Redpaper

Tracy Schramm

IBM i Reorganize Physical File Member

This IBM® Redpaper™ publication describes the IBM i Reorganize Physical File Member (RGZPFM) command and some of the recent enhancements to this command. Enhancements to the RGZPFM command continue to be made to meet client needs.

Reorganize Physical File Member (RGZPFM) command

The Reorganize Physical File Member (`RGZPFM`) command removes deleted records from (compresses) one member of a physical file in the database, and optionally reorganizes that member.

To see the evolution of this command, here is some background on it. There are two typical reorganize styles:

► Allow Cancel (`ALWCANCEL`) *NO, which is the original, "old" style reorganization that makes a copy of the records.

► Allow Cancel (`ALWCANCEL`) *YES, which moves the records around in place.

Along with the `ALWCANCEL` parameter, a few other parameters are relevant to this situation:

► Lock State (`LOCK`)

   Before IBM i 6.1, a lock was required at either the beginning or the end of the file to return the space. In IBM i 6.1 and later releases, this lock is no longer required.

   When using a `*SHRUPD` lock, the production jobs that are inserted in to the file affects the `RGZPFM` job and its ability to return space. After the reorganize job runs through the file one time, it looks to see whether any additional rows are inserted at the end of the file since the reorganization started. The job notes what rows must be moved and tries to move them to get as much space returned as possible. If a row is inserted at the end of the file after the last rows that must be moved were noted, the system cannot give back as much space. More activity on the file (inserts moved) can prevent some space from being returned.

Files set up with Reuse Deleted Records *NO requires all inserts to be placed at the end of the file, which increases the chance of limited space being returned. Files that are set up with Reuse Deleted Records *YES have a greater chance of production jobs inserting rows in the middle of the file and not impacting the RGZPFM job. However, there is still a chance that rows can be inserted at the end and prevent space being returned even with Reuse Deleted Records enabled.

For this reason, you might need to run **RGZPFM** several times to see all that available space that is returned. In some cases, a maintenance window is needed to run **RGZPFM** with no inserts happening. If so, let the reorganization run during the week and move as much as it can. You should perform this work as close to the maintenance window as you can and, if it ends earlier, start another one. Then, during the downtime, the reorganization has less work to do, and downtime should be decreased.

► Rebuild Access Paths (**RBDACCPTH**)

If **ALWCANCEL *YES** is specified, and **RBDACCPTH *NO** is specified, the access paths are maintained while **RGZPFM** is running. Reorganizing in the presence of LIFO, FIFO, or DATEFO indexes changes the order of duplicates unless that index is used as the **KEYFILE**.

► Key File (**KEYFILE**)

The **KEYFILE** specifies whether the physical file member has its arrival sequence changed to match its keyed sequence, is reorganized in the sequence of a logical file member, or is not reorganized (and is only compressed).

The following example illustrates these parameters more clearly.

For the initial setup purposes of this example, run the following commands in STRSQL:

```
create table tschram.t1 (col1 int primary key)
insert into tschramm.t1 values(9),(1),(8),(11),(7),(12),(6),(13),(5),(4),(3),(2)
```

Then, delete the following rows in preparation for doing RGZPFM by running the following commands:

```
delete from tschramm.t1 where col1 = 11
delete from tschramm.t1 where col1 = 12
delete from tschramm.t1 where col1 = 13
```

The data in the file before you run the **RGZPFM** command is shown in Example 1.

Example 1   Initial data in the file

```
....+....1....
        COL1
           9
           1
           8
           7
           6
           5
           4
           3
           2
********   End of data   ********
```

With the Key File parameter of Replace Deleted Records (**KEYFILE (*RPLDLTRCD)**) specified, the member is reorganized by replacing deleted records at the start of the file with valid records from the end of the file.

> **Tip:** If the rows must exactly match the current arrival sequence, do not use ***RPLDLTRCD**.

