云数据库 GaussDB
V2.0-8.201
集中式版工具参考

文档版本 01

发布日期 2024-10-30

华为云计算技术有限公司

版权所有 © 华为云计算技术有限公司 2024。 保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明

和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目 录

1 变更说明 1
2 简介 2
3 客户端工具 3
3.1 gsql 3
3.1.1 gsql 概述 3
3.1.2 使用指导 10
3.1.3 获取帮助 12
3.1.4 命令参考 13
3.1.5 元命令参考 19
3.1.6 常见问题处理 34
3.2 gs_loader 38
3.3 gs_dump 55
3.4 gs_dumpall 68
3.5 gs_restore 74
4 服务端工具 82
4.1 GaussRoach.py 82
4.1.1 Roach 简介 82
4.1.2 命令行参考 82
4.1.2.1 注意事项 82
4.1.2.2 备份命令 83
4.1.2.3 停止备份命令 100
4.1.2.4 备份删除命令 103
4.1.2.5 恢复命令 106
4.1.2.6 genstack 命令 127
4.1.2.7 启动命令 129
4.1.2.8 清除命令 135
4.1.2.9 配置命令 138
4.1.2.10 归档命令 140
4.1.2.11 快照恢复命令 147
4.1.2.12 版本命令 151
4.1.2.13 帮助命令 152
4.1.2.14 日志找回命令 157

 版权所有 © 华为云计算技术有限公司

4.1.2.15 Roach 操作参数 162
4.2 gs_blackbox 164
4.3 gs_cgroup 165
4.4 gs_check 174
4.5 gs_checkos 195
4.6 gs_collector 203
4.7 gs_encrypt 207
4.8 gs_guc 209
4.9 gs_om 217
4.10 gs_replace 230
4.11 gs_roach 238
4.12 gs_ssh 244
5 系统内部调用的工具 246
5.1 etcd 246
5.2 etcdctl 255
5.3 gaussdb 258
5.4 gs_backup 264
5.5 gs_ctl 266
5.6 gs_finishredo_retrieve 281
5.7 gs_expand 282
5.8 gs_hotpatch 291
5.9 gs_install 292
5.10 gs_initdb 296
5.10.1 gs_initdb 简介 297
5.10.2 使用指导 297
5.10.3 命令参考 298
5.11 gs_persist 304
5.12 gs_postuninstall 305
5.13 gs_preinstall 307
5.14 gs_redis 312
5.15 gs_redis_bucket 314
5.16 gs_rescue 314
5.17 gs_rescue_tui 327
5.18 gs_shrink 337
5.19 gs_sshexkey 344
5.20 gs_switch_ddb 346
5.21 gs_uninstall 347
5.22 gs_upgradectl 348
5.23 gns_ctl 356
5.24 gstrace 358
5.25 kadmin.local 360
5.26 kdb5_util 360

 版权所有 © 华为云计算技术有限公司

5.27 kdestroy..... 361
5.28 kinit..... 361
5.29 klist..... 361
5.30 krb5kdc..... 361
5.31 om_monitor..... 362
5.32 omRollback.py..... 362
5.33 gs_controldata..... 362
5.34 gs_resetxlog..... 363
5.35 pscp..... 364
5.36 pssh..... 365
5.37 py_pstree.py..... 367
5.38 LogicalStandbyDR.py..... 367
6 统一数据库管理工具..... 369
6.1 特性介绍..... 369
6.2 cm_ctl 工具介绍..... 390
6.3 cm_ctl 工具使用指南..... 404
6.3.1 查询集群状态..... 404
6.3.2 启停集群..... 409
6.3.3 数据库主备倒换..... 410
6.3.4 重建备 DN..... 411
6.3.5 检测实例进程运行状态..... 412
6.3.6 查看数据库实例配置文件..... 412
6.3.7 设置、获取 CM_Server 相关参数..... 413
6.3.8 CM 参数相关操作..... 414
6.3.9 CM_Server 热补丁升级操作..... 415
6.3.10 加密..... 416
6.3.11 ddb 相关操作..... 416
6.3.12 设置 DCF 相关内容..... 418
6.3.13 一键停止仲裁..... 419
6.3.14 其他..... 420
6.4 安全设计..... 420
6.4.1 手动替换证书步骤..... 420
6.4.2 证书使用指南..... 421
6.4.3 ifconfig 提权步骤..... 421
6.4.4 IP 命令提权步骤..... 421
6.5 CM 配置参数介绍..... 422
6.5.1 cm_agent 参数..... 422
6.5.2 cm_server 参数..... 433
6.6 CM 错误日志信息参考..... 446
7 FAQ..... 456
7.1 执行命令报错“Failed to obtain the GPHOME”..... 456

 版权所有 © 华为云计算技术有限公司

8 gs_collector 工具支持收集的系统表和视图列表.....457

 版权所有 © 华为云计算技术有限公司

1 变更说明

- 统一数据库管理工具，cm_agent参数新增“cm_hadr_process_type”。相关文档**cm_hadr_process_type**。
- 统一数据库管理工具，cm_server参数新增“cm_hadr_process_type”。相关文档**cm_hadr_process_type**。

版权所有 © 华为云计算技术有限公司

2 简介

GaussDB提供客户端和服务器端应用程序（工具），帮助用户更好地维护GaussDB。
工具位于安装数据库服务器的$GPHOME/script和$GAUSSHOME/bin路径下。

版权所有 © 华为云计算技术有限公司

3 客户端工具

数据库在部署成功后，需要通过一些工具来便捷地连接数据库，对数据库进行各种调试和操作。GaussDB提供了一些数据库连接工具。通过这些工具可以方便地连接数据库并对数据库进行数据操作。

3.1 gsql

gsql是GaussDB提供在命令行下运行的数据库连接工具，可以通过此工具连接服务器并对其进行操作和维护，除了具备操作数据库的基本功能，gsql还提供了若干**高级特性**，便于用户使用。

3.1.1 gsql 概述

基本功能

- **连接数据库：**默认只支持从服务器本机连接，如果需要连接到远端的数据库，必须在服务端进行配置。详细操作请参见《开发指南》中“数据库使用入门 > 连接数据库 > 使用gsql连接 > 远程连接数据库”章节。

  说明

  gsql创建连接时，会有5分钟超时时间。如果在这个时间内，数据库未正确地接受连接并对身份进行认证，gsql将超时退出。

  针对此问题，可以参考**常见问题处理**。

- **执行SQL语句：**支持交互式地键入并执行SQL语句，也可以执行一个文件中指定的SQL语句。
- **执行元命令：**元命令可以帮助管理员查看数据库对象的信息、查询缓存区信息、格式化SQL输出结果，以及连接到新的数据库等。元命令的详细说明请参见**元命令参考**。

高级特性

gsql的高级特性如**表3-1**所示。

版权所有 © 华为云计算技术有限公司

表 3-1 gsql 高级特性

特性名称	描述
变量	gsql提供类似于Linux的shell命令的变量特性，可以使用gsql的元命令\set设置一个变量，格式如下： **\set** *varname value* 删除由\set命令设置的变量请使用如下方式： **\unset** *varname* **说明** • 变量只是简单的名称/值对，这里的值可以是任意长度。 • 变量名称必须由字母（包括非拉丁字母）、数字和下划线组成，且对大小写敏感。 • 如果使用\set varname的格式（不带第二个参数），则只是设置这个变量而没有给变量赋值。 • 可以使用不带参数的\set来显示所有变量的值。 变量的示例和详细说明请参见**变量**。
SQL代换	利用gsql的变量特性，可以将常用的SQL语句设置为变量，以简化操作。 SQL代换的示例和详细说明请参见**SQL代换**。
自定义提示符	gsql使用的提示符支持用户自定义。可以通过修改gsql预留的三个变量PROMPT1、PROMPT2、PROMPT3来改变提示符。 这三个变量的值可以用户自定义，也可以使用gsql预定义的值。详细请参见**提示符**。
客户端操作历史记录	gsql支持客户端操作历史记录，当客户端连接时指定“-r”参数，此功能被打开。可以通过\set设置记录历史的条数，例如，\set HISTSIZE 50，将记录历史的条数设置为50，\set HISTSIZE 0，不记录历史。 **说明** • 客户端操作历史记录条数默认设置为32条，最多支持记录500条。当客户端交互式输入包含中文字符时，只支持UTF-8 的编码环境。 • 出于安全考虑，将包含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、SECRET_KEY等字符串（不区分大小写）的SQL语句记录识别为包含敏感信息的语句，不会记录到历史信息中，即不能通过上下翻回显。

- 变量

  可以使用gsql元命令\set设置一个变量。例如把变量foo的值设置为bar：

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

  要引用变量的值，在变量前面加冒号。例如查看变量的值：

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

  这种变量的引用方法适用于规则的SQL语句和除\copy、\ef、\help、\sf、\!以外的元命令。

  gsql预定义了一些特殊变量，同时也规划了变量的取值。为了保证和后续版本最大限度地兼容，请避免以其他目的使用这些变量。所有特殊变量见**表3-2**。

版权所有 © 华为云计算技术有限公司

说明

- 所有特殊变量都由大写字母、数字和下划线组成。
- 要查看特殊变量的默认值，请使用元命令**\echo :***varname*（例如\echo :DBNAME）。

表 3-2 特殊变量设置

变量	设置方法	变量说明
DBNAME	\set DBNAME *dbname*	当前连接的数据库的名称。每次连接数据库时都会被重新设置。
ECHO	\set ECHO all | queries	• 如果设置为all，只显示查询信息。等效于使用gsql连接数据库时指定-a参数。 • 如果设置为queries，显示命令行和查询信息。等效于使用gsql连接数据库时指定-e参数。
ECHO_HIDDEN	\set ECHO_HIDDEN on | off | noexec	当使用元命令查询数据库信息（例如\dg）时，此变量的取值决定了查询的行为： • 设置为on，先显示元命令实际调用的查询语句，然后显示查询结果。等效于使用gsql连接数据库时指定-E参数。 • 设置为off，则只显示查询结果。 • 设置为noexec，则只显示查询信息，不执行查询操作。
ENCODING	\set ENCODING *encoding*	当前客户端的字符集编码。
FETCH_COUNT	\set FETCH_COUNT *variable*	• 如果该变量的值为大于0的整数，假设为n，则执行SELECT语句时每次从结果集中取n行到缓存并显示到屏幕。 • 如果不设置此变量，或设置的值小于等于0，则执行SELECT语句时一次性把结果都取到缓存。 **说明** 设置合理的变量值，将减少内存使用量。一般来说，设为100到1000之间的值比较合理。
HISTCONTROL	\set HISTCONTROL ignorespace | ignoredups | ignoreboth | none	• ignorespace：以空格开始的行将不会写入历史列表。 • ignoredups：与以前历史记录里匹配的行不会写入历史记录。 • ignoreboth、none或者其他值：所有以交互模式读入的行都被保存到历史列表。 **说明** none表示不设置HISTCONTROL。
HISTFILE	\set HISTFILE *filename*	此文件用于存储历史名列表。缺省值是~/.bash_history。

版权所有 © 华为云计算技术有限公司

变量	设置方法	变量说明
HISTSIZE	\set HISTSIZE *size*	保存在历史命令里命令的个数。缺省值是500。
HOST	\set HOST *hostname*	已连接的数据库主机名称。
IGNOREEOF	\set IGNOREEOF *variable*	● 若设置此变量为数值，假设为10，则在gsql中输入的前9次EOF字符（通常是Ctrl+C）都会被忽略，在第10次按Ctrl+C才能退出gsql程序。 ● 若设置此变量为非数值，则缺省为10。 ● 若删除此变量，则向交互的gsql会话发送一个EOF终止应用。
LASTOID	\set LASTOID *oid*	最后影响的oid值，即为从一条INSERT或lo_import命令返回的值。此变量只保证在下一条SQL语句的结果显示之前有效。
ON_ERROR_ROLLBACK	\set ON_ERROR_ROLLBACK on | interactive | off	● 如果是on，当一个事务块里的语句产生错误的时候，这个错误将被忽略而事务继续。 ● 如果是interactive，这样的错误只是在交互的会话里忽略。 ● 如果是off（缺省），事务块里一个语句生成的错误将会回滚整个事务。on_error_rollback-on模式是通过在一个事务块的每个命令前隐含地发出一个SAVEPOINT的方式来工作的，在发生错误的时候回滚到该事务块。
ON_ERROR_STOP	\set ON_ERROR_STOP on | off	● on：命令执行错误时会立即停止，在交互模式下，gsql会立即返回已执行命令的结果。 ● off（缺省）：命令执行错误时将会跳过错误继续执行。
PORT	\set PORT *port*	当前连接数据库的端口号。
USER	\set USER *username*	当前用于连接的数据库用户。
VERBOSITY	\set VERBOSITY terse | default | verbose	这个选项可以设置为值terse、default、verbose之一以控制错误报告的冗余行。 ● terse：仅返回严重且主要的错误文本以及文本位置（一般适合于单行错误信息）。 ● default：返回严重且主要的错误文本及其位置，还包括详细的错误细节、错误提示（可能会跨越多行）。 ● verbose：返回所有的错误信息。

版权所有 © 华为云计算技术有限公司

- SQL代换

  像元命令的参数一样，gsql变量的一个关键特性是可以把gsql变量替换成正规的SQL语句。此外，gsql还提供为变量更换新的别名或其他标识符等功能。使用SQL代换方式替换一个变量的值可在变量前加冒号。例如：

```
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)
```

  执行以上命令，将会查询HR.areaS表。

须知

变量的值是逐字复制的，甚至可以包含不对称的引号或反斜杠命令。所以必须保证输入的内容有意义。

- 提示符

  通过表3-3的三个变量可以设置gsql的提示符，这些变量是由字符和特殊的转义字符所组成。

表 3-3 提示符变量

变量	描述	示例

PROMPT1	gsql请求一个新命令时使用的正常提示符。
	PROMPT1的默认值为：
	`%/%R%#`
		使用变量PROMPT1切换提示符：
		• 提示符变为[local]：
		`gaussdb=> \set PROMPT1 %M`
		`[local:/tmp/gaussdba_mppdb]`
		• 提示符变为name：
		`gaussdb=> \set PROMPT1 name`
		`name`
		• 提示符变为=：
		`gaussdb=> \set PROMPT1 %R`
		`=`

PROMPT2	在一个命令输入期待更多输入时（例如，查询没有用一个分号结束或者引号不完整）显示的提示符。	使用变量PROMPT2显示提示符：
		`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))`

版权所有 © 华为云计算技术有限公司

变量	描述	示例
PROMPT3	当执行COPY命令，并期望在终端输入数据时（例如，COPY FROM STDIN），显示提示符。	使用变量PROMPT3显示COPY提示符： 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 >>>>\.

提示符变量的值是按实际字符显示的，但是，当设置提示符的命令中出现“%”时，变量的值根据“%”后的字符，替换为已定义的内容，已定义的提示符请参见表3-4。

表 3-4 已定义的替换

符号	符号说明
%M	主机的全名（包含域名），若连接是通过Unix域套接字进行的，则全名为[local]，若Unix域套接字不是编译的缺省位置，就是[local:/dir/name]。
%m	主机名删去第一个点后面的部分。若通过Unix域套接字连接，则为[local]。
%>	主机正在侦听的端口号。
%n	数据库会话的用户名。
%/	当前数据库名称。
%~	类似 %/，如果数据库是缺省数据库时输出的是波浪线~。
%#	如果会话用户是数据库系统管理员，使用#，否则用>。
%R	• 对于PROMPT1通常是“=”，如果是单行模式则是“^”，如果会话与数据库断开（如果\connect失败可能发生）则是“!”。 • 对于PROMPT2该序列被“ -”、“ *”、单引号、双引号或“$”(取决于gsql是否等待更多的输入：查询没有终止、正在一个 /* ... */ 注释里、正在引号或者美元符扩展里)代替。
%x	事务状态： • 如果不在事务块里，则是一个空字符串。 • 如果在事务块里，则是“*”。 • 如果在一个失败的事务块里则是“!”。 • 如果无法判断事务状态时为“?”（比如没有连接）。
%digits	指定字节值的字符将被替换到该位置。
%:name	gsql变量“name”的值。

版权所有 © 华为云计算技术有限公司

符号	符号说明
%command	command的输出，类似于使用“^”替换。
%[. . . %]	提示可以包含终端控制字符，这些字符可以改变颜色、背景、提示文本的风格、终端窗口的标题。例如， gaussdb=> \set PROMPT1 '%[%033[1;33;40m%]%n@%/%R%[%033[0m%]%#' 这个句式的结果是在VT100兼容的可显示彩色的终端上的一个宽体（1;）黑底黄字（33;40）。

环境变量

表 3-5 与 gsql 相关的环境变量

名称	描述
COLUMNS	如果\set columns为0，则由此参数控制wrapped格式的宽度。这个宽度用于决定在自动扩展的模式下，是否要把宽输出模式变成竖线的格式。
PAGER	如果查询结果无法在一页显示，它们就会被重定向到这个命令。可以用\pset命令关闭分页器。典型的是用命令more或less来实现逐页查看。缺省值是平台相关的。 **说明** less的文本显示，受系统环境变量LC_CTYPE影响。
PSQL_EDITOR	\e和\ef命令使用环境变量指定的编辑器。变量是按照列出的先后顺序检查的。在Unix系统上默认的编辑工具是vi。
EDITOR	
VISUAL	
PSQL_EDITOR_LINENUMBER_ARG	当\e和\ef带上一行数字参数使用时，这个变量指定的命令行参数用于向编辑器传递起始行数。像Emacs或vi这样的编辑器，这只是个加号。如果选项和行号之间需要空白，在变量的值后加一个空格。例如： PSQL_EDITOR_LINENUMBER_ARG = '+' PSQL_EDITOR_LINENUMBER_ARG='--line ' Unix系统默认的是+。
PSQLRC	用户的.gsqlrc文件的交互位置。
SHELL	使用\!命令跟shell执行的命令是一样的效果。
TMPDIR	存储临时文件的目录。缺省是/tmp。

版权所有 © 华为云计算技术有限公司

3.1.2 使用指导

前提条件

- 连接数据库时使用的用户需要具备访问数据库的权限。
- gsql须与数据库版本配套。

背景信息

使用gsql命令可以连接本机的数据库服务，也可以连接远程数据库服务。连接远程数据库服务时，需要在服务器上设置允许远程连接，详细操作请参见《开发指南》中“数据库使用入门 > 连接数据库 > 使用gsql连接 > 远程连接数据库”章节。

操作步骤

步骤1 使用gsql连接到GaussDB服务器。

gsql工具使用-d参数指定目标数据库名、-U参数指定数据库用户名、-h参数指定主机名、-p参数指定端口号信息。

说明

若未指定数据库名称，则使用初始化时默认生成的数据库名称；若未指定数据库用户名，则默认使用当前操作系统用户作为数据库用户名；当某个值没有前面的参数（-d、-U等）时，若连接的命令中没有指定数据库名（-d）则该参数会被解释成数据库名；如果已经指定数据库名（-d）而没有指定数据库用户名（-U）时，该参数则会被解释成数据库用户名。

示例1，使用omm用户连接到本机gaussdb数据库的8000端口。

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

示例2，使用jack用户连接到远程主机gaussdb数据库的8000端口。

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

集中式数据库实例中，连接主DataNode时可以把DataNode的IP地址使用逗号分割全部添加到-h后，gsql将依次从前往后连接每个IP地址，查询当前DataNode是否为主DataNode，如果不是则断开连接尝试下一个IP地址，直到找到主DataNode为止。

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

示例3，参数gaussdb和omm不属于任何选项时，分别被解释成了数据库名和用户名。

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

等效于

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

详细的gsql参数请参见**命令参考**。

步骤2 执行SQL语句。

以创建数据库human_staff为例。

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

通常，输入的命令行在遇到分号的时候结束。如果输入的命令行没有错误，结果就会输出到屏幕上。

版权所有 © 华为云计算技术有限公司

步骤3 执行gsql元命令。

以列出GaussDB中所有的数据库和描述信息为例。

```
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)
```

