



Firebird Backup File Splitting Filter

Norman Dunbar

20 October 2009 – Document version 1.1

Table of Contents

Introduction	3
Gsplit Command line Options	3
Gsplit Command Parameters	4
Splitting Backups	5
Joining Backup Files	6
Appendix A: Document history	8
Appendix B: License notice	9

Introduction

In the past, many operating systems imposed a limit which defined how large a single file could become. This limit was 2 Gb on some systems, and 4 Gb on others. For example, on HP-UX 10.20 or 11.00 Unix systems, the maximum file size is 2 Gb unless the file system has the *largefiles* option turned on. This limit still exists on some operating systems.

Gsplit is a filter utility introduced with Interbase 5.0 which allows the output file from the gbak utility (when backing up a database) to be split into a number of chunks, or a number of chunks to be joined together and used to restore a database. Up until Interbase 5.0, dump files were limited to 2 Gb by the gbak utility itself - even if running on a system which allowed files to be 4 Gb.

Coming up in the remainder of this manual, we will discuss the following:

- Command line options for gsplit.
- Gsplit options and their parameters.
- Splitting backups using gsplit.
- Joining backup chunks using gsplit.

Note

From Interbase 6.0, gsplit is no longer required as gbak allows large files to be split directly. The details for gsplit given here are for reference only and you are advised to use gbak to split large backup files even if gsplit is supplied with your Firebird release.

Gsplit is only supplied with the Windows version of Firebird 1.5, it is not supplied with the Linux version. Linux doesn't require a separate utility as it can split files using the **split** command as well as using gbak.

Warning

In testing with Firebird 1.5 and Firebird 2.0 on Windows XP Home and Windows 2000, gsplit doesn't seem to work properly and always returns an error.

Because of the problems in getting gsplit to work correctly, as you shall see in the remainder of this manual, you are advised to use the splitting and joining facilities of the gbak utility itself rather than trying to make gsplit work for you.

Gsplit Command line Options

Gsplit has three command line options, although, strictly speaking, the **-help** option isn't really valid. These are:

- **-split_bk_file <parameters>**

Specifies that gsplit is to be used to split the output from gbak into a number of different files as part of a database backup. This option can be shortened as required, provided that at least **-s** is specified.

- **-join_bk_file <parameters>**

Specifies that gsplit is to be used to rejoin a number of files and use the result as input to gbak as part of a database restore. This option can be shortened as required, provided that at least `-j` is specified.

- **-help**

Using the `-help` option, specifying an illegal option, or omitting all options, displays the following information:

```
gsplit: invalid option '-help'
gsplit: Command Line Options Are:
gsplit -S[PLIT_BK_FILE] <file> <size>{k|m|g} [... <file> [<size>{k|m|g}]]
  or gsplit -J[OINT_BK_FILE] <file> [... <file>]
gsplit: option can be abbreviated to the unparenthesized characters
gsplit: Exiting before completion due to errors
```

Note

Take note of the error in the above help text. The correct command line option to join multiple sections of a large backup file is not `-JOINT_BK_FILE` as shown, but is in fact `-JOIN_BK_FILE`.

Gsplit Command Parameters

Both main gsplit command line options require parameters. When splitting a backup, the parameters are:

- `-S[PLIT_BK_FILE] <file> <size>{k|m|g} [... <file> [<size>{k|m|g}]]`

The first parameter is the first filename, followed by the maximum size it is allowed to be. You may specify the size in kilobytes, megabytes or gigabytes. There should not be any spaces between the digit(s) and the size specifier letter. There must be a space between the filename and the size however.

The remainder of the parameters specify the second and subsequent filenames and sizes, however, the final file need not have a size specified as it will be used to take up the remaining bytes after the other files have been filled to capacity. If there is a size specified, it is simply ignored but no error or warning messages are displayed.

If you have a backup file which will be 4 Gb in size and you only ask for two files, each 1 Gb in size, gsplit will ignore the size of the final file and just keep filling it up until the backup is complete.

The utility prevents files of less than 1 Mb and will produce an error if you attempt to specify a file smaller than this.

Note

Gsplit correctly specifies a Kilobyte as 1024 bytes, a Megabyte as 1024 Kilobytes and a Gigabyte as 1024 Megabytes.

- `-J[OIN_BK_FILE] <file> [... <file>]`

To join files together and use them to restore a database, you simply specify the filenames in the correct order. If they are not in the correct order, gsplit will complain and the restore will be abandoned.

Splitting Backups

To run `gsplit`, you need to use it as a filter on the command line for `gbak`, as the following example shows:

```
C:\>gbak -b norman.fdb stdout |
    gsplit -split norman_1.fbk 1m norman_2.fbk 1m norman_3.fbk
```

Warning

The command above assumes that `ISC_USER` and `ISC_PASSWORD` have been defined. For the sake of this demonstration, this is acceptable, but in a real system, consider the security implications before defining these variables.