After the rows are moved by using the **RGZPFM** command, Example 2 shows how the data looks according to ***RPLDTRCD** processing.

Example 2   Organization of data after RGZPFM KEYFILE (*RPLDLTRCD)

```
....+....1....
        COL1
           9
           1
           8
           2
           7
           3
           6
           4
           5
********  End of data   ****
```

With the Key File parameter of File (**KEYFILE (*FILE)**) specified, a physical file member has a keyed sequence access path, and the arrival sequence of the records in the member is changed to match their keyed sequence. In this example, after **RGZPFM KEYFILE(*FILE)** is run, the data is as shown in Example 3.

Example 3   Organization of data after using RGZPFM KEYFILE(*FILE)

```
....+....1....
        COL1
           1
           2
           3
           4
           5
           6
           7
           8
           9
 ********  End of data   ********
```

For a comparison of the different **RGZPFM** options, see the "Reorganization options" topic in the IBM i 7.1 Information Center found at the following website:

http://pic.dhe.ibm.com/infocenter/iseries/v7r1m0/index.jsp?topic=%2Fdbp%2Frbaforeo
rgtypes.htm

With **ALWCANCEL *YES**, journaling of the file is needed because commitment control is used when the job is deleting and reinserting the rows. It also moves a set of rows before a commit is issued. The number of rows that are moved per transaction varies from 1 to 4000. It is up to the system to adjust how many rows it moves before a commit is issued. Production jobs that are running with an isolation level of Cursor Stability or above might wait for rows that are moved until the commit is issued. The default record wait timeout is 60 seconds. If the record wait time is modified to be much smaller, lock timeout errors are possible.

To reduce the possibility of row lock waits by production jobs, have the reorganize job issue **OVRDBF FILE(FILE) WAITRCD(1)** before the reorganization. This action reduces the record wait time the reorganize job is waiting to lock a row and limit the time the set of records are locked. If the job cannot get a record lock during the one-second wait time, it not only skips that row and moves on, it also reduces the number of rows per transaction.

Here is an example of a query that can be used to give you an idea where the deleted rows are within a given file. This sample query shows you approximately where the deleted rows are in QSYS/QADBIFLD. It arbitrarily uses a group of 10000 rows. However, you can specify whatever group size you want.

> **Note:** The number of deleted rows in the last group is not accurate. To be accurate, get the total number of valid and deleted rows, then limit the number of deleted rows in the last group based on that total number.

```
With x as (select bigint(rrn(a)/10000) as rrngroup from qsys.qadbifld a)
select rrngroup, count(*) as Valid_in_Group, 10000-count(*) as Deleted_in_Group
from x group by rrngroup order by rrngroup;
```

It is also important to note in this example that QADBIFLD is a good example of a file on every system that has deleted records. However, you do not run **RGZPFM** against this particular file. If you are concerned about deleted records in this particular file, run **RCLSTG *DBXREF**, as described at the following website:

http://www-01.ibm.com/support/docview.wss?uid=nas8N1018121

> **Tip:** As a preferred practice, run **RGZPFM** only if a file contains 20% or more of deleted records.

Determining the status of a database reorganization

IBM i Navigator can show the status of any running database reorganization. With the IBM 7.1 technology refresh enhancements, you can see the history of canceled reorganizations and the status file that is kept for a reorganization that ran longer than 30 minutes.

To see the status of a database reorganization in IBM i Navigator, complete the following steps:

1. In IBM i Navigator, click **Databases** → **Database Maintenance** → **Table Reorganizations**, as shown in Figure 1.

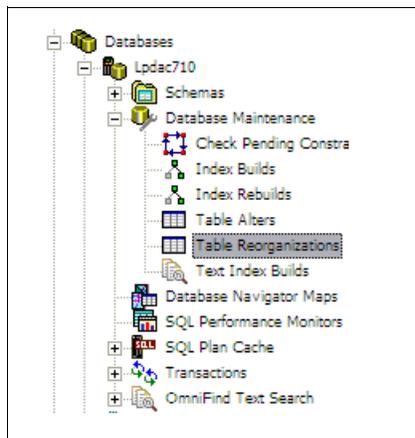> **Note:** The Database Maintenance folder is new in IBM i 7.1.

