to run the archiver task.

Section

Summary and references

38

This section covers the summary of the archiver feature discussed.

Summary and references

- WebSphere partner gateway V6.2 has a new improved archiver tool

- Archiver is a scheduled process

- Archiver configuration can be configured from the console. Reports are available from the console

- Archiver performance is dependent on the database processor availability and the database statistics being current

- Schedule the Archiver at idle server times or at low peak times

- Archived data can be restored and restored documents can be searched

Archiver

39

© 2009 IBM Corporation

WebSphere Partner Gateway V6.2 has a new improved archiver tool in terms of the performance and usability. The archiver runs as a scheduled task and the archiver configuration can be easily done from the console. Archiver reports are available on the console.

The archiver performance is dependent on the database processor availability and the database statistics being current. Ensure that "runstats" is done for the database before the archiver has run. Schedule the archiver at idle server times or at low peak times.

The new archiver restore feature allows the archived data to be restored and restored documents to be searched.

Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?
- Did it help you solve a problem or answer a question?
- Do you have suggestions for improvements?

Click to send e-mail feedback:

mailto:iea@us.ibm.com?subject=Feedback_about_WPG62Archiver.ppt

This module is also available in PDF format at: ../WPG62Archiver.pdf

You can help improve the quality of IBM Education Assistant content by providing feedback.

Trademarks, copyrights, and disclaimers