更多gsql元命令请参见**元命令参考**。

----结束

示例

以把一个查询分成多行输入为例。注意提示符的变化：

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

查看表的定义：

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

向HR.areaS表插入四行数据：

```
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
```

切换提示符：

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

查看表：

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

可以用\pset命令以不同的方法显示表：

 版权所有 © 华为云计算技术有限公司

```
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)
```

使用元命令：

```
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=#
```

3.1.3 获取帮助

操作步骤

- 连接数据库时，可以使用如下命令获取帮助信息。

  ```
  gsql --help
  ```

  显示如下帮助信息：

  ```
  ......
  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
  ......
  ```

- 连接到数据库后，可以使用如下命令获取帮助信息。

  ```
  help
  ```

  显示如下帮助信息：

  ```
  You are using gsql, the command-line interface to gaussdb.
  Type:  \copyright for distribution terms
         \h for help with SQL commands
  ```

 版权所有 © 华为云计算技术有限公司

```
\? for help with gsql commands
\g or terminate with semicolon to execute query
\q to quit
```

任务示例

步骤1 连接数据库，具体操作请参考《开发指南》中“使用数据库入门 > 连接数据库 > 使用gsql连接”章节。

步骤2 查看gsql的帮助信息。具体执行命令请参见表3-6。

表 3-6 使用 gsql 联机帮助

描述	示例
查看版权信息	\copyright

查看GaussDB支持的SQL语句的帮助	查看GaussDB支持的SQL语句的帮助
	例如，查看GaussDB支持的所有SQL语句：
	`gaussdb=# \h`
	`Available help:`
	`ABORT`
	`... ...`
	例如，查看CREATE DATABASE命令的参数可使用下面的命令：
	`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 ]}[...] ];`

查看gsql命令的帮助	例如，查看gsql支持的命令：
	`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`
	`... ...`

----结束

3.1.4 命令参考

详细的gsql参数请参见表3-7、表3-8、表3-9和表3-10。

版权所有 © 华为云计算技术有限公司

表 3-7 常用参数

参数	参数说明	取值范围
-c, --command=COMMAND	声明gsql要执行一条字符串命令然后退出。	-
-d, --dbname=DBNAME	指定想要连接的数据库名称。 另外，gsql允许使用扩展的DBNAME，即'postgres[ql]://[user[:password]@][netloc][:port][,...][/dbname][?param1=value1&...]'或'[key=value] [...]'形式的连接串作为DBNAME，gsql将从连接串中解析连接信息，并优先使用这些信息。 **注意** gsql使用扩展的DBNAME创建连接时，不支持指定replication参数。	字符串。
-f, --file=FILENAME	使用文件作为命令源而不是交互式输入。gsql将在处理完文件后结束。如果FILENAME是-（连字符），则从标准输入读取。	绝对路径或相对路径，且满足操作系统路径命名规则。
-l, --list	列出所有可用的数据库，然后退出。	-
-v, --set, --variable=NAME=VALUE	设置gsql变量NAME为VALUE。 变量的示例和详细说明请参见**变量**。	-
-X, --no-gsqlrc	不读取启动文件（系统范围的gsqlrc或者用户的~/.gsqlrc都不读取）。 **说明** 启动文件默认为~/.gsqlrc，或通过PSQLRC环境变量指定。	-
-1 ("one"), --single-transaction	当gsql使用-f选项执行脚本时，会在脚本的开头和结尾分别加上START TRANSACTION/COMMIT用以把整个脚本当作一个事务执行。这将保证该脚本完全执行成功，或者脚本无效。 **说明** 如果脚本中已经使用了START TRANSACTION，COMMIT，ROLLBACK，则该选项无效。	-
-y, --slash-command	将当前语句终结并发送到内核执行或者重新执行已经执行过的语句（不包括gsql元命令）。 **说明** 该功能仅支持A兼容模式数据库，而且不支持-c, --command参数的使用场景。 A兼容模式数据库下默认普通语句分隔符为;，PL/SQL为/，使用该参数时不支持更换分隔符。	-
-?, --help	显示关于gsql命令行参数的帮助信息然后退出。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-V, --version	打印gsql版本信息然后退出。	-

表 3-8 输入和输出参数

参数	参数说明	取值范围
-a, --echo-all	在读取行时向标准输出打印所有内容。 **注意** 使用此参数可能会暴露部分SQL语句中的敏感信息，如创建用户语句中的password信息等，请谨慎使用。	-
-e, --echo-queries	把所有发送给服务器的查询同时回显到标准输出。 **注意** 使用此参数可能会暴露部分SQL语句中的敏感信息，如创建用户语句中的password信息等，请谨慎使用。	-
-E, --echo-hidden	回显由\d和其他反斜杠命令生成的实际查询。	-
-k, --with-key=KEY	使用gsql对导入的加密文件进行解密。 **须知** • 对于本身就是shell命令中的关键字符如单引号（'）或双引号（"），Linux shell会检测输入的单引号（'）或双引号（"）是否匹配。如果不匹配，shell认为用户没有输入完毕，会一直等待用户输入，从而不会进入到gsql程序。 • 不支持解密导入存储过程和函数。	-
-L, --log-file=FILENAME	除了正常的输出源之外，把所有查询输出记录到文件FILENAME中。 **注意** • 使用此参数可能会暴露部分SQL语句中的敏感信息，如创建用户语句中的password信息等，请谨慎使用。 • 此参数只保留查询结果到相应文件中，主要目标是为了查询结果能够更好更准确地被其他调用者（例如自动化运维脚本）解析；而不是保留gsql运行过程中的相关日志信息。	绝对路径或相对路径，且满足操作系统路径命名规则。
-m, --maintenance	允许在两阶段事务恢复期间连接数据库。 **说明** 该选项是一个开发选项，禁止用户使用，只限专业技术人员使用，功能是：使用该选项时，gsql可以连接到备机，用于校验主备机数据的一致性。	-
-n, --no-libedit	关闭命令行编辑。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-o, --output=FILENAME	将所有查询输出重定向到文件FILENAME。	绝对路径或相对路径，且满足操作系统路径命名规则。
-q, --quiet	安静模式，执行时不会打印出额外信息。	缺省时gsql将打印许多其他输出信息。
-s, --single-step	单步模式运行。意味着每个查询在发往服务器之前都要提示用户，用这个选项也可以取消执行。此选项主要用于调试脚本。 **注意** 使用此参数可能会暴露部分SQL语句中的敏感信息，如创建用户语句中的password信息等，请谨慎使用。	-
-S, --single-line	单行运行模式，这时每个命令都将由换行符结束，像分号那样。	-
-C，-C1, --enable-client-encryption=1	当使用-C参数连接本地数据库或者连接远程数据库时，可通过该选项打开密态数据库开关，此开关为密态等值查询基本能力开关。	-
-C3，--enable-client-encryption=3	当使用-C参数连接本地数据库或者连接远程数据库时，可通过该选项打开内存解密逃生通道开关，支持密态等值查询基本能力以及内存解密逃生通道能力。	-

表 3-9 输出格式参数

参数	参数说明	取值范围
-A, --no-align	切换为非对齐输出模式。	缺省为对齐输出模式。
-F, --field-separator=STRING	设置域分隔符（默认为“|”）。	-
-H, --html	打开HTML格式输出。	-
-P, --pset=VAR[=ARG]	在命令行上以\pset的风格设置打印选项。 **说明** 这里必须用等号而不是空格分隔名称和值。例如，把输出格式设置为LaTeX，可以键入-P format=latex	-
-R, --record-separator=STRING	设置记录分隔符。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-r	开启在客户端操作中可以进行编辑的模式。	缺省为关闭。
-t, --tuples-only	只打印行。	-
-T, --table-attr=TEXT	允许声明放在HTML table标签里的选项。 使用时请搭配参数“-H,--html”，指定为HTML格式输出。	-
-x, --expanded	打开扩展表格式模式。	-
-z, --field-separator-zero	设置非对齐输出模式的域分隔符为空。 使用时请搭配参数“-A, --no-align”，指定为非对齐输出模式。	-
-0, --record-separator-zero	设置非对齐输出模式的记录分隔符为空。 使用时请搭配参数“-A, --no-align”，指定为非对齐输出模式。	-
-2, --pipeline	使用管道传输密码，禁止在终端使用，必须和-c或者-f参数一起使用。	-

版权所有 © 华为云计算技术有限公司

表 3-10 连接参数

参数	参数说明	取值范围
-h, --host=HOSTNAME	指定正在运行服务器的主机名、Unix域套接字的路径、或者域名。接受以“,”分割的字符串来指定多个主机地址，支持指定多个主机地址，支持指定IPv6主机地址。 当指定多个主机地址时，默认情况下会自动选择主节点地址进行连接。可通过设置PGTARGETSESSIONATTRS环境变量的值来选择连接到不同类型的节点，变量与节点类型对应的关系如下： **PGTARGETSESSIONATTRS环境变量的值--选择连接的节点类型** read-write--可读写的节点 read-only--只读节点 primary或者不设定--主节点 standby--备节点 prefer-standby--首选备节点，没有备节点则转为any any--不进行角色检查 **说明** 当-h只指定一个域名，但是该域名对应多个IP时，无法触发自动选主功能。	如果省略主机名，gsql将通过Unix域套接字与本地主机的服务器相连，或者在没有Unix域套接字的机器上，通过TCP/IP与localhost连接。
-p, --port=PORT	指定数据库服务器的端口号。可以配置一个或多个，当配置一个时，所有的主机地址都使用同一个端口连接；当配置多个时，顺序与主机地址顺序相同，个数必须与主机地址数相等，当不相等时会报错。 可以通过port参数修改默认端口号。	默认端口可通过编译参数来指定，不指定的话默认为5432。
-U, --username=USERNAME	指定连接数据库的用户。 **说明** • 通过该参数指定用户连接数据库时，需要同时提供用户密码用以身份验证。您可以通过交换方式输入密码，或者通过-W参数指定密码。 • 用户名中包含有字符$，需要在字符$前增加转义字符才可成功连接数据库。	字符串，默认使用与当前操作系统用户同名的用户。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-W, --password=PASSWORD	当使用-U参数连接本地数据库或者连接远端数据库时，可通过该选项指定密码。 **说明** • 登录数据库主节点所在服务器后连接本地数据库主节点实例时，默认使用trust连接，会忽略此参数。 • 用户密码中包含特殊字符“\”和"`"时，需要增加转义字符才可成功连接数据库。 • 如果用户未输入该参数，但是数据库连接需要用户密码，这时将出现交互式输入，请用户输入当前连接的密码。该密码最长长度为999字节，受限于GUC参数password_max_length的最大值。	字符串。

3.1.5 元命令参考

介绍使用GaussDB数据库命令行交互工具登录数据库后，gsql所提供的元命令。所谓元命令就是在gsql里输入的任何以不带引号的反斜杠开头的命令。

注意事项

- 一个gsql元命令的格式是反斜杠后面紧跟一个动词，然后是任意参数。参数命令动词和其他参数以任意个空白字符间隔。
- 要在参数里面包含空白，必须用单引号把它引起来。要在这样的参数里包含单引号，可以在前面加一个反斜杠。任何包含在单引号里的内容都会被进一步进行类似C语言的替换：\n（新行）、\t（制表符）、\b（退格）、\r（回车）、\f（换页）、\digits（八进制表示的字符）、\xdigits（十六进制表示的字符）。
- 用""包围的内容被当做一个命令行传入shell。该命令的输出（删除了结尾的新行）被当做参数值。
- 如果不带引号的参数以冒号（:）开头，它会被当做一个gsql变量，并且该变量的值最终会成为真正的参数值。
- 有些命令以一个SQL标识的名称（比如一个表）为参数。这些参数遵循SQL语法关于双引号的规则：不带双引号的标识强制转换成小写，而双引号保护字母不进行大小写转换，并且允许在标识符中使用空白。在双引号中，成对的双引号在结果名称中分析成一个双引号。比如，FOO"BAR"BAZ解析成fooBARbaz；而"Aweird""name"解析成A weird"name。
- 对参数的分析在遇到另一个不带引号的反斜杠时停止。这里会认为是一个新的元命令的开始。特殊的双反斜杠序列（\\）标识参数的结尾并将继续分析后面的SQL语句（如果存在）。这样SQL和gsql命令可以自由的在一行里面混合。但是在任何情况下，一条元命令的参数不能延续超过行尾。
- M-Compatibility模式数据库不支持\h元命令。

元命令

元命令的详细说明请参见**表3-11**、**表3-12**、**表3-13**、**表3-14**、**表3-16**、**表3-18**、**表3-19**、**表3-20**、**表3-22**和**表3-23**。

版权所有 © 华为云计算技术有限公司

须知

以下命令中所提到的FILE代表文件路径。此路径可以是绝对路径（如/home/gauss/file.txt），也可以是相对路径（file.txt，file.txt会默认在用户执行gsql命令所在的路径下创建）。

表 3-11 一般的元命令

参数	参数说明	取值范围
\copyright	显示GaussDB的版本和版权信息。	-
\g [FILE] or ;	执行查询（并将结果发送到文件或管道）。	-
\h(\help) [NAME]	给出指定SQL语句的语法帮助。	如果没有给出NAME，gsql将列出可获得帮助的所有命令。如果NAME是一个星号（*），则显示所有SQL语句的语法帮助。
\parallel [on [num]|off]	控制并发执行开关。 • on：打开控制并发执行开关，且最大并发数为num。 • off：关闭控制并发执行开关。 **说明** • 不支持事务中开启并发执行以及并发中开启事务。 • 不支持\d这类元命令的并发。 • 并发select返回结果混乱问题，此为客户可接受，core、进程停止响应不可接受。 • 不推荐在并发中使用set语句，否则导致结果与预期不一致。 • 不支持创建临时表！如需使用临时表，需要在开启parallel之前创建好，并在parallel内部使用。parallel内部不允许创建临时表。 • \parallel执行时最多会启动num个独立的gsql进程连接服务器。 • \parallel中所有作业的持续时间不能超过session_timeout，否则可能会导致并发执行过程中断连。 • 在\parallel on 之后一条或多条命令，会等到\parallel off执行后才会执行，因而，\parallel on之后需要有对应的\parallel off，否则\parallel on后的命令都无法执行。	num的默认值：1024。 **须知** • 服务器能接受的最大连接数受max_connection及当前已有连接数限制。 • 设置num时请考虑服务器当前可接受的实际连接数合理指定。
\q	退出gsql程序。在一个脚本文件里，只在脚本终止的时候执行。	-

版权所有 © 华为云计算技术有限公司

表 3-12 查询缓存区元命令

参数	参数说明
\e [FILE] [LINE]	使用外部编辑器编辑查询缓冲区（或者文件）。
\ef [FUNCNAME [LINE]]	使用外部编辑器编辑函数定义。如果指定了LINE（即行号），则光标会指到函数体的指定行。
\p	打印当前查询缓冲区到标准输出。
\r	重置（或清空）查询缓冲区。
\w FILE	将当前查询缓冲区输出到文件。

表 3-13 输入/输出元命令

参数	参数说明
\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]	在任何gsql客户端登录数据库成功后可以执行导入导出数据，这是一个运行SQL COPY命令的操作，但不是读取或写入指定文件的服务器，而是读取或写入文件，并在服务器和本地文件系统之间路由数据。这意味着文件的可访问性和权限是本地用户的权限，而不是服务器的权限，并且不需要数据库初始化用户权限。 **说明** • \COPY只适合小批量，格式良好的数据导入，导入数据应优先选择GDS或COPY。 • \COPY 可以指定数据导入时的客户端数量，从而实现数据文件的并行导入，目前并发数范围为[1, 8]。 • \COPY并行导入目前存在以下约束：临时表的并行导入不支持、在事务内的并行导入不支持、对二进制文件的并行导入不支持、数据导入支持AES128加密时不支持以及COPY选项中存在EOL。在这些情况下，即使指定了parallel参数，仍然会走非并行流程。 • \COPY的text格式和csv格式均支持header功能 • 其中LOAD功能为gs_loader进行语法转换后调用copy的标识，非主动调用功能。 • 其中LOAD_DISCARD功能为gs_loader解析后discard文件路径，非主动调用功能。
\echo [STRING]	把字符串写到标准输出。
prompt [STRING]	把字符串写到标准输出（等同于\echo）。
\i FILE	从文件FILE中读取内容，并将其当作输入，执行查询。
\i+ FILE KEY	执行加密文件中的命令。
\ir FILE	和\i类似，只是相对于存放当前脚本的路径。

版权所有 © 华为云计算技术有限公司

参数	参数说明
\ir+ FILE KEY	和\i+类似，只是相对于存放当前脚本的路径。
\o [FILE]	把所有的查询结果发送到文件里。
\qecho [STRING]	把字符串写到查询结果输出流里。

说明

表3-14中的选项S表示显示系统对象，+表示显示对象附加的描述信息。PATTERN用来指定要被显示的对象名称。

表 3-14 显示信息元命令

参数	参数说明	取值范围	示例
\d[S+]	列出当前search_path中模式下所有的表、视图和序列。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	-	列出当前search_path中模式下所有的表、视图和序列。 gaussdb=# \d
\d[S+] NAME	列出指定表、视图和索引的结构。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	-	假设存在表a，列出指定表a的结构。 gaussdb=# \dtable+ a
\d+ [PATTERN]	列出所有表、视图和索引。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的表、视图和索引。	列出所有名称以f开头的表、视图和索引。 gaussdb=# \d+ f*
\da[S] [PATTERN]	列出所有可用的聚集函数，以及它们操作的数据类型和返回值类型。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的聚集函数。	列出所有名称以f开头可用的聚集函数，以及它们操作的数据类型和返回值类型。 gaussdb=# \da f*
\db[+] [PATTERN]	列出所有可用的表空间。	如果声明了PATTERN，只显示名称匹配PATTERN的表空间。	列出所有名称以p开头的可用表空间。 gaussdb=# \db p*
\dc[S+] [PATTERN]	列出所有字符集之间的可用转换。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的转换。	列出所有字符集之间的可用转换。 gaussdb=# \dc *

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围	示例
\dC[+] [PATTERN]	列出所有类型转换。 PATTERN需要使用实际类型名，不能使用别名。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的转换。	列出所有名称以c开头的类型转换。 gaussdb=# \dC c*
\dd[S] [PATTERN]	显示所有匹配PATTERN的描述。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果没有给出参数，则显示所有可视对象。“对象”包括：聚集、函数、操作符、类型、关系(表、视图、索引、序列、大对象)、规则。	列出所有可视对象。 gaussdb=# \dd
\ddp [PATTERN]	显示所有默认的使用权限。	如果指定了PATTERN，只显示名称匹配PATTERN的使用权限。	列出所有默认的使用权限。 gaussdb=# \ddp
\dD[S+] [PATTERN]	列出所有可用域。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的域。	列出所有可用域。 gaussdb=# \dD
\det[+] [PATTERN]	列出所有的外部表。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的表。	列出所有的外部表。 gaussdb=# \det
\des[+] [PATTERN]	列出所有的外部服务器。	如果声明了PATTERN，只显示名称匹配PATTERN的服务器。	列出所有的外部服务器。 gaussdb=# \des
\deu[+] [PATTERN]	列出用户映射信息。	如果声明了PATTERN，只显示名称匹配PATTERN的信息。	列出用户映射信息。 gaussdb=# \deu
\dew[+] [PATTERN]	列出封装的外部数据。	如果声明了PATTERN，只显示名称匹配PATTERN的数据。	列出封装的外部数据。 gaussdb=# \dew

