GaussDB
8.x

Tool Reference for Primary+Standby Instances

Issue	01
Date	2024-06-06

HUAWEI CLOUD COMPUTING TECHNOLOGIES CO., LTD.

Huawei Cloud Computing Technologies Co., Ltd.

Address:	Huawei Cloud Data Center Jiaoxinggong Road Qianzhong Avenue Gui'an New District Gui Zhou 550029 People's Republic of China
Website:	https://www.huaweicloud.com/intl/en-us/

Contents

1 gsql

gsql, provided by GaussDB, is a database connection tool that runs in the command line. You can use **gsql** to connect to the server and perform operations and maintenance. In addition, **gsql** provides multiple **Advanced Features** for users.

1.1 Overview

Basic Features

- **Connect to the database**: By default, only the local server can be connected. To connect to a remote database, you must configure the server. For details, see "Database Quick Start > Connecting to a Database > Using gsql to Connect to a Database > Remotely Connecting to a Database" in *Developer Guide*.

  NOTE

  If gsql is used to connect to a database, the connection timeout interval will be 5 minutes. If the database has not correctly set up a connection and authenticated the identity of the client within this period, gsql will time out and exit.

  To resolve this problem, see **Troubleshooting**.
- **Run SQL statements**: Interactively entered SQL statements and specified SQL statements in a file can be run.
- **Run meta-commands**: Meta-commands help the administrator view database object information, query cache information, format SQL output, and connect to a new database. For details about meta-commands, see **Meta-Command Reference**.

Advanced Features

Table 1-1 lists the advanced features of gsql.

Table 1-1 Advanced features of gsql

Feature Name	Description

Variable	gsql provides a variable feature that is similar to the **shell** command of Linux. The following **\set** meta-command of gsql can be used to set a variable:
	`\set varname value`
	To delete the variables set by the **\set** command, run the following command:
	`\unset varname`
	NOTE
	● A variable is a simple name-value pair. The value can be any characters in any length.
	● Variable names must consist of case-sensitive letters (including non-Latin letters), digits, and underscores (_).
	● If the **\set** *varname* meta-command (without the second parameter) is used, the variable is set without a value specified.
	● If the **\set** meta-command without parameters is used, values of all variables are displayed.
	For details about variable examples and descriptions, see **Variables**.

SQL substitution	Common SQL statements can be set to variables using the variable feature of gsql to simplify operations. For details about examples and descriptions about SQL substitution, see **SQL substitution**.
Customized prompt	Prompts of gsql can be customized. Prompts can be modified by changing the reserved three variables of gsql: *PROMPT1*, *PROMPT2*, and *PROMPT3*. These variables can be set to customized values or the values predefined by gsql. For details, see **Prompt**.
Historical client operation records	gsql can record historical client operations. This function is enabled by specifying the **-r** parameter when a client is connected. The number of historical records can be set using the **\set** command. For example, **\set HISTSIZE 50** indicates that the number of historical records is set to **50**. **\set HISTSIZE 0** indicates that the operation history is not recorded. **NOTE** ● The default number of historical records is **32**. The maximum number of historical records is **500**. If interactively entered commands contain Chinese characters, only the UTF-8 encoding environment is supported. ● For security reasons, the records containing sensitive words (such as PASSWORD, IDENTIFIED, GS_ENCRYPT_AES128, GS_DECRYPT_AES128, GS_ENCRYPT, GS_DECRYPT, GS_ENCRYPT_BYTEA, GS_DECRYPT_BYTEA, PG_CREATE_PHYSICAL_REPLICATION_SLOT_EXTERN, SECRET_ACCESS_KEY, SECRETKEY, CREATE_CREDENTIAL, ACCESSKEY, and SECRET_KEY) are regarded sensitive and not recorded in historical information. This indicates that you cannot view these records in command output histories. Sensitive words are case insensitive.

- Variables

  To set a variable, run the **\set** meta-command of gsql. For example, to set variable *foo* to **bar**, run the following command:

  ```
  gaussdb=# \set foo bar
  ```

  To reference the value of a variable, add a colon (:) before the variable. For example, to view the value of variable *foo*, run the following command:

  ```
  gaussdb=# \echo :foo
  bar
  ```

  This variable reference method is applicable to regular SQL statements and meta-commands except **\copy**, **\ef**, **\help**, **\sf**, and **\!**.

  gsql pre-defines some special variables and plans the values of these variables. To ensure compatibility with later versions, do not use these variables for other purposes. For details about special variables, see **Table 1-2**.

  NOTE

  - All the special variables consist of upper-case letters, digits, and underscores (_).
  - To view the default value of a special variable, run the **\echo :***varname* meta-command, for example, **\echo :***DBNAME*.

Table 1-2 Settings of special variables

Variable	Setting Method	Description
DBNAME	`\set DBNAME dbname`	Name of the connected database. This variable is set again when a database is connected.
ECHO	`\set ECHO all | queries`	• If this variable is set to **all**, only the query information is displayed. This has the same effect as specifying the **-a** parameter when gsql is used to connect to a database. • If this variable is set to **queries**, the command line and query information are displayed. This has the same effect as specifying the **-e** parameter when gsql is used to connect to a database.

Variable	Setting Method	Description
ECHO_HIDDEN	`\set ECHO_HIDDEN  on | off | noexec`	When a meta-command (such as **\dg**) is used to query database information, the value of this variable determines the query behavior. ● If this variable is set to **on**, the query statements that are called by the meta-command are displayed, and then the query result is displayed. This has the same effect as specifying the **-E** parameter when gsql is used to connect to a database. ● If this variable is set to **off**, only the query result is displayed. ● If this variable is set to **noexec**, only the query information is displayed, and the query is not run.
ENCODING	\set ENCODING *encoding*	Character set encoding of the current client.
FETCH_COUNT	\set FETCH_COUNT *variable*	● If the value is an integer greater than **0**, for example, *n*, *n* lines will be selected from the result set to the cache and displayed on the screen when the **SELECT** statement is run. ● If this variable is not set or set to a value less than or equal to **0**, all results are selected at a time to the cache when the **SELECT** statement is run. **NOTE** A proper variable value helps reduce the memory usage. The recommended value range is from 100 to 1000.
HISTCONTROL	`\set HISTCONTROL ignorespace | ignoredups | ignoreboth | none`	● **ignorespace**: A line started with a space is not written to the historical record. ● **ignoredups**: A line that exists in the historical record is not written to the historical record. ● **ignoreboth**, **none**, or other values: All the lines read in interaction mode are saved in the historical record. **NOTE** **none** indicates that **HISTCONTROL** is not set.
HISTFILE	\set HISTFILE *filename*	Specifies the file for storing historical records. The default value is **~/.bash_history**.

Variable	Setting Method	Description
HISTSIZE	`\set HISTSIZE` *size*	Specifies the number of commands to store in the command history. The default value is **500**.
HOST	`\set HOST` *hostname*	Specifies the name of a connected host.
IGNOREEOF	`\set IGNOREEOF` *variable*	● If this variable is set to a number, for example, **10**, the first nine EOF characters (generally **Ctrl**+**C**) entered in gsql are neglected and the gsql program exits when the tenth **Ctrl**+**C** is entered. ● If this variable is set to a non-numeric value, the default value is **10**. ● If this variable is deleted, gsql exits when an EOF is entered.
LASTOID	`\set LASTOID` *oid*	Specifies the last OID, which is the value returned by an **INSERT** or **lo_import** command. This variable is valid only before the output of the next SQL statement is displayed.
ON_ERROR_ROLLBACK	`\set ON_ERROR_ROLLBACK on | interactive | off`	● If the value is **on**, an error that may occur in a statement in a transaction block is ignored and the transaction continues. ● If the value is **interactive**, the error is ignored only in an interactive session. ● If the value is **off** (default value), the error triggers the rollback of the transaction block. In **on_error_rollback-on** mode, a **SAVEPOINT** is set before each statement of a transaction block, and an error triggers the rollback of the transaction block.
ON_ERROR_STOP	`\set ON_ERROR_STOP on | off`	● **on**: specifies that the execution stops if an error occurs. In interactive mode, gsql returns the output of executed commands immediately. ● **off** (default value): specifies that an error, if occurring during the execution, is ignored, and the execution continues.
PORT	`\set PORT` *port*	Specifies the port number of a connected database.

Variable	Setting Method	Description
USER	`\set USER` *username*	Specifies the database user you are currently connected as.
VERBOSITY	`\set VERBOSITY   terse | default | verbose`	This variable can be set to **terse**, **default**, or **verbose** to control redundant lines of error reports. ● **terse**: Only critical and major error texts and text locations are returned (which is generally suitable for single-line error information). ● **default**: Critical and major error texts and text locations, error details, and error messages (possibly involving multiple lines) are all returned. ● **verbose**: All error information is returned.

- SQL substitution

  gsql, like a parameter of a meta-command, provides a key feature that enables you to substitute a standard SQL statement for a gsql variable. gsql also provides a new alias or identifier for the variable. To replace the value of a variable using the SQL substitution method, add a colon (:) before the variable. For example:

```
gaussdb=# \set foo 'HR.areaS'
gaussdb=# select * from :foo;
 area_id |       area_name
---------+------------------------
       4 | Middle East and Africa
       3 | Asia
       1 | Europe
       2 | Americas
(4 rows)
```

  The above command queries the HR.areaS table.

NOTICE

The value of the variable is copied literally, so it can even contain unbalanced quotation marks or backslash commands. Therefore, the input content must be meaningful.

- Prompt

  The gsql prompt can be set using the three variables in **Table 1-3**. These variables consist of characters and special escape characters.

Table 1-3 Prompt variables

Variable	Description	Example

PROMPT1	Specifies the normal prompt used when gsql requests a new command.
	The default value of *PROMPT1* is:
	`%/%R%#`
		PROMPT1 can be used to change the prompt.
		• Change the prompt to **[local]**:
		`gaussdb=> \set PROMPT1 %M`
		`[local:/tmp/gaussdba_mppdb]`
		• Change the prompt to **name**:
		`gaussdb=> \set PROMPT1 name`
		`name`
		• Change the prompt to **=**:
		`gaussdb=> \set PROMPT1 %R`
		`=`

PROMPT2	Specifies the prompt displayed when more input is expected because the command that is not terminated with a semicolon (;) or a quote (") is not closed.	*PROMPT2* can be used to display the prompt.
		`gaussdb=# \set PROMPT2 TEST`
		`gaussdb=# select * from HR.areaS TEST;`
		` area_id │       area_name`
		`---------+--------------------`
		`       1 │ Europe`
		`       2 │ Americas`
		`       4 │ Middle East and Africa`
		`       3 │ Asia`
		`(4 rows))`

PROMPT3	Specifies the prompt displayed when the **COPY** statement (such as **COPY FROM STDIN**) is run and data input is expected.	*PROMPT3* can be used to display the COPY prompt.
		`gaussdb=# \set PROMPT3 '>>>>'`
		`gaussdb=# copy HR.areaS from STDIN;`
		`Enter data to be copied followed by a newline.`
		`End with a backslash and a period on a line by itself.`
		`>>>>1 aa`
		`>>>>2 bb`
		`>>>>\.`

The value of the selected prompt variable is printed literally. However, a value containing a percent sign (%) is replaced by the predefined contents depending on the character following the percent sign (%). For details about the defined substitutions, see **Table 1-4**.

Table 1-4 Defined substitutions

Symbol	Description
%M	Replaced with the full host name (with domain name). The full name is **[local]** if the connection is over a Unix domain socket, or **[local:/dir/name]** if the Unix domain socket is not at the compiled default location.
%m	Replaced with the host name truncated at the first dot. It is **[local]** if the connection is over a Unix domain socket.

Symbol	Description
%>	Replaced with the number of the port that the host is listening on.
%n	Replaced with the database session username.
%/	Replaced with the name of the current database.
%~	Similar to **%/**. However, the output is tilde (~) if the database is your default database.
%#	Uses **#** if the session user is the database administrator. Otherwise, uses **>**.
%R	● In *PROMPT1* normally **=**, but **^** if in single-line mode, or **!** if the session is disconnected from the database (which can happen if **\connect** fails). ● In *PROMPT2* %R is replaced with a hyphen (-), an asterisk (*), a single or double quotation mark, or a dollar sign ($), depending on whether gsql expects more input because the query is inside a /*...*/ comment or inside a quoted or dollar-escaped string.
%x	Replaced with the transaction status. ● An empty string when it is not in a transaction block ● An asterisk (*) when it is in a transaction block ● An exclamation mark (!) when it is in a failed transaction block ● A question mark (?) when the transaction status is indefinite (for example, because there is no connection).
%digits	Replaced with the character with the specified byte.
%:name	Replaced with the value of the *name* variable of gsql.
%command	Replaced with the command output, similar to substitution with the "^" symbol.
%[. . . %]	Prompts may contain terminal control characters which, for example, change the color, background, or style of the prompt text, or change the title of the terminal window. For example: `gaussdb=> \set PROMPT1 '%[%033[1;33;40m%]%n@%/%R%[%033[0m%]%#'` The output is a boldfaced (1;) yellow-on-black (33;40) prompt on VT100-compatible, color-capable terminals.

Environment Variables

Table 1-5 Environment variables related to gsql

Name	Description
COLUMNS	If **\set columns** is set to **0**, this parameter controls the width of the wrapped format. This width determines whether to change the wide output mode into the vertical output mode if automatic expansion is enabled.
PAGER	If the query results do not fit on the screen, they are redirected through this command. You can use the **\pset** command to disable the pager. Typically, the **more** or **less** command is used for viewing the query result page by page. The default is platform-dependent. **NOTE** Display of the **less** command is affected by the *LC_CTYPE* environment variable.
PSQL_EDITOR	The **\e** and **\ef** commands use the editor specified by the environment variables. The variables are examined in the order listed. The default editor on Unix is vi.
EDITOR	
VISUAL	

PSQL_EDITOR_LINENUMBER_ARG	When the **\e** or **\ef** command is used with a line number parameter, this variable specifies the command-line parameter used to pass the starting line number to the editor. For editors, such as Emacs or vi, this is a plus sign. Include a space in the value of the variable if space is needed between the option name and the line number. For example:
	`PSQL_EDITOR_LINENUMBER_ARG = '+'`
	`PSQL_EDITOR_LINENUMBER_ARG='--line '`
	A plus sign (+) is used by default on Unix.

PSQLRC	Specifies the location of the user's .gsqlrc file.
SHELL	Has the same effect as the **\!** command.
TMPDIR	Specifies the directory for storing temporary files. The default value is **/tmp**.

1.2 Usage Guidelines

Prerequisites

- The user using **gsql** must have the permission to access the database.
- The gsql version must match the database version.

Background

You can use the **gsql** command to connect to the local database or remote database. When connecting to the remote database, enable remote connection on the server. For details, see "Database Quick Start > Connecting to a Database > Using gsql to Connect to a Database > Remotely Connecting to a Database" in the *Developer Guide*.

Procedure

Step 1 Connect to the GaussDB server using the **gsql** tool.

The **gsql** tool uses the **-d** parameter to specify the target database name, the **-U** parameter to specify the database username, the **-h** parameter to specify the host name, and the **-p** parameter to specify the port number.

NOTE

If the database name is not specified, the default database name generated during initialization will be used. If the database username is not specified, the current OS username will be used by default. If a variable does not belong to any parameter (such as **-d** and **-U**), and **-d** is not specified, the variable will be used as the database name. If **-d** is specified but **-U** is not specified, the variable will be used as the database username.

Example 1: Connect to the 8000 port of the local **gaussdb** database as user **omm**.

```
gsql -d gaussdb -p 8000
```

Example 2: Connect to the 8000 port of the remote **gaussdb** database as user **jack**.

```
gsql -h 10.180.123.163 -d gaussdb -U jack -p 8000
```

In a centralized database instance, when connecting to the primary DN, you can use commas (,) to separate the IP addresses of DNs and add them to the end of **-h**. **gsql** connects to each IP address in sequence to check whether the current DN is the primary DN. If the current DN is not the primary DN, **gsql** disconnects from the current IP address and attempts to connect to the next IP address until the primary DN is found.

```
gsql -h 10.180.123.163,10.180.123.164,10.180.123.165 -d gaussdb -U jack -p 8000
```

Example 3: If **gaussdb** and **omm** are not parameter values, they are interpreted as the database name and the username, respectively.

```
gsql gaussdb omm -p 8000
```

Equals

```
gsql -d gaussdb -U omm -p 8000
```

For details about the **gsql** parameters, see **Command Reference**.

Step 2 Run a SQL statement.

The following takes creating database **human_staff** as an example:

```
CREATE DATABASE human_staff;
CREATE DATABASE
```

Ordinarily, input lines end when a command-terminating semicolon is reached. If the command is sent and executed without any error, the command output is displayed on the screen.

Step 3 Execute gsql meta-commands.

The following takes all GaussDB databases and description information as an example:

```
gaussdb=# \l
                                List of databases
      Name       |  Owner   | Encoding  | Collate | Ctype |   Access privileges
----------------+----------+-----------+---------+-------+-----------------------
 human_resource | omm | SQL_ASCII | C       | C     |
 postgres       | omm | SQL_ASCII | C       | C     |
 template0      | omm | SQL_ASCII | C       | C     | =c/omm         +
                |          |           |         |       | omm=CTc/omm
 template1      | omm | SQL_ASCII | C       | C     | =c/omm          +
                |          |           |         |       | omm=CTc/omm
 human_staff    | omm | SQL_ASCII | C       | C     |
(5 rows)
```

For details about **gsql** meta-commands, see **Meta-Command Reference**.

----End

Example

The example shows how to spread a command over several lines of input. Note the prompt change:

```
gaussdb=# CREATE TABLE HR.areaS(
gaussdb(# area_ID   NUMBER,
gaussdb(# area_NAME VARCHAR2(25)
gaussdb-# )tablespace EXAMPLE;
CREATE TABLE
```

Query the table definition:

```
gaussdb=# \d HR.areaS
               Table "hr.areas"
  Column   |         Type          | Modifiers
-----------+-----------------------+-----------
 area_id   | numeric               | not null
 area_name | character varying(25) |
```

Insert four lines of data into **HR.areaS**.

```
gaussdb=# INSERT INTO HR.areaS (area_ID, area_NAME) VALUES (1, 'Europe');
INSERT 0 1
gaussdb=# INSERT INTO HR.areaS (area_ID, area_NAME) VALUES (2, 'Americas');
INSERT 0 1
gaussdb=# INSERT INTO HR.areaS (area_ID, area_NAME) VALUES (3, 'Asia');
INSERT 0 1
gaussdb=# INSERT INTO HR.areaS (area_ID, area_NAME) VALUES (4, 'Middle East and Africa');
INSERT 0 1
```

Change the prompt.

```
gaussdb=# \set PROMPT1 '%n@%m %~%R%#'
omm@[local] gaussdb=#
```

Query the table:

```
omm@[local] gaussdb=# SELECT * FROM HR.areaS;
 area_id |       area_name
---------+------------------------
```

```
       1 | Europe
       4 | Middle East and Africa
       2 | Americas
       3 | Asia
(4 rows)
```

Use the **\pset** command to display the table in different ways:

```
omm@[local] gaussdb=# \pset border 2
Border style is 2.
omm@[local] gaussdb=# SELECT * FROM HR.areaS;
+---------+------------------------+
| area_id |       area_name        |
+---------+------------------------+
|       1 | Europe                 |
|       2 | Americas               |
|       3 | Asia                   |
|       4 | Middle East and Africa |
+---------+------------------------+
(4 rows)
omm@[local] gaussdb=# \pset border 0
Border style is 0.
omm@[local] gaussdb=# SELECT * FROM HR.areaS;
area_id       area_name
------- ----------------------
      1 Europe
      2 Americas
      3 Asia
      4 Middle East and Africa
(4 rows)
```

Use the meta-command:

```
omm@[local] gaussdb=# \a \t \x
Output format is unaligned.
Showing only tuples.
Expanded display is on.
omm@[local] gaussdb=# SELECT * FROM HR.areaS;
area_id|2
area_name|Americas

area_id|1
area_name|Europe

area_id|4
area_name|Middle East and Africa

area_id|3
area_name|Asia
omm@[local] gaussdb=#
```

1.3 Obtaining Help Information

Procedure

- When connecting to the database, run the following command to obtain the help information:

  ```
  gsql --help
  ```

  The following help information is displayed:

  ```
  ......
  Usage:
    gsql [OPTION]... [DBNAME [USERNAME]]

  General options:
    -c, --command=COMMAND    run only single command (SQL or internal) and exit
  ```

```
 -d, --dbname=DBNAME       database name to connect to (default: "omm")
 -f, --file=<FILE_NAME>      execute commands from file, then exit
......
```

- When connecting to the database, run the following command to obtain the help information:

  ```
  help
  ```

  The following help information is displayed:

  ```
  You are using gsql, the command-line interface to gaussdb.
  Type:  \copyright for distribution terms
         \h for help with SQL commands
         \? for help with gsql commands
         \g or terminate with semicolon to execute query
         \q to quit
  ```

Examples

Step 1 Connect to the database. For details, see "Database Quick Start > Connecting to a Database > Using gsql to Connect to a Database" in the *Developer Guide*.