In addition, the above command line has been split over two lines to allow the pdf generation of this manual to work. In reality, the command must be typed on a single line.

It is unfortunate that the utility supplied with Firebird 1.5 *doesn't work*, as the following shows:

```
C:\>gbak -b norman.fdb stdout |
    gsplit -split norman_1.fbk 1m norman_2.fbk 1m norman_3.fbk

fail to read input from ib_stdin, errno = 9
gsplit: program fails to generate multi-volumn back-up files

Done with volume #0, "stdout"
    Press return to reopen that file, or type a new
    name followed by return to open a different file.
Name:^C
```

If you type a filename at the prompt it will simply be used as a destination for a full dump of the database, so be careful not to overwrite anything important. I tend to hit **CTRL-C** at this point to avoid problems.

The utility has actually created the first file in the above list, `norman_1.fbk`, and written 100 bytes to a special header which identifies it as a `gsplit` created file.

Warning

The command above assumes that `ISC_USER` and `ISC_PASSWORD` have been defined. For the sake of this demonstration, this is acceptable, but in a real system, consider the security implications before defining these variables.

In addition, the above command line has been split over two lines to allow the pdf generation of this manual to work. In reality, the command must be typed on a single line.

Note

The spelling errors in 'program' and 'volumn' above are as produced by the utility.

The command does work under Firebird 2, however, it only creates the first file and then prompts for a new filename for the second file as the following shows:

```
C:\>gbak -b norman.fdb stdout |
    gsplit -split norman_1.fbk 1m norman_2.fbk 1m norman_3.fbk
```

Done with volume #1, "stdout" Press return to reopen that file, or type a new name followed by return to open a different file. Name:

If you press return at the prompt, the second file will be named `stdout` and not `norman_2.fbk`. If, on the other hand, you supply the filename `norman_2.fbk` then that file will be written to. Unfortunately you cannot specify how large the file is allowed to be, if you try, the file size becomes part of the file name.

Warning

I would say that as with Firebird 1.5's version of `gsplit`, using it to split a backup file is pointless. Use the `gbak` method to be safe.

Joining Backup Files

Had the above backup actually worked, the command to restore a backup from a number of files created by `gsplit` would be as follows:

```
C:\>gsplit -join norman_1.fbk norman_2.fbk norman_3.fbk |
    gbak -c stdin create_norman.fdb
```

Warning

The above command line has been split over two lines to allow the pdf generation of this manual to work. In reality, the command must be typed on a single line.

If you have a number of split backup files created using `gbak` itself and not filtered through `gsplit`, you cannot use `gsplit` to stitch them together for a restore as the following example shows:

```
C:\>gsplit -join norman_1.fbk norman_2.fbk norman_3.fbk |
    gbak -c stdin create_norman.fdb
gsplit: expected GSPLIT description record
gsplit: Exiting before completion due to errors
gsplit: program fails to join multi-volume back-up files
gbak: ERROR: expected backup description record
gbak: Exiting before completion due to errors
```

Warning

The above command line has been split over two lines to allow the pdf generation of this manual to work. In reality, the command must be typed on a single line.

It appears that `gsplit` and `gbak` have different header information in the backup files and the two are not compatible.

Note

The spelling errors in 'program' and 'volume' above are as produced by the utility.

Even with the version of gsplit supplied with Firebird 2, joining it's own partial files doesn't work:

```
C:\>gsplit -join norman_1.fbk norman_2.fbk norman_3.fbk |
gbak -c stdin create_norman.fdb

gbak:do not recognize domain attribute 13 -- continuing
gbak: ERROR:do not recognize record type 11
gbak:Exiting before completion due to errors

gsplit: expected GSPLIT description record
gsplit: Exiting before completion due to errors
gsplit: program fails to join multi-volumn back-up files
```

Appendix A: Document history

The exact file history is recorded in the manual module in our CVS tree; see http://sourceforge.net/cvs/?group_id=9028. The full URL of the CVS log for this file can be found at http://firebird.cvs.sourceforge.net/viewvc/firebird/manual/src/docs/firebirddocs/fbutil_gsplit.xml?view=log

Revision History

1.0	21 December 2004	ND	Created as a chapter in the Command Line Utilities manual.
1.1	20 October 2009	ND	Some updates for Firebird 2 and converted to a stand alone manual.

Appendix B: License notice

The contents of this Documentation are subject to the Public Documentation License Version 1.0 (the “License”); you may only use this Documentation if you comply with the terms of this License. Copies of the License are available at <http://www.firebirdsql.org/pdfmanual/pdl.pdf> (PDF) and <http://www.firebirdsql.org/manual/pdl.html> (HTML).

The Original Documentation is titled *Firebird Backup File Splitting Filter*.

The Initial Writer of the Original Documentation is: Norman Dunbar.

Copyright (C) 2004–2009. All Rights Reserved. Initial Writer contact: NormanDunbar at users dot sourceforge dot net.