Figure 1   IBM i Navigator

2. If there are active database reorganizations, they are displayed. Right-click the reorganization of interest and select **Show Status**, as shown in Figure 2.

Figure 2   Displaying the status of an active database reorganization

3. A status window opens that is refreshed automatically with the status of the database reorganization, as shown in Figure 3.
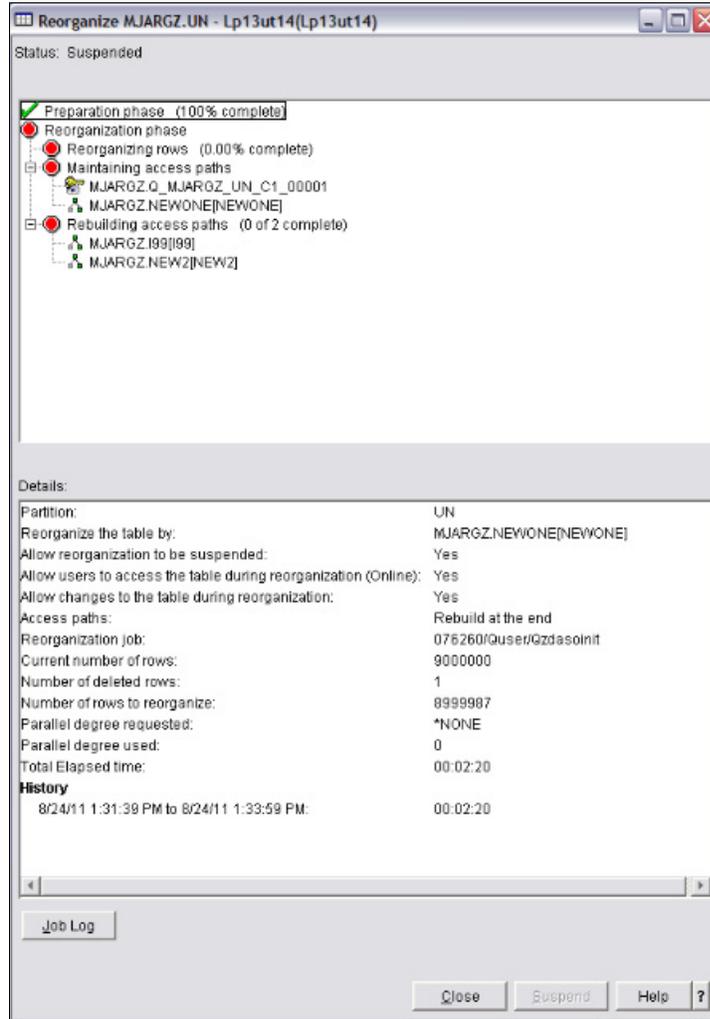

Figure 3   Status of database reorganization

4. From IBM i Navigator, click **Database**, select your database, click **Schemas**, and select the schema that contains the file that is rebuilding.

5. Right-click the file and select **Data** → **Reorganize**.

6. Click **Yes** to respond to the message shown in Figure 4. The status information for `RGZPFM` is displayed.
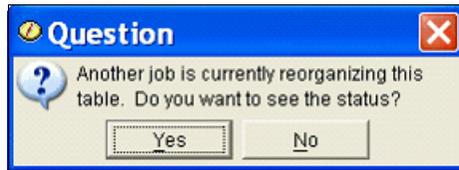