Step 2 View the gsql help information. For details, see **Table 1-6**.

Table 1-6 gsql online help

Description	Example
Query the copyright.	\copyright

View help information about SQL statements supported by GaussDB.	View help information about SQL statements supported by GaussDB.
	For example, view all SQL statements supported by GaussDB.
	`gaussdb=# \h`
	`Available help:`
	`  ABORT`
	`  ALTER APP WORKLOAD GROUP`
	`  ALTER APP WORKLOAD GROUP MAPPING`
	`... ...`
	For example, view parameters of the **CREATE DATABASE** command:
	`gaussdb=# \help CREATE DATABASE`
	`Command:     CREATE DATABASE`
	`Description: create a new database`
	`Syntax:`
	`CREATE DATABASE database_name`
	`    [ [ WITH ] {[ OWNER [=] user_name ]|`
	`          [ TEMPLATE [=] template ]|`
	`          [ ENCODING [=] encoding ]|`
	`          [ LC_COLLATE [=] lc_collate ]|`
	`          [ LC_CTYPE [=] lc_ctype ]|`
	`          [ DBCOMPATIBILITY [=] compatibility_type ]|`
	`          [ TABLESPACE [=] tablespace_name ]|`
	`          [ CONNECTION LIMIT [=] connlimit ]}[...] ];`

Description	Example

View the help information about gsql commands.	For example, view commands supported by gsql.
	`gaussdb=# \?`
	`General`
	`  \copyright            show GaussDB Kernel usage and distribution terms`
	`  \g [FILE] or ;         execute query (and send results to file or |pipe)`
	`  \h(\help) [NAME]              help on syntax of SQL commands, * for all commands`
	`  \q                     quit gsql`
	`... ...`

----End

1.4 Command Reference

For details about gsql parameters, see **Table 1-7**, **Table 1-8**, **Table 1-9**, and **Table 1-10**.

Table 1-7 Common parameters

Parameter	Description	Value Range
-c, --command=COMMAND	Specifies that gsql is to run a string command and then exit.	-
-d, --dbname=DBNAME	Specifies the name of the database to connect to. In addition, gsql allows you to use extended DBNAMEs, that is, connection strings in the format of **'postgres[ql]://[user[:password]@][netloc][:port][, ...][/dbname][?param1=value1&...]'** or **'[key=value] [...]'** as DBNAMEs. gsql parses connection information from the connection strings and preferentially uses the information. **CAUTION** When gsql uses the extended DBNAME to create a connection, the **replication** parameter cannot be specified.	String
-f, --file=FILENAME	Specifies that files are used as the command source instead of interactively-entered commands. After the files are processed, gsql exits. If *FILENAME* is - (hyphen), then standard input is read.	An absolute path or relative path that meets the OS path naming convention
-l, --list	Lists all available databases and then exits.	-

Parameter	Description	Value Range
-v, --set, --variable=NAME=VALUE	Sets gsql variable *NAME* to *VALUE*. For details about variable examples and descriptions, see **Variables**.	-
-X, --no-gsqlrc	Does not read the startup file (neither the system-wide **gsqlrc** file nor the user's **~/.gsqlrc** file). **NOTE** The startup file is **~/.gsqlrc** by default or it can be specified by the environment variable *PSQLRC*.	-
-1 ("one"), --single-transaction	When gsql uses the **-f** parameter to execute a script, **START TRANSACTION** and **COMMIT** are added to the start and end of the script, respectively, so that the script is executed as one transaction. This ensures that the script is executed successfully. If the script cannot be executed, the script is invalid. **NOTE** If the script has used **START TRANSACTION**, **COMMIT**, or **ROLLBACK**, this parameter is invalid.	-
-?, --help	Displays help information about gsql command parameters, and exits.	-
-V, --version	Prints the gsql version information and exits.	-

Table 1-8 Input and output parameters

Parameter	Description	Value Range
-a, --echo-all	Prints all input lines to standard output as they are read. **CAUTION** When this parameter is used in some SQL statements, the sensitive information, such as user password, may be disclosed. Use this parameter with caution.	-
-e, --echo-queries	Displays all queries sent to the server to the standard output as well. **CAUTION** When this parameter is used in some SQL statements, the sensitive information, such as user password, may be disclosed. Use this parameter with caution.	-

Parameter	Description	Value Range
-E, --echo-hidden	Echoes the actual queries generated by **\d** and other backslash commands.	-
-k, --with-key=KEY	Uses gsql to decrypt imported encrypted files. **NOTICE** ● For key characters, such as the single quotation mark (') or double quotation mark (") in shell commands, Linux shell checks whether the input single quotation mark (') or double quotation mark (") matches. If no match is found, Linux shell does not enter the gsql program until input is complete. ● Stored procedures and functions cannot be decrypted and imported.	-
-L, --log-file=FILENAME	Writes normal output source and all query output into the **FILENAME** file. **CAUTION** ● When this parameter is used in some SQL statements, the sensitive information, such as user password, may be disclosed. Use this parameter with caution. ● This parameter retains only the query result in the corresponding file, so that the result can be easily found and parsed by other invokers (for example, automatic O&M scripts). Logs about gsql operations are not retained.	An absolute path or relative path that meets the OS path naming convention
-m, --maintenance	Allows connections to the database during two-phase transaction recovery. **NOTE** The parameter is for engineers only. When this parameter is used, gsql can be connected to the standby node to check data consistency between the primary and standby nodes.	-
-n, --no-libedit	Closes command line editing.	-
-o, --output=FILENAME	Puts all query output into the **FILENAME** file.	An absolute path or relative path that meets the OS path naming convention
-q, --quiet	Indicates the quiet mode and no additional information will be printed.	By default, gsql displays various information.

Parameter	Description	Value Range
-s, --single-step	Runs in single-step mode. It indicates that the user is prompted before each command is sent to the server. This option can be also used for canceling execution. Use this option to debug scripts. **CAUTION** When this parameter is used in some SQL statements, the sensitive information, such as user password, may be disclosed. Use this parameter with caution.	-
-S, --single-line	Runs in single-line mode where a line break terminates a command, as a semicolon does.	-
-C, -C1, --enable-client-encryption=1	Enables the encrypted database function when the **-C** parameter is used to connect to a local or remote database. The encrypted equality query is supported.	-
-C3, --enable-client-encryption=3	Enables a memory decryption emergency channel when the **-C** parameter is used to connect to a local or a remote database. The encrypted equality query and the memory decryption emergency channel are supported.	-

Table 1-9 Output format parameters

Parameter	Description	Value Range
-A, --no-align	Switches to unaligned output mode.	The default output mode is aligned.
-F, --field-separator=STRING	Specifies the field separator. The default is the vertical bar (|).	-
-H, --html	Turns on the HTML tabular output.	-
-P, --pset=VAR[=ARG]	Specifies the print option in the **\pset** format in the command line. **NOTE** The equal sign (=), instead of the space, is used here to separate the name and value. For example, enter **-P format=latex** to set the output format to **LaTeX**.	-
-R, --record-separator=STRING	Sets the record separator.	-

Parameter	Description	Value Range
-r	Enables the editing mode on the client.	This function is disabled by default.
-t, --tuples-only	Prints only tuples.	-
-T, --table-attr=TEXT	Specifies options to be placed within the HTML table tag. Use this parameter with the **-H,--html** parameter to specify the output to the HTML format.	-
-x, --expanded	Turns on the expanded table formatting mode.	-
-z, --field-separator-zero	Sets the field separator in the unaligned output mode to be blank. Use this parameter with the **-A, --no-align** parameter to switch to unaligned output mode.	-
-0, --record-separator-zero	Sets the record separator in the unaligned output mode to be blank. Use this parameter with the **-A, --no-align** parameter to switch to unaligned output mode.	-
-2, --pipeline	Uses a pipe to transmit the password. This parameter cannot be used on devices and must be used together with the **-c** or **-f** parameter.	-

Table 1-10 Connection parameters

Parameter	Description	Value Range
-h, --host=HOSTNAME	Specifies the host name of the running server, the path of the Unix domain socket, or the domain name. You can specify multiple host addresses by using character strings separated by commas (,), or specify an IPv6 host address. If multiple host addresses are specified, the primary node address is automatically selected for connection by default. You can set the *PGTARGETSESSIONATTRS* environment variable to connect to different types of nodes. The mapping between variables and node types is as follows: A value of the *PGTARGETSESSIONATTRS* environment variable indicates a type of node to be connected. **read-write**: readable and writable node. **read-only**: read-only node. **primary** or not set: primary node. **standby**: standby node. **prefer-standby**: preferred standby node. If there is no standby node, **any** is used. **any**: The role is not checked. **NOTE** If **-h** specifies only one domain name but the domain name corresponds to multiple IP addresses, the automatic primary node selection function cannot be triggered.	If the host name is omitted, gsql connects to the server of the local host over the Unix domain socket or over TCP/IP to connect to local host without the Unix domain socket.
-p, --port=PORT	Specifies the port number of the database server. You can configure one or more port numbers. When one port number is configured, all host addresses use the same port for connection. When multiple port numbers are configured, the sequence is the same as the host address sequence, and the number of port numbers must be the same as the number of host addresses. If they are different, an error is reported. You can modify the default port number using the **-p, --port=PORT** parameter.	The default port number can be specified by using compilation parameters. If the port number is not specified, the default port number **5432** is used.

Parameter	Description	Value Range
-U, --username=USERNAME	Specifies the user who connects to the database. **NOTE** ● If this parameter is specified, you also need to enter your password for identity authentication when connecting to the database. You can enter the password interactively or use the **-W** parameter to specify a password. ● To connect to a database, add an escape character before any dollar sign ($) in the username.	String. The default user is the current user who operates the system.
-W, --password=PASSWORD	Specifies a password when the **-U** parameter is used to connect to the local database or a remote database. **NOTE** ● When the server where the primary database node is located connects to the local primary database node instance, the trust connection is used by default and this parameter is ignored. ● To connect to a database, add an escape character before any backslash (\) or back quote (`) in the password. ● If this parameter is not specified but database connection requires your password, you will be prompted to enter your password in interactive mode. The maximum length of the password is 999 bytes, which is restricted by the maximum value of the GUC parameter **password_max_length**.	String

1.5 Meta-Command Reference

This section describes meta-commands provided by gsql after the GaussDB database CLI tool is used to connect to a database. A gsql meta-command can be anything that you enter in gsql and begins with an unquoted backslash.

Precautions

- The format of the gsql meta-command is a backslash (\) followed by a command verb, and then a parameter. The parameters are separated from the command verb and from each other by any number of whitespace characters.
- To include whitespace characters into an argument, you must quote them with a single straight quotation mark. To include a single straight quotation mark into such an argument, precede it by a backslash. Anything contained in single quotation marks is furthermore subject to C-like substitutions for \n (new line), \t (tab), \b (backspace), \r (carriage return), \f (form feed), \digits (octal), and \xdigits (hexadecimal).

- Within a parameter, text enclosed in double quotation marks ("") is taken as a command line input to the shell. The output of the command (with any trailing newline removed) is taken as the argument value.
- If an unquoted argument begins with a colon (:), the argument is taken as a gsql variable and the value of the variable is used as the argument value instead.
- Some commands take an SQL identifier (such as a table name) as a parameter. These parameters follow the SQL syntax rules: Unquoted letters are forced to lowercase, while double quotation marks ("") protect letters from case conversion and allow incorporation of whitespace into the identifier. Within double quotation marks, paired double quotation marks reduce to a single double quotation mark in the result name. For example, **FOO"BAR"BAZ** is interpreted as **fooBARbaz**, and **"Aweird""name"** becomes **A weird"name**.
- Parsing for arguments stops when another unquoted backslash is found. This is taken as the beginning of a new meta-command. The special sequence \\ (two backslashes) marks the end of parameters and continues parsing SQL statements if any. In this way, SQL and gsql commands can be freely mixed in a line. But in any case, the arguments of a meta-command cannot continue beyond the end of the line.
- M-compatible databases do not support the **\h** meta-command.

Meta-commands

For details about meta-commands, see **Table 1-11**, **Table 1-12**, **Table 1-13**, **Table 1-14**, **Table 1-16**, **Table 1-18**, **Table 1-19**, **Table 1-20**, **Table 1-22**, and **Table 1-23**.

NOTICE

FILE mentioned in the following commands indicates a file path. This path can be an absolute path such as **/home/gauss/file.txt** or a relative path, such as **file.txt**. By default, a **file.txt** is created in the path where the user runs gsql commands.

Table 1-11 Common meta-commands

Parameter	Description	Value Range
\copyright	Displays GaussDB version and copyright information.	-
\g [FILE] or ;	Performs a query operation and sends the result to a file or pipe.	-

Parameter	Description	Value Range
\h(\help) [NAME]	Provides syntax help on the specified SQL statement.	If the name is not specified, then gsql will list all the commands for which syntax help is available. If the name is an asterisk (*), syntax help on all SQL statements is displayed.
\parallel [on [num]|off]	Controls the parallel execution function. ● **on**: The switch is enabled and the maximum number of concurrently executed tasks is **num**. ● **off**: The switch is disabled. **NOTE** ● Parallel execution is not allowed in a running transaction and a transaction is not allowed to be started during parallel execution. ● Parallel execution of **\d** meta-commands is not allowed. ● If **SELECT** statements are run concurrently, customers can accept the problem that the return results are displayed randomly but they cannot accept it if a core dump or process response failure occurs. ● It is not recommended that you run **SET** statements in concurrent tasks because they may cause unexpected results. ● Temporary tables cannot be created in parallel. If temporary tables are required, create them before parallel execution is enabled, and use them only in the parallel execution. Temporary tables cannot be created in parallel execution. ● When **\parallel** is executed, *num* independent gsql processes can be connected to the database server. ● The total duration of all **\parallel** tasks cannot exceed **session_timeout**. Otherwise, the connection may fail during concurrent execution. ● One or more commands following **\parallel on** will be executed only after **\parallel off** is executed. Therefore, **\parallel on** must be followed by **\parallel off**. Otherwise, the commands following **\parallel on** cannot be executed.	The default value of *num* is **1024**. **NOTICE** ● The maximum number of connections allowed by the server is determined based on **max_connection** and the number of current connections. ● Set the value of *num* based on the allowed number of connections.

Parameter	Description	Value Range
\q	Exits the gsql program. This command is executed only when a script terminates in a script file.	-

Table 1-12 Query buffer meta-commands

Parameter	Description
\e [FILE] [LINE]	Uses an external editor to edit the query buffer or file.
\ef [FUNCNAME [LINE]]	Edits the function definition using an external editor. If **LINE** is specified, the cursor will point to the specified line of the function body.
\p	Prints the current query buffer to the standard output.
\r	Resets or clears the query buffer.
\w FILE	Outputs the current query buffer to a file.

Table 1-13 Input/Output commands

Parameter	Description
\copy { table [(column_list)] | (query) } { from | to } { filename | stdin | stdout | pstdin | pstdout }[LOAD] [LOAD_DISCARD 'string'] [with] [binary] [oids] [delimiter [as] 'character'] [useeof] [null [as] 'string'] [csv [header] [quote [as] 'character'] [escape [as] 'character'] [force quote column_list | *] [force not null column_list]] [parallel integer]	After logging in to the database on any gsql client, you can import and export data. This is an operation of running the **SQL COPY** command, but not the server that reads or writes data to a specified file. Instead, data is transferred between the server and the local file system. In this way, local user permissions instead of server permissions are required for file access, and the user permissions do not need to be initialized. **NOTE** ● \COPY only applies to small-batch data import with uniform formats. GDS or COPY is preferred for data import. ● \COPY specifies the number of clients to import data to implement parallel import of data files. Currently, the value range is [1,8]. ● The parallel import using \COPY has the following constraints: Parallel import of temporary tables is not supported. Parallel import within transactions is not supported. Parallel import of binary files is not supported. Parallel import of data encrypted using AES-128 is not supported. The COPY option contains EOL. In these cases, even if the parallel parameter is specified, a non-parallel process is performed. ● Both the TEXT and CSV formats of \COPY support the header function. ● The LOAD function is used by gs_loader to call COPY after syntax conversion. It is not an active calling function. ● The LOAD_DISCARD function is used by gs_loader to discard file path after parsing. It is not an active calling function.
\echo [STRING]	Displays a character string as the standard output.
\i FILE	Reads content from *FILE* and uses them as the input for a query.
\i+ FILE KEY	Runs commands in an encrypted file.
\ir FILE	Similar to **\i**, but resolves relative path names differently.
\ir+ FILE KEY	Similar to **\i+**, but resolves relative path names differently.
\o [FILE]	Saves all query results to a file.

Parameter	Description
\qecho [STRING]	Writes character strings to the query output flow.

NOTE

In **Table 1-14**, **S** indicates displaying the system object and **+** indicates displaying the additional description information of the object. **PATTERN** specifies the name of an object to be displayed.