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围	示例
\df[antw][S+][PATTERN]	列出所有可用函数，以及它们的参数和返回的数据类型。a代表聚集函数，n代表普通函数，t代表触发器，w代表窗口函数。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的函数。	列出所有可用函数，以及它们的参数和返回的数据类型。 gaussdb=# \df
\dF[+][PATTERN]	列出所有的文本搜索配置信息。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的配置信息。	列出所有的文本搜索配置信息。 gaussdb=# \dF+
\dFd[+][PATTERN]	列出所有的文本搜索字典。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的字典。	列出所有的文本搜索字典。 gaussdb=# \dFd
\dFp[+][PATTERN]	列出所有的文本搜索分析器。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的分析器。	列出所有的文本搜索分析器。 gaussdb=# \dFp
\dFt[+][PATTERN]	列出所有的文本搜索模板。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只显示名称匹配PATTERN的模板。	列出所有的文本搜索模板。 gaussdb=# \dFt
\dg[+][PATTERN]	列出所有数据库角色。 **说明** 因为用户和群组的概念被统一为角色，所以这个命令等价于\du。为了和以前兼容，所以保留两个命令。	如果指定了PATTERN，只显示名称匹配PATTERN的角色。	列出名称为“j?e”所有数据库角色（“?”表示任一字符）。 gaussdb=# \dg j?e
\dl	\lo_list的别名，显示一个大对象的列表。	-	列出所有的大对象。 gaussdb=# \dl
\dL[S+][PATTERN]	列出可用的程序语言。	如果指定了PATTERN，只列出名称匹配PATTERN的语言。	列出可用的程序语言。 gaussdb=# \dL

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围	示例
\dm[S+] [PATTERN]	列出物化视图。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果指定了PATTERN，只列出名称匹配PATTERN的物化视图。	列出物化视图。 gaussdb=# \dm
\dn[S+] [PATTERN]	列出所有模式（名称空间）。如果向命令追加+，会列出每个模式相关的权限及描述。	如果声明了PATTERN，只列出名称匹配PATTERN的模式名。缺省时，只列出用户创建的模式。	列出所有名称以d开头的模式以及相关信息。 gaussdb=# \dn+ d*
\do[S] [PATTERN]	列出所有可用的操作符，以及它们的操作数和返回的数据类型。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只列出名称匹配PATTERN的操作符。缺省时，只列出用户创建的操作符。	列出所有可用的操作符，以及它们的操作数和返回的数据类型。 gaussdb=# \do
\dO[S+] [PATTERN]	列出排序规则。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果声明了PATTERN，只列出名称匹配PATTERN的规则。缺省时，只列出用户创建的规则。	列出排序规则。 gaussdb=# \dO
\dp [PATTERN]	列出一列可用的表、视图以及相关的权限信息。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。 \dp显示结果如下： rolename=xxxx/yyyy --赋予一个角色的权限 =xxxx/yyyy --赋予public的权限 xxxx表示赋予的权限，yyyy表示授予这个权限的角色。权限的参数说明请参见**表3-15**。	如果指定了PATTERN，只列出名称匹配PATTERN的表、视图。	列出一列可用的表、视图以及相关的权限信息。 gaussdb=# \dp
\drds [PATTERN1 [PATTERN2]]	列出所有修改过的配置参数。这些设置可以是针对角色的、针对数据库的或者同时针对两者的。PATTERN1和PATTERN2表示要列出的角色PATTERN和数据库PATTERN。	如果声明了PATTERN，只列出名称匹配PATTERN的规则。缺省或指定*时，则会列出所有设置。	列出数据库所有修改过的配置参数。 gaussdb=# \drds * dbname

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围	示例
\dT[S+] [PATTERN]	列出所有的数据类型。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。	如果指定了PATTERN，只列出名称匹配PATTERN的类型。	列出所有的数据类型。 gaussdb=# \dT
\du[+] [PATTERN]	列出所有数据库角色。 **说明** 因为用户和群组的概念被统一为角色，所以这个命令等价于\dg。为了和以前兼容，所以保留两个命令。	如果指定了PATTERN，则只列出名称匹配PATTERN的角色。	列出所有数据库角色。 gaussdb=# \du
\dE[S+] [PATTERN] \di[S+] [PATTERN] \ds[S+] [PATTERN] \dt[S+] [PATTERN] \dv[S+] [PATTERN]	这一组命令，字母E，i，s，t和v分别代表着外部表，索引，序列，表和视图。可以以任意顺序指定其中一个或者它们的组合来列出这些对象。当search_path中不同模式存在同名对象时，只显示search_path中位置靠前模式下的同名对象。例如：\dit列出所有的索引和表。在命令名称后面追加+，则每一个对象的物理尺寸以及相关的描述也会被列出。	如果指定了PATTERN，只列出名称匹配该PATTERN的对象。默认情况下只会显示用户创建的对象。通过PATTERN或者S修饰符可以把系统对象包括在内。	列出所有的索引和视图。 gaussdb=# \div
\dx[+] [PATTERN]	列出安装数据库的扩展信息。	如果指定了PATTERN，则只列出名称匹配PATTERN的扩展信息。	列出安装数据库的扩展信息。 gaussdb=# \dx
\l[+]	列出服务器上所有数据库的名称、所有者、字符集编码以及使用权限。	-	列出服务器上所有数据库的名称、所有者、字符集编码以及使用权限。 gaussdb=# \l
\sf[+] FUNCNAME	显示函数的定义。 **说明** 对于带圆括号的函数名，需要在函数名两端添加双引号，并且在双引号后面加上参数类型列表。参数类型列表两端添加圆括号。 如果存在同名的函数，则会返回多个函数的定义。	-	假设存在函数function_a和函数名带圆括号的函数func()name，列出函数的定义。 gaussdb=# \sf function_a gaussdb=# \sf "func()name"(argtype1, argtype2)

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围	示例
\z [PATTERN]	列出数据库中所有表、视图和序列，以及它们相关的访问特权。	如果给出任何pattern ，则被当成一个正则表达式，只显示匹配的表、视图、序列。	列出数据库中所有表、视图和序列，以及它们相关的访问特权。 gaussdb=# \z

表 3-15 权限的参数说明

参数	参数说明
r	SELECT：允许对指定的表、视图读取数据。
w	UPDATE：允许对指定表更新字段。
a	INSERT：允许对指定表插入数据。
d	DELETE：允许删除指定表中的数据。
D	TRUNCATE：允许清理指定表中的数据。
x	REFERENCES：允许创建外键约束。
t	TRIGGER：允许在指定表上创建触发器。
X	EXECUTE：允许使用指定的函数，以及利用这些函数实现的操作符。
U	USAGE： ● 对于过程语言，允许用户在创建函数时，指定过程语言。 ● 对于模式，允许访问包含在指定模式中的对象。 ● 对于序列，允许使用nextval函数。
C	CREATE： ● 对于数据库，允许在该数据库里创建新的模式。 ● 对于模式，允许在该模式中创建新的对象。 ● 对于表空间，允许在其中创建表，以及允许创建数据库和模式的时候把该表空间指定为其缺省表空间。
c	CONNECT：允许用户连接到指定的数据库。
T	TEMPORARY：允许创建临时表。
A	ALTER：允许用户修改指定对象的属性。
P	DROP：允许用户删除指定的对象。
m	COMMENT：允许用户定义或修改指定对象的注释。
i	INDEX：允许用户在指定表上创建索引。

版权所有 © 华为云计算技术有限公司

参数	参数说明
v	VACUUM：允许用户对指定的表执行ANALYZE和VACUUM操作。
*	给前面权限的授权选项。

表 3-16 格式化元命令

参数	参数说明
\a	对齐模式和非对齐模式之间的切换。
\C [STRING]	把正在打印的表的标题设置为一个查询的结果或者取消这样的设置。
\f [STRING]	对于不对齐的查询输出，显示或者设置域分隔符。
\H	• 若当前模式为文本格式，则切换为HTML输出格式。 • 若当前模式为HTML格式，则切换回文本格式。
\pset NAME [VALUE]	设置影响查询结果表输出的选项。NAME的取值见**表 3-17**。
\t [on|off]	切换输出的字段名的信息和行计数脚注。
\T [STRING]	指定在使用HTML输出格式时放在table标签里的属性。如果参数为空，不设置。
\x [on|off|auto]	切换扩展行格式。

表 3-17 可调节的打印选项

选项	选项说明	取值范围
border	value必须是一个数字。通常这个数字越大，表的边界就越宽线就越多，但是这个取决于特定的格式。	• 在HTML格式下，取值范围为大于0的整数。 • 在其他格式下，取值范围： – 0：无边框 – 1：内部分隔线 – 2：台架

版权所有 © 华为云计算技术有限公司

选项	选项说明	取值范围
expanded (或x)	在正常和扩展格式之间切换。	● 当打开扩展格式时，查询结果用两列显示，字段名称在左、数据在右。这个模式在数据无法放进通常的“水平”模式的屏幕时很有用。 ● 在正常格式下，当查询输出的格式比屏幕宽时，用扩展格式。正常格式只对aligned和wrapped格式有用。
fieldsep	声明域分隔符来实现非对齐输出。这样就可以创建其他程序希望的制表符或逗号分隔的输出。要设置制表符域分隔符，键入\pset fieldsep '\t'。缺省域分隔符是 '|' (竖条符)。	-
fieldsep_zero	声明域分隔符来实现非对齐输出到零字节。	-
footer	用来切换脚注。	-
format	设置输出格式。允许使用唯一缩写（这意味着一个字母就够了）。	取值范围： ● unaligned：写一行的所有列在一条直线上中，当前活动字段分隔符分隔。 ● aligned：此格式是标准的，可读性好的文本输出。 ● wrapped：类似aligned，但是包装跨行的宽数据值，使其适应目标字段的宽度输出。 ● html：把表输出为可用于文档里的对应标记语言。输出不是完整的文档。 ● latex：把表输出为可用于文档里的对应标记语言。输出不是完整的文档。 ● troff-ms：把表输出为可用于文档里的对应标记语言。输出不是完整的文档。
null	打印一个字符串，用来代替一个null值。	缺省是什么都不打印，这样很容易和空字符串混淆。
numericlocale	切换分隔小数点左边的数值的区域相关的分组符号。	● on：显示指定的分隔符。 ● off：不显示分隔符。 忽略此参数，显示默认的分隔符。

版权所有 © 华为云计算技术有限公司

选项	选项说明	取值范围
pager	控制查询和gsql帮助输出的分页器。如果设置了环境变量PAGER，输出将被指向到指定程序，否则使用系统缺省。	• on：当输出到终端且不适合屏幕显示时，使用分页器。 • off：不使用分页器。 • always：当输出到终端无论是否符合屏幕显示时，都使用分页器。
recordsep	声明在非对齐输出格式时的记录分隔符。	-
recordsep_zero	声明在非对齐输出到零字节时的记录分隔符。	-
tableattr（或T）	声明放在html输出格式中HTML table标签的属性（例如：cellpadding或bgcolor）。注意：这里可能不需要声明border，因为已经在\pset border里用过了。如果没有给出value，则不设置表的属性。	-
title	为随后打印的表设置标题。这个可以用于给输出一个描述性标签。如果没有给出value，不设置标题。	-
tuples_only（或者t）	在完全显示和只显示实际的表数据之间切换。完全显示将输出像列头、标题、各种脚注等信息。在tuples_only模式下，只显示实际的表数据。	-
feedback	切换是否输出结果行数	-

表 3-18 连接元命令

参数	参数说明	取值范围
\c[onnect] [DBNAME|- USER|- HOST|- PORT|-]	连接到一个新的数据库。当数据库名称长度超过63个字节时，默认前63个字节有效，连接到前63个字节对应的数据库，但是gsql的命令提示符中显示的数据库对象名仍为截断前的名称。 **说明** 重新建立连接时，如果切换数据库登录用户，将可能会出现交互式输入，要求输入新用户的连接密码。该密码最长长度为999字节，受限于GUC参数password_max_length的最大值。	-
\encoding [ENCODING]	设置客户端字符编码格式。	不带参数时，显示当前的编码格式。
\conninfo	输出当前连接的数据库的信息。	-

版权所有 © 华为云计算技术有限公司

表 3-19 操作系统元命令

参数	参数说明	取值范围
\cd [DIR]	切换当前的工作目录。	绝对路径或相对路径，且满足操作系统路径命名规则。
\setenv NAME [VALUE]	设置环境变量NAME为VALUE，如果没有给出VALUE值，则不设置环境变量。	-
\timing [on|off]	以毫秒为单位显示每条SQL语句的执行时间（不包括屏显打印时间）。	• on表示打开显示。 • off表示关闭显示。
\! [COMMAND]	返回到一个单独的Unix shell或者执行Unix命令COMMAND。	-

表 3-20 变量元命令

参数	参数说明
\prompt [TEXT] NAME	提示用户用文本格式来指定变量名称。
\set [NAME [VALUE]]	设置内部变量NAME为VALUE或者如果给出了多于一个值，设置为所有这些值的连接结果。如果没有给出第二个参数，只设变量不设值。 有一些常用变量被gsql特殊对待，它们是一些选项设置，通常所有特殊对待的变量都是由大写字母组成(可能还有数字和下划线)。 表3-21是一个所有特殊对待的变量列表。
\unset NAME	不设置（或删除）gsql变量名。

表 3-21 \set 常用命令

名称	命令说明	取值范围
\set VERBOSITY value	这个选项可以设置为值default，verbose，terse之一以控制错误报告的冗余行。	value取值范围：default，verbose，terse
\set ON_ERROR_STOP value	如果设置了这个变量，脚本处理将马上停止。如果该脚本是从另外一个脚本调用的，该脚本也会按同样的方式停止。如果最外层的脚本不是从一次交互的gsql会话中调用的而是用-f选项调用的，gsql将返回错误代码3，以示这个情况与致命错误条件的区别（错误代码为1）。	value取值范围为：on/off，true/false，yes/no，1/0

版权所有 © 华为云计算技术有限公司

名称	命令说明	取值范围
\set AUTOCOMMIT [on|off]	设置当前gsql连接的自动提交行为，on为打开自动提交，off为关闭自动提交。默认情况下，gsql连接处于自动提交模式，每个单独的语句都被隐式提交。如果基于性能或者其它方面考虑，需要关闭自动提交时，需要用户自己显示的发出COMMIT命令来保证事务的提交。例如，在指定的业务SQL执行完之后发送COMMIT语句显式提交，特别是gsql客户端退出之前务必保证所有的事务已经提交。 **说明** gsql默认使用自动提交模式，若关闭自动提交，将会导致后面执行的语句都受到隐式事务包裹，数据库中不支持在事务中执行的语句不能在此模式下执行。	• on表示打开自动提交。 • off表示关闭自动提交。

表 3-22 大对象元命令

参数	参数说明
\lo_list	显示一个目前存储在该数据库里的所有GaussDB大对象和提供给大对象的注释。

表 3-23 全密态元命令

参数	参数说明
\send_token	全密态功能，传输密钥到服务端缓存，只在开启内存解密逃生通道的情况下使用。
\st	全密态功能，传输密钥到服务端缓存，只在开启内存解密逃生通道的情况下使用。
\clear_token	全密态功能，销毁服务端缓存的密钥，只在开启内存解密逃生通道的情况下使用。
\ct	全密态功能，销毁服务端缓存的密钥，只在开启内存解密逃生通道的情况下使用。
\key_info KEY_INFO	在全密态数据库特性中，用于设置访问外部密钥管理者的参数。

说明

M-Compatibility暂不支持全密态数据库。

PATTERN

很多\d命令都可以用一个PATTERN参数来指定要被显示的对象名称。在最简单的情况下，PATTERN正好就是该对象的准确名称。在PATTERN中的字符通常会被变成小写形

版权所有 © 华为云计算技术有限公司

式（就像在SQL名称中那样），例如\dt FOO将会显示名为foo的表。就像在SQL名称中那样，把PATTERN放在双引号中可以阻止它被转换成小写形式。如果需要在一个PATTERN中包括一个真正的双引号字符，则需要把它写成两个相邻的双引号，这同样是符合SQL引用标识符的规则。例如，\dt "FOO""BAR"将显示名为FOO"BAR（不是foo"bar）的表。和普通的SQL名称规则不同，不能只在PATTERN的一部分周围放上双引号，例如\dt FOO"FOO"BAR将会显示名为fooFOObar的表。

不使用PATTERN参数时，\d命令会显示当前schema搜索路径中可见的全部对象——这等价于用*作为PATTERN。所谓对象可见是指可以直接用名称引用该对象，而不需要用schema来进行限定。要查看数据库中所有的对象而不管它们的可见性，可以把*.*用作PATTERN。

如果放在一个PATTERN中，*将匹配任意字符序列（包括空序列），而?会匹配任意的单个字符（这种记号方法就像 Unix shell 的文件名PATTERN一样）。例如，\dt int*会显示名称以int开始的表。但是如果被放在双引号内，*和?就会失去这些特殊含义而变成普通的字符。

包含一个点号（.）的PATTERN被解释为一个schema名称模式后面跟上一个对象名称模式。例如，\dt foo*.*bar*会显示名称以foo开始的schema中所有名称包括bar的表。如果没有出现点号，那么模式将只匹配当前schema搜索路径中可见的对象。同样，双引号内的点号会失去其特殊含义并且变成普通的字符。

高级用户可以使用字符类等正则表达式记法，如[0-9]可以匹配任意数字。所有的正则表达式特殊字符都按照POSIX正则表达式所说的工作。以下字符除外：

- .会按照上面所说的作为一种分隔符。
- *会被翻译成正则表达式记号.*。
- ?会被翻译成.。
- $则按字面意思匹配。

根据需要，可以通过书写?、(*R*+|)、(*R*|)和*R*?来分别模拟PATTERN字符.、*R**和*R*?。$不需要作为一个正则表达式字符，因为PATTERN必须匹配整个名称，而不是像正则表达式的常规用法那样解释（换句话说，$会被自动地追加到PATTERN上）。如果不希望该PATTERN的匹配位置被固定，可以在开头或者结尾写上*。注意在双引号内，所有的正则表达式特殊字符会失去其特殊含义并且按照其字面意思进行匹配。另外，在操作符名称PATTERN中（即\do的PATTERN参数），正则表达式特殊字符也按照字面意思进行匹配。

DELIMITER

更改SQL语句之间分隔符命令，分隔符默认值为“;”。

DELIMITER命令用来为客户端设置一个分隔符。当用户设置分隔符后，gsql客户端识别到分隔符时，会立即将SQL语句发送到服务端执行，但是服务端仍然将“;”看做SQL语句分隔符，并相应的处理SQL语句。

注意事项：

- delimiter符号目前不是自由设定的，结束符范围有限制，目前接受大小写字母组合或特殊字符组合（~/!/@/#/^/&/`/?/+/-/*//（除号）/%/</>/=），其中常见的用法是“//”。
- 符号组合中尽量使用无歧义符号组合，特殊符号组合（注释符：“*”、“--”，以加号“+”或减号“-”结尾的符号组合）目前不支持用于delimiter命名。
- delimiter长度范围：0~15。

 版权所有 © 华为云计算技术有限公司

- 设置的结束符的级别是会话级别的，当切换数据库时delimiter_name会设置为默认值“;”。
- 用户如果想使用其他字符组合例如“adbc $$”，可以使用引号包含，例如delimiter "adbc $$"，但使用时也需要使用引号包含，例如：select 1"adbc $$"。
- delimiter分隔符只有sql_compatibility = 'B'时支持。

3.1.6 常见问题处理

连接性能问题

- 开启log_hostname，但是配置错误的DNS导致的连接性能问题。

  连接数据库，通过“show log_hostname”语句，检查数据库中是否开启log_hostname参数。

  如果开启了相关参数，那么数据库内核通过DNS反查客户端所在机器的主机名。如果数据库配置了不正确的/不可达的DNS服务器，导致数据库建立连接过程较慢。此参数的更多信息详见GUC参数**log_hostname**。

- 数据库内核执行初始化语句较慢导致的性能问题。

  此种情况定位较难，可以尝试使用Linux的命令：strace。

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

  在屏幕上打印出数据库的连接过程。比如较长时间停留在下面的操作上：

  ```
  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}])
  ```

  可以确定是数据库执行“SELECT VERSION()”语句较慢。

  在连接上数据库后，便可以通过执行“explain performance select version()”语句确定初始化语句执行较慢的原因。更多信息，详见《开发指南》中“SQL调优指南 > SQL执行计划介绍”章节。

  另外还有一种场景不太常见：由于DN所在机器的磁盘满或故障，此时所查询等受影响，无法进行用户认证，导致连接过程挂起。解决此问题清理DN的数据盘空间便可。

- TCP连接创建较慢问题。

  此问题可以参考上面的初始化语句较慢排查的做法，通过strace侦听，如果长时间停留在：

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

  或者

  ```
  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)
  ```

  说明客户端与数据库端建立物理连接过慢，此时应当检查网络是否存在不稳定、网络吞吐量太大的问题。

- 资源负载满导致连接较慢的问题。

  原因分析：当CPU、内存、I/O负载中的任意一项接近100%时，会出现gsql连接慢的现象。

  问题解决：

  a. 通过top命令等确认CPU使用率；通过free命令确认内存使用情况；通过iostat命令确认I/O负载；还可以通过cm_agent中的监控日志，以及数据库运维平台中的监测记录进行检查。

 版权所有 © 华为云计算技术有限公司

b. 针对短时间内大量慢查询导致的峰值负载场景，可通过[数据库服务器的端口号+1]端口连接，查询pg_stat_activity视图；针对慢查询，可以使用系统函数pg_terminate_backend进行查杀会话。

c. 针对业务量长期超负载情况（即无明显慢查询，或慢查询查杀后但新的查询依然会变成慢查询），应考虑降低业务负载、增加数据库资源的方式进行优化。

创建连接故障

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

  此问题一般是指定了不可达的地址或者端口导致的。请检查-h参数与-p参数是否添加正确。

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

  此问题一般是输入了错误的用户名和密码导致的，请联系数据库管理员，确认用户名和密码的正确性。

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

  数据库由于安全问题，禁止远程登录时使用trust模式。这时需要修改gs_hba.conf里的连接认证信息。请联系管理员处理。

  说明

  请不要修改gs_hba.conf中数据库主机的相关设置，否则可能导致数据库功能故障。建议业务应用部署在数据库之外，而非数据库内部。

- 在DN连接数据库，添加“-h 127.0.0.1”可以连接，去掉后无法连接问题。

  通过执行SQL语句“show unix_socket_directory”检查DN使用的Unix套接字目录，是否与shell中的环境变量$PGHOST一致。

  如果检查结果不一致，那么修改PGHOST环境变量到GUC参数unix_socket_directory指向的目录。

  关于unix_socket_directory的更多信息，请联系管理员获取。

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

  此问题是由于环境中使用的libpq.so的版本与gsql的版本不匹配导致的，请通过“ldd gsql”命令确认当前加载的libpq.so的版本，并通过修改LD_LIBRARY_PATH环境变量来加载正确的libpq.so。

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

  此问题是由于环境中使用的libpq.so的版本与gsql的版本不匹配导致的（也有可能是环境中存在PostgreSQL的libpq.so），请通过“ldd gsql”命令确认当前加载的libpq.so的版本，并通过修改LD_LIBRARY_PATH环境变量来加载正确的libpq.so。

- 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?