Figure 4   Status message for database being reorganized

If `RGZPFM` cannot start from where it left off, you see the message that is shown in Figure 5.

```
Message ID . . . . . . :   CPD3199        Severity . . . . . . . :    10
Message type . . . . . :   Diagnostic
Date sent  . . . . . . :   06/10/08       Time sent  . . . . . . :    21:17:14


Message . . . . :   Canceled reorganize operation not continued on file UN in
library UN, member MJARGZ.
Cause . . . . . :   A reorganize operation on file UN in library MJARGZ, member
UN detected that a previous reorganize operation was canceled. The reorganize
operation will start over and will not continue from where the canceled
reorganize operation ended for reason code 2. The reasons codes are as follows:
    1 -- The parameters specified on the current reorganize (*N) did not
exactly match the parameters from the canceled reorganize (*N).
    2 -- The member has undergone significant changes since the canceled
reorganize operation. Restore, delete file, remove member, rename file, rename
member, move file, rename library, clear physical file member, change end of
data, ALTER TABLE, and certain CHGPF operations are all examples of significant
changes to the member that will prevent a canceled reorganize operation from
being continued.
    3 -- The reorganize status file associated with the canceled reorganize did
not exist or was not valid.
Recovery  . . . :   No action is necessary. The reorganize operation will start
over.
```

Figure 5   Message that is shown when RGZPFM cannot start from where it left off

The benefit to using database reorganization while it is active is that it can take less time if ***RPLDLTRCD** is used, and you can do the reorganization while applications are running.

The disadvantage to database reorganization while it is active is that `RGZPFM` can take more time if ***NONE** or ***KEYFILE** is used. Also, because records are being moved, the record can disappear for a brief moment. In addition, journaling is required. Therefore, if replication software is used, there can be increased network traffic as journal entries are replicated.

From Record (FROMRCD) parameter

With the technology refreshes in IBM i 6.1 (SF99601 level 29) and IBM i 7.1 (SF99701 level 21), `RGZPFM` has a new From Record (`FROMRCD`) parameter that allows the initialization of the reorganization starting at a point within the table instead of starting at the beginning of the file and proceeding to the end of the file.

Here are the options for the `FROMRCD` parameter:

► `*START`

All records in the file are reorganized.

► `*PRVRGZ`

If the previous reorganization of this file could not remove some or all of the deleted records, the reorganization begins at the record that would have been the last record in the file if a concurrent insertion had not prevented the deleted records from being removed. If the previous reorganization completed normally and was able to remove the deleted records, the reorganization begins with the first record in the file.

`*PRVRGZ` is ignored if the reorganization is continued from a previously canceled reorganization.

> **Note:** If `*PRVRGZ` is specified, `ALWCANCEL(*YES)` must be specified and either `KEYFILE(*RPLDLTRCD)` or `KEYFILE(*NONE)` must be specified.

► 1 - 4294967288

Specifies the record number of the first record to be reorganized. If a record number greater than 1 is specified, `ALWCANCEL(*YES)` must be specified and either `KEYFILE(*RPLDLTRCD)` or `KEYFILE(*NONE)` must be specified.

The From Record (`FROMRCD`) parameter makes it easier to finish incomplete reorganizations or target the reorganization at the most fragmented portion of the table. This option is helpful in cases where most of the deleted rows are at the end of the table (especially when reorganizing to preserve the arrival sequence).