Table 1-14 Information display meta-commands

Parameter	Description	Value Range	Example
\d[S+]	Lists all tables, views, and sequences of all schemas in **search_path**. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	-	List all tables, views, and sequences of all schemas in **search_path**. `gaussdb=# \d`
\d[S+] NAME	Lists the structure of specified tables, views, and indexes. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	-	List the structure of table **a**. `gaussdb=# \dtable+ a`
\d+ [PATTERN]	Lists all tables, views, and indexes. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only tables, views, and indexes whose names match **PATTERN** are displayed.	List all tables, views, and indexes whose names start with **f**. `gaussdb=# \d+ f*`

Parameter	Description	Value Range	Example
\da[S] [PATTERN]	Lists all available aggregate functions, together with the data type they perform operations on and the return value types. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only aggregate functions whose names match **PATTERN** are displayed.	List all available aggregate functions whose names start with **f**, together with their return value types and the data types. `gaussdb=# \da f*`
\db[+] [PATTERN]	Lists all available tablespaces.	If **PATTERN** is specified, only tablespaces whose names match **PATTERN** are displayed.	List all available tablespaces whose names start with **p**. `gaussdb=# \db p*`
\dc[S+] [PATTERN]	Lists all available conversions between character-set encodings. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only conversions whose names match **PATTERN** are displayed.	List all available conversions between character-set encodings. `gaussdb=# \dc *`
\dC[+] [PATTERN]	Lists all available type conversions. PATTERN must be the actual type name and cannot be an alias. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only conversions whose names match **PATTERN** are displayed.	List all type conversions whose patten names start with **c**. `gaussdb=# \dC c*`

Parameter	Description	Value Range	Example
\dd[S] [PATTERN]	Lists descriptions about objects matching **PATTERN**. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If no parameter is specified, all visible objects are displayed. The objects include aggregations, functions, operators, types, relations (tables, views, indexes, sequences, and large objects), and rules.	List all visible objects. `gaussdb=# \dd`
\ddp [PATTERN]	Lists all default permissions.	If **PATTERN** is specified, only permissions whose names match **PATTERN** are displayed.	List all default permissions. `gaussdb=# \ddp`
\dD[S+] [PATTERN]	Lists all available domains. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only domains whose names match **PATTERN** are displayed.	List all available domains. `gaussdb=# \dD`
\det[+] [PATTERN]	Lists all foreign tables. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only tables whose names match **PATTERN** are displayed.	List all foreign tables. `gaussdb=# \det`
\des[+] [PATTERN]	Lists all foreign servers.	If **PATTERN** is specified, only servers whose names match **PATTERN** are displayed.	List all foreign servers. `gaussdb=# \des`

Parameter	Description	Value Range	Example
\deu[+] [PATTERN]	Lists all user mappings.	If **PATTERN** is specified, only information whose name matches **PATTERN** is displayed.	List all user mappings. `gaussdb=# \deu`
\dew[+] [PATTERN]	Lists all encapsulated external data.	If **PATTERN** is specified, only data whose name matches **PATTERN** is displayed.	List all encapsulated external data. `gaussdb=# \dew`
\df[antw][S+] [PATTERN]	Lists all available functions, together with their parameters and return types. **a** indicates an aggregate function, **n** indicates a common function, **t** indicates a trigger, and **w** indicates a window function. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only functions whose names match **PATTERN** are displayed.	List all available functions, together with their parameters and return types. `gaussdb=# \df`
\dF[+] [PATTERN]	Lists all text search configuration information. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only configurations whose names match **PATTERN** are displayed.	List all text search configuration information. `gaussdb=# \dF+`
\dFd[+] [PATTERN]	Lists all text search dictionaries. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only dictionaries whose names match **PATTERN** are displayed.	List all text search dictionaries. `gaussdb=# \dFd`

Parameter	Description	Value Range	Example
\dFp[+] [PATTERN]	Lists all text search analyzers. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only analyzers whose names match **PATTERN** are displayed.	List all text search analyzers. `gaussdb=# \dFp`
\dFt[+] [PATTERN]	Lists all text search templates. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only templates whose names match **PATTERN** are displayed.	List all text search templates. `gaussdb=# \dFt`
\dg[+] [PATTERN]	Lists all database roles. **NOTE** Since the concepts of "users" and "groups" have been unified into "roles", this command is now equivalent to **\du**. Both commands are retained to ensure compatibility with earlier versions.	If **PATTERN** is specified, only roles whose names match **PATTERN** are displayed.	Lists all database roles named **j?e** (the question mark (?) indicates any character). `gaussdb=# \dg j?e`
\dl	This is an alias for **\lo_list**, which shows a list of large objects.	-	List all large objects. `gaussdb=# \dl`
\dL[S+] [PATTERN]	Lists all available program languages.	If **PATTERN** is specified, only languages whose names match **PATTERN** are displayed.	List all available program languages. `gaussdb=# \dL`
\dm[S+] [PATTERN]	Lists materialized views. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only materialized views whose names match **PATTERN** are displayed.	List materialized views. `gaussdb=# \dm`

Parameter	Description	Value Range	Example
\dn[S+] [PATTERN]	Lists all schemas (namespace). If **+** is added to the command, the permission and description of each schema are listed.	If **PATTERN** is specified, only schemas whose names match the pattern are shown. By default, only schemas you created are displayed.	List information about all schemas whose names start with **d**. `gaussdb=# \dn+ d*`
\do[S] [PATTERN]	Lists available operators with their operand and return types. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only operators whose names match **PATTERN** are displayed. By default, only the operators created by the user are listed.	List available operators with their operand and return types. `gaussdb=# \do`
\dO[S+] [PATTERN]	Lists collation rules. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only rules whose names match **PATTERN** are displayed. By default, only user-created rules are shown.	List collation rules. `gaussdb=# \dO`

Parameter	Description	Value Range	Example

\dp [PATTERN]	Lists tables, views, and related permissions. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.
	The following result about **\dp** is displayed:
	`rolename=xxxx/yyyy  -- Assigns permissions to a role.`
	`=xxxx/yyyy -- Assigns permissions to public.`
	xxxx indicates the assigned permissions, and *yyyy* indicates the roles with the assigned permissions. For details about permission descriptions, see **Table 1-15**.
		If **PATTERN** is specified, only tables and views whose names match the pattern are shown.	List tables, views, and related permissions.
			`gaussdb=# \dp`

\drds [PATTERN1 [PATTERN2]]	Lists all parameters that have been modified. These settings can be for roles, for databases, or for both. **PATTERN1** and **PATTERN2** indicate a role pattern and a database pattern, respectively.	If **PATTERN** is specified, only collations rules whose names match **PATTERN** are displayed. If the default value is used or ***** is specified, all settings are listed.	Lists all the modified configuration parameters of the database. `gaussdb=# \drds * dbname`
\dT[S+] [PATTERN]	Lists all data types. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed.	If **PATTERN** is specified, only types whose names match **PATTERN** are displayed.	List all data types. `gaussdb=# \dT`
\du[+] [PATTERN]	Lists all database roles. **NOTE** Since the concepts of "users" and "groups" have been unified into "roles", this command is now equivalent to **\dg**. Both commands are retained to ensure compatibility with earlier versions.	If **PATTERN** is specified, only roles whose names match **PATTERN** are displayed.	List all database roles. `gaussdb=# \du`

Parameter	Description	Value Range	Example
\dE[S+] [PATTERN] \di[S+] [PATTERN] \ds[S+] [PATTERN] \dt[S+] [PATTERN] \dv[S+] [PATTERN]	In this group of commands, the letters E, i, s, t, and v stand for foreign table, index, sequence, table, and view, respectively. You can specify any or a combination of these letters sequenced in any order to obtain an object list. When objects with the same name exist in different schemas in **search_path**, only the object in the schema that ranks first in **search_path** is displayed. For example, **\dit** lists all indexes and tables. If + is added to the end of a command name, the physical size and related description of each object are also listed.	If **PATTERN** is specified, only objects whose names match **PATTERN** are displayed. By default, only objects you created are displayed. You can specify **PATTERN** or **S** to view other system objects.	List all indexes and views. `gaussdb=# \div`
\dx[+] [PATTERN]	Lists installed extensions.	If **PATTERN** is specified, only extensions whose names match **PATTERN** are displayed.	List installed extensions. `gaussdb=# \dx`
\l[+]	Lists the names, owners, character set encodings, and permissions of all the databases in the server.	-	List the names, owners, character set encodings, and permissions of all the databases in the server. `gaussdb=# \l`

Parameter	Description	Value Range	Example

\sf[+] FUNCNAME	Displays the definition of a function.
	NOTE
	If the function name contains parentheses, enclose the function name with double quotation marks and add the parameter type list following the double quotation marks. Also enclose the list with parentheses.
	If a function with the same name exists, the definitions of multiple functions are returned.
		-	Assume a function **function_a** and a function **func()name**. This parameter will be as follows:
			`gaussdb=# \sf function_a`
			`gaussdb=# \sf "func()name"(argtype1, argtype2)`

\z [PATTERN]	Lists all tables, views, and sequences in the database and their access permissions.	If a pattern is given, it is a regular expression, and only matched tables, views, and sequences are shown.	Lists all tables, views, and sequences in the database and their access permissions. `gaussdb=# \z`

Table 1-15 Description of permissions

Parameter	Description
r	**SELECT**: allows users to read data from specified tables and views.
w	**UPDATE**: allows users to update columns for specified tables.
a	**INSERT**: allows users to insert data to specified tables.
d	**DELETE**: allows users to delete data from specified tables.
D	**TRUNCATE**: allows users to delete all data from specified tables.
x	**REFERENCES**: allows users to create foreign key constraints.
t	**TRIGGER**: allows users to create a trigger on specified tables.
X	**EXECUTE**: allows users to use specified functions and the operators that are realized by the functions.

Parameter	Description
U	**USAGE**: ● For procedural languages, allows users to specify a procedural language when creating a function. ● For schemas, allows users to access objects included in specified schemas. ● For sequences, allows users to use the nextval function.
C	**CREATE**: ● For databases, allows new schemas to be created within the database. ● For schemas, allows users to create objects in a schema. ● For tablespaces, allows users to create tables in a tablespace and set the tablespace to default one when creating databases and schemas.
c	**CONNECT**: allows users to connect to specified databases.
T	**TEMPORARY**: allows users to create temporary tables.
A	**ALTER**: allows users to modify the attributes of a specified object.
P	**DROP**: allows users to delete specified objects.
m	**COMMENT**: allows users to define or modify comments of a specified object.
i	**INDEX**: allows users to create indexes on specified tables.
v	**VACUUM**: allows users to perform ANALYZE and VACUUM operations on specified tables.
*	Authorization options for preceding permissions.

Table 1-16 Formatting meta-commands

Parameter	Description
\a	Switches between aligned and unaligned table output formats.
\C [STRING]	Sets the title of any table being printed as the result of a query or unsets any such title.
\f [STRING]	Sets the field separator for unaligned query outputs.

Parameter	Description
\H	● If the text format schema is used, switches to the HTML format. ● If the HTML format schema is used, switches to the text format.
\pset NAME [VALUE]	Sets options affecting the output of query result tables. For details about the value of **NAME**, see **Table 1-17**.
\t [on|off]	Switches the display of output name information and row count footer.
\T [STRING]	Specifies attributes to be placed within the table tag in HTML output format. If this parameter is empty, no attribute is specified.
\x [on|off|auto]	Switches expanded table formatting mode.

Table 1-17 Adjustable printing options

Option	Description	Value Range
border	The value must be a number. In general, the larger the number, the more borders and lines the tables will have, but this depends on the particular format.	● The value is an integer greater than 0 in HTML format. ● The value range in other formats is as follows: – **0**: no border – **1**: internal dividing line – **2**: table frame
expanded (or x)	Switches between regular and expanded formats.	● When the expanded format is enabled, query results are displayed in two columns, with the column name on the left and the data on the right. This format is useful if the data does not fit the screen in the normal "horizontal" format. ● Use the expanded format when the query output format is wider than the screen in regular format. The regular format is effective only in the aligned and wrapped formats.

Option	Description	Value Range
fieldsep	Specifies the field separator to be used in unaligned output mode. In this way, you can create tab- or comma-separated output required by other programs. To set a tab as field separator, type **\pset fieldsep '\t'**. The default field separator is a vertical bar (|).	-
fieldsep_zero	Sets the field separator to use in unaligned output format to a zero byte.	-
footer	Switches the display of the default footer.	-
format	Selects the output format. Unique abbreviations are allowed (this indicates that one letter is enough).	Value range: ● **unaligned**: Write all columns of a row on one line, separated by the currently active column separator. ● **aligned**: This format is standard and human-readable. ● **wrapped**: This format is similar to **aligned**, but includes the packaging cross-line width data value to suit the width of the target field output. ● **html**: This format outputs tables to the markup language for a document. The output is not a complete document. ● **latex**: This format outputs tables to the markup language for a document. The output is not a complete document. ● **troff-ms:** This format outputs tables to the markup language for a document. The output is not a complete document.
null	Sets a character string to be printed in place of a null value.	The default is to print nothing, which can be easily mistaken for an empty string.

Option	Description	Value Range
numericlo cale	Switches the display of a locale-aware character to separate groups of digits to the left of the decimal marker.	• **on**: The specified separator is displayed. • **off**: The specified separator is not displayed If this parameter is ignored, the default separator is displayed.
pager	Controls the use of a pager for query and gsql help outputs. If the **PAGER** environment variable is set, the output is redirected to the specified program. Otherwise, the platform-dependent default value is used.	• **on**: The pager is used for terminal output that does not fit the screen. • **off**: The pager is not used. • **always**: The pager is used for all terminal output regardless of whether it fits the screen.
recordsep	Specifies the record separator to use in unaligned output mode.	-
recordsep _zero	Sets the record separator to use in unaligned output format to a zero byte.	-
tableattr (or T)	Specifies attributes to be placed inside the HTML table tag in HTML output format (such as cellpadding or bgcolor). Note that you do not need to specify **border** here because it has been used by **\pset border**. If no value is given, the table attributes do not need to be set.	-
title	Sets the table title for any subsequently printed tables. This can be used to give your output descriptive tags. If no value is given, the title does not need to be set.	-
tuples_onl y (or t)	Enables or disables the tuples-only mode. Full display may show extra information, such as column headers, titles, and various footers. In tuples-only mode, only the table data is shown.	-
feedback	Specifies whether to output the number of result lines.	-

Table 1-18 Connection meta-commands

Parameter	Description	Value Range
\c[onnect] [DBNAME|- USER|- HOST|- PORT|-]	Connects to a new database. If a database name contains more than 63 bytes, only the first 63 bytes are valid and are used for connection. However, the database name displayed in the command line of gsql is still the name before the truncation. **NOTE** If the database login user is changed during reconnection, you need to enter the password of the new user. The maximum length of the password is 999 bytes, which is restricted by the maximum value of the GUC parameter **password_max_length**.	-
\encoding [ENCODING]	Sets the client character set encoding.	Without an argument, this command shows the current encoding.
\conninfo	Prints information about the current connected database.	-

Table 1-19 OS meta-commands

Parameter	Description	Value Range
\cd [DIR]	Changes the current working directory.	An absolute path or relative path that meets the OS path naming convention
\setenv NAME [VALUE]	Sets the **NAME** environment variable to **VALUE**. If **VALUE** is not provided, do not set the environment variable.	-
\timing [on|off]	Toggles a display of how long each SQL statement takes, in milliseconds (exclude the time of screen displaying).	● **on**: specifies that the display is enabled. ● **off**: indicates that the display is disabled.
\! [COMMAND]	Escapes to a separate Unix shell or runs a Unix command.	-

Table 1-20 Variable meta-commands

Parameter	Description
\prompt [TEXT] NAME	Prompts the user to use texts to specify a variable name.
\set [NAME [VALUE]]	Sets the *NAME* internal variable to **VALUE**. If more than one value is provided, *NAME* is set to the concatenation of all of them. If no second argument is given, the variable is just set with no value. Some common variables are processed differently in gsql and they are combinations of uppercase letters, numbers and underscores. **Table 1-21** describes a list of variables that are processed in a way different from other variables.
\unset NAME	Deletes the variable name of gsql.

Table 1-21 Common **\set** commands

Command	Description	Value Range
\set VERBOSITY value	This variable can be set to **default**, **verbose**, or **terse** to control redundant lines of error reports.	Value range: **default**, **verbose**, and **terse**
\set ON_ERROR_STOP value	If this variable is set, the script execution stops immediately. If this script is called from another script, that script will be stopped immediately as well. If the outermost script is called using the **-f** option rather than from an interactive gsql session, gsql will return error code **3**, indicating the difference between the current error and critical errors. (The error code for critical errors is **1**.)	Value range: **on**/**off**, **true**/**false**, **yes**/**no**, or **1**/**0**

Command	Description	Value Range
\set AUTOCOMMIT [on|off]	Sets the auto commit behavior of the current gsql connection. **on** indicates that auto commit is enabled, and **off** indicates that auto commit is disabled. By default, the gsql connection is automatically committed, and each individual statement is implicitly committed. If auto commit is disabled for performance or other purposes, you need to explicitly run the COMMIT command to ensure that transactions are committed. For example, execute the COMMIT statement to explicitly commit transactions after a specified service SQL statement is executed. Particularly, ensure that all transactions are committed before the gsql client exits. **NOTE** The auto commit is enabled in gsql by default. If you disable it, all the statements executed later will be packaged in implicit transactions, and you cannot execute statements that cannot be executed within transactions.	• **on**: The auto commit is enabled. • **off**: The auto commit is disabled.

Table 1-22 Large object meta-commands

Parameter	Description
\lo_list	Displays a list of all GaussDB large objects stored in the database, along with the comments provided for the large objects.

Table 1-23 Fully-encrypted meta-commands

Parameter	Description
\send_token	Sends keys to the server for caching. This function is used only when the memory decryption emergency channel is enabled. This is a fully-encrypted function.
\st	Sends keys to the server for caching. This function is used only when the memory decryption emergency channel is enabled. This is a fully-encrypted function.
\clear_token	Destroys the keys cached on the server. This function is used only when the memory decryption emergency channel is enabled. This is a fully-encrypted function.

Parameter	Description
\ct	Destroys the keys cached on the server. This function is used only when the memory decryption emergency channel is enabled. This is a fully-encrypted function.
\key_info KEY_INFO	In the fully-encrypted database features, this parameter is used to set the parameters for accessing the external key manager.

NOTE

Currently, fully-encrypted databases are not M-compatible.

PATTERN

The various **\d** commands accept a **PATTERN** parameter to specify the object name to be displayed. In the simplest case, PATTERN is the exact name of the object. Characters in **PATTERN** are usually converted to lowercase (as in SQL names), for example, **\dt FOO** will display a table named **foo**. As in SQL names, placing double quotation marks (") around a pattern prevents them being folded to lower case. If you need to include a double quotation mark (") in a pattern, write it as a pair of double quotation marks ("") within a double-quote sequence, which is in accordance with the rules for SQL quoted identifiers. For example, **\dt "FOO""BAR"** will be displayed as a table named **FOO"BAR** instead of **foo"bar**. You cannot put double quotation marks around just part of a pattern, which is different from the normal rules for SQL names. For example, **\dt FOO"FOO"BAR** will be displayed as a table named **fooFOObar** if just part of a pattern is quoted.

Whenever the **PATTERN** parameter is omitted completely, the **\d** commands display all objects that are visible in the current schema search path, which is equivalent to using an asterisk (*) as the pattern. An object is regarded to be visible if it can be referenced by name without explicit schema qualification. To see all objects in the database regardless of their visibility, use a dot within double quotation marks (*.*) as the pattern.

Within a pattern, the asterisk (*) matches any sequence of characters (including no characters) and a question mark (?) matches any single character. This notation is comparable to Unix shell file name patterns. For example, **\dt int*** displays tables whose names start with **int**. But within double quotation marks, the asterisk (*) and the question mark (?) lose these special meanings and are just matched literally.

A pattern that contains a dot (.) is interpreted as a schema name pattern followed by an object name pattern. For example, **\dt foo*.*bar*** displays all tables (whose names include **bar**) in schemas starting with **foo**. If no dot appears, then the pattern matches only visible objects in the current schema search path. Likewise, the dot within double quotation marks loses its special meaning and becomes an ordinary character.

Senior users can use regular-expression notations, such as character classes. For example [0-9] can be used to match any digit. All regular-expression special characters work as specified in POSIX. The following characters are excluded:

- A dot (.) is used as a separator.
- An asterisk (*) is translated into an asterisk prefixed with a dot (.*), which is a regular-expression marking.
- A question mark (?) is translated into a dot (.).
- A dollar sign ($) is matched literally.

You can write ?, (*R*+|), (*R*|), and *R* to the following pattern characters: ., *R**, and *R*?. The dollar sign ($) does not need to be used as a regular expression character because **PATTERN** must match the entire name instead of being interpreted as a regular expression (in other words, $ is automatically appended to **PATTERN**). If you do not expect a pattern to be anchored, write an asterisk (*) at its beginning or end. All regular-expression special characters within double quotation marks lose their special meanings and are matched literally. Regular-expression special characters in operator name patterns (such as the **\do** parameter) are also matched literally.

DELIMITER

The DELIMITER command is used to change the delimiter between SQL statements. The default delimiter is a semicolon (;).

Using the DELIMITER command, you can set a delimiter for the client. When a delimiter is set, the gsql client sends the SQL statements to the server for execution immediately after identifying the delimiter. However, the server still considers the semicolon (;) as the SQL statement delimiter and processes the SQL statements accordingly.

Precautions:

- Currently, delimiters cannot be set freely. The terminator can be a combination of uppercase and lowercase letters or a combination of special characters (**~ ! @ # ^ & ` ? + - * / % < > =**). The common delimiter is **//**.
- The combination of special characters should be unambiguous. Ambiguous combinations, such as comment characters ***** and **--** and combinations ending with a plus sign (+) or minus sign (-), cannot be used for delimiter naming.
- The delimiter length ranges from 0 to 15.
- The level of the terminator is session-level. When the database is switched, **delimiter_name** is set to the default value semicolon (;).
- To use other combinations, you can add quotation marks (for example, **delimiter "adbc $$"**). The quotation marks are required in statements, for example, **select 1"adbc $$"**.
- The delimiter is supported only when **sql_compatibility** is set to **'B'**.

1.6 Troubleshooting

Low Connection Performance

- **log_hostname** is enabled, but DNS is incorrect.

  Connect to the database, and run **show log_hostname** to check whether **log_hostname** is enabled in the database.