  此问题是由于网络连接故障造成。请检查客户端与数据库服务器间的网络连接。如果发现从客户端无法PING到数据库服务器端，则说明网络连接出现故障。请联系网络管理人员排查解决。

  ```
  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
  ```

 版权所有 © 华为云计算技术有限公司

```
--- 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.

  此问题是由于用户不具备访问该数据库的权限，可以使用如下方法解决。

  a. 使用管理员用户dbadmin连接数据库。

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

  b. 赋予该用户访问数据库的权限。

  GRANT CONNECT ON DATABASE gaussdb TO user1;

  说明

  实际上，常见的许多错误操作也可能产生用户无法连接上数据库的现象。如用户连接的数据库不存在，用户名或密码输入错误等。这些错误操作在客户端工具也有相应的提示信息。

  ```
  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.

  此问题是由于系统连接数量超过了最大连接数量。请联系数据库DBA进行会话连接数管理，释放无用会话。

  关于查看用户会话连接数的方法如表3-24。

  会话状态可以在视图PG_STAT_ACTIVITY中查看。无用会话可以使用函数pg_terminate_backend进行释放。

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

  其中pid的值即为该会话的线程ID。根据线程ID结束会话。

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

  显示类似如下信息，表示结束会话成功。

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

表 3-24 查看会话连接数

描述	命令
查看指定用户的会话连接数上限。	执行如下命令查看连接到指定用户USER1的会话连接数上限。其中-1表示没有对用户user1设置连接数的限制。 SELECT ROLNAME,ROLCONNLIMIT FROM PG_ROLES WHERE ROLNAME='user1'; rolname | rolconnlimit ---------+-------------- user1 | -1 (1 row)

 版权所有 © 华为云计算技术有限公司

描述	命令
查看指定用户已使用的会话连接数。	执行如下命令查看指定用户USER1已使用的会话连接数。其中，1表示USER1已使用的会话连接数。 SELECT COUNT(*) FROM dv_sessions WHERE USERNAME='user1'; count ------- 1 (1 row)
查看指定数据库的会话连接数上限。	执行如下命令查看连接到指定数据库gaussdb的会话连接数上限。其中-1表示没有对数据库gaussdb设置连接数的限制。 SELECT DATNAME,DATCONNLIMIT FROM PG_DATABASE WHERE DATNAME='gaussdb'; datname | datconnlimit ----------+-------------- gaussdb | -1 (1 row)
查看指定数据库已使用的会话连接数。	执行如下命令查看指定数据库gaussdb上已使用的会话连接数。其中，1表示数据库gaussdb上已使用的会话连接数。 SELECT COUNT(*) FROM PG_STAT_ACTIVITY WHERE DATNAME='gaussdb'; count ------- 1 (1 row)
查看所有用户已使用会话连接数。	执行如下命令查看所有用户已使用的会话连接数。 SELECT COUNT(*) FROM dv_sessions; count ------- 10 (1 row)

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

  gsql在向数据库发起连接的时候，会有5分钟超时机制，如果在这个超时时间内，数据库未能正常的对客户端请求进行校验和身份认证，那么gsql会退出当前会话的连接过程，并报出如上错误。

  一般来说，此问题是由于连接时使用的-h参数及-p参数指定的连接主机及端口有误（即错误信息中的xxx部分），导致通信故障；极少数情况是网络故障导致。要排除此问题，请检查数据库的主机名及端口是否正确。

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

  同时，检查DN日志中出现类似如下日志“ FATAL: cipher file "/data/coordinator/server.key.cipher" has group or world access”，一般是由于数据目录或部分关键文件的权限被误操作篡改导致。请参照其他正常实例下的相关文件权限修改。

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

  目标DN的gs_hba.conf里配置了当前客户端IP使用"gss"方式来做认证，该认证算法不支持用作客户端的身份认证，请修改到"sha256"后再试。具体操作请联系管理员处理。

 版权所有 © 华为云计算技术有限公司

说明

- 请不要修改gs_hba.conf中数据库主机的相关设置，否则可能导致数据库功能故障。
- 建议业务应用部署在数据库之外，而非数据库内部。

其他故障

- 出现因“总线错误”（Bus error）导致的core dump或异常退出。

  一般情况下出现此种问题，是进程运行过程中加载的共享动态库（在Linux为.so文件）出现变化；或者进程二进制文件本身出现变化，导致操作系统加载机器的执行码或者加载依赖库的入口发生变化，操作系统出于保护目的将进程终止，产生core dump文件。

  解决此问题，请重试。同时请尽可能避免在升级等运维操作过程中，在数据库内部运行业务程序，避免升级时因替换文件产生此问题。

  说明

  此故障的core dump文件的可能堆栈是dl_main及其子调用，它是操作系统用来初始化进程做共享动态库加载的。如果进程已经初始化，但是共享动态库还未加载完成，严格意义上来说，进程并未完全启动。

3.2 gs_loader

概述

gs_loader工具用于进行数据导入。gs_loader将控制文件支持的语法转换为\COPY语法，然后利用已有的\COPY功能，做主要数据导入工作，同时gs_loader将\COPY结果记录到日志中。

使用gs_loader前请确保gs_loader版本与gsql版本、数据库版本保持一致。

说明

gs_loader工具当前不支持M-Compatibility数据库。

gs_loader工具当前支持PDB。

安装部署

在存放数据源文件的服务器上，安装并配置gs_loader客户端工具，方便使用gs_loader工具进行数据的导入。

步骤1 创建用于存放gs_loader工具包的目录。

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

步骤2 将gsql工具包上传至新创建的目录中。

以上传EULER Linux版本的工具包为例，将软件安装包中的gsql工具包“GaussDB-Kernel_数据库版本号_操作系统版本号_64bit_gsql.tar.gz”上传至新创建的目录中。

步骤3 在工具包所在的目录下，解压工具包。

```
cd /opt/bin
tar -zxvf GaussDB-Kernel_数据库版本号_操作系统版本号_64bit_gsql.tar.gz
source gsql_env.sh
```

步骤4 验证工具位置及版本信息。

版权所有 © 华为云计算技术有限公司

```
which gs_loader
```

步骤5 验证客户端版本信息。

gs_loader工具版本与gsql工具版本相对应，直接查询gsql客户端版本即可验证客户端版本信息。

```
gsql -V
```

步骤6 验证数据库版本信息，确保与客户端工具版本保持一致。

使用gsql工具成功连接数据库后输入：

```
select version();
```

----结束

日志等级配置

设置日志级别，可以供开发者查看。设置后会在控制台打印工具运行的相应信息。

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

使用权限

使用场景分为三权分立场景下及非三权分立场景下的使用。使用者可以选择将guc参数enableSeparationOfDuty设置为on或者off来控制三权分立功能的开启或关闭。

GUC参数enable_copy_error_log是控制是否使用错误表pgxc_copy_error_log的参数，默认为off、即不使用错误表，错误记录直接记录到gs_loader的bad文件中。如果该参数设置为on，则会使用错误表pgxc_copy_error_log，将错误记录插入错误表。

默认场景，不开启三权分立（即enableSeparationOfDuty=off）时，使用者可以是数据库普通用户或管理员用户。当使用者为普通用户的时候，需要管理员用户对普通用户赋权。管理员账户可以直接使用。错误表pgxc_copy_error_log通过GUC参数enable_copy_error_log控制开启和关闭，默认关闭。

1. 使用管理员用户创建新的用户：

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

2. 给新用户授权pg_catalog.gs_copy_summary表：

   说明

   gs_loader使用的public.gs_copy_summary变更为pg_catalog.gs_copy_summary，原public.gs_copy_summary表废弃。由于public.gs_copy_summary可能与用户表同名，因此需要用户进行识别，确认不是用户表后，将数据迁移到pg_catalog.gs_copy_summary表里之后，删除public.gs_copy_summary。由于存在pg_catalog.gs_copy_summary，gsql元命令\d(+)优先找到pg_catalog.gs_copy_summary，而因为是系统表的缘故不显示，因此列表中找不到gs_copy_summary，但是实际上public.gs_copy_summary是存在的。可以直接通过public.gs_copy_summary引用。

   ```
   GRANT INSERT,SELECT ON  pg_catalog.gs_copy_summary To load_user;
   ```

3. （可选）给新用户授权错误表pgxc_copy_error_log：

 版权所有 © 华为云计算技术有限公司

说明

gs_loader使用的public.pgxc_copy_error_log变更为pg_catalog.pgxc_copy_error_log，原public.pgxc_copy_error_log表废弃。由于public.pgxc_copy_error_log可能与用户表同名，因此需要用户进行识别，确认不是用户表后，将数据迁移到pg_catalog.pgxc_copy_error_log表里之后，删除public.pgxc_copy_error_log。由于存在pg_catalog.pgxc_copy_error_log，gsql元命令\d(+)优先找到pg_catalog.pgxc_copy_error_log，而因为是系统表的缘故不显示，因此列表中找不到pgxc_copy_error_log，但是实际上public.pgxc_copy_error_log是存在的。可以直接通过public.pgxc_copy_error_log引用。

```
GRANT INSERT,SELECT,DELETE ON  pg_catalog.pgxc_copy_error_log To load_user;
```

开启三权分立（即enableSeparationOfDuty=on）时，使用者可以是数据库普通用户或管理员用户。使用前需要到各自的schema下创建pgxc_copy_error_log表以及gs_copy_summary这两张表并添加索引，不需要再进行授权。

说明

因为pg_catalog模式下始终存在pgxc_copy_error_log以及gs_copy_summary两张表，而pg_catalog模式在search_path中优先级高于用户模式，因此在三权分立场景下无法直接通过gsql元命令\d(+)查询到，使用这两张表需要显式指定用户模式。

1. 使用初始用户创建新用户：
   ```
   CREATE USER load_user WITH PASSWORD '********';
   ```
2. 从初始用户切换为新用户：
   ```
   \c - load_user
   ```
3. 创建gs_copy_summary表并添加索引：
   ```
   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. （可选）创建pgxc_copy_error_log表并添加索引：

   说明

   1. 如果guc参数enable_copy_error_log未设置（默认为off），或者设置为off，则无需使用错误表，无需创建。否则需要创建该错误表。

   ```
   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);
   ```

使用环境

由用户自己将工具路径添加到PATH中。gs_loader支持SSL加密通信，使用方式同gsql方式。

使用指导

步骤1 （非三权分立）仅对于普通用户。

1. （在管理员用户下）创建用户：
   ```
   CREATE USER load_user WITH PASSWORD '************';
   ```
2. （在管理员用户下）给用户授权gs_copy_summary表：
   ```
   GRANT INSERT,SELECT ON  pg_catalog.gs_copy_summary To load_user;
   ```
3. 切换用户。
   ```
   \c - load_user
   ```

步骤2 （三权分立）对于普通用户和管理员用户。

 版权所有 © 华为云计算技术有限公司

1. （在初始用户下）创建用户：
   ```
   CREATE USER load_user WITH PASSWORD '********';
   ```
2. （在初始用户下）切换为load_user用户：
   ```
   \c - load_user
   ```
3. 创建gs_copy_summary表并添加索引。
   ```
   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);
   ```

步骤3 创建表和控制文件，准备数据文件。

创建表loader_tbl。

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

（在gs_loader客户端机器上）创建控制文件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
)
```

（在gs_loader客户端机器上）创建guc参数文件guc.txt。

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

（在gs_loader客户端机器上）创建数据文件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
```

步骤4 进行导入。

（在gs_loader客户端机器上）执行导入前，先确认gs_loader工具有可执行权限。确保当前路径有文件写入权限（gs_loader在处理过程中会生成一些临时文件，导入完成后自动删除）。

```
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
```

执行结果：

 版权所有 © 华为云计算技术有限公司

```
gs_loader: version 0.1

 5 Rows successfully loaded.

log file is:
 loader.log
```

----结束

说明

gs_copy_summary用于记录调用的copy语法及其详细情况，*[badfile]_bad.log*文件用于记录错误数据及其详细情况。为防止上一次导入时记录的错误数据以及详细情况被覆盖，建议每次执行导入时使用不同的bad参数。如果使用错误表pgxc_copy_error_log记录错误数据以及详细情况，请开启GUC参数enable_copy_error_log。如需删除表中的数据，可以对上述表执行truncate或者delete操作。

参数说明

表 3-25 gs_loader 参数说明

参数	参数说明	参数类型：取值范围
help	查看帮助信息。	-
user	数据库链接用户（与-U等价）。	字符串
-U	数据库链接用户（与user等价）。	字符串
passwd	用户密码（与-W等价）。	字符串
-W	用户密码（与passwd等价）。	字符串
db	数据库名称（必选，与-d等价）。	字符串
-d	数据库名称（必选与db等价）。	字符串
host	指定正在运行服务器的主机名、Unix域套接字的路径、或者域名。接受以“,”分隔的字符串来指定多个主机地址，支持指定多个主机地址（与-h等价）。 当指定多个主机地址时，默认选择连接到主节点。	参考gsql --host参数
-h	指定正在运行服务器的主机名、Unix域套接字的路径、或者域名。接受以“,”分隔的字符串来指定多个主机地址，支持指定多个主机地址（与host等价）。 当指定多个主机地址时，默认选择连接到主节点。	参考gsql --host参数
port	指定数据库服务器的端口号。可以配置一个或多个，当配置一个时，所有的IP都使用同一个端口连接；当配置多个时，顺序与IP顺序相同，个数必须与IP数相等，当不相等时会报错（与-p等价）。	参考gsql --port参考

版权所有 © 华为云计算技术有限公司

参数	参数说明	参数类型：取值范围
-p	指定数据库服务器的端口号。可以配置一个或多个，当配置一个时，所有的IP都使用同一个端口连接；当配置多个时，顺序与IP顺序相同，个数必须与IP数相等，当不相等时会报错（与port等价）。	参考gsql --port参考
create	是否创建pgxc_copy_error_log和gs_copy_summary表。由于当前版本会默认创建这两张表，因此该参数失去意义。保留仅为了兼容性。	[true, false]，默认true
clean	是否清除本次错误记录。	[true, false]，默认false
data	数据文件，可以指定多个，或者通配符多字符通配(*)以及单字符通配(?)（必选）。	字符串
control	控制文件名称（必选）。	字符串
log	日志文件名称。	字符串
bad	出错行以及详细情况记录文件名称，也可以指定目录，不指定时根据数据文件名生成。	字符串
discard	WHEN匹配失败行记录文件名称，也可以指定目录，根据数据文件名生成。	字符串
errors	允许数据文件中出现多少出错行。	整数，默认0
skip	允许跳过数据文件的前多少行。	整数，默认0
limit	指定最多导入的行数。	整数，默认无限大
bindsize	仅做语法兼容不实现功能。	-
rows	多行提交参数，指定导入多少行数据后进行一次提交。	整数，取值范围[1, 2147483647]
compatible_nul	是否开启数据中nul字符(0x00)兼容，开启后当数据文件中存在nul字符时，会先将nul字符转换成为空格字符' '(0x20)，再进行判断，加工和导入。	[true, false]，默认true。
compatible_illegal_chars	是否开启非法字符容错功能，与COPY语法中的COMPATIBLE_ILLEGAL_CHARS容错规则和限制相同，详见《开发指南》的“SQL 语法 > COPY”章节的COPY_OPTION中COMPATIBLE_ILLEGAL_CHARS参数说明。	[true, false]，默认false。

版权所有 © 华为云计算技术有限公司

参数	参数说明	参数类型：取值范围
parallel	指定并行导入的并发度。当并发度大于1时，表示开启并行导入模式。当并发度等于1时，转化为串行导入。并发度最大值限定不超过客户端CPU核数的两倍。当客户端运行在容器中时，由于获取到的是主机的cpu数量，可能比容器实际能使用的多，并发度建议由用户自行控制在实际能使用的客户端cpu数量的两倍范围内。另外该能力通过多线程并发多个事务实现，实际并发度会受限于服务端的线程池模型，也同时会给服务端增加压力，请按实际情况合理设置并发度。	整数，取值范围[1, CPU核数两倍]，默认1
binary	数据文件是否为通过copy的binary模式导出的二进制文件。	[true, false]，默认false。

版权所有 © 华为云计算技术有限公司

⚠ 注意

- 参数均为小写，不支持大写，同时兼容gsql登录方式：-p端口号，-h主机，-d数据库，-U用户名，-W密码方式。
- 使用rows参数时，提交次数不要超过1000次，否则会对性能产生影响。提交次数约等于数据文件中数据行数除以rows参数取值。不指定rows参数时，rows无默认取值，表现为只进行一次提交，即所有数据都导入表中后进行一次事务提交。
- 小数据量频繁的提交会影响导入数据的性能，推荐合理配置rows参数的取值，保证每次提交的数据量大于5MB。对于常用的16U128G规格机器，一主两备部署场景下，向5个字段的表内导入13GB数据，排除网络影响，多次提交和单次提交（每次提交5MB数据）的速率基本持平，为10MB/s左右。
- compatible_nul参数实际控制guc参数loader_support_nul_character值的设置：
  - compatible_nul=true对应session级set loader_support_nul_charchter='s2'。
  - compatible_nul=false对应session级set loader_support_nul_character='s1'。

    建议通过命令行设置此参数且通过compatible_nul设置优先级高于guc_param中设置。
- 当前gs_loader仅支持数据文件中存在nul字符时的兼容，不支持ctl控制文件中存在nul字符。ctl文件中存在nul字符会存在不可预期的问题。
- gs_loader导入过程中，不需要转码场景下，单行数据最大不超过1GB-1；转码场景下单行数据最大不超过256MB-1。
- 如果是基本的数据迁移场景，建议先删除表上的索引，禁用表上的触发器，待数据迁移完成后再重建索引，恢复表上的触发器。对导入性能会有所提升。
- 指定binary参数为true后，有以下行要求：
  - 数据文件必须为通过\COPY中BINARY模式导出的二进制格式数据文件，但是该模式导出的数据文件通常兼容性及可移植性较差，建议直接使用\COPY语句进行导入。
  - gs_loader会将控制文件中语法转换为\COPY中BINARY模式下最简单的语法，即\COPY table_name FROM 'binary_file_path' BINARY; 语句。只解析控制文件中导入模式，表名信息和命令行中的control、data、binary、guc_param及数据库连接参数信息，不对其他参数语法进行解析和生效。
  - 对于gs_loader的命令行及控制文件中有以下要求：
    - 不支持字符集配置。
    - 不支持WHEN条件过滤及DISCARD生成。
    - 不支持enable_copy_error_log = off下将错误数据直接写入bad文件，enable_copy_error_log=on下将错误数据直接写入错误表。
    - 不支持配置CSV模式，不支持指定分隔符及包裹符，不支持TRAILING NULLCOLS语法。
    - 不支持数据类型配置、POSITION配置及列表达式使用。
    - 不支持FILLER、CONSTANT、SEQUENCE、NULLIF参数。
    - 不支持skip、rows、compatible_nul、compatible_illegal_chars参数。
- 指定parallel大于1时：
  - 当同时设置binary参数为true时，parallel参数失效，按串行导入。
  - 不支持在控制文件中设置OPTIONALLY ENCLOSED BY或者FIELDS CSV。

版权所有 © 华为云计算技术有限公司

- 不支持在控制文件中设置SEQUENCE列。
- 无法保证数据按数据文件中的顺序导入。如果表中存在自增列，导入后自增列值的顺序无法保证与数据文件中顺序一致。
- 同时使用errors参数时，errors参数的意义为每个子任务允许出现的最大错误行数。
- 同时使用skip参数时，skip参数的意义是在整个数据文件开头跳过的行数。
- 同时使用rows参数时，分批提交的批次各子任务独立计算。
- 在客户端CPU、内存和服务端CPU、内存、空闲线程以及网络带宽不存在瓶颈，并且bad/discard总数不超过1%时，相比于串行导入，并发度为2/4/8时的性能提升不低于1.5/3/5倍。
- 并发度每增加1，大约增加客户端10MB内存，服务端大约35MB内存。

- 当设置GUC参数support_zero_character为on时，表示数据库支持0x00字符的写入和读取，gs_loader导入数据时，会将0x00按照原始样式导入，而不是受其他兼容性参数影响转换成0x20。对于0字符处理场景，参数优先级为support_zero_character > copy_special_character_version > compatible_illegal_chars > loader_support_nul_character。
- 当设置了copy_special_character_version='no_error'时，优先级高于compatible_illegal_chars，非法编码的数据会原样导入，而不是发生转换。

控制文件

- 语法说明：