The example that is shown in Figure 6 on page 9 shows an `RGZPFM` worst case example, and how the `FROMRCD` parameter can help improve the process of deleting records. The red bars in Figure 6 on page 9 are valid records.
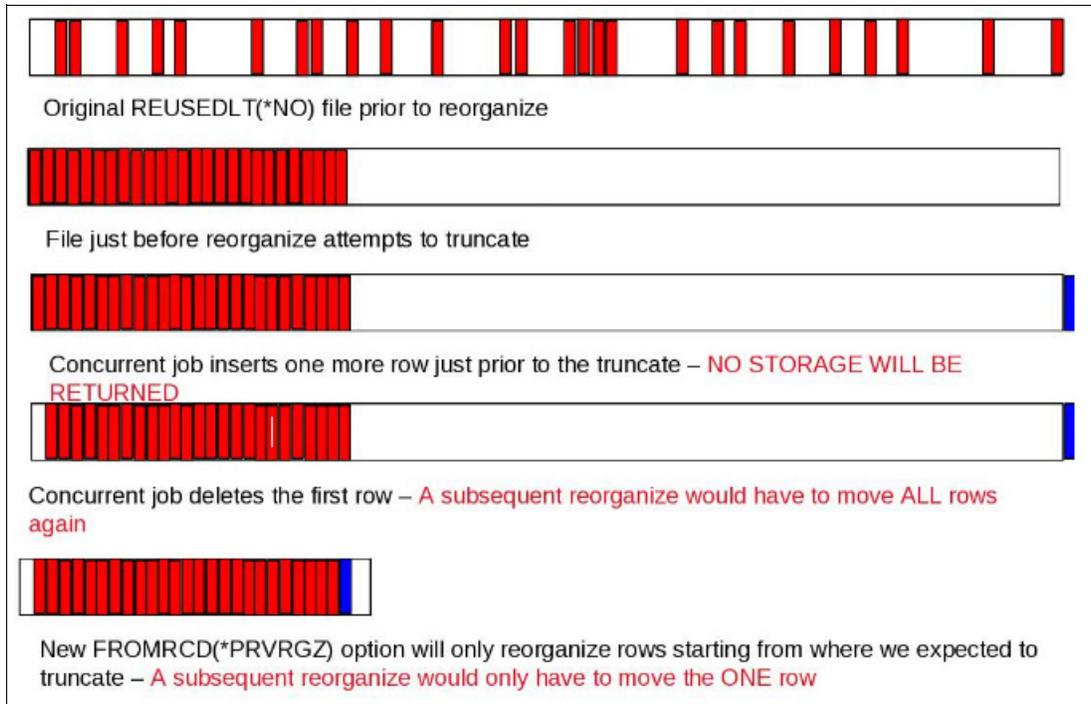
Figure 6   Database reorganization worst case example

Additional IBM i 7.1 TR 7 enhancements for RGZPFM

In addition to the new `FROMRCD` parameter, IBM i 7.1 Technology Refresh (TR) 7 has additional `RGZPFM` enhancements. Delayed maintenance indexes are preserved when the `RBDACCPTH(*NO)` and `ALWCANCEL(*YES)` parameters are used, and the unique or referential indexes are not rebuilt if they were maintained when `RBDACCPTH(*YES)`, `LOCK(*SHRUPD)`, and `ALWCANCEL(*YES)` were specified.

For more information about the `RGZPFM` command, see the "Reorganize Physical File Mbr (RGZPFM)" topic in the IBM i 7.1 Information Center found at the following website:

http://pic.dhe.ibm.com/infocenter/iseries/v7r1m0/topic/cl/rgzpfm.htm?resultof=%22%72%67%7a%70%66%6d%22%20

Author

This paper was produced by a technical specialist working at the International Technical Support Organization, Rochester Center.

Tracy Schramm is a Staff Software Engineer in the IBM i Global Support Center specializing in the database area. Before her assignment in database, she also worked on Print and Fax. In addition, she has had temporary assignments in development and an international assignment at IBM Mainz. Tracy holds a Bachelor of Science in Computer Science and a Bachelor of Arts in Spanish from Southwest Minnesota State University. In addition, she holds a Master of Information Systems / Management from the University of Phoenix. Tracy can be reached through email at tschramm@us.ibm.com.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new IBM Redbooks® publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document REDP-5083-00 was created or updated on March 10, 2014.

Send us your comments in one of the following ways:
- ► Use the online **Contact us** review Redbooks form found at:
  ibm.com/redbooks
- ► Send your comments in an email to:
  redbooks@us.ibm.com
- ► Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

IBM ®

Redpaper ™

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

IBM®	Redpaper™
Redbooks®	Redbooks (logo) ®

The following terms are trademarks of other companies:

Other company, product, or service names may be trademarks or service marks of others.