  If it is enabled, the database kernel will use DNS to check the name of the host where the client is deployed. If the host where the database is configured with an incorrect or unreachable DNS, the database connection will take a long time to set up. For more details about **log_hostname**, see the section "GUC Parameters".

- The database kernel slowly runs the initialization statement.

  Problems are difficult to locate in this scenario. Try using the **strace** Linux trace command.

  ```
  strace gsql -U MyUserName -d gaussdb -h 127.0.0.1 -p 23508 -r -c '\q'
  Password for MyUserName:
  ```

  The database connection process will be printed on the screen. If the following statement takes a long time to run:

  ```
  sendto(3, "Q\0\0\0\25SELECT VERSION()\0", 22, MSG_NOSIGNAL, NULL, 0) = 22
  poll([{fd=3, events=POLLIN|POLLERR}], 1, -1) = 1 ([{fd=3, revents=POLLIN}])
  ```

  It can be determined that the database executes the SELECT VERSION() statement slowly.

  After the database is connected, you can run the **explain performance select version()** statement to find the reason why the initialization statement was run slowly. For more information, see "SQL Optimization > Introduction to the SQL Execution Plan" in the *Developer Guide*.

  An uncommon scenario is that the disk of the machine where the DN resides is full or faulty, affecting queries and leading to user authentication failures. As a result, the connection process is suspended. To solve this problem, simply clear the data disk space of the DN.

- TCP connection is set up slowly.

  Adapt the steps of troubleshooting slow initialization statement execution. Use **strace**. If the following statement is run slowly:

  ```
  connect(3, {sa_family=AF_FILE, path="/home/test/tmp/gaussdb_llt1/.s.PGSQL.61052"}, 110) = 0
  ```

  Or,

  ```
  connect(3, {sa_family=AF_INET, sin_port=htons(61052), sin_addr=inet_addr("127.0.0.1")}, 16) = -1
  EINPROGRESS (Operation now in progress)
  ```

  It indicates that the physical connection between the client and the database is set up slowly. In this case, check whether the network is unstable or has high throughput.

- The connection is slow because the resource load is full.

  Cause analysis: When the CPU, memory, or I/O load is close to 100%, the gsql connection is slow.

  Solution:

  a. Run the **top** command to check the CPU usage. Run the **free** command to check the memory usage. Run the **iostat** command to check the I/O load. You can also check the monitor logs in the CM Agent and the monitoring records on the database O&M platform.

b. For peak load scenarios caused by a large number of slow queries in a short period of time, you can use the port specified by [*Port number of the database server* + 1] to query the pg_stat_activity view. For slow queries, you can use the system function pg_terminate_backend to kill sessions.

c. If service overloading exists for a long time (that is, there is no obvious slow query, or new queries still become slow after slow queries are killed), reduce the service load and increase database resources.

Problems in Setting Up Connections

- gsql: could not connect to server: No route to host

  This problem occurs generally because an unreachable IP address or port number was specified. Check whether the values of **-h** and **-p** parameters are correct.

- gsql: FATAL: Invalid username/password,login denied.

  This problem occurs generally because an incorrect username or password was entered. Contact the database administrator to check whether the username and password are correct.

- gsql: FATAL: Forbid remote connection with trust method!

  For security purposes, remote login in trust mode is forbidden. In this case, you need to modify the connection authentication information in the **gs_hba.conf** file. Contact the administrator.

  NOTE

  Do not modify the configurations of database hosts in the **gs_hba.conf** file. Otherwise, the database may become faulty. It is recommended that service applications be deployed outside the database instead of inside the database.

- The DN can connect to the database if **-h 127.0.0.1** is specified, and the connection will fail if **-h 127.0.0.1** is removed.

  Run the SQL statement **show unix_socket_directory** to check whether the **unix socket directory** used by the DN is the same as that specified by the environment variable *$PGHOST* in the **shell** directory.

  If they are different, set *$PGHOST* to the directory specified by **unix_socket_directory**.

  For more details about **unix_socket_directory**.

- The "libpq.so" loaded mismatch the version of gsql, please check it.

  This problem occurs because the version of **libpq.so** used in the environment does not match that of gsql. Run the **ldd gsql** command to check the version of the loaded **libpq.so**, and then load correct **libpq.so** by modifying the environment variable *LD_LIBRARY_PATH*.

- gsql: symbol lookup error: xxx/gsql: undefined symbol: libpqVersionString

  This problem occurs because the version of **libpq.so** used in the environment does not match that of gsql (or the PostgreSQL **libpq.so** exists in the environment). Run the **ldd gsql** command to check the version of the loaded **libpq.so**, and then load correct **libpq.so** by modifying the environment variable *LD_LIBRARY_PATH*.

- gsql: connect to server failed: Connection timed out

  Is the server running on host "xx.xxx.xxx.xxx" and accepting TCP/IP connections on port xxxx?

  This problem is caused by network connection faults. Check the network connection between the client and the database server. If you cannot ping from the client to the database server, the network connection is abnormal. Contact network management personnel for troubleshooting.

  ```
  ping -c 4 10.10.10.1
  PING 10.10.10.1 (10.10.10.1) 56(84) bytes of data.
  From 10.10.10.1: icmp_seq=2 Destination Host Unreachable
  From 10.10.10.1 icmp_seq=2 Destination Host Unreachable
  From 10.10.10.1 icmp_seq=3 Destination Host Unreachable
  From 10.10.10.1 icmp_seq=4 Destination Host Unreachable
  --- 10.10.10.1 ping statistics ---
  4 packets transmitted, 0 received, +4 errors, 100% packet loss, time 2999ms
  ```

- gsql: FATAL: permission denied for database "gaussdb"

  DETAIL: User does not have CONNECT privilege.

  This problem occurs because the user does not have the permission to access the database. To solve this problem, perform the following steps:

  a. Connect to the database as the system administrator **dbadmin**.
     ```
     gsql -d gaussdb -U dbadmin -p 8000
     ```
  b. Grant the user with the permission to access the database.

     GRANT CONNECT ON DATABASE gaussdb TO user1;

     NOTE

     Actually, some common misoperations may also cause a database connection failure, for example, entering an incorrect database name, username, or password. In this case, the client tool will display the corresponding error messages.

     ```
     gsql -d gaussdb -p 8000
     gsql: FATAL:  database "gaussdb" does not exist

     gsql -d gaussdb -U user1 -p 8000
     Password for user user1:
     gsql: FATAL:  Invalid username/password, login denied.
     ```

- gsql: FATAL: sorry, too many clients already, active/non-active: 197/3.

  This problem occurs because the number of system connections exceeds the allowed maximum. Contact the DBA database administrator to release unnecessary session connections.

  You can check the number of session connections as described in **Table 1-24**.

  You can view the session status in the PG_STAT_ACTIVITY view. To release unnecessary sessions, use the pg_terminate_backend function.

  ```
  select datid,pid,state from pg_stat_activity;
   datid |       pid       | state
  -------+-----------------+--------
   13205 | 139834762094352 | active
   13205 | 139834759993104 | idle
  (2 rows)
  ```

  The value of **pid** is the thread ID of the session. Terminate the session using its thread ID.

  ```
  SELECT PG_TERMINATE_BACKEND(139834759993104);
  ```

  If a command output similar to the following is displayed, the session is successfully terminated.

```
PG_TERMINATE_BACKEND
----------------------
 t
(1 row)
```

Table 1-24 Viewing the number of session connections

Description	Command

View the maximum number of session connections of a specific user.	Run the following command to view the maximum number of **USER1**'s session connections, where **–1** indicates that no upper limit is set for the number of **USER1**'s session connections:
	`SELECT ROLNAME,ROLCONNLIMIT FROM PG_ROLES WHERE ROLNAME='user1';`
	` rolname | rolconnlimit`
	`---------+--------------`
	` user1    |           -1`
	`(1 row)`

View the number of session connections that have been used by a specified user.	Run the following command to view the number of session connections that have been used by **USER1**, where **1** indicates the number of session connections that have been used by **USER1**:
	`SELECT COUNT(*) FROM dv_sessions WHERE USERNAME='user1';`
	
	` count`
	`-------`
	`     1`
	`(1 row)`

View the maximum number of session connections of a specific database.	Run the following command to view the maximum number of **gaussdb**'s session connections, where **–1** indicates that no upper limit is set for the number of **gaussdb**'s session connections:
	`SELECT DATNAME,DATCONNLIMIT FROM PG_DATABASE WHERE DATNAME='gaussdb';`
	
	` datname  | datconnlimit`
	`----------+--------------`
	` gaussdb |           -1`
	`(1 row)`

View the number of session connections that have been used by a specified database.	Run the following command to view the number of session connections that have been used by **gaussdb**, where **1** indicates the number of session connections that have been used by **gaussdb**:
	`SELECT COUNT(*) FROM PG_STAT_ACTIVITY WHERE DATNAME='gaussdb';`
	` count`
	`-------`
	`     1`
	`(1 row)`

Description	Command

View the number of session connections that have been used by all users.	Run the following command to view the number of session connections that have been used by all users:
	`SELECT COUNT(*) FROM dv_sessions;`
	
	` count`
	`-------`
	`     10`
	`(1 row)`

- gsql: wait xxx.xxx.xxx.xxx:xxxx timeout expired

  When gsql initiates a connection request to the database, a 5-minute timeout period is used. If the database cannot correctly authenticate the client request and client identity within this period, gsql will exit the connection process for the current session, and will report the above error.

  Generally, this problem is caused by the incorrect host and port (that is, the *xxx* part in the error information) specified by the **-h** and **-p** parameters. As a result, the communication fails. Occasionally, this problem is caused by network faults. To resolve this problem, check whether the host name and port number of the database are correct.

- gsql: could not receive data from server: Connection reset by peer.

  Check whether DN logs contain information similar to "FATAL: cipher file "/data/coordinator/server.key.cipher" has group or world access". This error is usually caused by incorrect tampering with the permissions for data directories or some key files. For details about how to correct the permissions, see related permissions for files on other normal instances.

- gsql: FATAL: GSS authentication method is not allowed because XXXX user password is not disabled.

  In **gs_hba.conf** of the target DN, the authentication mode is set to **gss** for authenticating the IP address of the current client. However, this authentication algorithm cannot authenticate clients. Change the authentication algorithm to **sha256** and try again. For details, contact the administrator.

  NOTE

  - Do not modify the configurations of database hosts in the **gs_hba.conf** file. Otherwise, the database may become faulty.
  - It is recommended that service applications be deployed outside the database instead of inside the database.

Other Faults

- There is a core dump or abnormal exit due to the bus error.

  Generally, this problem is caused by changes to the shared dynamic library (.so file in Linux) loaded during process running. Alternatively, if the process binary file changes, the execution code for the OS to load machines or the entry for loading a dependent library will change accordingly. In this case, the OS terminates the process for protection purposes, generating a core dump file.

To solve this problem, try again. In addition, do not run service programs in a database during O&M operations, such as an upgrade, preventing such a problem caused by file replacement during the upgrade.

NOTE

A possible stack of the core dump file contains dl_main and its function calling. The file is used by the OS to initialize a process and load the shared dynamic library. If the process has been initialized but the shared dynamic library has not been loaded, the process cannot be considered completely started.

2 gs_loader

Overview

gs_loader is used to import data. gs_loader converts the syntax supported by the control file to the \COPY syntax, uses the existing \COPY function to import data, and records the \COPY result in logs.

Before using gs_loader, ensure that the gs_loader version is consistent with the gsql version and database version.

NOTE

gs_loader does not support M-compatible databases.

Installation and Deployment

Install and configure the gs_loader client tool on the server where source data files are stored so that you can use the gs_loader tool to import data.

Step 1 Create a directory for storing the gs_loader tool package.

```
mkdir -p /opt/bin
```

Step 2 Upload the gsql tool package to the created directory.

For example, upload the gsql tool package **GaussDB-Kernel_***Database version number_OS version number*_**64bit_gsql.tar.gz** in the software installation package to the created directory.

Step 3 Go to the directory where the tool package is located and decompress the package.

```
cd /opt/bin
tar -zxvf GaussDB-Kernel_Database version number_OS version number_64bit_gsql.tar.gz
source gsql_env.sh
```

Step 4 Verify the tool location and version information.

```
which gs_loader
```

Step 5 Verify the client version information.

The gs_loader tool version number corresponds to the gsql tool version number. You can directly query the gsql client version number to verify the client version information.

```
gsql -V
```

Step 6 Verify that the database version is the same as the client tool version.

Use gsql to connect to the database and run the following command:

```
select version();
```

----End

Log Level Configuration

Set the log level for developers to view. After the setting, the tool running information is printed on the console.

```
export gs_loader_log_level=debug
export gs_loader_log_level=info
export gs_loader_log_level=warning
export gs_loader_log_level=error
```

Permission

The application scenarios are classified into separation-of-duties and non-separation-of-duties scenarios. You can set **enableSeparationOfDuty** to **on** or **off** to enable or disable the separation of duties function.

The **enable_copy_error_log** parameter specifies whether to use the error table pgxc_copy_error_log. The default value is **off**, indicating that the error table is not used and error records are directly recorded in the .bad file of gs_loader. If this parameter is set to **on**, the error table pgxc_copy_error_log is used and error records are inserted into the error table.

By default, if **enableSeparationOfDuty** is set to **off**, the user can be a common database user or an administrator. If the user is a common user, the administrator needs to grant permissions to the common user. The administrator account can be used directly. The GUC parameter **enable_copy_error_log** determines whether to enable the error table pgxc_copy_error_log. By default, it is disabled.

1. Create a user as an administrator.
   ```
   CREATE USER load_user WITH PASSWORD '************';
   ```
2. Grant the public schema permission to the new user.
   ```
   GRANT ALL ON SCHEMA public TO load_user;
   ```
3. Create the **gs_copy_summary** table and grant permissions to the new user.

   NOTE

   The gs_copy_summary table cannot contain objects that may cause privilege escalation, such as RULE, TRIGGER, index functions, row-level security, CHECK constraints, GENERATED columns, DEFAULT columns, and ON UPDATE columns; otherwise, the system considers that the user is created by a malicious user, reports an error, and exits.

   ```
   SELECT copy_summary_create() WHERE NOT EXISTS(SELECT * FROM pg_tables WHERE
   schemaname='public' AND tablename='gs_copy_summary');
   GRANT INSERT,SELECT ON  public.gs_copy_summary To load_user;
   ```
4. (Optional) Create an error table **pgxc_copy_error_log** and grant permissions to the new user.

NOTE

- If the GUC parameter **enable_copy_error_log** is not set (**off** by default) or is set to **off**, you do not need to use the error table and do not need to create it. Otherwise, you need to create the error table.
- The pgxc_copy_error_log table cannot contain objects that may cause privilege escalation, such as RULE, TRIGGER, index functions, row-level security, CHECK constraints, GENERATED columns, DEFAULT columns, and ON UPDATE columns; otherwise, the system considers that the user is created by a malicious user, reports an error, and exits.

```
SELECT copy_error_log_create() WHERE NOT EXISTS(SELECT * FROM pg_tables WHERE
schemaname='public' AND tablename='pgxc_copy_error_log');
GRANT INSERT,SELECT,DELETE ON  public.pgxc_copy_error_log To load_user;
```

If **enableSeparationOfDuty** is set to **on**, the user can be a common database user or an administrator. Create the pgxc_copy_error_log and gs_copy_summary tables in their respective schemas and add indexes. No permission granting is required.

1. Create a user as the initial user.
```
CREATE USER load_user WITH PASSWORD '********';
```
2. Switch to the new user as the initial user.
```
\c - load_user
```
3. Create a gs_copy_summary table and add an index.
```
CREATE TABLE load_user.gs_copy_summary(relname varchar, begintime timestamptz, endtime
timestamptz, id bigint, pid bigint, readrows bigint, skiprows bigint, loadrows bigint, errorrows bigint,
whenrows bigint, allnullrows bigint, detail text);
CREATE INDEX gs_copy_summary_idx ON load_user.gs_copy_summary(id);
```
4. (Optional) Create a pgxc_copy_error_log table and add an index.

NOTE

1. If the GUC parameter **enable_copy_error_log** is not set (**off** by default) or is set to **off**, you do not need to use the error table and do not need to create it. Otherwise, you need to create the error table.

```
CREATE TABLE load_user.pgxc_copy_error_log (relname varchar, begintime timestamptz, filename
varchar, lineno int8, rawrecord text, detail text);
CREATE INDEX copy_error_log_relname_idx ON load_user.pgxc_copy_error_log(relname);
```

Usage Environment

You need to add the tool path to *PATH*. gs_loader supports SSL encrypted communication. The method of using gs_loader is the same as that of using gsql.

Adding System Catalogs

The gs_copy_summary table is added to record the COPY execution result summary, including the number of successful rows, number of error rows, number of ignored rows, and number of empty rows.

The copy_summary_create function is added to create the gs_copy_summary table.

The format of the gs_copy_summary table is as follows:

```
relname     | public.sqlldr_tbl
begintime   | 2021-09-03 16:00:11.7129-04
endtime     | 2021-09-03 16:00:15.259908-04
id          | 21870
pid         | 47582725060352
readrows    | 100000
skiprows    | 0
```

```
loadrows    | 111
errorrows   | 0
whenrows    | 99889
allnullrows | 0
detail      | 111 Rows successfully loaded.
            | 0 Rows not loaded due to data errors.
            | 99889 Rows not loaded because all WHEN clauses were failed.
            | 0 Rows not loaded because all fields were null.
            |
```

Columns in the gs_copy_summary System Catalog

Table 2-1 gs_copy_summary columns

Column	Description
relname	Name of the target table to be imported.
begintime	Start time of an import task.
endtime	End time of an import task.
id	ID of the transaction to be imported.
pid	ID of the worker thread for the current import.
readrows	Total number of data rows read by the import task.
skiprows	Total number of data rows skipped in the import task.
loadrows	Number of data rows successfully imported in the current import task.
errorrows	Number of error data rows in the current import task.
whenrows	Number of data rows that violate the WHEN filter criterion in the current import task.
allnullrows	Number of data rows where all columns are empty.
detail	Summary of the import task, including the number of successfully imported rows, number of error data rows, number of rows that violate the WHEN condition, and number of blank rows.

Usage Guidelines

Step 1 (If the separation of duties function is disabled) For common users only:

1. Create a user (as an administrator).
```
CREATE USER load_user WITH PASSWORD '************';
```
2. Grant the public schema permission to the user (as an administrator).
```
GRANT ALL ON SCHEMA public TO load_user;
```
3. Create the gs_copy_summary table and grant table permissions to the user (as an administrator).
```
SELECT copy_summary_create() WHERE NOT EXISTS(SELECT * FROM pg_tables WHERE
schemaname='public' AND tablename='gs_copy_summary');
GRANT ALL PRIVILEGES ON  public.gs_copy_summary To load_user;
```

4. (Optional) Create the pgxc_copy_error_log table and grant table permissions to the user (as an administrator).

   NOTE

   If the GUC parameter **enable_copy_error_log** is not set (**off** by default) or is set to **off**, you do not need to use the error table and do not need to create it. Otherwise, you need to create the error table.

   ```
   SELECT copy_error_log_create() WHERE NOT EXISTS(SELECT * FROM pg_tables WHERE
   schemaname='public' AND tablename='pgxc_copy_error_log');
   GRANT ALL PRIVILEGES ON  public.pgxc_copy_error_log To load_user;
   ```

5. Switch to another user.

   ```
   \c - load_user
   ```

Step 2 (If the separation of duties function is enabled) For common users and administrators:

1. Create a user (as the initial user).

   ```
   CREATE USER load_user WITH PASSWORD '********';
   ```

2. Switch to the **load_user** user (as the initial user).

   ```
   \c - load_user
   ```

3. Create a gs_copy_summary table and add an index.

   ```
   CREATE TABLE load_user.gs_copy_summary(relname varchar, begintime timestamptz, endtime
   timestamptz, id bigint, pid bigint, readrows bigint, skiprows bigint, loadrows bigint, errorrows bigint,
   whenrows bigint, allnullrows bigint, detail text);
   CREATE INDEX gs_copy_summary_idx ON load_user.gs_copy_summary(id);
   ```

4. (Optional) Create a pgxc_copy_error_log table and add an index.