```
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)]
[, ...]
)]
```

- 参数说明：
  - **CHARACTERSET**

    字符集。

    取值范围：字符串，目前可指定为，'AL32UTF8','zhs16gbk','zhs32gb18030'。

    注意：控制文件中**CHARACTERSET**指定的字符集，应该和文件的编码格式保持一致，否则会报错或者导入数据乱码。
  - **INFILE**

    当前关键字无效，并在控制文件中需要单独占一行，运行时候会忽略该关键字。需要用户在gs_loader命令行参数中指定对应的数据文件。
  - **BADFILE**

 版权所有 © 华为云计算技术有限公司

当前关键字无效，运行时候会忽略该关键字，如果gs_loader 命令行参数没有指定badfile，则会根据对应控制文件名称生成对应的badfile文件。

- **OPTIONS**

  其中只有skip和rows功能生效，skip=n为导入时跳过前n条数据，rows=n为导入多少行数据后进行一次提交。命令行和控制文件同时指定时，命令行优先级更高。

- **INSERT | APPEND | REPLACE | TRUNCATE**

  导入模式。

  INSERT：如果表中有数据，则报错。

  APPEND：直接插入数据。

  REPLACE：如果表中有数据，则全部删除，然后再插入。

  TRUNCATE：如果表中有数据，则全部删除，然后再插入。

  说明

  - 在写控制文件(.ctl)文件时，在INTO TABLE table_name语句前后都可以指定（导入模式，INSERT | APPEND | REPLACE | TRUNCATE），使用优先级为：在INTO TABLE table_name语句后面指定导入模式优先级高于在INTO TABLE table_name语句前面指定导入模式，在INTO TABLE table_name语句后面指定导入模式会覆盖在前面指定的导入模式。
  - 当开启多个gs_loader会话，并发地向同一张表中导入数据时，推荐以APPEND的方式进行导入，以INSERT|REPLACE|TRUNCATE的方式会出现导入报错或数据导入不全的问题。

- **FIELDS CSV**

  标识使用copy的CSV模式。在CSV模式下分隔符缺省值为逗号，引号字符的缺省值为双引号。

  说明

  - 当前CSV模式下，被双引号包含的换行符被视为字段数据的一部分。
  - CSV模式下，设置了GUC参数a_format_copy_version为's1'时，会跳过字段开头空格。并且当某个字段第一个非空格字符不是enclosed字符时，忽略enclosed设置。当未匹配到关闭enclosed字符，先匹配到行末时，会进行报错。
  - CSV模式下，在不打开0字符GUC开关support_zero_characters时，如果使用了compatible_nul或者compatible_illegal_chars参数对0x00字符进行兼容，由于0x00转换0x20的行为发生在跳过开头空格的行为之前，因此字段开头的0x00字符会被当作0x20处理被删除掉。

- **table_name**

  表的名称（可以有模式修饰）。

  取值范围：已存在的表名。

- **TERMINATED [BY] { 'string' }**

  在文件中分隔各个字段的字符串，分隔符最大长度不超过10个字节。

  取值范围：不允许包含\.abcdefghijklmnopqrstuvwxyz0123456789中的任何一个字符。不支持将nul字符设置为分隔符。

  缺省值：在文本模式下，缺省是水平制表符，在CSV模式下是一个逗号。

版权所有 © 华为云计算技术有限公司

⚠ 注意

开启nul字符兼容，即compatible_nul=true，如果指定分隔符为' '空格字符(0x20)时需要注意，所判断的分隔符为数据文件中已存在的空格字符，并非nul字符转换而来的空格字符。

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

  CSV格式文件下的引号字符。

  仅在使用FIELDS CSV参数明确说明的CSV模式下缺省值：双引号。

  其余模式下无缺省值。

  说明

  - 设置**OPTIONALLY ENCLOSED BY { 'string' }**时，数据左边可以不带引号字符，如果有引号字符，数据左右都必须为奇数个，但个数不必相等。
  - 当前仅CSV模式支持**OPTIONALLY ENCLOSED BY { 'string' }**。当指定**OPTIONALLY ENCLOSED BY { 'string' }**时，默认进入CSV模式。

- **TRAILING NULLCOLS**

  当数据加载时，若数据源文件中一行的多个字段缺失的处理方式。

  当一行数据的最后存在一个或多个字段为空时，按照空值处理将其导入到表中。不设置则会报错字段为空，将这行数据当作错误数据处理。

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

  对行中的start到end之间的字符串，或者根据列名进行行过滤。

  取值范围：字符串。

  说明

  - 当GUC参数enable_copy_when_filler=on（默认）时，支持根据FILLER类型列进行过滤。当GUC参数enable_copy_when_filler=off时，则不支持。
  - WHEN条件后的常量字符串中不支持'\0'、'\r'等特殊字符。

- **POSITION ({ start:end })**

  对列进行处理，根据start到end范围获取对应字符串。

须知

使用POSITION用法表示在转换后的\COPY语句中使用FIXED FORMATTER的定长模式，会无法正确处理字段数据中的换行符。

- **"sql_string"**

  对列进行处理，列表达式，根据表达式计算列的取值。详见**•列表达式**。

  取值范围：字符串。

- **FILLER**

  对列进行处理，如果出现FILLER，则这个字段跳过。

  说明

  当前不支持FILLER与POSITION **({ start:end })**同时使用。

- **column_type [external]**

版权所有 © 华为云计算技术有限公司

在导入数据时，根据不同的数据类型对数据进行处理。详见•**数据类型**。

- **CONSTANT**

  对列进行处理，将插入的对应字段设置为常量。

  取值范围：字符串。

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

  对列进行处理，生成对应的序列值。

  - COUNT：表示根据表中数据的行数开始计算。
  - MAX：表示根据表中这一列的最大值开始计算。
  - integer：表示从用户指定的值开始计算。
  - incr：表示每次递增多少。

- **NULLIF**

  在设置a_format_copy_version等于's1'时，当指定列的数据只包含空白字符时返回NULL，否则返回trim(COL)，等价于列表达式 "nullif(trim(COL), '')"。

  在设置a_format_copy_version不等于's1'时，对列进行处理，在多行导入场景中，若列名后未指定sysdate、constant、position、列表达式等运算时，执行导入操作，表现为未指定NULLIF关键字的列字段设置为空。

  当前只支持COL POSITION() CHAR NULLIF (COL=BLANKS)语法。具体使用详见•**NULLIF使用用例**。

⚠ 注意

- 不支持OPTIONS、INFILE、BADFILE，仅在特定场景下不报语法错误。
- gs_loader使用bad文件来记录出错数据，如果设置guc参数enable_copy_error_log开启错误表，该数据来自错误表的rawrecord字段，由于错误表对于以某种编码无法读起的错误不记录rawrecord，因此bad文件中遇到此情况时记录空行。
- gs_loader在设置guc参数a_format_load_with_constraints_violation开启支持约束冲突不回滚场景时，如果表带有BEFORE/AFTER ROW INSERT触发器，则每次提交行数不能超过1000万行。
- gs_loader在设置guc参数a_format_load_with_constraints_violation开启支持约束冲突不回滚场景时，不支持语句级触发器。

- bad文件对应数据为空的需要对应错误表的内容参考源文件和行号（不识别某种编码序列，不写bad文件内容，只记录空行）。

```
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)
//如上例子对于loader对应的文件，查找数据文本第一行找出源数据
```

- NULLIF使用用例

```
// 建表
create table gsloader_test_nullif(
```


版权所有 © 华为云计算技术有限公司

```
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)
);
// 数据文件 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
// 控制文件 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))
// 执行导入
gs_loader -p xxx host=xxx control=test.ctl  data=test.csv -d testdb -W xxx
// 导入结果：导入成功
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)
```

从导入表中的数据可以看出在使用NULLIF关键字后，除指定NULLIF运算和sysdate运算的列执行导入操作后导入字段正常，其余未指定运算的列表现为导入字段为空。

- 列表达式

  gs_loader支持对指定列进行表达式转换和场景扩展：

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

  其中data_type指定该列在表达式参数中的数据类型；transform_expr为目标表达式，返回与表中目标列数据类型一致的结果值。

  示例：

  – ctl文件中不指定列类型，源数据不满足表中列限制（数据类型限制、数据长度限制）。

    ```
    // 建表
    create table t_test(id int, text varchar(5));
    // 数据文件 test.csv
    addf2,bbbbaaa,20220907,
    ```

 版权所有 © 华为云计算技术有限公司

```
// 控制文件 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';
// 执行导入
gs_loader -p xxx host=xxx control=test.ctl  data=test.csv -d testdb -W xxx guc_param=test_guc.txt
// 导入结果：导入成功
select * from t_test;
 id | text
----+-------
  5 | aaaaa
(1 row)
```

- ctl文件中不指定列类型，隐式类型转换（涉及隐式类型转换，建议加上兼容性参数）。

```
// 建表
create table test(mes int, mes1 text, mes2 float8, mes3 timestamp with time zone, mes4
INTEGER);
// 数据文件
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
// 控制文件
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';
//执行导入
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
// 导入结果：导入成功
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
```

- 数据类型

  对应控制文件中的column_type [external]，在加载数据时，根据不同的数据类型对数据进行处理。gs_loader中可以将数据类型分为普通数据类型和特殊数据类型。

  - 普通数据类型

    - CHAR [(length)]：

      按照字段分隔符读取数据，并转换使用CHAR类型来保存值。length表示单条数据的最大长度，以字节为单位，通常一个字符占用一个字节，并且可以缺省，分为以下几种场景：

 版权所有 © 华为云计算技术有限公司

- 缺省对length长度的声明时，length的值会根据POSITION的声明来继承最大长度值。
- 声明了length的长度，则它会覆盖POSITION中对于最大长度的声明。
- 缺省了length的声明，同时也缺省了POSITION的声明，length的长度会根据分隔符间长度进行设置。
- 对于长度声明的优先级：length > POSITION > 分隔符。
- 缺省length，POSITION，分隔符的声明时，会从当前位置读到行结束符为止。
- 如果实际数据长度超过了length声明的最大长度，会报错。

■ INTEGER external [(length)]：

按照字段分隔符读取数据，并转换使用INTEGER类型来保存值。length的使用规则与CHAR类型中相同。

■ FLOAT external [(length)]：

按照字段分隔符读取数据，并转换使用FLOAT类型来保存值。length的使用规则与CHAR类型中相同。

■ DECIMAL external (length)：

按照字段分隔符读取数据，并转换使用DECIMAL类型来保存值。length的使用规则与CHAR类型中相同

■ TIMESTAMP：

按照字段分隔符读取数据，并转换使用TIMESTAMP类型来保存值。

■ DATE：

按照字段分隔符读取数据，并转换使用DATE类型来保存值。

■ DATE external：

按照字段分隔符读取数据，并转换使用DATE类型来保存值。

■ SYSDATE：

在数据库执行对应的插入时，取系统时间。该字段对应对应的值无法被引用使用，被引用使用的内容为SYSDATE字符串。

– 特殊数据类型

■ INTEGER：

无视字段分隔符读取四个字节长度的字符，按小端存储逻辑保存，然后将每个字符解析成十六进制ASCII码值，最后将整体转换为十进制数来保存值。

■ SMALLINT：

无视字段分隔符读取两个字节长度的字符，按小端存储逻辑保存，然后将每个字符解析成十六进制ASCII码值，最后将整体转换为十进制数来保存值。

示例：

```
// 建表
create table t_spec(col1 varchar(10), col2 varchar(10));
// 数据文件
```

 版权所有 © 华为云计算技术有限公司

```
cat t_spec.txt
1234,5678,
// 控制文件
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';
// 执行导入
gs_loader -p xxx host=xxx control=t_spec.ctl data=t_spec.txt -d testdb -W xxx
guc_param=test_guc.txt
// 导入结果：导入成功
select * from t_spec;
   col1    | col2
-----------+-------
 741618482 | 13612
(1 row)
```

- RAW：

  会把每个字符解析成ASCII码值保存，转义字符“\”不执行转义操作。

  限制：RAW不能使用分隔符。

  示例：

```
// 建表
create table t_raw(col raw(50));
// 数据文件
cat t_raw.txt
12\n\x78!<~?'k^(%s)>/c[$50]
// 控制文件
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';
// 执行导入
gs_loader -p xxx host=xxx control=t_raw.ctl data=t_raw.txt -d testdb -W xxx
guc_param=test_guc.txt
// 导入结果：导入成功
select * from t_raw;
                            col
------------------------------------------------------------
 31325C6E5C783738213C7E3F276B5E282573293E2F635B2435305D
(1 row)
```

 版权所有 © 华为云计算技术有限公司

⚠ 注意

- 在多列导入场景中，不指定guc参数时，部分position与分隔符不能同时使用。
- 在多列导入场景中，SYSDATE和CONSTANT运算不能和POSITION运算同时使用。
- 指定数据类型导入时，包含普通数据类型需要通过guc_param设置a_format_copy_version参数，包含特殊数据类型则需要通过guc_param设置a_format_copy_version和a_format_dev_version及a_format_version参数。
- 列表达式涉及到系统函数时，需要根据对应功能通过guc_param设置合适的a_format_dev_version及a_format_version参数。
- 带length数据类型的使用，length需指定为大于0的整数；RAW数据类型作为特殊类型，RAW(length)的使用区别于普通类型的使用，如INTEGER EXTERNAL(length)的使用，当不指定position时，INTEGER EXTERNAL(length)表现为，当length小于文本文件 (.csv/.txt等)中对应列数据长度时报错；当length大于文本文件 (.txt)中对应列数据长度时，输出INTEGER EXTERNAL类型的结果。RAW(length)当不指定position时表现为读取length个字符。
- POSITION使用时，POSITION(start:end)，start需设置为大于0的整数，且end值应大于等于start的值。
- 指定POSITION时，在处理字段内容时不会省略尾部的空格；不指定POSITION时，处理字段内容时会省略尾部的空格，如果需要保留空格，需要在guc_param所指定的文件中，已设置好a_format_version的前提下，添加 set behavior_compat_options='char_coerce_compat'; 详细内容请参考管理员指南中behavior_compat_options的设置。
- 并发导入时，若多个gs_loader的discard文件名或bad文件名指向同一目录同名文件，则后一个执行的gs_loader会中止报错。若前一个已经导入完成，则文件被覆盖。

  报错如下：

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

  log文件中：

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

- 控制文件中对于字段值的部分若不为空且不使用本字段内容，则不占用数据文件的位置。

  比如控制文件如下：

  ```
  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)"
  )
  ```

  数据文件如下：

  ```
  11,22你好,33,
  ```

  导入结果为：

  ```
  loader=# select * from gsloader;
  id |  text  |     gmt_create      | create_str
  ```


版权所有 © 华为云计算技术有限公司