   NOTE

   If the GUC parameter **enable_copy_error_log** is not set (**off** by default) or is set to **off**, you do not need to use the error table and do not need to create it. Otherwise, you need to create the error table.

   ```
   CREATE TABLE load_user.pgxc_copy_error_log (relname varchar, begintime timestamptz, filename
   varchar, lineno int8, rawrecord text, detail text);
   CREATE INDEX copy_error_log_relname_idx ON load_user.pgxc_copy_error_log(relname);
   ```

Step 3 Create a table and a control file, and prepare a data file.

Create the **loader_tbl** table.

```
CREATE TABLE  loader_tbl
(
    ID   NUMBER,
    NAME VARCHAR2(20),
    CON  VARCHAR2(20),
    DT   DATE
);
```

(On the gs_loader client) Create the control file **loader.ctl**.

```
LOAD DATA
truncate into table loader_tbl
WHEN (2:2) = ','
fields terminated by ','
trailing nullcols
(
    id integer external,
    name char(32),
    con ":id || '-' || :name",
    dt date
)
```

(On the gs_loader client) Create the GUC parameter file **guc.txt**.

```
set a_format_copy_version='s1';
```

(On the gs_loader client) Create the data file **data.csv**.

```
1,OK,,2007-07-8
2,OK,,2008-07-8
3,OK,,2009-07-8
4,OK,,2007-07-8
43,DISCARD,,2007-07-8
,,,
32,DISCARD,,2007-07-8
a,ERROR int,,2007-07-8
8,ERROR date,,2007-37-8
,,,,
 ,
8,ERROR fields,,2007-37-8
,,,
5,OK,,2021-07-30
```

Step 4 Import the data.

(On the gs_loader client) Before importing data, ensure that the gs_loader tool has the execute permission on the gs_loader tool. Ensure that the current path has the write permission on files. (The gs_loader generates some temporary files during the processing and automatically deletes them after the import is complete.)

```
gs_loader control=loader.ctl data=data.csv db=testdb bad=loader.bad guc_param=guc.txt errors=5
port=8000 passwd=************ user=load_user
```

Execution result:

```
gs_loader: version 0.1

 5 Rows successfully loaded.

log file is:
 loader.log
```

----End

NOTE

gs_copy_summary is used to record the called COPY syntax and its details. *[badfile]*_**bad.log** is used to record error data and its details. To prevent the error data and details recorded during the last import from being overwritten, you are advised to use different bad parameters for each import. If the error table pgxc_copy_error_log is used to record error data and details, enable the GUC parameter **enable_copy_error_log**. To delete data from a table, perform the TRUNCATE or DELETE operation on the table.

Parameters

Table 2-2 gs_loader parameters

Parameter	Description	Parameter Type: Value Range
help	Displays help information.	-
user	Database connection user (equivalent to **-U**).	String

Parameter	Description	Parameter Type: Value Range
-U	Database connection user (equivalent to **user**).	String
passwd	User password (equivalent to **-W**).	String
-W	User password (equivalent to **passwd**).	String
db	Database name. This parameter is required and is equivalent to **-d**.	String
-d	Database name. This parameter is required and is equivalent to **db**.	String
host	Host name of the running server, the path of the Unix domain socket, or the domain name. You can specify multiple host addresses by using character strings separated by commas (,). This parameter is equivalent to **-h**. If multiple host addresses are specified, the primary node is connected by default.	See the **gsql --host** parameter.
-h	Host name of the running server, the path of the Unix domain socket, or the domain name. You can specify multiple host addresses by using character strings separated by commas (,). This parameter is equivalent to **host**. If multiple host addresses are specified, the primary node is connected by default.	See the **gsql --host** parameter.
port	Port number of the database server. One or more port numbers can be configured. When one port number is configured, all IP addresses use the same port for connection. When multiple port numbers are configured, the sequence is the same as the IP address sequence, and the number of port numbers must be the same as the number of IP addresses. If they are different, an error is reported. This parameter is equivalent to **-p**.	See the **gsql --port** parameter.

Parameter	Description	Parameter Type: Value Range
-p	Port number of the database server. One or more port numbers can be configured. When one port number is configured, all IP addresses use the same port for connection. When multiple port numbers are configured, the sequence is the same as the IP address sequence, and the number of port numbers must be the same as the number of IP addresses. If they are different, an error is reported. This parameter is equivalent to **port**.	See the **gsql --port** parameter.
create	Specifies whether to create the pgxc_copy_error_log and gs_copy_summary tables.	The value can be **true** or **false**. The default value is **true**.
clean	Specifies whether to clear the error record.	The value can be **true** or **false**. The default value is **false**.
data	(Required) Data file. You can specify multiple data files or use wildcards (*) and question marks (?) to represent multiple data files.	String
control	(Required) Name of a control file.	String
log	Name of a log file.	String
bad	Name of the file that records the error lines and details. You can also specify a directory. If you do not specify a directory, the file is generated based on the data file name.	String
discard	Name of the file recording the lines that fail to be matched by WHEN. You can also specify a directory to generate the file name based on the data file name.	String
errors	Maximum number of error lines in a data file.	Integer. Default value: **0**.
skip	Number of first lines that can be skipped in a data file.	Integer. Default value: **0**.
bindsize	Only syntax compatibility is implemented, but functions are not implemented.	-

Parameter	Description	Parameter Type: Value Range
rows	Number of rows of data to be imported before a commit.	The value is an integer, in the range [1,2147483647].
compatible_nul	Specifies whether to enable the compatibility of null characters (0x00) in data. After this function is enabled, if null characters exist in a data file, the null characters are converted to space characters (0x20), and then the data file is processed and imported.	The value can be **true** or **false**. The default value is **true**.
binary	Specifies whether the binary file is exported in COPY binary mode.	The value can be **true** or **false**. The default value is **false**.

CAUTION

- All parameters are in lowercase and are compatible with the gsql login mode, including **-p** port number, **-h** host, **-d** database, **-U** username, and **-W** password.
- When the **rows** parameter is specified, the number of commit times cannot exceed 1000. Otherwise, the performance is affected. The number of commit times is approximately equal to the number of data rows in the data file divided by the value of **rows**. If the **rows** parameter is not specified, there is no default value for **rows**. In this case, the transaction is committed only once after all data is imported to the table.
- Frequent commit of a small amount of data affects the data import performance. You are advised to set the **rows** parameter properly to ensure that the amount of data committed each time is greater than 5 MB. For common 16U 128 GB servers, in the scenario where one primary node and two standby nodes are deployed and 13 GB data is imported to a table with five columns. The rate of multiple commits is about 10 MB/s, which is basically the same as that of a single commit (5 MB data is committed each time).
- The **compatible_nul** parameter controls the value of the GUC parameter **loader_support_nul_character**.
  - **compatible_nul=true** corresponds to session-level **set loader_support_nul_charchter='s2'**.
  - **compatible_nul=false** corresponds to session-level **set loader_support_nul_character='s1'**.

    You are advised to set this parameter using the CLI. The priority of setting this parameter using **compatible_nul** is higher than that of setting this parameter using guc_param.
- Currently, gs_loader supports compatibility only when data files contain nul characters. It does not support nul characters in .ctl control files. If the .ctl file contains the nul character, unexpected problems may occur.
- After the **binary** parameter is set to **true**, the following requirements must be met:
  - The data file must be a binary file exported in \COPY binary mode. However, the data file exported in this mode has poor compatibility and portability. You are advised to use \COPY to import the data file.
  - gs_loader converts the syntax in the control file to the simplest syntax in \COPY binary mode, that is, \COPY *table_name* FROM *'binary_file_path'* BINARY;. Only the import mode, table name, as well as control, data, binary, guc_param, and database connection parameters in the control file are parsed. Other parameters are not parsed and do not take effect.
  - The command lines and control files of gs_loader must meet the following requirements:
    - Character set configuration is not supported.
    - The WHEN filter and DISCARD operation are not supported.
    - Error data cannot be directly written to BAD files when **enable_copy_error_log** is set to **off**. The default value of **errors** is **unlimited**, indicating that encoding exception data is recorded by default.

- The CSV mode is not supported, delimiters and wrappers cannot be specified, and the TRAILING NULLCOLS syntax is not supported.
- Data type configuration, POSITION configuration, and column expression usage are not supported.
- The **FILLER**, **CONSTANT**, **SEQUENCE**, and **NULLIF** parameters are not supported.
- The **skip**, **rows**, and **compatible_nul** parameters are not supported.

Control Files

- Syntax

```
LOAD [ DATA ]
[CHARACTERSET char_set_name]
[INFILE [directory_path] [filename ] ]
[BADFILE [directory_path] [filename ] ]
[OPTIONS(name=value)]
[{ INSERT | APPEND | REPLACE | TRUNCATE }]
INTO TABLE table_name
[{ INSERT | APPEND | REPLACE | TRUNCATE }]
[FIELDS CSV]
[TERMINATED [BY] { 'string' }]
[OPTIONALLY ENCLOSED BY { 'string' }]
[TRAILING NULLCOLS]
[ WHEN { (start:end) | column_name } {= | !=} 'string' ]
[(
col_name [ [ POSITION ({ start:end }) ]  ["sql_string"] ] | [ FILLER [column_type [external] ] ] |
[ CONSTANT "string" ] | [ SEQUENCE ( { COUNT | MAX | integer } [, incr] ) ]|[NULLIF (COL=BLANKS)]
[, ...]
)]
```

- Parameter description:
  - **CHARACTERSET**

    Character set.

    Value range: a string. Currently, the value can be **'AL32UTF8'**, **'zhs16gbk'**, or **'zhs32gb18030'**.

    Note: The character set specified by **CHARACTERSET** in the control file must be the same as the encoding format of the file. Otherwise, an error is reported or garbled characters are displayed in the imported data.
  - **INFILE**

    The current keyword is invalid and needs to occupy a separate line in the control file. The keyword is ignored during running. You need to specify the corresponding data file in the gs_loader command line parameters.
  - **BADFILE**

    The current keyword is invalid and will be ignored during running. If no .bad file is specified in the gs_loader command, a .bad file will be generated based on the name of the corresponding control file.
  - **OPTIONS**

    Only the **skip** and **rows** parameters take effect. **skip=***n* indicates that the first *n* records are skipped during import, and **rows=***n* indicates the number of rows to be imported before a commit. If both the command line and control file are specified, the command line has a higher priority.
  - **INSERT | APPEND | REPLACE | TRUNCATE**

    Import mode.

INSERT: If the table contains data, an error is reported.

APPEND: Data is inserted directly.

REPLACE: If the table contains data, all data is deleted and then inserted.

TRUNCATE: If the table contains data, all data is deleted and then inserted.

NOTE

- When writing a control file (.ctl), you can specify the import mode (**INSERT | APPEND | REPLACE | TRUNCATE**) before and after the INTO TABLE table_name statement. The priority is as follows: The import mode specified after the statement takes precedence over and overwrites that specified before the statement.
- When multiple gs_loader sessions are started to concurrently import data to the same table, you are advised to use the **APPEND** mode. If you use the **INSERT**, **REPLACE**, or **TRUNCATE** mode, an import error may occur or the imported data may be incomplete.

– **FIELDS CSV**

  Specifies that the CSV mode of COPY is used. In CSV mode, the default separator is a comma (,), and the default quotation mark is a double quotation mark (").

  NOTE

  In the current CSV mode, newlines that are enclosed with double quotation marks are considered as part of the column data.

– **table_name**

  Specifies the name (possibly schema-qualified) of an existing table.

  Value range: an existing table name

– **TERMINATED [BY] { 'string' }**

  The string that separates columns within each row (line) of the file, and it cannot be larger than 10 bytes.

  Value range: The value cannot include any of the following characters: \.abcdefghijklmnopqrstuvwxyz0123456789 The nul character cannot be set as a separator.

  Value range: The default value is a tab character in text format and a comma in CSV format.

  CAUTION

  After enabling nul character compatibility (**compatible_nul=true**), if the specified separator is a space character (0x20), note that the separator is the space character that exists in the data file instead of the space character converted from the nul character.

– **OPTIONALLY ENCLOSED BY { 'string' }**

  Specifies a quoted character string for a CSV file.

  The default value is double quotation marks (") only in CSV mode that is explicitly specified by the **FIELDS CSV** parameter.

  In other modes, there is no default value.

NOTE

- When you set **OPTIONALLY ENCLOSED BY { 'string' }**, there should be no quotation mark on the left of the data; otherwise, the number of quotation marks on either left or right must be an odd number but they do not have to be equal.
- Currently, **OPTIONALLY ENCLOSED BY { 'string' }** is supported only in CSV mode. If **OPTIONALLY ENCLOSED BY { 'string' }** is specified, the system enters the CSV mode by default.

– **TRAILING NULLCOLS**

  Specifies how to handle the problem that multiple columns of a row in a source data file are lost during data import.

  If one or more columns at the end of a row are null, the columns are imported to the table as null values. If this parameter is not set, an error message is displayed, indicating that the error column is null. In this case, the data in this row is processed as an error.

– **WHEN { (start:end) | column_name } {= | !=}**

  Filters rows by character string between **start** and **end** or by column name.

  Value range: a string.

  NOTE

  - When the GUC parameter **enable_copy_when_filler** is set to **on** (default value), data can be filtered based on the **FILLER** column. When the GUC parameter **enable_copy_when_filler** is set to **off**, this usage is not supported.
  - The WHEN condition cannot be followed by special characters such as '\0' and '\r'.

– **POSITION ({ start:end })**

  Processes columns and obtain the corresponding character strings between **start** and **end**.

– **"sql_string"**

  Processes columns and calculates column values based on column expressions. For details, see **• Column expression**.

  Value range: a string.

– **FILLER**

  Processes columns. If FILLER occurs, this column is skipped.

  NOTE

  Currently, FILLER and POSITION ({ *start:end* }) cannot be used at the same time.

– **column_type [external]**

  Processes the imported data according to different data types. For details, see **• Data types**.

– **CONSTANT**

  Processes columns and sets the inserted columns to constants.

  Value range: a string.

– **SEQUENCE ({ COUNT | MAX | integer } [, incr])**

  Processes columns to generate the corresponding sequence values.

- **COUNT**: The count starts based on the number of rows in the table.
- **MAX**: The count starts from the maximum value of this column in the table.
- **integer**: The count starts from the specified value.
- **incr**: indicates the increment each time.

– **NULLIF**

Processes columns. In multi-row import scenarios, if sysdate, constant, position, or column expression is not specified after a column name, the column whose NULLIF keyword is not specified is left empty.

Currently, only the COL POSITION() CHAR NULLIF (COL=BLANKS) syntax is supported. For details, see **• NULLIF use cases**.

CAUTION

- OPTIONS, INFILE, and BADFILE are not supported. Syntax errors are not reported only in specific scenarios.
- gs_loader uses a .bad file to record errors from the rawrecord column in an error table if the guc parameter **enable_copy_error_log** is set to enable the error table. The error table does not record rawrecord if an error cannot be read by certain code. In this case, a blank line is recorded in the .bad file.
- When gs_loader sets the GUC parameter **a_format_load_with_constraints_violation** to support non-rollback upon constraint conflicts, if a table has a BEFORE/AFTER ROW INSERT trigger, a maximum of 10,000,000 rows can be committed at a time.
- gs_loader does not support statement-level triggers when the GUC parameter **a_format_load_with_constraints_violation** is set to support non-rollback upon constraint conflicts.

- If the data in the .bad file is empty, refer to the source file and row number in the error table. (The code sequence is not identified, the .bad file content is not written, and only blank rows are recorded.)

```
loader=# select * from pgxc_copy_error_log;
       relname        |          begintime            | filename | lineno | rawrecord |
detail
---------------------+-----------------------------+----------+--------+-----------
+-----------------------------------------------
 public.test_gsloader | 2023-02-09 09:20:33.646843-05 | STDIN    |      1 |           | invalid byte sequence
for encoding "UTF8": 0xb4
(1 row)
// In the preceding example, for the file corresponding to the loader, search for the first row of the
data text to find the source data.
```

- NULLIF use cases

```
// Create a table.
create table gsloader_test_nullif(
col1   varchar2(100) not null enable,
col2   number(5,0) not null enable,
col3   varchar2(200) not null enable,
col4   varchar2(34) not null enable,
col5   varchar2(750),
col6   number(20,0),
col7   varchar2(4000),
col8   varchar2(200)
);
// Data file test.csv
6007 17060072021-09-0360070001102010000000230         1        600700010000218              0
1        1         229465        3
6007 17060072021-09-0360070001102010000000299         1        600700010000282              0
1        1         230467        3
6007 17060072021-09-0360070001102010000000242         1        600700010000255              0
1        1         226400        3
6007 17060072021-09-0360070001102010000000202         1        600700010000288              0
1        1         219107        3
6007 17060072021-09-0360070001102010000000294         1        600700010000243              0
1        1         204404        3
6007 17060072021-09-0360070001102010000000217         1        600700010000270              0
1        1         226644        3
// Control file test.ctl
LOAD DATA
CHARACTERSET UTF8
TRUNCATE
INTO TABLE gsloader_test_nullif
TRAILING NULLCOLS
(COL1 POSITION(1:10) CHAR NULLIF (COL1 = BLANKS),
COL2  POSITION(11:14) CHAR NULLIF (COL2 = BLANKS),
COL3  POSITION(21:30) CHAR NULLIF (COL3 = BLANKS),
COL4  POSITION(31:40) CHAR NULLIF (COL4 = BLANKS),
COL5  sysdate,
COL6,
COL7,
COL8 POSITION(71:80) CHAR NULLIF (COL8 = BLANKS))
// Import data.
gs_loader -p xxx host=xxx control=test.ctl  data=test.csv -d testdb -W xxx
// Result: Imported.
loader=# select * from gsloader_test_nullif;
    col1    | col2 |    col3    |    col4    |        col5         | col6 | col7 |   col8
------------+------+------------+------------+---------------------+------+------+-----------
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000218
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000282
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000255
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000288
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000243
 6007 17060 |  720 | 0360070001 | 1020100000 | 2023-05-17 20:32:03 |      |      | 010000270
(6 rows)
```

According to the data in the imported table, after the NULLIF keyword is used, the imported columns are normal except for the columns with the specified NULLIF and sysdate calculations. The imported columns without specified calculations are empty.

- Column expression

  gs_loader supports expression conversion and scenario extension for specified columns.

  ```
  ({ column_name [ data_type ] [ AS transform_expr ] } [, ...])
  ```

  data_type specifies the data type of the column in the expression parameter. **transform_expr** specifies the target expression and returns the result value whose data type is the same as that of the target column in the table.

  Example:

  – The column type is not specified in the .ctl file, and the source data does not meet the column restrictions (data type and length restrictions) in the table.

    ```
    // Create a table.
    create table t_test(id int, text varchar(5));
    // Data file test.csv
    addf2,bbbbaaa,20220907,
    // Control file test.ctl
    ```

```
Load Data
TRUNCATE INTO TABLE t_test
fields terminated by ','
TRAILING NULLCOLS(
id "length(trim(:id))",
text "replace(trim(:text),'bbbb','aa')"
)
// guc_param file
set a_format_copy_version='s1';
// Import data.
gs_loader -p xxx host=xxx control=test.ctl  data=test.csv -d testdb -W xxx guc_param=test_guc.txt
// Result: Imported.
select * from t_test;
 id | text
----+-------
  5 | aaaaa
(1 row)
```

– The column type is not specified in the .ctl file, and the implicit type conversion is performed. (You are advised to add compatibility parameters because the implicit type conversion is involved.)

```
// Create a table.
create table test(mes int, mes1 text, mes2 float8, mes3 timestamp with time zone, mes4
INTEGER);
// Data file
cat load_support_transform.data
1,mmoo,12.6789,Thu Jan 01 15:04:28 1970 PST,32767
2,yyds,180.883,Thu Jun 21 19:00:00 2012 PDT,32768
// Control file
cat load_support_transform.ctl
Load Data
TRUNCATE INTO TABLE test
fields terminated by ','
TRAILING NULLCOLS(
mes,
mes1 "mes1 || mes2",
mes2 "mes2 + 1",
mes3 "date_trunc('year', mes3)",
mes4
)
// guc_param file
cat test_guc.txt
set a_format_copy_version='s1';
set a_format_dev_version='s2';
set a_format_version='10c';
// Import data.
gs_loader -p xxx host=xxx control=load_support_transform.ctl data=load_support_transform.data
-d testdb -W xxx guc_param=test_guc.txt
// Result: Imported.
select * from test;
 mes |    mes1     |  mes2   |          mes3          | mes4
-----+-------------+---------+------------------------+-------
   1 | mmoo12.6789 | 13.6789 | 1970-01-01 00:00:00+08 | 32767
   2 | yyds180.883 | 181.883 | 2012-01-01 00:00:00+08 | 32768
```

- Data types

  Correspond to **column_type [external]** in the control file. During data loading, data is processed based on the data type. gs_loader classifies data types into common and special data types.

  – Common data types

    ■ CHAR [(length)]:

      Reads data based on a column separator and converts the value to the CHAR type. **length** indicates the maximum length of a single piece of data, in bytes. Generally, one character occupies one byte. The value can be left blank. The scenarios are as follows:

- If a length is not declared, the value inherits the maximum length value declared by **POSITION**.
- If a length is declared, it overwrites the maximum length declared by **POSITION**.
- If neither **length** nor **POSITION** declares a length, the value is set based on the length between separators.
- The priority of the length declaration is as follows: length > POSITION > separator.
- The default length, POSITION, and delimiter are read from the current position to the line terminator.
- If the actual data length exceeds the maximum value declared by **length**, an error is reported.

- INTEGER external [(length)]:

  Reads data based on a column separator and converts the value to the INTEGER type. The rules for using **length** are the same as those described in "CHAR [(length)]."

- FLOAT external [(length)]:

  Reads data based on a column separator and converts the value to the FLOAT type. The rules for using **length** are the same as those described in "CHAR [(length)]."

- DECIMAL external (length):

  Reads data based on a column separator and converts the value to the DECIMAL type. The rules for using **length** are the same as those described in "CHAR [(length)]."

- TIMESTAMP:

  Reads data based on a column separator and converts the value to the TIMESTAMP type.

- DATE:

  Reads data based on a column separator and converts the value to the DATE type.

- DATE external:

  Reads data based on a column separator and converts the value to the DATE type.

- SYSDATE:

  Obtains the system time when the corresponding insertion is performed in the database. The value cannot be referenced. The referenced content is the SYSDATE character string.

– Special data types

- INTEGER:

  Ignores the column separator, reads four-byte characters, saves them based on the little-endian storage logic, parses each character into a hexadecimal ASCII code value, and converts the value into a decimal number.

- SMALLINT:

  Ignores the column separator, reads two-byte characters, saves them based on the little-endian storage logic, parses each character into a hexadecimal ASCII code value, and converts the value into a decimal number.

  Example:

```
// Create a table.
create table t_spec(col1 varchar(10), col2 varchar(10));
// Data file
cat t_spec.txt
1234,5678,
// Control file
cat t_spec.ctl
Load Data
TRUNCATE INTO TABLE t_spec
fields terminated by ','
TRAILING NULLCOLS(
col1 position(2:6) integer,
col2 position(5:8) smallint
)
// guc_param file
cat test_guc.txt
set a_format_copy_version='s1';
set a_format_dev_version='s2';
set a_format_version='10c';
// Import data.
gs_loader -p xxx host=xxx control=t_spec.ctl data=t_spec.txt -d testdb -W xxx
guc_param=test_guc.txt
// Result: Imported.
select * from t_spec;
   col1    | col2
-----------+-------
 741618482 | 13612
(1 row)
```

- RAW:

  Parses each character into an ASCII code value. The backslash (\) is not used as an escape character.

  Restriction: Separators cannot be used in RAW data.

  Example:

```
// Create a table.
create table t_raw(col raw(50));
// Data file
cat t_raw.txt
12\n\x78!<~?'k^(%s)>/c[$50]
// Control file
cat t_raw.ctl
Load Data
TRUNCATE INTO TABLE t_raw
TRAILING NULLCOLS(
col position(1:50) raw
)
// guc_param file
cat test_guc.txt
set a_format_copy_version='s1';
set a_format_dev_version='s2';
set a_format_version='10c';
// Import data.
gs_loader -p xxx host=xxx control=t_raw.ctl data=t_raw.txt -d testdb -W xxx
guc_param=test_guc.txt
// Result: Imported.
select * from t_raw;
                          col
--------------------------------------------------------
```

```
 31325C6E5C783738213C7E3F276B5E282573293E2F635B2435305D
(1 row)
```

CAUTION

- In the multi-column import scenario, if the GUC parameter is not specified, some positions and separators cannot be used at the same time.
- In the multi-column import scenario, the SYSDATE and CONSTANT operations cannot be used together with the POSITION operation.
- When importing data of a specified data type, you need to use **guc_param** to set **a_format_copy_version** for common data types and use **guc_param** to set **a_format_copy_version**, **a_format_dev_version**, and **a_format_version** for special data types.
- If a column expression involves a system function, you need to use **guc_param** to set **a_format_dev_version** and **a_format_version** based on the corresponding function.
- If the data type contains **length**, the value of **length** must be set to an integer greater than 0. The special data type RAW(length) is used differently from common types. For example, if **POSITION** is not specified for the common type **INTEGER EXTERNAL(length)**, an error is reported when the value of **length** is less than the data length of the corresponding column in a text file (such as .csv or .txt). If the value of **length** is greater than the data length of the corresponding column in a text file (such as .txt), the result of the INTEGER EXTERNAL type is output. If **POSITION** is not specified for the special data type **RAW(length)**, the first *length* characters are read.
- If **POSITION(start:end)** is specified, the value of **start** must be set to an integer greater than 0, and the value of **end** must be greater than or equal to the value of **start**.
- When POSITION is specified, spaces at the end of a column are not omitted when the column content is processed. If POSITION is not specified, spaces at the end of the column content will be omitted. To retain spaces, ensure that **a_format_version** has been set and add **set behavior_compat_options='char_coerce_compat';** to the file specified by **guc_param**. For details, see "behavior_compat_options" in the *Administrator Guide*.
- During concurrent import, if multiple names of files specified by **discard** or **bad** point to files with the same name in the same directory, gs_loader stops importing the next file and reports an error. If a previous file has been imported, the file will be overwritten.

  The following error is reported:

  ```
  ERROR: An error occurred. Please check logfile.
  ```

  In the log file:

  ```
  …lock failed: Resource temporarily unavailable…
  ```

- If the column value in the control file is not empty and the column content is not used, the location of the data file is not occupied.

  For example, the control file is as follows:

  ```
  Load Data
  TRUNCATE INTO TABLE gsloader
  fields terminated by ','
  TRAILING NULLCOLS(
  id "trim(:id)",
  ```

```
text "to_char(SYSDATE,'yyyymmdd')",
gmt_create  "trim(:gmt_create)",
create_str "trim(:create_str)"
)
```

The data file is as follows:

```
11,22,33,
```

The import result is as follows:

```
loader=# select * from gsloader;
id |  text  |      gmt_create       | create_str
----+--------+---------------------+------------
11 | 2023-02-08 16:00:54 | 22 |  33
```

3 gs_dump

Context

gs_dump, provided by GaussDB, is used to export database information. You can export a database or its objects (such as schemas, tables, and views). The database can be the default **postgres** database or a user-specified database.

When gs_dump is used to export data, other users can still access (read or write) the database.

gs_dump can export complete, consistent data. For example, if gs_dump is started to export database A at T1, data of the database at that time point will be exported, and modifications on the database after that time point will not be exported.

The generated columns are not dumped during gs_dump is used.

gs_dump supports the export of text files that are compatible with the V1 database.

gs_dump can export database information to a plain-text SQL script file or archive file.

- Plain-text SQL script: It contains the SQL statements required to restore the database. You can use **gsql** to execute the SQL script. With only a little modification, the SQL script can rebuild a database on other hosts or database products.
- Archive file: It contains data required to restore the database. It can be a tar-, directory-, or custom-format archive. For details, see **Table 3-1**.

gs_dump supports SSL encrypted communication. The method is the same as that of using gsql.

Before using gs_dump, ensure that the gs_dump version is consistent with the database version. gs_dump of a later version may not be fully compatible with kernel data of an earlier version.

Main Functions

gs_dump can create export files in four formats, which are specified by [**-F** or **--format=**], as listed in **Table 3-1**.

Table 3-1 Formats of exported files

Format	Value of -F	Description	Suggestion	Import Tool
Plain-text	p	A plain-text script file containing SQL statements and commands. The commands can be executed on gsql, a command line terminal, to rebuild database objects and load table data.	You are advised to use plain-text export files for small databases.	Before using **gsql** to restore database objects, you can use a text editor to edit the plain-text export file as needed.
Custom	c	A binary file that allows the restoration of all or selected database objects from an exported file.	You are advised to use custom-format archive files for medium-or large-sized database.	You can use **gs_restore** to import database objects from a custom-, directory-, or TAR-format archive.
Directory	d	A directory containing directory files and the data files of tables and BLOB objects.	-	
.tar archive	t	A tar-format archive that allows the restoration of all or selected database objects from an exported file. It cannot be further compressed and has an 8-GB limitation on the size of a single table.	-	

NOTE

To reduce the size of an exported file, you can use the gs_dump tool to compress it to a directory archive file or custom-format file. When a directory archive or custom-format archive is generated, medium-level compression is applied by default. Archived exported files cannot be compressed using gs_dump.

Precautions

- Do not modify the files and contents exported using the **-F c/d/t** formats. Otherwise, the restoration may fail. For files exported using the **-F p** format, you can edit the exported files with caution if necessary.

- To ensure the data consistency and integrity, gs_dump acquires a share lock on a table to be dumped. If a share lock has been set for the table in other transactions, gs_dump locks the table after it is released. If the table cannot be locked within the specified time, the dump fails. You can customize the timeout interval to wait for lock release by specifying the **--lock-wait-timeout** parameter.
- Storing procedures and functions cannot be exported in encrypted mode.
- For materialized views, this tool supports only definition export. After importing materialized views, you need to manually run the REFRESH statement to restore data.
- For temporary objects, this tool can export only global temporary tables.
- This tool cannot be used on standby nodes.
- When gs_dump is used to export partitioned indexes, the attributes of some index partitions cannot be exported, for example, the unusable status of index partitions. You can query the PG_PARTITION system catalog or ADM_IND_PARTITIONS or ADM_IND_SUBPARTITIONS view to obtain the attributes of an index partition. You can run the ALTER INDEX statement to manually set the attributes of an index partition.
- For scheduled tasks, this tool can export only scheduled tasks created using CREATE EVENT or non-periodic scheduled tasks created using advanced packages from a B-compatible database.

NOTE

Common users cannot export subscriptions, directories, and synonyms. If a common user attempts to export related data, the message "WARNING: xx not dumped because current user is not a superuser" is displayed.

Syntax

```
gs_dump [OPTION]... [DBNAME]
```

NOTE

DBNAME does not follow a short or long option. It specifies the database to connect to.

For example:

Specify DBNAME without a -d option preceding it.

```
gs_dump -p port_number  testdb -f dump1.sql
```

Or,

```
export PGDATABASE=testdb
 gs_dump -p port_number -f dump1.sql
```

Environment variable: *PGDATABASE*

Parameters

Common parameters:

- -f, --file=<FILE_NAME>

  Sends the output to the specified file or directory. If this parameter is omitted, the standard output is generated. If the output format is (-F c/-F d/-F t), the -f parameter must be specified. If the value of the **-f** parameter contains a directory, the directory has the read and write permissions to the current user.

- -F, --format=c|d|t|p

  Selects an exported file format. The formats are as follows:

  – p|plain: generates a text SQL script file. This is the default value.
  – c|custom: outputs a custom-format archive as a directory to be used as the input of gs_restore. This is the most flexible output format in which users can manually select it and reorder the archived items during the restoration process. An archive in this format is compressed by default.
  – d|directory: creates a directory containing directory files and the data files of tables and BLOB objects.
  – t|tar: outputs a tar-format archive as the input of gs_restore. The .tar format is compatible with the directory format. Extracting a .tar archive generates a valid directory-format archive. However, the .tar archive cannot be further compressed and has an 8-GB limitation on the size of a single table. The order of table data items cannot be changed during restoration.

- -v, --verbose

  Specifies the verbose mode. If it is specified, gs_dump writes detailed object comments and number of startups/stops to the dump file, and progress messages to standard errors.

- -V, --version

  Prints the gs_dump version and exits.

- -Z, --compress=0-9

  Specifies the used compression level.

  Value range: 0 to 9

  – 0 indicates no compression.
  – 1 indicates the lowest compression ratio and the fastest processing speed.
  – 9 indicates the highest compression ratio and the slowest processing speed.

  For the custom-format archive, this option specifies the compression level of a single table data segment. By default, data is compressed at a medium level. The .tar archive format and plain-text format do not support compression currently.

- --lock-wait-timeout=TIMEOUT

  Do not keep waiting to obtain shared table locks at the beginning of the dump. Consider it as failed if you are unable to lock a table within the specified time. The timeout interval can be specified in any of the formats accepted by SET statement_timeout.

- -?, --help

  Displays help information about gs_dump parameters and exits.

Dump parameters:

- -a, --data-only

  Generates only the data, not the schema (data definition). Dumps the table data, big objects, and sequence values.

- -b, --blobs

  Reserved for function extension. The option is not recommended.

- -c, --clean

  Before writing the command of creating database objects into the backup files, writes the command of cleaning (deleting) database objects to the backup files. (If no objects exist in the target database, gsql or gs_restore probably displays some error information.)

  This parameter is used only for the plain-text format. For the archive format, you can specify the option when using gs_restore.

- -C, --create

  Specifies that the backup starts from the command for creating and connecting to a database. (If a command script is executed in this mode, it does not matter which database is connected before the script is executed.)

  This parameter is used only for the plain-text format. For the archive format, you can specify the option when using gs_restore.

- -E, --encoding=ENCODING

  Creates the dump in the specified character set encoding. By default, the dump file is created in the database encoding. (Alternatively, you can set the environment variable PGCLIENTENCODING to the required dump encoding.)

- -n, --schema=SCHEMA

  Dumps only schemas matching the schema names. This option contains the schema and all its contained objects. If this option is not specified, all non-system schemas in the target database will be dumped. Multiple schemas can be selected by specifying multiple -n options. The schema parameter is interpreted as a pattern according to the same rule used by the \d command of gsql. Therefore, multiple schemas can also be selected by writing wildcard characters in the pattern. When you use wildcard characters, quote patterns to prevent the shell from expanding the wildcard characters.

  NOTE

  - If -n is specified, gs_dump does not dump any other database objects that the selected schemas might depend upon. Therefore, there is no guarantee that the results of a specific-schema dump can be automatically restored to an empty database.
  - If -n is specified, the non-schema objects are not dumped.
  - In M-compatibility mode, if a database with **templatem** is created by running **CREATE DATABASE**, data is exported by specifying **db_name**; if a database is created by running **CREATE DATABASE** *db_name*, data is exported by specifying **-n** because such database is equivalent to a schema.

  Multiple schemas can be dumped. Entering **-n** *schemaname* multiple times dumps multiple schemas.

  For example:

  ```
  gs_dump -h host_name -p port_number testdb -f backup/bkp_shl2.sql -n sch1 -n sch2
  ```

  In the preceding example, sch1 and sch2 are dumped.

- -N, --exclude-schema=SCHEMA

  Does not dump any schemes matching the schema pattern. The pattern is interpreted according to the same rule as for **-n**. **-N** can be specified multiple times to exclude schemas matching any of the specified patterns.

  When both **-n** and **-N** are specified, the schemas that match at least one **-n** option but no **-N** is dumped. If **-N** is specified and **-n** is not, the schemas matching **-N** are excluded from what is normally dumped.

Dump allows you to exclude multiple schemas during dumping.

Specifies **-N** exclude schema name to exclude multiple schemas while dumping.

For example:

```
gs_dump -h host_name -p port_number testdb -f backup/bkp_shl2.sql -N sch1 -N sch2
```

In the preceding example, sch1 and sch2 will be excluded during the dumping.

- -o, --oids

  Dumps OIDs as parts of the data in each table. Use this option if your application references the OID columns in some way (for example, in a foreign key constraint). If the preceding situation does not occur, do not use this parameter.

- -O, --no-owner

  Do not output commands to set ownership of objects to match the original database. By default, gs_dump issues the ALTER OWNER or SET SESSION AUTHORIZATION command to set ownership of created database objects. These statements will fail when the script is running unless it is started by a system administrator (or the same user who owns all of the objects in the script). To make a script that can be stored by any user and give the user ownership of all objects, specify **-O**.

  This parameter is used only for the plain-text format. For the archive format, you can specify the option when using gs_restore.

- -s, --schema-only

  Dumps only the object definition (schema) but not data.

- -S, --sysadmin=NAME

  Reserved for function extension. The option is not recommended.

- -t, --table=TABLE

  Specifies a list of tables, views, sequences, or foreign tables not to be dumped. You can use multiple **-t** parameters or wildcard characters to specify tables.

  When you use wildcard characters, quote patterns to prevent the shell from expanding the wildcard characters.

  The **-n** and **-N** options have no effect when **-t** is used, because tables selected by using **-t** will be dumped regardless of those options, and non-table objects will not be dumped.

NOTE

- The number of **-t** parameters must be less than or equal to 100.
- If the number of **-t** parameters is greater than 100, you are advised to use the **--include-table-file** parameter to replace some **-t** parameters.
- If **-t** is specified, gs_dump does not dump any other database objects that the selected tables might depend upon. Therefore, there is no guarantee that the results of a specific-table dump can be automatically restored to an empty database.
- **-t** *tablename* only dumps visible tables in the default search path. **-t *.***tablename* dumps tablename tables in all the schemas of the dumped database. **-t** *schema.table* dumps tables in a specific schema.
- **-t** *tablename* does not export trigger information from a table.
- If the table name contains uppercase letters, you need to add \" to the table name when using the **-t** parameter to specify the export. To export the **"abC"** table, specify **-t \"abC\"**. To export the **schema."abC"** table, specify **-t schema.\"abC\"**.
- **-t ""** does not match any table.

For example:

```
gs_dump -h host_name -p port_number testdb -f backup/bkp_shl2.sql -t schema1.table1 -t
schema2.table2
```

In the preceding example, schema1.table1 and schema2.table2 are dumped.

- --include-table-file=<FILE_NAME>

  Specifies the table file to be dumped.

- -T, --exclude-table=TABLE

  Specifies a list of table, view, sequence, or foreign table objects not to be dumped. You can use multiple **-T** parameters or wildcard characters to specify multiple lists.

  When both **-t** and **-T** are specified, it will dump the objects in the **-t** list, but not those in the **-T** list.

  For example:

  ```
  gs_dump -h host_name -p port_number testdb -f backup/bkp_shl2.sql -T table1 -T table2
  ```

  In the preceding example, table1 and table2 are excluded from the dumping.

- --exclude-table-file=<FILE_NAME>

  Specifies the table file not to be dumped.

  NOTE

  Same as **--include-table-file**, the content format of this parameter is as follows:

  schema1.table1

  schema2.table2

  ......

- -x, --no-acl

  Prevents the dumping of access permissions (grant/revoke commands). Only ACL objects are affected. Privilege objects are not affected.

- -q, --target

  Exports text files compatible with databases of other versions. Currently, only the **v1** and **v5** can be specified. If other parameter values are specified, no error is reported but the setting does not take effect. The **v1** value means to export GaussDB v5 data as a text file compatible with GaussDB v1. The **v5** value means to export GaussDB v5 data as a text file, reducing errors that may occur when data is imported to a GaussDB v5 database.

When **v1** is specified, you are advised to use it along with parameters such as **--exclude-guc="enable_cluster_resize"**, **--exclude-function**, and **--exclude-with**. Otherwise, an error may be reported during data import to a GaussDB v1 database.

- -g, --exclude-guc

  Reserved for function extension. The option is not recommended.

- --exclude-function

  Specifies that functions and stored procedures are not exported.

- --exclude-with

  Specifies that the description such as **WITH(orientation=row, compression=on)** is not added to the end of the exported table definition.