```
----+---------+---------------------+------------
11 | 2023-02-08 16:00:54 | 22你好 |  33
```

3.3 gs_dump

背景信息

gs_dump是GaussDB用于导出数据库相关信息的工具，用户可以自定义导出一个数据库或其中的对象（模式、表、视图等）。支持导出的数据库可以是默认数据库postgres，也可以是自定义数据库。

gs_dump支持导出PDB。当导出PDB时，用户可以自定义导出一个PDB或其中的对象（模式、表、视图等）。

gs_dump工具在进行数据导出时，其他用户可以访问数据库（读或写）。

gs_dump工具支持导出完整一致的数据。例如，T1时刻启动gs_dump导出A数据库，那么导出数据结果将会是T1时刻A数据库的数据状态，T1时刻之后对A数据库的修改不会被导出。

gs_dump时生成列不会被转储。

gs_dump时HTAP表创建的IMCV元信息（gs_imcv系统表）不会被转储。

gs_dump支持导出兼容v1数据库的文本格式文件。

gs_dump支持将数据库信息导出至纯文本格式的SQL脚本文件或其他归档文件中。

- 纯文本格式的SQL脚本文件：包含将数据库恢复为其保存时的状态所需的SQL语句。通过**gsql**运行该SQL脚本文件，可以恢复数据库。即使在其他主机和其他数据库产品上，只要对SQL脚本文件稍作修改，也可以用来重建数据库。
- 归档格式文件：包含将数据库恢复为其保存时的状态所需的数据，可以是tar格式、目录归档格式或自定义归档格式，详见**表3-26**。

gs_dump支持SSL加密通信，使用方式同gsql方式。

使用gs_dump前请确保gs_dump版本与数据库版本保持一致，高版本gs_dump不保证完全兼容低版本内核数据。

gs_dump不适合库中对象数量过多的场景。当库中对象数量过多，或者对象间依赖关系过于复杂时，gs_dump导出时间会很长。

主要功能

gs_dump可以创建四种不同的导出文件格式，通过**[**-F或者--format=**]**选项指定，具体如**表3-26**所示。

版权所有 © 华为云计算技术有限公司

表 3-26 导出文件格式

格式名称	-F的参数值	说明	建议	对应导入工具
纯文本格式	p	纯文本脚本文件包含SQL语句和命令。命令可以由gsql命令行终端程序执行，用于重新创建数据库对象并加载表数据。	小型数据库，一般推荐纯文本格式。	使用**gsql**工具恢复数据库对象前，可根据需要使用文本编辑器编辑纯文本导出文件。
自定义归档格式	c	一种二进制文件。支持从导出文件中恢复所有或所选数据库对象。	中型或大型数据库，推荐自定义归档格式。	使用**gs_restore**可以选择要从自定义归档/目录归档/tar归档导出文件中导入相应的数据库对象。
目录归档格式	d	该格式会创建一个目录，该目录包含两类文件，一类是目录文件，另一类是每个表和blob对象对应的数据文件。	-	
tar归档格式	t	tar归档文件支持从导出文件中恢复所有或所选数据库对象。tar归档格式不支持压缩且对于单独表大小应小于8GB。	-	

说明

可以使用gs_dump工具将文件压缩为目录归档或自定义归档导出文件，减少导出文件的大小。生成目录归档或自定义归档导出文件时，默认进行中等级别的压缩。gs_dump程序无法压缩已归档导出文件。

注意事项

- 禁止修改-F c/d/t 格式导出的文件和内容，否则可能无法恢复成功。对于-F p 格式导出的文件，如有需要，可谨慎编辑导出的文件。
- 为了保证数据一致性和完整性，gs_dump会对需要转储的表设置共享锁。如果表在别的事务中设置了共享锁，gs_dump会等待锁释放后锁定表。如果无法在指定时间内锁定某个表，转储会失败。用户可以通过指定--lock-wait-timeout选项，自定义等待锁超时时间。
- 不支持加密导出存储过程和函数。
- 对于物化视图，本工具仅支持物化视图定义的导出，在导入后需手动执行REFRESH命令来进行数据恢复。
- 对于临时对象，本工具仅支持导出全局临时表。
- 本工具不支持在备机上使用。
- gs_dump导出分区索引时，部分索引分区的属性无法导出，比如索引分区的unusable状态。可以通过查询系统表PG_PARTITION或者查询视图

版权所有 © 华为云计算技术有限公司

ADM_IND_PARTITIONS/ADM_IND_SUBPARTITIONS获取索引分区的具体属性，通过ALTER INDEX命令可以手动设置索引分区属性。

- 对于定时任务，本工具仅支持导出在B兼容性数据库中，通过CREATE EVENT创建的定时任务或通过高级包创建的非周期性定时任务。
- gs_dump不支持导出自定义Tokenweight分词词典，可以根据报错WARNING: dictionary xx cannot be automatically exported, please create it manually手动创建对应分词词典。
- 在多租场景下，使用gs_dump导出时，不支持导出模板PDB，也不支持导出关闭的PDB。
- 在多租场景下，普通用户使用gs_dump导出时，只能导出该用户有权限的数据库对象和数据。
- 当未开启多租时，gs_dump不支持导出PDB及其内的对象。
- gs_dump时HTAP表创建IMCV元信息(gs_imcv系统表)不会被转储。
- 如果数据库中存在初始用户创建的表且表上有含用户自定义函数的表达式索引，系统管理员使用gs_dump导出后，需要使用初始用户通过gsql或gs_restore进行导入。否则会因为安全原因，导致创建索引失败。

说明

普通用户不支持导出DIRECTORY、SYNONYM，若普通用户进行相关导出，会提示“WARNING: xx not dumped because current user is not a superuser”。

语法

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

说明

“dbname”前面不需要加短或长选项。“dbname”指定要连接的数据库。

例如：

不需要-d，直接指定“dbname”。

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

或者

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

环境变量：PGDATABASE

参数说明

通用参数：

- -f, --file=<FILE_NAME>

  将输出发送至指定文件或目录。如果省略该参数，则使用标准输出。如果输出格式为(-F c/-F d/-F t)时，必须指定-f参数。如果-f的参数值含有目录，要求目录对当前用户具有读写权限。

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

  选择输出格式。格式如下：

  - p|plain：输出一个文本SQL脚本文件（默认）。
  - c|custom：输出一个自定义格式的归档，并且以目录形式输出，作为gs_restore输入信息。该格式是最灵活的输出格式，因为能手动选择，而且能在恢复过程中将归档项重新排序。该格式默认状态下会被压缩。

版权所有 © 华为云计算技术有限公司

  - d|directory：该格式会创建一个目录，该目录包含两类文件，一类是目录文件，另一类是每个表和blob对象对应的数据文件。
  - t|tar：输出一个tar格式的归档形式，作为gs_restore输入信息。tar格式与目录格式兼容；tar格式归档形式在提取过程中会生成一个有效的目录格式归档形式。但是，tar格式不支持压缩且对于单独表有8GB的大小限制。此外，表数据项的相应排序在恢复过程中不能更改。

- -v, --verbose

  指定verbose模式。该选项将导致gs_dump向转储文件输出详细的对象注解和启动/停止次数，向标准错误流输出处理信息。

- -V, --version

  打印gs_dump版本，然后退出。

- -Z, --compress=0-9

  指定使用的压缩比级别。

  取值范围：0~9

  - 0表示无压缩。
  - 1表示压缩比最小，处理速度最快。
  - 9表示压缩比最大，处理速度最慢。

  针对自定义归档格式，该选项指定单个表数据片段的压缩，默认方式是以中等级别进行压缩。tar归档格式和纯文本格式目前不支持压缩。

- --lock-wait-timeout=TIMEOUT

  请勿在转储刚开始时一直等待以获取共享表锁。如果无法在指定时间内锁定某个表，就选择失败。可以以任何符合SET statement_timeout的格式指定超时时间。

- -?, --help

  显示gs_dump命令行参数帮助，然后退出。

转储参数：

- -a, --data-only

  只输出数据，不输出模式(数据定义)。转储表数据、大对象和序列值。

- -b, --blobs

  该参数为扩展预留接口，不建议使用。

- -c, --clean

  在将创建数据库对象的指令输出到备份文件之前，先将清理（删除）数据库对象的指令输出到备份文件中。（如果目标数据库中没有任何对象，gsql或gs_restore工具可能会输出一些提示性的错误信息。）

  该选项只对文本格式有意义。针对归档格式，可以在调用gs_restore时指定选项。

- -C, --create

  备份文件以创建数据库和连接到创建的数据库的命令开始（如果命令脚本是这种方式执行，无所谓在运行脚本之前连接的是哪一个数据库）。

  该选项只对文本格式有意义。针对归档格式，可以在调用gs_restore时指定选项。

  说明

  - 在多租场景下，使用gs_dump导出指定PDB时，不支持使用该选项。
  - 在M-Compatibility模式数据库下，不支持使用该选项。必须先在目标实例上创出M-Compatibility模式数据库，然后在源实例上导出，最后在目标实例连接新创出的M-Compatibility模式数据库导入。

 版权所有 © 华为云计算技术有限公司

- -E, --encoding=ENCODING

  以指定的字符集编码创建转储。默认情况下，以数据库编码创建转储。（得到相同结果的另一个办法是将环境变量“PGCLIENTENCODING”设置为所需的转储编码）

  说明

  当指定转储编码存在转码场景时，且表中的数据存在非法编码的数据，导出会报错invalid byte sequence，建议使用gs_dump的-s参数只导出定义，并单独使用COPY打开编码容错进行数据的导出与导入。

- -n, --schema=SCHEMA

  只转储与模式名称匹配的模式，此选项包括模式本身和所有它包含的对象。如果该选项没有指定，所有在目标数据库中的非系统模式将会被转储。写入多个-n选项来选择多个模式。此外，根据gsql的\d命令所使用的相同规则，模式参数可被理解成一个pattern，所以多个模式也可以通过在该pattern中写入通配符来选择。使用通配符时，注意给pattern打引号，防止shell扩展通配符。

  说明

  - 当-n已指定时，gs_dump不会转储已选模式所附着的任何其他数据库对象。因此，无法保证某个指定模式的转储结果能够自行成功地储存到一个空数据库中。
  - 当-n指定时，非模式对象不会被转储。
  - M-Compatibility兼容模式下， 通过CREATE DATABASE带templatem创建的数据库直接通过指定db_name导出数据； 而通过CREATE DATABASE db_name创建出来的DATABASE与Schema等效，只能通过-n导出数据。
  - GaussDB会自动将对象名称中的大写字母转为小写，当模式名称中包含大写字母时，需要添加额外的引号，如：-n '"Sch1"' 或 -n "\"Sch1\""
  - M-Compatibility兼容模式下，该参数值受GUC参数lower_case_table_names影响。比如在大小写敏感模式下（lower_case_table_names=0），该参数值也需要大小写敏感，如果包含大写字母，则需要添加额外的引号，否则效果相当于小写。在大小写不敏感模式下（lower_case_table_names=1），该参数值需要传小写名称。

  转储支持多个模式的转储。多次输入-n schemaname转储多个模式。

  例如：

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

  在上面这个例子中，sch1和sch2会被转储。

- -N, --exclude-schema=SCHEMA

  不转储任何与模式pattern匹配的模式。Pattern将参照针对-n的相同规则来理解。可以通过输入多次-N，不转储与任何pattern匹配的模式。

  当同时输入-n和-N时，会转储与至少一个-n选项匹配、与-N选项不匹配的模式。如果有-N没有-n，则不转储常规转储中与-N匹配的模式。

  转储过程支持排除多个模式。

  在转储过程中，输入-N exclude schema name排除多个模式。

  例如：

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

  在上面这个例子中，sch1和sch2在转储过程中会被排除。

  说明

  GaussDB会自动将对象名称中的大写字母转为小写，当模式名称中包含大写字母时，需要添加额外的引号，如：-N '"Sch1"' 或 -N "\"Sch1\""

- -o, --oids

 版权所有 © 华为云计算技术有限公司

转储每个表的对象标识符（OIDs），作为表的一部分数据。该选项用于应用以某种方式（例如：外键约束方式）参照了OID列的情况。如果不是以上这种情况，请勿使用该选项。

- -O, --no-owner

  不输出设置对象的归属这样的命令，以匹配原始数据库。默认情况下，gs_dump会发出ALTER OWNER或SET SESSION AUTHORIZATION语句设置所创建的数据库对象的归属。如果脚本正在运行，该语句不会执行成功，除非是由系统管理员触发（或是拥有脚本中所有对象的同一个用户）。通过指定-O，编写一个任何用户都能存储的脚本，且该脚本会授予该用户拥有所有对象的权限。

  该选项只对文本格式有意义。针对归档格式，可以在调用gs_restore时指定选项。

- -s, --schema-only

  只转储对象定义（模式），而非数据。

- -S, --sysadmin=NAME

  该参数为扩展预留接口，不建议使用。

- -t, --table=TABLE

  指定转储的表（或视图、或序列、或外表）对象列表，可以使用多个-t选项来选择多个表，也可以使用通配符指定多个表对象。

  当使用通配符指定多个表对象时，注意给pattern打引号，防止shell扩展通配符。

  当使用-t时，-n和-N没有任何效应，这是因为由-t选择的表的转储不受那些选项的影响。

  说明

  - -t参数选项个数必须小于等于100。
  - 如果-t参数选项个数大于100，建议使用参数--include-table-file来替换。
  - 当-t已指定时，gs_dump不会转储已选表所附着的任何其他数据库对象。因此，无法保证某个指定表的转储结果能够自行成功地储存到一个空数据库中。
  - -t tablename只转储在默认搜索路径中可见的表。-t *.tablename转储数据库下所有模式下的tablename表。-t schema.table转储特定模式中的表。
  - -t tablename不会导出表上的触发器信息。
  - 对于表名中包含大写字母的表，在使用-t参数指定导出时需对表名添加\"来导出。如对于表"abC"，导出需指定-t \"abC\"；如对于表schema."abC"，导出需指定-t schema.\"abC\"。
  - -t "" 不匹配任何表
  - M-Compatibility兼容模式下，该参数值受GUC参数lower_case_table_names影响。比如在大小写敏感模式下（lower_case_table_names=0），该参数值也需要大小写敏感，如果包含大写字母，则需要添加额外的引号，否则效果相当于小写。在大小写不敏感模式下（lower_case_table_names=1），该参数值需要传小写名称。

  例如：

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

  在上面这个例子中，schema1.table1和schema2.table2会被转储。

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

  指定需要dump的表文件。

- -T, --exclude-table=TABLE

  不转储的表（视图、序列、外表）对象列表，可以使用多个-T选项来选择多个表，也可以使用通配符指定多个表对象。

  当同时输入-t和-T时，会转储在-t列表中，而不在-T列表中的表对象。

 版权所有 © 华为云计算技术有限公司

例如：

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

在上面这个例子中，table1和table2在转储过程中会被排除。

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

  指定不需要dump的表文件。

  说明

  同--include-table-file，其内容格式如下：

  schema1.table1

  schema2.table2

  .....

- -x, --no-acl

  防止转储访问权限（授权/撤销命令），只影响acl对象，不影响privilege对象。

- -q, --target

  指定导出兼容其他版本数据库的文本文件，目前支持v1和v5参数，指定其他参数不会报错，但不会生效。v1参数用于导出GaussDB v5版本数据库的数据为兼容GaussDB v1版本的文本文件。v5参数用于导出GaussDB v5版本数据库的数据为GaussDB v5版本数据库格式的文本文件，减少了导入GaussDB v5版本数据库时的可能的报错情况。

  在使用v1参数时，建议和--exclude-guc="enable_cluster_resize"，--exclude-function，--exclude-with等选项共用，否则导入到GaussDB v1版本数据库时可能报错。

- -g, --exclude-guc

  该参数为扩展预留接口，不建议使用。

- --exclude-function

  不导出函数和存储过程。

- --exclude-with

  导出的表定义，末尾不添加WITH(orientation=row，compression=on）这样的描述。

- --binary-upgrade

  该参数为扩展预留接口，不建议使用。

  说明

  M-Compatibility模式数据库不支持该选项。

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

  该参数为扩展预留接口，不建议使用。

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

  以INSERT命令带列名（INSERT INTO表（列、…）值…）方式导出数据。这会导致恢复缓慢。但是由于该选项会针对每行生成一个独立分开的命令，所以在重新加载某行时出现的错误只会导致对应的一行数据丢失，而非整个表内容。

  说明

  M-Compatibility模式数据库不支持该选项。

- --disable-dollar-quoting

 版权所有 © 华为云计算技术有限公司

该选项将禁止在函数体前使用美元符号$，并强制使用SQL标准字符串语法对其进行引用。

- --include-alter-table

  dump后的表删除列。

- --disable-triggers

  该参数为扩展预留接口，不建议使用。

- --exclude-table-data=TABLE

  指定不转储任何匹配表pattern的表这方面的数据。依照针对-t的相同规则理解该pattern。

  可多次输入--exclude-table-data来排除匹配任何pattern的表。当用户需要特定表的定义但不需要其中的数据时，这个选项很有帮助。

  排除数据库中所有表的数据，请参见**-s, --schema-only**。

- --inserts

  发出INSERT命令（而非COPY命令）时转储数据。这会导致恢复缓慢。

  但是由于该选项会针对每行生成一个独立分开的命令，所以在重新加载某行时出现的错误只会导致对应的一行数据丢失，而非整个表内容。注意如果重排列顺序，可能会导致恢复整个失败。列顺序改变时，--column-inserts选项不受影响，虽然会更慢。

  说明

  M-Compatibility模式数据库不支持该选项。

- --no-security-labels

  该参数为扩展预留接口，不建议使用。

- --no-tablespaces

  不输出选择表空间的命令。使用该选项，无论默认表空间是哪一个，在恢复过程中所有对象都会被创建。

  该选项只对文本格式有意义。针对归档格式，可以在调用gs_restore时指定选项。

- --no-unlogged-table-data

  该参数为扩展预留接口，不建议使用。

- --non-lock-table

  该参数仅供软件间接口调用。

- --quote-all-identifiers

  强制对所有标识符加引号。为了向后续版本迁移，且其中可能涉及引入额外关键词，在转储相应数据库时该选项会有帮助。

  说明

  M-Compatibility模式数据库不支持该选项。

- --section=SECTION

  指定已转储的名称区段（pre-data、data、和post-data）。

- --serializable-deferrable

  转储过程中使用可串行化事务，以确保所使用的快照与之后的数据库状态一致；要实现该操作需要在无异常状况的事务流中等待某个点，因为这样才能保证转储成功，避免引起其他事务出现serialization_failure要重新再做。

 版权所有 © 华为云计算技术有限公司

但是该选项对于灾难恢复没有益处。对于在原始数据库进行升级的时候，加载一个数据库的复制作为报告或其他只读加载共享的转储是有帮助的。没有这个选项，转储会反映一个与任何事务最终提交的序列化执行不一致的状态。

如果当gs_dump启动时，读写事务仍处于非活动状态，即便使用该选项也不会对其产生影响。如果读写事务处于活动状态，转储的开始时间可能会延迟一段不确定的时间。

- --use-set-session-authorization

  输出符合SQL标准的SET SESSION AUTHORIZATION命令而不是ALTER OWNER命令来确定对象所有权。这样令转储更加符合标准，但是如果转储文件中的对象的历史有些问题，那么可能不能正确恢复。并且，使用SET SESSION AUTHORIZATION的转储需要数据库系统管理员的权限才能转储成功，而ALTER OWNER需要的权限则低得多。但是SET SESSION AUTHORIZATION部分支持直接使用密文密码，因为使用此参数导出的脚本可能无法正常恢复，不建议使用此参数导出。

  说明

  SET SESSION AUTHORIZATION使用范围：

  - 系统管理员可以通过SET SESSION AUTHORIZATION语句切换到普通用户，无法切换到初始用户, 其他sysadmin、opradmin、monadmin、poladmin和auditadmin。
  - 其他用户无法通过SET SESSION AUTHORIZATION语句切换用户。

- --with-encryption=AES128

  指定转储数据需用AES128进行加密。

- --with-key=KEY

  AES128密钥规则如下：

  - 密钥长度为8~16个字符。
  - 至少包含大写字母（A-Z），小写字母（a-z），数字（0-9），非字母数字字符（限定为~!@#$%^&*()-_=+\|[{}];:,<.>/?）四类字符中的三类字符。

  说明

  不支持加密导出存储过程和函数。

- --with-salt=RANDVALUES

  gs_dumpall使用此参数传递随机值。

- --include-extensions

  在转储中包含扩展。

  须知

  扩展功能为内部使用功能，不建议用户使用。

- --include-depend-objs

  备份结果包含依赖于指定对象的对象信息。该参数需要同-t/--include-table-file参数关联使用才会生效。

- --exclude-self

  备份结果不包含指定对象自身的信息。该参数需要同-t/--include-table-file参数关联使用才会生效。

 版权所有 © 华为云计算技术有限公司

- --pipeline

  使用管道传输密码，禁止在终端使用。

- --dont-overwrite-file

  文本、tar、以及自定义格式情况下会重写现有文件。这对目录格式不适用。

  例如：

  设想这样一种情景，即当前目录下backup.sql已存在。如果在输入命令中输入-f backup.sql选项时，当前目录恰好也生成backup.sql，文件就会被重写。

  如果备份文件已存在，且输入--dont-overwrite-file选项，则会报告附带‘转储文件已经存在’信息的错误。

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

说明

- -s/--schema-only和-a/--data-only不能同时使用。
- -c/--clean和-a/--data-only不能同时使用。
- --inserts/--column-inserts和-o/--oids不能同时使用，因为INSERT命令不能设置OIDS。
- --role和--rolepassword必须一起使用。
- --binary-upgrade-usermap和--binary-upgrade必须一起使用。
- --include-depend-objs/--exclude-self需要同-t/--include-table-file参数关联使用才会生效。
- --exclude-self必须同--include-depend-objs一起使用。
- --with-encryption=AES128仅支持-F p/plain。
- --with-key=KEY仅支持-F p/plain。
- --with-salt=RANDVALUES由gs_dumpall调用，不需要用户手动输入。

连接参数：

- -h, --host=HOSTNAME

  指定主机名称。如果数值以斜杠开头，则被用作到Unix域套接字的路径。缺省从PGHOST环境变量中获取（如果已设置），否则，尝试一个Unix域套接字连接。

  该参数只针对数据库外，对数据库内本机IPv4用127.0.0.1，IPv6用::1。

  例如：主机名

  环境变量：PGHOST

- -p, --port=PORT

  指定主机端口。在开启线程池情况下，建议使用 pooler port，即主机端口+1。

  环境变量：PGPORT

- -U, --username=NAME

  指定所连接主机的用户名，跨节点执行不支持使用初始用户。

  环境变量：PGUSER

- -w, --no-password

  不出现输入密码提示。如果主机要求密码认证并且密码没有通过其它形式给出，则连接尝试将会失败。 该选项在批量工作和不存在用户输入密码的脚本中很有帮助。

- -W, --password=PASSWORD

  指定用户连接的密码。如果主机的认证策略是trust，则不会对系统管理员进行密码验证，即无需输入-W选项；如果不加此参数，并且不是系统管理员，则会提示交互式输入，为了系统安全，推荐使用交互式输入密码方式。

 版权所有 © 华为云计算技术有限公司

- --role=ROLENAME

  指定创建转储使用的角色名。选择该选项，会使gs_dump连接数据库后，发起一个SET ROLE角色名命令。当所授权用户（由-U指定）没有gs_dump要求的权限时，该选项会起到作用，即切换到具备相应权限的角色。某些安装操作规定不允许直接以超系统管理员身份登录，而使用该选项能够在不违反该规定的情况下完成转储。

- --rolepassword=ROLEPASSWORD

  指定角色名的密码。

说明

如果某数据库有任何本地数据要添加到template1数据库，请谨慎将gs_dump的输出恢复到一个真正的空数据库中，否则可能会因为被添加对象的定义被复制，出现错误。要创建一个无本地添加的空数据库，需从template0而非template1复制，例如：

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

tar归档形式的文件大小不得超过8GB（tar文件格式的固有限制）。tar文档整体大小和任何其他输出格式没有限制，操作系统可能对此有要求。

由gs_dump生成的转储文件不包含优化程序用来做执行计划决定的统计数据。因此，建议从某转储文件恢复之后运行ANALYZE以确保最佳效果。转储文件不包含任何ALTER DATABASE…SET命令，这些设置由gs_dumpall转储，还有数据库用户和其他完成安装设置。

示例

使用gs_dump转储数据库为SQL文本文件或其它格式的操作，如下所示。

示例中“backup/MPPDB_backup.sql”表示导出的文件，其中backup表示相对于当前目录的相对目录；“37300”表示数据库服务器端口；“testdb”表示要访问的数据库名。

说明

导出操作时，请确保该目录存在并且当前的操作系统用户对其具有读写权限。

示例1：执行gs_dump，导出testdb数据库全量信息，导出的MPPDB_backup.sql文件格式为纯文本格式。