- --binary-upgrade

  Reserved for function extension. The option is not recommended.

  NOTE

  M-compatible databases do not support this option.

- --binary-upgrade-usermap="USER1=USER2"

  Reserved for function extension. The option is not recommended.

- --column-inserts/--attribute-inserts

  Exports data by running the INSERT command with explicit column names {INSERT INTO table (column, ...) VALUES ...}. This will cause a slow restoration. However, since this option generates an independent command for each row, an error in reloading a row causes only the loss of the row rather than the entire table content.

  NOTE

  M-compatible databases do not support this option.

- --disable-dollar-quoting

  Disables the use of dollar sign ($) for function bodies, and forces them to be quoted using the SQL standard string syntax.

- --include-alter-table

  Deletes columns from the dumped table.

- --disable-triggers

  Reserved for function extension. The option is not recommended.

- --exclude-table-data=TABLE

  Does not dump data that matches any of table patterns. The pattern is interpreted according to the same rule as for **-t**.

  --exclude-table-data can be entered more than once to exclude tables matching any of several patterns. When the user needs the specified table definition rather than data in the table, this option is helpful.

  To exclude data of all tables in the database, see **-s, --schema-only**.

- --inserts

Dumps data when the INSERT statement (rather than COPY) is issued. This will cause a slow restoration.

However, since this option generates an independent command for each row, an error in reloading a row causes only the loss of the row rather than the entire table content. The restoration may fail if you rearrange the column order. The **--column-inserts** option is unaffected against column order changes, though even slower.

NOTE

M-compatible databases do not support this option.

- --no-security-labels

  Reserved for function extension. The option is not recommended.

- --no-tablespaces

  Does not select any tablespaces. All the objects will be created during the restoration process, no matter which tablespace is selected when using this option.

  This parameter is used only for the plain-text format. For the archive format, you can specify the option when using gs_restore.

- --no-unlogged-table-data

  Reserved for function extension. The option is not recommended.

- --non-lock-table

  This parameter is used only for inter-software API calling.

- --quote-all-identifiers

  Forcibly quotes all identifiers. This parameter is useful when you dump a database for migration to a later version, in which additional keywords may be introduced.

  NOTE

  M-compatible databases do not support this option.

- --section=SECTION

  Specifies dumped name sections (pre-data, data, or post-data).

- --serializable-deferrable

  Uses a serializable transaction for the dump to ensure that the used snapshot is consistent with later database status. Perform this operation at a time point in the transaction flow, at which everything is normal. This ensures successful transaction and avoids serialization failures of other transactions, which requires serialization again.

  This option has no benefits to disaster recovery. During the upgrade of the original database, loading a database copy as a report or other shared read-only dump is helpful. The option does not exist, dump reveals a status which is different from the submitted sequence status of any transaction.

  This option will make no difference if there are no active read-write transactions when gs_dump is started. If the read-write transactions are active, the dump start time will be delayed for an uncertain period.

- --use-set-session-authorization

  Specifies that the standard SQL SET SESSION AUTHORIZATION command rather than ALTER OWNER is generated to determine the object ownership.

This makes dumping more standard. However, if a dump file contains objects that have historical problems, restoration may fail. A dump using SET SESSION AUTHORIZATION requires the system administrator permissions, whereas ALTER OWNER requires lower permissions. However, the SET SESSION AUTHORIZATION statement supports ciphertext passwords. The script exported using this parameter may not be restored. Therefore, you are advised not to use this parameter to export the script.

NOTE

Application scope of SET SESSION AUTHORIZATION:

- The system administrator can switch to a common user through the SET SESSION AUTHORIZATION statement, but cannot switch to an initial user, sysadmin, opradmin, monadmin, poladmin, or auditadmin.
- Other users cannot switch to another user through the SET SESSION AUTHORIZATION statement.

- --with-encryption=AES128

  Specifies that dumping data needs to be encrypted using AES128.

- --with-key=KEY

  The AES128 key rules are as follows:

  - Consists of 8 to 16 characters.
  - Contains at least three of the following character types: uppercase characters, lowercase characters, digits, and special characters (limited to ~!@#$ %^&*()-_=+\|[{}];:,<.>/?).

  NOTE

  Storing procedures and functions cannot be exported in encrypted mode.

- --with-salt=RANDVALUES

  gs_dumpall uses this parameter to transfer a random value.

- --include-extensions

  Includes extensions in the dump.

  NOTICE

  The extended function is for internal use only. You are advised not to use it.

- --include-depend-objs

  Includes information about the objects that depend on the specified object in the backup result. This parameter takes effect only if the **-t** or **--include-table-file** parameter is specified.

- --exclude-self

  Excludes information about the specified object from the backup result. This parameter takes effect only if the **-t** or **--include-table-file** parameter is specified.

- --pipeline

  Uses a pipe to transmit the password. This parameter cannot be used on devices.

- --dont-overwrite-file

  The existing files in plain-text, .tar, or custom format will be overwritten. This option is not applicable to the directory format.

  For example:

  Assume that the backup.sql file exists in the current directory. If you specify **-f backup.sql** in the input command, and the backup.sql file is generated in the current directory, the original file will be overwritten.

  If the backup file exists and **--dont-overwrite-file** is specified, an error will be reported with the message that the dump file exists.

  ```
  gs_dump -p port_number testdb -f backup.sql -F plain --dont-overwrite-file
  ```

NOTE

- The **-s/--schema-only** and **-a/--data-only** options do not coexist.
- The **-c/--clean** and **-a/--data-only** parameters do not coexist.
- **--inserts/--column-inserts** and **-o/--oids** do not coexist, because OIDS cannot be set using the INSERT statement.
- **--role** must be used with **--rolepassword**.
- **--binary-upgrade-usermap** must be used with **--binary-upgrade**.
- **--include-depend-objs**/**--exclude-self** takes effect only when **-t**/**--include-table-file** is specified.
- **--exclude-self** must be used with **--include-depend-objs**.
- **--with-encryption=AES128** supports only **-F p/plain**.
- **--with-key=KEY** supports only **-F p/plain**.
- **--with-salt=RANDVALUES** is called by gs_dumpall and does not require manual input.

Connection parameters:

- -h, --host=HOSTNAME

  Specifies the host name. If the value begins with a slash (/), it is used as the directory for the Unix domain socket. The default is taken from the PGHOST environment variable (if available). Otherwise, a Unix domain socket connection is attempted.

  This parameter is valid only for the external database. For the local host in the database, only **127.0.0.1** can be used.

  Example: host name

  Environment variable: PGHOST

- -p, --port=PORT

  Specifies the host port number. If the thread pool is enabled, you are advised to use the pooler port, that is, the host port number plus 1.

  Environment variable: PGPORT

- -U, --username=NAME

  Specifies the username for connecting to a host. The initial user cannot be used for cross-node execution.

  Environment variable: PGUSER

- -w, --no-password

  Never issues a password prompt. The connection attempt fails if the host requires password for authentication and the password is not provided in other ways. This parameter is useful in batch jobs and scripts in which no user password is required.

- -W, --password=PASSWORD

  Specifies the user password for connection. If the authentication policy of the host is **trust**, the password of the system administrator is not verified. That is, you do not need to enter the **-W** option. If this parameter is not specified and you are not a system administrator, the system prompts you to enter the password in interactive mode. To ensure system security, you are advised to enter the password in interactive mode.

- --role=ROLENAME

  Specifies a role name to be used for creating the dump. If this option is selected, the SET ROLE command will be issued after the database is connected to gs_dump. It is useful when the authenticated user (specified by **-U**) lacks the permissions required by gs_dump. It allows the user to switch to a role with the required permissions. Some installations have a policy against logging in directly as a system administrator. This option allows dumping data without violating the policy.

- --rolepassword=ROLEPASSWORD

  Password of the role

Description

If your database has any local additions to the **template1** database, restore the output of the gs_dump tool into an empty database with caution. Otherwise, you are likely to obtain errors due to duplicate definitions of the added objects. To create an empty database without any local additions, copy data from template0 rather than template1. For example:

```
CREATE DATABASE foo WITH TEMPLATE template0;
```

The tar file size must be smaller than 8 GB. (This is the tar file format limitations.) The total size of a .tar archive and any of the other output formats are not limited, except possibly by the OS.

The dump file generated by gs_dump does not contain the statistics used by the optimizer to make execution plans. Therefore, you are advised to run **ANALYZE** after restoring from a dump file to ensure optimal performance. The dump file does not contain any ALTER DATABASE ... SET commands; these settings are dumped by gs_dumpall, along with database users and other installation settings.

Example

Use the gs_dump to dump a database as an SQL text file or a file in other formats.

In the following examples, **backup/MPPDB_backup.sql** indicates an exported file where backup indicates the relative path of the current directory. **37300** indicates the port ID of the database server. **testdb** indicates the name of the database to be accessed.

NOTE

Before exporting files, ensure that the directory exists and you have the read and write permissions on the directory.

Example 1: Use gs_dump to export the full information of the **testdb** database. The exported **MPPDB_backup.sql** file is in plain-text format.

```
gs_dump -U omm -f backup/MPPDB_backup.sql -p 37300 testdb -F p
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 09:49:17]: The total objects number is 356.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 09:49:17]: [100.00%] 356 objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 09:49:17]: dump database testdb successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 09:49:17]: total time: 1274  ms
```

Use gsql to import data from the export plain-text file.

Example 2: Use gs_dump to export the full information of the **testdb** database. The exported **MPPDB_backup.tar** file is in .tar format.

```
gs_dump -U omm -f backup/MPPDB_backup.tar -p 37300 testdb -F t
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:02:24]: The total objects number is 1369.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:02:53]: [100.00%] 1369 objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:02:53]: dump database testdb successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:02:53]: total time: 50086  ms
```

Example 3: Use gs_dump to export the full information of the **testdb** database. The exported **MPPDB_backup.dmp** file is in custom format.

```
gs_dump -U omm -f backup/MPPDB_backup.dmp -p 37300 testdb -F c
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:05:40]: The total objects number is 1369.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:06:03]: [100.00%] 1369 objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:06:03]: dump database testdb successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:06:03]: total time: 36620  ms
```

Example 4: Use gs_dump to export the full information of the **testdb** database. The exported **MPPDB_backup** file is in directory format.

```
gs_dump -U omm -f backup/MPPDB_backup -p 37300  testdb -F d
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:16:04]: The total objects number is 1369.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:16:23]: [100.00%] 1369 objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:16:23]: dump database testdb successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:16:23]: total time: 33977  ms
```

Example 5: Use gs_dump to export the information of the **testdb** database, excluding the information of the table specified in the **/home/MPPDB_temp.sql** file. The exported MPPDB_backup.sql file is in plain-text format.

```
gs_dump -U omm -p 37300 testdb --exclude-table-file=/home/MPPDB_temp.sql -f backup/MPPDB_backup.sql
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:37:01]: The total objects number is 1367.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:37:22]: [100.00%] 1367 objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:37:22]: dump database testdb successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-27 10:37:22]: total time: 37017  ms
```

Example 6: Use gs_dump to export only the information about the views that depend on the testtable table. Create another **testtable** table, and then restore the views that depend on it.

Back up only the views that depend on the **testtable** table.

```
gs_dump -U omm -s -p 37300 testdb -t PUBLIC.testtable --include-depend-objs --exclude-self -f backup/
MPPDB_backup.sql -F p
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-15 14:12:54]: The total objects number is
331.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-15 14:12:54]: [100.00%] 331 objects have
been dumped.
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-15 14:12:54]: dump database testdb
successfully
gs_dump[user='omm'][localhost][port='37300'][testdb][2018-06-15 14:12:54]: total time: 327  ms
```

Change the name of the **testtable** table.

```
gsql -p 37300 testdb -r -c "ALTER TABLE PUBLIC.testtable RENAME TO testtable_bak;"
```

Create another **testtable** table.

```
CREATE TABLE PUBLIC.testtable(a int, b int, c int);
```

Restore the views for the new **testtable** table.

```
gsql -p 37300 testdb -r -f backup/MPPDB_backup.sql
```

Related Commands

gs_dumpall

4 gs_dumpall

Context

gs_dumpall, provided by GaussDB, is used to export all database information, including the data of the default **postgres** database, data of user-specified databases, and global objects of all databases.

When gs_dumpall is used to export data, other users can still access (read or write) the database.

gs_dumpall can export complete, consistent data. For example, if gs_dumpall is started to export entire database at T1, data of the databases at that time point will be exported, and modifications on the databases after that time point will not be exported.

The generated columns are not dumped during gs_dumpall is used.

gs_dumpall exports all databases in two parts:

- gs_dumpall exports all global objects, including information about database users and groups, tablespaces, and attributes (for example, global access permissions).
- gs_dumpall calls gs_dump to export SQL scripts from each database, which contain all the SQL statements required to restore databases.

The exported files are both plain-text SQL scripts. Use **gsql** to execute them to restore databases.

gs_dumpall supports SSL encrypted communication. The method is the same as that of using gsql.

Before using gs_dumpall, ensure that the gs_dumpall version is consistent with the gs_dump version and database version. gs_dump and gs_dump of a later version may not be fully compatible with the kernel data of an earlier version.

Precautions

- Do not modify any exported file. Otherwise, restoration may fail.
- To ensure the data consistency and integrity, gs_dumpall sets a share lock for a table to be dumped. If a share lock has been set for the table in other transactions, gs_dumpall locks the table after it is released. If the table cannot

be locked within the specified time, the dump fails. You can customize the timeout interval to wait for lock release by specifying the **--lock-wait-timeout** parameter.

- During an export, gs_dumpall reads all tables in a database. Therefore, you need to connect to the database as a database administrator to export a complete file. When you use gsql to execute SQL scripts, administrator permissions are also required to add users and user groups, and create databases. Before importing a backup, you need to verify the security to prevent administrator permissions from being exploited.
- If you use gs_dumpall to export all database objects and want to import them to a new instance environment, ensure that the names and permissions of the users used for the export and import are the same. Otherwise, an error message will be displayed, indicating that the names are inconsistent or the permissions are insufficient.
- When files exported using gs_dumpall are imported to a new instance, the connected database must be the default database. Otherwise, syntax incompatibility may occur, resulting in import errors.
- Only one M-compatible database can be created globally. If an M-compatible database is exported, ensure that no other M-compatible database exists in the target instance environment.
- For scheduled tasks, this tool can export only scheduled tasks created using CREATE EVENT or non-periodic scheduled tasks created using advanced packages from a B-compatible database.

Syntax

```
gs_dumpall [OPTION]...
```

Parameters

Common parameters:

- -f, --filename=<FILE_NAME>

  Sends the output to the specified file. If this option is omitted, the standard output is used.
- -v, --verbose

  Specifies the verbose mode. This will cause gs_dumpall to output detailed object comments and start/stop times to the dump file, and progress messages to standard errors.
- -V, --version

  Prints the gs_dumpall version and exits.
- --lock-wait-timeout=TIMEOUT

  Do not keep waiting to obtain shared table locks at the beginning of the dump. Consider it as failed if you are unable to lock a table within the specified time. The timeout interval can be specified in any of the formats accepted by SET statement_timeout.
- -?, --help

  Displays help about the command line parameters for gs_dumpall and exits.

Dump parameters:

- -a, --data-only

  Dumps only the data, not the schema (data definition).
- -c, --clean

  Runs SQL statements to delete databases before rebuilding them. Statements for dumping roles and tablespaces are added.
- -g, --globals-only

  Dumps only global objects (roles and tablespaces) but no databases.
- -o, --oids

  Dumps OIDs as parts of the data in each table. Use this option if your application references the OID columns in some way (for example, in a foreign key constraint). If the preceding situation does not occur, do not use this parameter.
- -O, --no-owner

  Do not output commands to set ownership of objects to match the original database. By default, gs_dumpall issues the ALTER OWNER or SET SESSION AUTHORIZATION statement to set ownership of created schema elements. These statements will fail when the script is running unless it is started by a system administrator (or the same user who owns all of the objects in the script). By specifying **-O**, you can compile a script that can be stored by any user. The script grants the user the permission to own all objects because the ALTER OWNER or SET SESSION AUTHORIZATION statement is not used and the execute user permission is always used during the import. Therefore, before importing the dump file, check whether there are risks in the dump file. For example, check whether the dump file contains a privilege escalation statement and whether the statement is known to the administrator.
- -r, --roles-only

  Dumps only roles but not databases or tablespaces.
- -s, --schema-only

  Dumps only the object definition (schema) but not data.
- -S, --sysadmin=NAME

  Reserved for function extension. The option is not recommended.
- -t, --tablespaces-only

  Dumps only tablespaces but not databases or roles.
- -x, --no-privileges

  Prevents the dumping of access permissions (grant/revoke commands).
- --column-inserts/--attribute-inserts

  Exports data by running the INSERT command with explicit column names {INSERT INTO table (column, ...) VALUES ...}. This will cause a slow restoration. However, since this option generates an independent command for each row, an error in reloading a row causes only the loss of the row rather than the entire table content.

  NOTE

  M-compatible databases do not support this option. This option is skipped when an M-compatible database is exported.
- --disable-triggers

Reserved for function extension. The option is not recommended.

- --inserts

  Dumps data when the INSERT statement (rather than COPY) is issued. This will cause a slow restoration. The restoration may fail if you rearrange the column order. The --column-inserts parameter is safer against column order changes, though even slower.

  NOTE

  M-compatible databases do not support this option. This option is skipped when an M-compatible database is exported.

- --no-security-labels

  Reserved for function extension. The option is not recommended.

- --no-tablespaces

  Do not output statements to create tablespaces or select tablespaces for objects. All the objects will be created during the restoration process, no matter which tablespace is selected when using this option.

- --no-unlogged-table-data

  Reserved for function extension. The option is not recommended.

- --include-alter-table

  Exports information about deleted columns in the table.

- --quote-all-identifiers

  Forcibly quotes all identifiers. This parameter is useful when you dump a database for migration to a later version, in which additional keywords may be introduced.

  NOTE

  M-compatible databases do not support this option. This option is skipped when an M-compatible database is exported.

- --dont-overwrite-file

  Do not overwrite the current file.

- --use-set-session-authorization

  Specifies that the standard SQL SET SESSION AUTHORIZATION command rather than ALTER OWNER is generated to determine the object ownership. This makes dumping more standard. However, if a dump file contains objects that have historical problems, restoration may fail. A dump using SET SESSION AUTHORIZATION requires the system administrator permissions, whereas ALTER OWNER requires lower permissions. However, the SET SESSION AUTHORIZATION statement supports user and permission switching using a ciphertext password. The script exported using this parameter may not be restored. Therefore, you are advised not to use this parameter to export the script.

NOTE

Application scope of SET SESSION AUTHORIZATION:

- The system administrator can switch to a common user through the SET SESSION AUTHORIZATION statement, but cannot switch to an initial user, sysadmin, opradmin, monadmin, poladmin, or auditadmin.
- Other users cannot switch to another user through the SET SESSION AUTHORIZATION statement.

- --with-encryption=AES128

  Specifies that dumping data needs to be encrypted using AES128.
- --with-key=KEY

  The AES128 key rules are as follows:
  - Consists of 8 to 16 characters.
  - Contains at least three of the following character types: uppercase characters, lowercase characters, digits, and special characters (limited to ~!@#$ %^&*()-_=+\|[{}];:,<.>/?).
- --include-extensions

  Backs up all CREATE EXTENSION statements if the include-extensions parameter is set.

NOTICE

The extended function is for internal use only. You are advised not to use it.

- --include-templatedb

  Includes template databases during the dump.
- --binary-upgrade

  Reserved for function extension. The option is not recommended.

  NOTE

  M-compatible databases do not support this option. This option is skipped when an M-compatible database is exported.
- --binary-upgrade-usermap="USER1=USER2"
- Reserved for function extension. The option is not recommended.
- --non-lock-table

  This parameter is used only for inter-software API calling.
- --tablespaces-postfix

  Reserved for function extension. The option is not recommended.
- --parallel-jobs

  Specifies the number of concurrent backup processes. The value range is 1–1000.
- --pipeline

  Uses a pipe to transmit the password. This parameter cannot be used on devices.

NOTE

- The -g/--globals-only and -r/--roles-only parameters do not coexist.
- The **-g/--globals-only** and **-t/--tablespaces-only** parameters do not coexist.
- The -r/--roles-only and -t/--tablespaces-only parameters do not coexist.
- The -s/--schema-only and -a/--data-only parameters do not coexist.
- The -r/--roles-only and -a/--data-only parameters do not coexist.
- The -t/--tablespaces-only and -a/--data-only parameters do not coexist.
- The **-g/--globals-only** and **-a/--data-only parameters** do not coexist.
- --tablespaces-postfix must be used with --binary-upgrade.
- --binary-upgrade-usermap must be used with --binary-upgrade.
- --parallel-jobs must be used with -f/--file.

Connection parameters:

- -h, --host=HOSTNAME

  Specifies the host name. If the value begins with a slash (/), it is used as the directory for the Unix domain socket. The default value is taken from the PGHOST environment variable. If it is not set, a Unix domain socket connection is attempted.

  This parameter is valid only for the external database. For the local host in the database, only 127.0.0.1 can be used.

  Environment variable: *PGHOST*
- -l, --database=DATABASENAME

  Specifies the name of the database connected to dump all objects and discover other databases to be dumped. If this parameter is not specified, the **postgres** database will be used. If the **postgres** database does not exist, **template1** will be used.

  NOTE

  No matter which database is specified as the database to be connected for export, you must connect to **postgres** when using gsql to import data to a new instance. Otherwise, syntax incompatibility may occur, resulting in import errors. In particular, if the M-compatible database is incorrectly connected for import, the creation fails because CREATE DATABASE on the M-compatible database is equivalent to CREATE SCHEMA.
- -p, --port=PORT

  Specifies the TCP port listened on by the server or the local Unix domain socket file name extension to ensure a correct connection. The default value is the PGPORT environment variable.

  If the thread pool is enabled, you are advised to use the pooler port, that is, the listening port number plus 1.

  Environment variable: *PGPORT*
- -U, --username=NAME

  Specifies the name of the connected user. The initial user cannot be used for cross-node execution.

  Environment variable: *PGUSER*
- -w, --no-password

Never issues a password prompt. The connection attempt fails if the server requires password for authentication and the password is not provided in other ways. This parameter is useful in batch jobs and scripts in which no user password is required.

- -W, --password=PASSWORD

  Specifies the user password for connection. If the authentication policy of the host is **trust**, the password of the system administrator is not verified. That is, you do not need to enter the **-W** option. If this parameter is not specified and you are not a system administrator, the system prompts you to enter the password in interactive mode. To ensure system security, you are advised to enter the password in interactive mode.

- --role=ROLENAME

  Specifies a role name to be used for creating the dump. This option causes gs_dumpall to issue the SET ROLE statement after connecting to the database. It is useful when the authenticated user (specified by **-U**) lacks the permissions required by gs_dumpall. It allows the user to switch to a role with the required permissions. Some installations have a policy against logging in directly as a system administrator. This option allows dumping data without violating the policy.

- --rolepassword=ROLEPASSWORD

  Specifies the password of the specific role.

Description

- gs_dumpall internally calls gs_dump. For details about the diagnosis information, see **gs_dump**.
- Once gs_dumpall is restored, it is advised to run **ANALYZE** on each database so that the optimizer can provide useful statistics.
- gs_dumpall requires all needed tablespace directories to be empty before the restoration. Otherwise, database creation will fail if the databases are in non-default locations.

Example

Use gs_dumpall to export all databases at a time.

NOTE

gs_dumpall supports only plain-text format export. Therefore, only gsql can be used to restore a file exported using gs_dumpall.

```
gs_dumpall -U omm -f backup/bkp2.sql -p 37300
gs_dump[user='omm'][localhost][port='37300'][dbname='testdb'][2018-06-27 09:55:09]: The total objects
number is 2371.
gs_dump[user='omm'][localhost][port='37300'][dbname='testdb'][2018-06-27 09:55:35]: [100.00%] 2371
objects have been dumped.
gs_dump[user='omm'][localhost][port='37300'][dbname='testdb'][2018-06-27 09:55:46]: dump database
dbname='testdb' successfully
gs_dump[user='omm'][localhost][port='37300'][dbname='testdb'][2018-06-27 09:55:46]: total time: 55567
ms
gs_dumpall[user='omm'][localhost][port='37300'][2018-06-27 09:55:46]: dumpall operation successful
gs_dumpall[user='omm'][localhost][port='37300'][2018-06-27 09:55:46]: total time: 56088  ms
In the preceding command, backup/bkp2.sql indicates the exported file, 37300 indicates the port number
of the database server, and omm indicates the username.
```

Related Commands

gs_dump

5 gs_restore

Context

gs_restore, provided by GaussDB, is used to import data that was exported using gs_dump. It can also be used to import files exported by gs_dump.

It has the following functions:

- Imports data to the database.

  If a database is specified, data is imported to the database. For parallel import, the password for connecting to the database is required. During data import, the generated columns are automatically updated and saved as common columns.
- Imports data to an archive file.

  If the **-l** parameter is specified, an archive file containing a brief summary of the data is generated.

gs_restore supports SSL encrypted communication. The method is the same as that of using gsql.

Before using gs_restore, ensure that the gs_restore version is consistent with the gs_dump version and database version.

Format

```
gs_restore [OPTION]... FILE
```

NOTE

- The **FILE** does not have a short or long option. It is used to specify the location for the archive files.
- The **dbname** or **-l** option is required as prerequisites. Users cannot enter **dbname** and **-l** parameters at the same time.
- gs_restore incrementally imports data by default. To prevent data exception caused by consecutive imports, use the **-e** and **-c** parameters for each import. **-c** indicates that the database objects that already exist in the database to be restored are cleared (deleted) before the database objects are rebuilt. **-e** indicates that if an error occurs when an SQL statement is sent to the database, the system exits. By default, the system continues to import data and displays a series of error information after the import is complete.
- If the owner of the schema object has the OPRADMIN system permission, the initial user is required for the import.

Parameters

Common parameters:

- -d, --dbname=NAME

  Connects to the **dbname** database and imports data to the database.
- -f, --file=<FILE_NAME>

  Specifies the output file for the generated archive, or the output file in the list specified using **-l**.

  The default is the standard output.

  NOTE

  -f cannot be used with **-d**.
- -F, --format=c|d|t

  Specifies the format of the archive. The format does not need to be specified because the gs_restore determines the format automatically.

  Value range:

  - **c/custom**: The archive form is the customized format in gs_dump.
  - **d/directory**: The archive format is **directory**.
  - **t/tar**: The archive format is **tar**.
- -l, --list

  Lists the formats of the archive. The operation output can be used for the input of the **-L** option. If filtering parameters, such as **-n** or **-t**, are used together with **-l**, they will restrict the listed items.
- -v, --verbose

  Specifies the verbose mode.
- -V, --version

  Prints the gs_restore version and exits.
- -?, --help

  Displays help information about gs_restore parameters and exits.

Import parameters:

- -a, -data-only

  Imports only the data, not the schema (data definition). gs_restore incrementally imports data.
- -c, --clean

  Cleans (deletes) existing database objects in the database to be restored before recreating them. If the target database does not contain the objects involved in the deletion operation, some promptive error information may be displayed.
- -C, --create

  Before data is imported to a database, CREATE DATABASE is used to create the database. (After this option is specified, the database specified by **-d** is only used to execute the **CREATE DATABASE** command, and all data is still imported to the created database.)

- -e, --exit-on-error

  Exits if an error occurs when you send the SQL statement to the database. If you do not exit, the commands will still be sent and error information will be displayed when the import ends.

- -I, --index=NAME

  Imports only the definition of the specified index. Multiple indexes can be imported. Enter *-I index* multiple times to import multiple indexes.

  For example:

  ```
  gs_restore -h host_name -p port_number -d testdb -I Index1 -I Index2 backup/MPPDB_backup.tar
  ```

  In this example, **Index1** and **Index2** will be imported.

- -j, --jobs=NUM

  Specifies the number of concurrent, the most time-consuming jobs of gs_restore (such as loading data, creating indexes, or creating constraints). This parameter can greatly reduce the time to import a large database to a server running on a multi-processor machine.

  Each task may be a process or a thread, which is determined by the OS. Each task is connected to the server separately.

  The optimal value for this option depends on the server hardware settings, the client, the network, the number of CPU cores, and disk settings. It is recommended that the parameter be set to the number of CPU cores on the server. In addition, a larger value can also lead to faster import in many cases. However, an overly large value will lead to decreased performance because of thrashing.

  This option supports the customized archive format only. The input file must be a regular file (not the pipe file). This parameter can be ignored when you select the script method rather than connecting to a database server. In addition, multiple jobs cannot be used with the **--single-transaction** option.

  NOTE

  This parameter applies to multi-table, multi-index, and multi-constraint scenarios. In practice, the number of created processes (or threads) is related to the number of tables, indexes, and constraints. The maximum number of concurrent jobs does not exceed the specified number of jobs.

- -L, --use-list=<FILE_NAME>

  Imports only archive elements that are listed in **list-file** and imports them in the order that they appear in the file. If filtering parameters, such as **-n** or **-t**, are used with **-L**, they will further limit the items to be imported.

  list-file is normally created by editing the output of a previous **-l** parameter. File lines can be modified or removed, and can also be commented out by placing a semicolon (;) at the beginning of the row. An example is provided in this document.

- -n, --schema=NAME

  Restores only objects that are listed in schemas.

  This parameter can be used with the **-t** parameter to import a specific table.

  Entering **-n** *schemaname* multiple times can import multiple schemas.

  For example:

  ```
  gs_restore -h host_name -p port_number -d testdb -n sch1 -n sch2 backup/MPPDB_backup.tar
  ```

  In this example, **sch1** and **sch2** will be imported.

NOTE

- In M-compatibility mode, if a database with **templatem** is created by running **CREATE DATABASE**, data is imported by specifying **db_name** using the **-d** option; if a database is created by running **CREATE DATABASE** *db_name*, data is imported by specifying **-n** because such database is equivalent to a schema.
- The specified schema name must exist in the archive file that is used as the input. If a schema name that does not exist in the archive file is specified, the import of the schema does not take effect.

- -O, --no-owner

  Do not output commands to set ownership of objects to match the original database. By default, gs_restore issues the ALTER OWNER or SET SESSION AUTHORIZATION statement to set ownership of created schema elements. Unless the system administrator or the user who has all the objects in the script initially connects to the database. Otherwise, the statement will fail. Any username can be used for the initial connection using **-O**, and this user will own all the created objects.

- -P, --function=NAME(args)

  Imports only listed functions. You need to correctly spell the function name and the parameter based on the contents of the dump file in which the function exists.

  Entering **-P** alone means importing all function-name(args) functions in a file. Entering **-P** with **-n** means importing the function-name(args) functions in a specified schema. Entering **-P** multiple times and using **-n** once means that all imported functions are in the **-n** schema by default.

  You can enter **-n schema-name -P 'function-name(args)'** multiple times to import functions in specified schemas.

  For example:

  ```
  ./gs_restore -h host_name -p port_number -d testdb -n test1 -P 'Func1(integer)' -n test2 -P
  'Func2(integer)' backup/MPPDB_backup.tar
  ```

  In this example, both **Func1 (i integer)** in the **test1** schema and **Func2 (j integer)** in the **test2** schema will be imported.

- -s, --schema-only

  Imports only schemas (data definitions), instead of data (table content). The current sequence value will not be imported.

- -S, --sysadmin=NAME

  Reserved for function extension. The option is not recommended.

- -t, --table=NAME

  Imports only listed table definitions or data, or both. This parameter can be used with the **-n** parameter to specify a table object in a schema. When **-n** is not entered, the default schema is **PUBLIC**. Entering **-n** *schemaname* **-t** *tablename* multiple times can import multiple tables in a specified schema.

  For example:

  Import **table1** in the **PUBLIC** schema.

  ```
  gs_restore -h host_name -p port_number -d testdb -t table1 backup/MPPDB_backup.tar
  ```

  Import **test1** in the **test1** schema and **test2** in the **test2** schema.

  ```
  gs_restore -h host_name -p port_number -d testdb -n test1 -t test1 -n test2 -t test2 backup/
  MPPDB_backup.tar
  ```

  Import **table1** in the **PUBLIC** schema and **test1** in the **test1** schema.

```
gs_restore -h host_name -p port_number -d testdb -n PUBLIC -t table1 -n test1 -t table1 backup/
MPPDB_backup.tar
```

NOTICE

- **-t** does not support the *schema_name.table_name* input format. If this format is specified, no error is reported but the setting does not take effect.
- When **-t** is specified, gs_restore does not import any other database objects that are attached to the selected table. Therefore, there is no guarantee that the results of a specific-table dump can be automatically imported to an empty database.
- **-t tablename** does not import trigger information from a table.

- -T, --trigger=NAME

  This parameter is reserved for extension.
- -x, --no-privileges/--no-acl

  Prevents the import of access permissions (**GRANT**/**REVOKE** commands).
- -1, --single-transaction

  Executes import as a single transaction (that is, commands are wrapped in **BEGIN**/**COMMIT**).

  This option ensures that either all the commands are completed successfully, or no application is changed. This option means **--exit-on-error**.
- --disable-triggers

  Reserved for function extension. The option is not recommended.
- --no-data-for-failed-tables

  By default, table data will be imported even if the statement to create a table fails (for example, the table exists). Data in such table is skipped using this parameter. This operation is useful if the target database already contains the desired table contents.

  This parameter takes effect only when you import data directly to a database, not when you output SQL scripts.
- --no-security-labels

  Reserved for function extension. The option is not recommended.
- --no-tablespaces

  Does not select any tablespaces. All the objects will be created during the import process, no matter which tablespace is selected when using this option.
- --section=SECTION

  Imports the listed sections (such as pre-data, data, or post-data).
- --use-set-session-authorization

  This option is used for backing up the plain-text format.

  Generates the **SET SESSION AUTHORIZATION** statement as the output instead of the **ALTER OWNER** statement to determine object ownership. This parameter makes dump more standards-compatible. If the records of objects in exported files are referenced, import may fail. Only administrators can use the **SET SESSION AUTHORIZATION** statement to dump data, and the administrators must manually change and verify the passwords of exported files by referencing the **SET SESSION AUTHORIZATION** statement before import. The **ALTER OWNER** statement requires lower permissions.

- --pipeline

  Uses a pipe to transmit the password. This parameter cannot be used on devices.

NOTICE

- If any local additions need to add to the template1 database during the installation, restore the output of gs_restore into an empty database with caution. Otherwise, you are likely to obtain errors due to duplicate definitions of the added objects. To create an empty database without any local additions, copy data from template0 rather than template1. For example:

```
CREATE DATABASE foo WITH TEMPLATE template0;
```

- gs_restore cannot import large objects selectively. For example, it can only import the objects of a specified table. If an archive form contains large objects, all large objects will be imported. If this archive object is excluded by **-L**, **-t**, or other options, none of the large objects will be imported.

NOTE

1. The **-d/--dbname** and **-f/--file** parameters do not coexist.
2. The **-s/--schema-only** and **-a/--data-only** parameters do not coexist.
3. The **-c/--clean** and **-a/--data-only** parameters do not coexist.
4. When the **--single-transaction** option is used, **-j/--jobs** must be a single job.
5. --**role** must be used with --**rolepassword**.

Connection parameters:

- -h, --host=HOSTNAME

  Specifies the host name. If the value begins with a slash (/), it is used as the directory for the Unix domain socket. The default value is taken from the *PGHOST* environment variable. If it is not set, a Unix domain socket connection is attempted.

  This parameter is valid only for the external database. For the local host in the database, only 127.0.0.1 can be used.

  Environment variable: *PGHOST*

- -p, --port=PORT

  Specifies the TCP port listened on by the server or the local Unix domain socket file name extension to ensure a correct connection. The default value is the *PGPORT* environment variable.

  If the thread pool is enabled, you are advised to use pooler port, that is, the listening port number plus 1.

  Environment variable: *PGPORT*

- -U, --username=NAME

  Specifies the name of the connected user. The initial user cannot be used for cross-node execution.

Environment variable: *PGUSER*

- -w, --no-password

  Never issues a password prompt. The connection attempt fails if the server requires password for authentication and the password is not provided in other ways. This parameter is useful in batch jobs and scripts in which no user password is required.

- -W, --password=PASSWORD

  Specifies the user password for connection. If the authentication policy of the host is **trust**, the password of the system administrator is not verified. That is, you do not need to enter the **-W** parameter. If this parameter is not specified and you are not a system administrator, the system prompts you to enter the password in interactive mode. To ensure system security, you are advised to enter the password in interactive mode.

- --role=ROLENAME

  Specifies a role name for the import operation. If this parameter is selected, the SET ROLE statement will be issued after gs_restore connects to the database. It is useful when the authenticated user (specified by **-U**) lacks the permissions required by gs_restore. This parameter allows the user to switch to a role with the required permissions. Some installations have a policy against logging in directly as the initial user. This parameter allows data to be imported without violating the policy.

- --rolepassword=ROLEPASSWORD

  Specifies the password of the specific role.

Example

Special case: Execute the gsql tool. Import the **MPPDB_backup.sql** file in the export folder (in plain-text format) generated by gs_dump or gs_dumpall to the **testdb** database.

```
gsql -d testdb -p 8000 -f /home/omm/test/MPPDB_backup.sql
SET
SET
SET
SET
SET
ALTER TABLE
ALTER TABLE
ALTER TABLE
ALTER TABLE
ALTER TABLE
CREATE INDEX
CREATE INDEX
CREATE INDEX
SET
CREATE INDEX
REVOKE
REVOKE
GRANT
GRANT
total time: 30476  ms
In the example, the file after -f is the exported file, and 8000 indicates the port number of the database
server. testdb indicates the name of the database to be accessed.
```

gs_restore is used to import the files exported by gs_dump.

Example 1: Execute the gs_restore tool to import the exported **MPPDB_backup.dmp** file (custom format) to the **testdb** database.

```
gs_restore backup/MPPDB_backup.dmp -p 8000 -d testdb
restore operation successful
total time: 13053  ms
```

Example 2: Execute the gs_restore tool to import the exported **MPPDB_backup.tar** file (.tar format) to the **testdb** database.

```
gs_restore backup/MPPDB_backup.tar -p 8000 -d testdb
restore operation successful
total time: 21203  ms
```

Example 3: Execute the gs_restore tool to import the exported **MPPDB_backup** file (directory format) to the **testdb** database.

```
gs_restore backup/MPPDB_backup -p 8000 -d testdb
restore operation successful
total time: 21003  ms
```

Example 4: Execute the gs_restore tool and import the **MPPDB_backup.dmp** file (in custom format). Specifically, import all the object definitions and data in the **PUBLIC** schema. Existing objects are deleted from the target database before the import. If an existing object references to an object in another schema, you need to manually delete the referenced object first.

```
gs_restore backup/MPPDB_backup.dmp -p 8000 -d testdb -e -c -n PUBLIC
Error while PROCESSING TOC:
Error from TOC entry 313; 1259 337399 TABLE table1 gaussdba
could not execute query: ERROR:  cannot drop table table1 because other objects depend on it
DETAIL:  view t1.v1 depends on table table1
HINT:  Use DROP ... CASCADE to drop the dependent objects too.
   Command was: DROP TABLE IF EXISTS public.table1;
```

Manually delete the referenced object and create it again after the import is complete.

```
gs_restore backup/MPPDB_backup.dmp -p 8000 -d testdb -e -c -n PUBLIC
restore operation successful
total time: 2203  ms
```

Example 5: Execute the gs_restore tool and import the **MPPDB_backup.dmp** file (in custom format). Specifically, import only the definition of **table1** in the **PUBLIC** schema.

```
gs_restore backup/MPPDB_backup.dmp -p 8000 -d testdb -e -c -s -n PUBLIC -t table1
restore operation successful
total time: 21000  ms
```

Example 6: Execute the gs_restore tool and import the **MPPDB_backup.dmp** file (in custom format). Specifically, import only the data of **table1** in the **PUBLIC** schema.

```
gs_restore backup/MPPDB_backup.dmp -p 8000 -d testdb -e -a -n PUBLIC -t table1
restore operation successful
total time: 20203  ms
```

Related Commands

gs_dump and **gs_dumpall**