```
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
```

使用gsql程序从纯文本导出文件中导入数据。

示例2：执行gs_dump，导出testdb数据库全量信息，导出的MPPDB_backup.tar文件格式为tar格式。

```
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.
```

 版权所有 © 华为云计算技术有限公司

```
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
```

示例3：执行gs_dump，导出testdb数据库全量信息，导出的MPPDB_backup.dmp文件格式为自定义归档格式。

```
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
```

示例4：执行gs_dump，导出testdb数据库全量信息，导出的MPPDB_backup文件格式为目录格式。

```
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
```

示例5：执行gs_dump，导出testdb数据库信息，但不导出/home/MPPDB_temp.sql中指定的表信息。导出的MPPDB_backup.sql文件格式为纯文本格式。

```
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
```

示例6：执行gs_dump，仅导出依赖于指定表testtable的视图信息。然后创建新的testtable表，再恢复依赖其上的视图。

备份仅依赖于testtable的视图

```
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
```

修改testtable名称

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

创建新的testtable表

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

还原依赖于testtable的视图

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

 版权所有 © 华为云计算技术有限公司

示例7：在多租场景下，执行gs_dump，导出名称为testpdb的PDB的全量信息，导出的backup_pdb.sql文件格式为纯文本格式。

```
gs_dump -U omm testpdb -f backup/backup_pdb.sql -p 20000 -F p
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:41:21]: The total objects number is 459.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:41:21]: [100.00%] 459 objects have been dumped.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:41:21]: dump database testpdb successfully
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:41:21]: total time: 5427  ms
```

示例8：在多租场景下，执行gs_dump，导出名称为testpdb的PDB的全量信息，导出的backup_pdb_t.tar文件格式为tar格式。

```
gs_dump -U omm testpdb -p 20000 -f backup/backup_pdb_t.tar -F t
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:02:40]: The total objects number is 459.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:02:40]: [100.00%] 459 objects have been dumped.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:02:40]: dump database testpdb successfully
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:02:40]: total time: 5506  ms
```

示例9：在多租场景下，执行gs_dump，导出名称为testpdb的PDB的全量信息，导出的backup_pdb_c文件格式为自定义归档格式。

```
gs_dump -U omm testpdb -p 20000 -f backup/backup_pdb_c -F c
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 16:57:19]: The total objects number is 459.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 16:57:19]: [100.00%] 459 objects have been dumped.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 16:57:19]: dump database testpdb successfully
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 16:57:19]: total time: 5622  ms
```

示例10：在多租场景下，执行gs_dump，导出名称为testpdb的PDB的全量信息，导出的backup_pdb_dir文件格式为目录格式。

```
gs_dump -U omm testpdb -p 20000 -f backup/backup_pdb_dir -F d
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:05:46]: The total objects number is 459.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:05:46]: [100.00%] 459 objects have been dumped.
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:05:46]: dump database testpdb successfully
gs_dump[user='omm'][localhost][port='20000'][testpdb][2024-04-26 17:05:46]: total time: 5680  ms
```

示例11：在多租场景下，执行gs_dump，使用-C, --create选项导出名称为testpdb的PDB的全量信息，gs_dump报错退出

```
gs_dump -U omm testpdb -C -p 20000 -f backup/backup_pdb_dir -F d
gs_dump unsupport the '-C, --create' option for pdb.
```

相关命令

gs_dumpall

 版权所有 © 华为云计算技术有限公司

3.4 gs_dumpall

背景信息

gs_dumpall是GaussDB用于导出所有数据库相关信息工具，它可以导出数据库的所有数据，包括默认数据库postgres的数据、自定义数据库的数据、以及所有数据库公共的全局对象。

gs_dumpall工具在进行数据导出时，其他用户可以访问数据库（读或写）。

gs_dumpall工具支持导出完整一致的数据。例如，T1时刻启动gs_dumpall导出整个数据库，那么导出数据结果将会是T1时刻该数据库的数据状态，T1时刻之后对数据库的修改不会被导出。

gs_dumpall时生成列不会被转储。

gs_dump时HTAP表创建IMCV元信息（gs_imcv系统表）不会被转储。

gs_dumpall在导出所有数据库时分为两部分：

- gs_dumpall自身对所有数据库公共的全局对象进行导出，包括有关数据库用户和组，表空间以及属性（例如，适用于数据库整体的访问权限）信息。
- gs_dumpall通过调用gs_dump来完成数据库中各数据库的SQL脚本文件导出，该脚本文件包含将数据库恢复为其保存时的状态所需要的全部SQL语句。

以上两部分导出的结果为纯文本格式的SQL脚本文件，使用**gsql**运行该脚本文件可以恢复数据库。

gs_dumpall支持SSL加密通信，使用方式同gsql方式。

使用gs_dumpall前请确保gs_dumpall版本与gs_dump版本、数据库版本保持一致，高版本gs_dump、gs_dump不保证完全兼容低版本内核数据。

注意事项

- 禁止修改导出的文件和内容，否则可能无法恢复成功。
- 为了保证数据一致性和完整性，gs_dumpall会对需要转储的表设置共享锁。如果某张表在别的事务中设置了共享锁，gs_dumpall会等待此表的锁释放后锁定此表。如果无法在指定时间内锁定某张表，转储会失败。用户可以通过指定--lock-wait-timeout选项，自定义等待锁超时时间。
- 由于gs_dumpall读取所有数据库中的表，因此必须以数据库管理员身份进行连接，才能导出完整文件。在使用gsql执行脚本文件导入时，同样需要管理员权限，以便添加用户和组，以及创建数据库。导入备份前，验证其安全性，防止管理员权限被利用。
- 使用gs_dumpall导出所有数据库对象，并希望在新的实例环境上进行导入时，需要保证导出和导入时使用用户的名称和权限相同，否则会出现名称不一致或权限不足的报错。
- gs_dumpall导出的文件，在新的实例环境上进行导入时，连接的数据库必须是默认database。否则可能出现语法不兼容的情况，导致导入出错。
- 由于M-compatibility数据库全局只能创建一个，当导出的数据库中包含M-compatibility数据库时，须保证导入的目标实例环境上不存在其他的M-compatibility模式数据库。

版权所有 © 华为云计算技术有限公司

- 对于定时任务，本工具仅支持导出在B兼容性数据库中，通过CREATE EVENT创建的定时任务或通过高级包创建的非周期性定时任务。
- gs_dumpall不支持导出自定义Tokenweight分词词典，可以根据报错WARNING: dictionary xx cannot be automatically exported, please create it manually手动创建对应分词词典。
- 当导出的数据库中包含PDB时，使用gs_dumpall不会导出模板PDB及创建的PDB。
- gs_dump时HTAP表创建IMCV元信息（gs_imcv系统表）不会被转储。
- 当指定转储编码存在转码场景时，且表中的数据存在非法编码的数据，导出会报错invalid byte sequence，建议使用gs_dump的-s参数只导出定义，并单独使用COPY打开编码容错进行数据的导出与导入。

语法

gs_dumpall [OPTION]...

参数说明

通用参数：

- -f, --filename=<FILE_NAME>

  将输出发送至指定文件。如果这里省略，则使用标准输出。
- -v, --verbose

  指定verbose模式。该选项将导致gs_dumpall向转储文件输出详细的对象注解和启动/停止次数，向标准错误流输出处理信息。
- -V, --version

  打印gs_dumpall版本，然后退出。
- --lock-wait-timeout=TIMEOUT

  请勿在转储刚开始时一直等待以获取共享表锁。如果无法在指定时间内锁定某个表，就选择失败。可以以任何符合SET statement_timeout的格式指定超时时间。
- -?, --help

  显示gs_dumpall命令行参数帮助，然后退出。

转储参数：

- -a, --data-only

  只转储数据，不转储模式（数据定义）。
- -c, --clean

  在重新创建数据库之前，执行SQL语句清理（删除）这些数据库。针对角色和表空间的转储命令已添加。
- -g, --globals-only

  只转储全局对象（角色和表空间），无数据库。
- -o, --oids

  转储每个表的对象标识符（OIDs），作为表的一部分数据。该选项用于应用以某种方式（例如：外键约束方式）参照了OID列的情况。如果不是以上这种情况，请勿使用该选项。
- -O, --no-owner

 版权所有 © 华为云计算技术有限公司

不输出设置对象的归属这样的命令，以匹配原始数据库。默认情况下，gs_dumpall会发出ALTER OWNER或SET SESSION AUTHORIZATION语句设置所创建的模式元素的所属。如果脚本正在运行，该语句不会执行成功，除非是由系统管理员触发（或是拥有脚本中所有对象的同一个用户）。通过指定-O，编写一个任何用户都能存储的脚本，且该脚本会授予该用户拥有所有对象的权限，因为没有使用ALTER OWNER或SET SESSION AUTHORIZATION语句，导入时一直使用执行用户权限，所以在导入前请确认转储文件内是否存在风险。例如，检查转储文件中是否包含提权语句，且该语句管理员是否知晓。

- -r, --roles-only

  只转储角色，不转储数据库或表空间。

- -s, --schema-only

  只转储对象定义（模式），而非数据。

- -S, --sysadmin=NAME

  该参数为扩展预留接口，不建议使用。

- -t, --tablespaces-only

  只转储表空间，不转储数据库或角色。

- -x, --no-privileges

  防止转储访问权限（授权/撤销命令）。

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

  以INSERT命令带列名（INSERT INTO表（列、…）值…）方式导出数据。这会导致恢复缓慢。但是由于该选项会针对每行生成一个独立分开的命令，所以在重新加载某行时出现的错误只会导致失败语句对应的数据丢失，而非整个表内容。

  说明

  M-Compatibility模式数据库不支持该选项，导出M-Compatibility模式数据库时会跳过该选项。

- --disable-triggers

  该参数为扩展预留接口，不建议使用。

- --inserts

  发出INSERT命令（而非COPY命令）时转储数据。这会导致恢复缓慢。注意如果重排列顺序，可能会导致恢复整个失败。--column-inserts选项更加安全，虽然可能更慢些。

  说明

  M-Compatibility模式数据库不支持该选项，导出M-Compatibility模式数据库时会跳过该选项。

- --no-security-labels

  该参数为扩展预留接口，不建议使用。

- --no-tablespaces

  请勿输出创建表空间的命令，也请勿针对对象选择表空间。使用该选项，无论默认表空间是哪一个，在恢复过程中所有对象都会被创建。

- --no-unlogged-table-data

  该参数为扩展预留接口，不建议使用。

- --include-alter-table

  导出表中已删除的列信息。

版权所有 © 华为云计算技术有限公司

- --quote-all-identifiers

  强制对所有标识符加引号。为了向后续版本迁移，且其中可能涉及引入额外关键词，在转储相应数据库时该选项会有帮助。

  说明

  M-Compatibility模式数据库不支持该选项，导出M-Compatibility模式数据库时会跳过该选项。

- --dont-overwrite-file

  不重写当前文件。

- --use-set-session-authorization

  输出符合SQL标准的SET SESSION AUTHORIZATION命令而不是ALTER OWNER命令来确定对象所有权。这样令转储更加符合标准，但是如果转储文件中的对象的历史有些问题，那么可能不能正确恢复。并且，使用SET SESSION AUTHORIZATION的转储需要数据库系统管理员的权限才能转储成功，而ALTER OWNER需要的权限则低得多。但是SET SESSION AUTHORIZATION部分支持使用密文密码随意切换用户及权限，因为使用此参数导出的脚本可能无法正常恢复，不建议使用此参数导出。

  说明

  SET SESSION AUTHORIZATION使用范围：

  - 系统管理员可以通过SET SESSION AUTHORIZATION语句切换到普通用户，无法切换到初始用户, 其他sysadmin、opradmin、monadmin、poladmin和auditadmin。
  - 其他用户无法通过SET SESSION AUTHORIZATION语句切换用户。

- --with-encryption=AES128

  指定转储数据需用AES128进行加密。

- --with-key=KEY

  AES128密钥规则如下：

  - 密钥长度为8~16个字符。
  - 至少包含大写字母（A-Z），小写字母（a-z），数字（0-9），非字母数字字符（限定为~!@#$%^&*()-_=+\|[{}];:,<.>/?）四类字符中的三类字符。

- --include-extensions

  如果include-extensions参数被设置，将备份所有的CREATE EXTENSION语句。

  须知

  扩展功能为内部使用功能，不建议用户使用。

- --include-templatedb

  转储过程中包含模板库。

- --binary-upgrade

  该参数为扩展预留接口，不建议使用。

  说明

  M-Compatibility数据库不支持该选项，导出M-Compatibility数据库时会跳过该选项。

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

版权所有 © 华为云计算技术有限公司

- 该参数为扩展预留接口，不建议使用。
- --non-lock-table

  该参数仅供软件间接口调用。
- --tablespaces-postfix

  该参数为扩展预留接口，不建议使用。
- --parallel-jobs

  指定备份进程并发数，取值范围为1~1000。
- --pipeline

  使用管道传输密码，禁止在终端使用。

说明

- -g/--globals-only和-r/--roles-only不能同时使用。
- -g/--globals-only和-t/--tablespaces-only不能同时使用。
- -r/--roles-only和-t/--tablespaces-only不能同时使用。
- -s/--schema-only和-a/--data-only不能同时使用。
- -r/--roles-only和-a/--data-only不能同时使用。
- -t/--tablespaces-only和-a/--data-only不能同时使用。
- -g/--globals-only和-a/--data-only不能同时使用。
- --tablespaces-postfix和--binary-upgrade必须一起使用。
- --binary-upgrade-usermap和--binary-upgrade必须一起使用。
- --parallel-jobs和-f/--file必须一起使用。

连接参数：

- -h, --host=HOSTNAME

  指定主机的名称。如果取值是以斜线开头，它将用作Unix域套接字的目录。默认值取自PGHOST环境变量；如果没有设置，将启动某个Unix域套接字建立连接。

  该参数只针对数据库外，对数据库内本机IPv4用127.0.0.1，IPv6用::1。

  环境变量：PGHOST
- -l, --database=DATABASENAME

  指定所连接的转储全局对象的数据库名称，并去寻找还有其他哪些数据库需要被转储。如果没有指定，会使用postgres数据库，如果postgres数据库不存在，会使用template1。

  说明

  无论指定哪一个数据库作为导出时连接的数据库，在新实例环境上使用gsql进行导入时，必须连接postgres，否则可能出现语法不兼容的情况，导致导入出错。特别地，如果错误地连接M-compatibility数据库进行导入时，由于M-compatibility数据库下CREATE DATABASE等价与CREATE SCHEMA会导致创建失败。
- -p, --port=PORT

  指定服务器所侦听的TCP端口或本地Unix域套接字后缀，以确保连接。默认值设置为PGPORT环境变量。

  在开启线程池情况下，建议使用 pooler port，即侦听端口+1。

  环境变量：PGPORT
- -U, --username=NAME

 版权所有 © 华为云计算技术有限公司

所连接的用户名，跨节点执行不支持使用初始用户。

环境变量：PGUSER

- -w, --no-password

  不出现输入密码提示。如果服务器要求密码认证并且密码没有通过其它形式给出，则连接尝试将会失败。 该选项在批量工作和不存在用户输入密码的脚本中很有帮助。

- -W, --password=PASSWORD

  指定用户连接的密码。如果主机的认证策略是trust，则不会对系统管理员进行密码验证，即无需输入-W选项；如果不加此参数，并且不是系统管理员，则会提示交互式输入，为了系统安全，推荐使用交互式输入密码方式。

- --role=ROLENAME

  指定创建转储使用的角色名。选择该选项，会使gs_dumpall连接数据库后，发起一个SET ROLE角色名命令。当所授权用户（由-U指定）没有gs_dumpall要求的权限时，该选项会起到作用，即切换到具备相应权限的角色。某些安装操作规定不允许直接以系统管理员身份登录，而使用该选项能够在不违反该规定的情况下完成转储。

- --rolepassword=ROLEPASSWORD

  指定具体角色用户的角色密码。

说明

- 由于gs_dumpall内部调用gs_dump，所以一些诊断信息请参见**gs_dump**。
- 一旦恢复，建议在每个数据库上运行ANALYZE，优化程序提供有用的统计数据。
- gs_dumpall恢复前需要所有必要的表空间目录为空；否则，对于处在非默认位置的数据库，数据库创建会失败。

示例

使用gs_dumpall一次导出数据库的所有数据库。

说明

gs_dumpall仅支持纯文本格式导出。所以只能使用gsql恢复gs_dumpall导出的转储内容。

```
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
```

其中：backup/bkp2.sql表示导出的文件，37300表示数据库服务器端口，omm为用户名。

相关命令

gs_dump

版权所有 © 华为云计算技术有限公司

3.5 gs_restore

背景信息

gs_restore是GaussDB提供的针对gs_dump导出数据的导入工具。通过此工具可将gs_dump导出生成的文件进行导入。

主要功能包含：

- 导入到数据库

  如果连接参数中指定了数据库，则数据将被导入到指定的数据库中。其中，并行导入必须指定连接的密码。导入时生成列会自动更新，并像普通列一样保存。

  如果连接参数中指定了PDB，则数据将被导入到指定的PDB中。

  说明

  - 如果导入至指定的PDB，则不支持使用-C, --create选项。
  - 如果导入指定的PDB处于close状态时，导入操作会执行失败。
  - 如果导入指定的PDB设置事务为只读事务时，导入操作会执行失败。
- 导入到归档文件

  如果参数指定"-l"，则生成归档文件，其中包含简略的数据总结。

gs_restore支持SSL加密通信，使用方式同gsql方式。

使用gs_restore前请确保gs_restore版本与gs_dump版本、数据库版本保持一致。

命令格式

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

说明

- FILE没有短选项或长选项。用来指定归档文件所处的位置。
- 作为前提条件，需输入dbname或-l选项。不允许用户同时输入dbname和-l选项。
- gs_restore默认是以追加的方式进行数据导入。为避免多次导入造成数据异常，在进行导入时，建议选择使用“-c”和“-e”参数。“-c”表示在重新创建数据库对象前，清理（删除）已存在于将要还原的数据库中的数据库对象；“-e”表示当发送SQL语句到数据库时如果出现错误请退出，默认状态下会继续，且在导入后会显示一系列错误信息。
- 在进行导入时，如果schema对象的owner拥有OPRADMIN的系统权限，那么在导入时需要使用初始用户。
- 在进行导入时，如果数据中存在非法编码的数据并且不需要进行转码，可以配置数据库兼容性参数copy_special_character_version为'no_error'将非法编码的数据进行导入，否则会报错，但不会中断导入。

参数说明

通用参数：

- -d, --dbname=NAME

  连接数据库dbname并直接导入到该数据库中。
- -f, --file=<FILE_NAME>

  指定生成归档的输出文件，使用-l时列表的输出文件。

版权所有 © 华为云计算技术有限公司

默认是标准输出。

说明

-f不能同-d一起使用。

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

  指定归档格式。由于gs_restore会自动决定格式，因此不需要指定格式。

  取值范围：

  - c/custom：该归档形式为gs_dump的自定义格式。
  - d/directory：该归档形式是一个目录归档形式。
  - t/tar：该归档形式是一个tar归档形式。

- -l, --list

  列出归档形式内容。这一操作的输出可用作-L选项的输入。注意如果像-n或-t的过滤选项与-l使用，过滤选项将会限制列举的项目（即归档形式内容）。

- -v, --verbose

  指定verbose模式。

- -V, --version

  打印gs_restore版本，然后退出。

- -?, --help

  显示gs_restore命令行参数帮助，然后退出。

导入参数：

- -a, -data-only

  只导入数据，不导入模式（数据定义）。gs_restore的导入是以追加方式进行的。

- -c, --clean

  在重新创建数据库对象前，清理（删除）已存在于将要还原的数据库中的数据库对象。如果目标数据库中没有删除操作涉及的对象，可能会输出一些提示性的错误信息。

- -C, --create

  导入数据库之前会先使用CREATE DATABASE创建数据库（指定该选项后，-d指定的数据库仅用以执行CREATE DATABASE命令，所有数据依然会导入到创建的数据库中）。

  说明

  - 在多租场景下，使用gs_restore导入数据至PDB时，不支持使用该选项。
  - 在M-Compatibility模式数据库下，不支持使用该选项。必须先在目标实例上创出M-Compatibility模式数据库，然后在源实例上导出，最后在目标实例连接新创出的M-Compatibility模式数据库导入。

- -e, --exit-on-error

  当发送SQL语句到数据库时如果出现错误，请退出。默认状态下会继续，且在导入后会显示一系列错误信息。

- -I, --index=NAME

  只导入已列举的index的定义。允许导入多个index。如果多次输入-I index导入多个index。

  例如：

 版权所有 © 华为云计算技术有限公司

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

在上面这个例子中，Index1和Index2会被导入。

- -j, --jobs=NUM

  运行gs_restore最耗时的部分（如加载数据、创建index、或创建约束）使用并发任务。该选项能大幅缩短导入时间，即将一个大型数据库导入到某一多处理器的服务器上。

  每个任务可能是一个进程或一个线程，这由操作系统决定。每个任务与服务器进行单独连接。

  该选项的最优值取决于服务器的硬件设置、客户端、以及网络。还包括这些因素，如CPU核数量、硬盘设置。建议是从增加服务器上的CPU核数量入手，更大的值（服务器上CPU核数量）在很多情况下也能导致数据文件更快地被导入。相应的，过高的值会由于超负荷反而导致性能降低。

  该选项只支持自定义归档格式。输入文件必须是常规文件（不能是像pipe的文件）。如果是通过脚本文件，而非直接连接数据库服务器，该选项可忽略。而且，多任务不能与--single-transaction选项一起使用。

  说明

  - 此参数适用于多表/多索引/多约束的情况。实际使用过程中，创建的进程数（或线程数）与表、索引、约束等的数量有关，最高并发不会超过给定的jobs数。
  - 当gs_restore导入目标数据库是PDB时，该选项的最优值取决于分配给该PDB的资源规格。

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

  只导入列举在list-file中的那些归档形式元素，导入顺序以它们在文件中的顺序为准。注意如果像-n或-t的过滤选项与-L使用，它们将会进一步限制导入的项目。

  一般情况下，list-file是通过编辑前面提到的某个-l参数的输出创建的。文件行的位置可更改或直接删除行，也可使用分号（;）在行的开始注出。见下文的举例。

- -n, --schema=NAME

  只导入已列举的模式中的对象。

  该选项可与-t选项一起用以导入某个指定的表。

  多次输入-n *schemaname*可以导入多个模式。

  例如：

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

  在上面这个例子中，sch1和sch2会被导入。

  说明

  - M-Compatibility兼容模式下，通过CREATE DATABASE带templatem创建的数据库通过-d指定db_name导入数据，而通过CREATE DATABASE db_name创建出来的database等价于Schema，只能通过-n导入数据。
  - 指定的模式名必须存在于作为输入的归档文件中，如果指定了归档文件中不存在的模式名，该模式的导入不生效。
  - 与gs_dump不同，在gs_restore中该参数为全匹配，当模式名称中包含大写字母时，不需要添加额外的引号，如：-n "Sch1"
  - M-Compatibility兼容模式下，该参数值受GUC参数lower_case_table_names影响。比如在大小写敏感模式下（lower_case_table_names=0），该参数值也需要大小写敏感，如果包含大写字母，不需要添加额外的引号。在大小写不敏感模式下（lower_case_table_names=1），该参数值需要传小写名称。

- -O, --no-owner

 版权所有 © 华为云计算技术有限公司

不输出设置对象的归属这样的命令，以匹配原始数据库。默认情况下，gs_restore会发出ALTER OWNER或SET SESSION AUTHORIZATION语句设置所创建的模式元素的所属。除非是由系统管理员（或是拥有脚本中所有对象的同一个用户）进行数据库首次连接的操作，否则语句会失败。使用-O选项，任何用户名都可用于首次连接，且该用户拥有所有已创建的对象。

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

  只导入已列举的函数。请按照函数所在转储文件中的目录，准确拼写函数名称和参数。

  当-P单独使用时，表示导入文件中所有'function-name(args)'函数；当-P同-n一起使用时，表示导入指定模式下的'function-name(args)'函数；多次输入-P，而仅指定一次-n，表示所有导入的函数默认都是位于-n模式下的。

  可以多次输入-n schema-name -P 'function-name(args)'同时导入多个指定模式下的函数。

  例如：

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

  在上面这个例子中，test1模式下的函数Func1(i integer)和test2模式下的函数Func2(j integer)会被一起导入。

- -s, --schema-only

  只导入模式（数据定义），不导入数据（表内容）。当前的序列值也不会导入。

- -S, --sysadmin=NAME

  该参数为扩展预留接口，不建议使用。

- -t, --table=NAME

  只导入已列举的表定义、数据或定义和数据。该选项与-n选项同时使用时，用来指定某个模式下的表对象。-n参数不输入时，默认为PUBLIC模式。多次输入-n <schemaname> -t <tablename>可以导入指定模式下的多个表。

  例如：

  导入PUBLIC模式下的table1

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

  导入test1模式下的test1和test2模式下test2

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

  导入PUBLIC模式下的table1和test1 模式下test1

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

 版权所有 © 华为云计算技术有限公司

须知

- -t不支持schema_name.table_name的输入格式，指定此格式不会报错，但不会生效。
- 当-t已指定时，gs_restore不会导入已选表所附着的任何其他数据库对象。因此，无法保证某个指定表的转储结果能够自行成功地导入到一个空数据库中。
- -t tablename不会导入表上的触发器信息。
- 与gs_dump不同，在gs_restore中该参数为全匹配，当表名称中包含大写字母时，不需要添加额外的引号，如：-t "Tbl1"
- M-Compatibility兼容模式下，该参数值受GUC参数lower_case_table_names影响。比如在大小写敏感模式下（lower_case_table_names=0），该参数值也需要大小写敏感，如果包含大写字母，不需要添加额外的引号。在大小写不敏感模式下（lower_case_table_names=1），该参数值需要传小写名称。

- -T, --trigger=NAME

  该参数为扩展预留接口。

- -x, --no-privileges/--no-acl

  防止导入访问权限（grant/revoke命令）。

- -1, --single-transaction

  执行导入作为一个单独事务（即把命令包围在BEGIN/COMMIT中）。

  该选项确保要么所有命令成功完成，要么没有改变应用。该选项意为--exit-on-error。

- --disable-triggers

  该参数为扩展预留接口，不建议使用。

- --no-data-for-failed-tables

  默认状态下，即使创建表的命令失败（如表已经存在），表数据仍会被导入。使用该选项，像这种表的数据会被跳过。如果目标数据库已包含想要的表内容，这种行为会有帮助。

  该选项只有在直接导入到某数据库中时有效，不针对生成SQL脚本文件输出。

- --no-security-labels

  该参数为扩展预留接口，不建议使用。

- --no-tablespaces

  不输出选择表空间的命令。使用该选项，无论默认表空间是哪一个，在导入过程中所有对象都会被创建。

- --section=SECTION

  导入已列举的区段（如pre-data、data、或post-data）。

- --use-set-session-authorization

  该选项用来进行文本格式的备份。

  输出SET SESSION AUTHORIZATION命令，而非ALTER OWNER命令，用以决定对象归属。该选项使转储更加兼容标准，但通过参考转储中对象的记录，导入过程可能会有问题。使用SET SESSION AUTHORIZATION的转储要求必须是系统管理员，同时在导入前还需参考"SET SESSION AUTHORIZATION"，手工对导出文件的密码进行修改验证，只有这样才能进行正确的导入操作，相比之下，ALTER OWNER对权限要求较低。

版权所有 © 华为云计算技术有限公司

- --pipeline

  使用管道传输密码，禁止在终端使用。

须知

- 如果安装过程中有任何本地数据要添加到template1数据库，请谨慎将gs_restore的输出载入到一个真正的空数据库中；否则可能会因为被添加对象的定义被复制，而出现错误。要创建一个无本地添加的空数据库，需从template0而非template1复制，例如：

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

- gs_restore不能选择性地导入大对象；例如只能导入那些指定表的对象。如果某个归档形式包含大对象，那所有大对象都会被导入。如果此归档对象通过-L、-t或其他选项被排除，那么所有大对象一个都不会被导入。

说明

1.-d/--dbname 和 -f/--file 不能同时使用。

2. -s/--schema-only 和 -a/--data-only不能同时使用。

3. -c/--clean 和 -a/--data-only不能同时使用。

4. 使用--single-transaction时，-j/--jobs必须为单任务。

5. --role 和 --rolepassword必须一起使用。

连接参数：

- -h, --host=HOSTNAME

  指定的主机名称。如果取值是以斜线开头，他将用作Unix域套接字的目录。默认值取自PGHOST环境变量；如果没有设置，将启动某个Unix域套接字建立连接。

  该参数只针对数据库外，对数据库内本机IPv4用127.0.0.1，IPv6用::1。

  环境变量：PGHOST

- -p, --port=PORT

  指定服务器所侦听的TCP端口或本地Unix域套接字后缀，以确保连接。默认值设置为PGPORT环境变量。

  在开启线程池情况下，建议使用 pooler port，即侦听端口+1。

  环境变量：PGPORT

- -U, --username=NAME

  所连接的用户名，跨节点执行不支持使用初始用户。

  环境变量：PGUSER

- -w, --no-password

  不出现输入密码提示。如果服务器要求密码认证并且密码没有通过其它形式给出，则连接尝试将会失败。 该选项在批量工作和不存在用户输入密码的脚本中很有帮助。

- -W, --password=PASSWORD

  指定用户连接的密码。如果主机的认证策略是trust，则不会对系统管理员进行密码验证，即无需输入-W参数；如果不加此参数，并且不是系统管理员，则会提示交互式输入，为了系统安全，推荐使用交互式输入密码方式。

 版权所有 © 华为云计算技术有限公司

- --role=ROLENAME

  指定导入操作使用的角色名。选择该参数，会使gs_restore连接数据库后，发起一个SET ROLE角色名命令。当所授权用户（由-U指定）没有gs_restore要求的权限时，该参数会起到作用，即切换到具备相应权限的角色。某些安装操作规定不允许直接以初始用户身份登录，而使用该参数能够在不违反该规定的情况下完成导入。

- --rolepassword=ROLEPASSWORD

  指定具体角色用户的角色密码。

示例

特例：执行gsql程序，使用如下选项导入由gs_dump/gs_dumpall生成导出文件夹（纯文本格式）的MPPDB_backup.sql文件到testdb数据库。

```
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
```

示例中“-f”后的是导出的文件，“8000”表示数据库服务器端口；“testdb”表示要访问的数据库名。

gs_restore用来导入由gs_dump生成的导出文件。

示例1：执行gs_restore，将导出的MPPDB_backup.dmp文件（自定义归档格式）导入到testdb数据库。

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

示例2：执行gs_restore，将导出的MPPDB_backup.tar文件（tar格式）导入到testdb数据库。

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

示例3：执行gs_restore，将导出的MPPDB_backup文件（目录格式）导入到testdb数据库。

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

示例4：执行gs_restore，使用自定义归档格式的MPPDB_backup.dmp文件来进行如下导入操作。 导入PUBLIC模式下所有对象的定义和数据。在导入时会先删除已经存在的对象，如果原对象存在跨模式的依赖则需手工强制干预。

 版权所有 © 华为云计算技术有限公司

```
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;
```

手工删除依赖，导入完成后再重新创建。

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

示例5：执行gs_restore，使用自定义归档格式的MPPDB_backup.dmp文件来进行如下导入操作。只导入PUBLIC模式下表table1的定义。

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

示例6：执行gs_restore，使用自定义归档格式的MPPDB_backup.dmp文件来进行如下导入操作。只导入PUBLIC模式下表table1的数据。

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

示例7：在多租场景下，执行gs_restore，导入指定PDB_backup.dmp文件（自定义归档格式）到名称为testpdb的PDB。

```
gs_restore backup/PDB_backup.dmp -p 20000 -d testpdb
restore operation successful
total time: 371  ms
```

示例8：在多租场景下，执行gs_restore，导入指定PDB_backup.tar文件（tar归档格式）到名称为testpdb的PDB。

```
gs_restore backup/PDB_backup.tar -p 20000 -d testpdb
restore operation successful
total time: 367  ms
```

示例9：在多租场景下，执行gs_restore，导入指定PDB_backup目录文件（目录归档格式）到名称为testpdb的PDB。

```
gs_restore backup/PDB_backup -p 20000 -d testpdb
restore operation successful
total time: 370  ms
```

示例10：在多租场景下，执行gs_retsore，使用-C, --create选项来进行导入操作，gs_restore报错退出。

```
gs_restore backup/PDB_backup -C -p 20000 -d testpdb
gs_restore unsupport the '-C, --create' option for pdb.
```

相关命令

gs_dump，**gs_dumpall**

 版权所有 © 华为云计算技术有限公司

4 服务端工具

在使用GaussDB过程中，经常需要对数据库进行安装、升级，以及对数据库进行健康管理。为了简单、方便的维护数据库，GaussDB提供了一系列的数据库管理工具。

说明

沙箱部署场景，所有工具需要进入到沙箱内部才能正确执行。
进入沙箱命令参考：chroot /var/chroot

4.1 GaussRoach.py

4.1.1 Roach 简介

GaussRoach.py工具是GaussDB提供的用于备份和恢复的实用工具。可对整个数据库中的数据、WAL归档日志和运行日志进行备份。

GaussRoach.py工具是一款数据库高可用性以及容灾恢复策略的备份管理工具。使用该工具可以备份恢复数据库；不仅可以备份到物理磁盘，也可以备份到OBS、NAS和REMOTE。数据库级备份包含数据库静态配置文件（cluster_static_config）、数据库动态配置文件（cluster_dynamic_config）、数据节点DN（Datanode）及其备实例。

4.1.2 命令行参考

4.1.2.1 注意事项

以下是使用Roach命令行界面的注意事项：

- 除了章节**版本命令**和**帮助命令**的版本和帮助命令外，gs_roach命令的名称区分大小写，但命令的取值不区分大小写。
- 不支持的参数和命令行参数都将被忽略。

  在如下示例中，由于数据库备份命令不支持，**--agent-port**会被忽略，但命令会成功执行。

  ```
  python3 GaussRoach.py -t backup --metadata-destination /home/userB/meta --media-destination /home/userB/backup --media-type DISK --master-port 2000 --agent-port 1234
  ```

- 如果选项的缩写形式唯一，或者与定义的选项完全匹配，长选项名可在Linux平台上进行缩写。这不适用于章节**版本命令**和**帮助命令**的版本命令和帮助命令。

版权所有 © 华为云计算技术有限公司

在如下示例中，--log-file存在歧义，因为它对应两个选项--log-filesize和--log-filecount。因此，工具显示错误消息。

```
python3 GaussRoach.py -t backup --media-destination /home/userB/backup --media-type DISK --master-port 2000 --log-file 5
./gs_roach: option '--log-file' is ambiguous
ERROR: Try " --help" for more information.
```

在如下示例中，--log-files是--log-filesize的唯一缩写。因此，该命令可成功执行。

```
python3 GaussRoach.py -t backup --media-destination /home/userB/backup --media-type DISK --master-port 2000 --log-files 5
```

- Metadata destination为必选参数。该参数仅在Metadata destination路径设为环境变量时可跳过。
- 如果media type参数取值为REMOTE时，需要设置以下三个参数**--remote-ip**、**--remote-port**和**--remote-config**。media-destination需要在值里面包含数据库实例标识clusterid。

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/gauss/clusterid/xxx --media-type REMOTE --metadata-destination /home/userA/metadata  --remote-ip x.x.x.x --remote-port "8009" --remote-config "null"
```

说明

- 如果备份与元数据目标选项一起运行，那么恢复和删除操作都应指定元数据目标参数。否则，操作将失败，并且Roach会显示错误信息。
- 删除操作中，如果元数据目标路径和介质目标路径相同，备份将与元数据文件一同删除。
- 删除操作中，如果元数据目标路径和介质目标路径不同，元数据目标文件夹不会删除但介质目标路径下的文件夹会删除。
- OBS/NAS/REMOTE备份操作中，如果元数据目标路径中包含其它非元数据的文件时，也会被打包上传至OBS/NAS/REMOTE中，为了保证备份集的大小，建议将元数据目标路径配置成专用的路径。
- 恢复和启动操作时必须是管理员用户，否则会因为权限问题启动失败。
- 在恢复期间会恢复libgcc_s.so、 libstdc++.so动态库，在此时调用对这两个动态库有依赖的二进制时可能会失败。
- 数据库内核为Paxos时进行Roach本地盘备份，需要传入配置静态参数-X config_name.xml，当前版本暂不支持此备份方式。
- 不支持对M-Compatibility数据库的表进行细粒度表级备份恢复，建议对M-Compatibility数据库的表进行实例级备份恢复。

4.1.2.2 备份命令

功能

backup命令用于备份数据库或表数据，其中数据库备份支持备份到OBS、DISK、NAS、REMOTE第三方介质，物理表级备份支持备份到OBS、NAS和REMOTE。在多租实例中，物理PDB级备份支持备份到OBS、NAS。

命令格式

```
python3 GaussRoach.py
-t backup
--master-port <master-port>
--media-type <media-type>
--media-destination <media-destination-path>
--metadata-destination <metadata-path>
[--block-dw-flush-timeout] <timeout>
[--bucket-name <bucketname>]
[--buffer-block-size <block-size>]
```

 版权所有 © 华为云计算技术有限公司

```
[--buffer-size <buffer-size>]
[--cbm-recycle-level <level>]
[--check-balance-retry-times]
[--cluster-unique-id <value>]
[--compression-level <compression-level>]
[--compression-type <type>]
[--conf-location <path-name>]
[--cpu-relinquish-size <cpu-relinquish-size-value>]
[--cpu-relinquish-time <cpu-relinquish-time-value>]
[--enable-logging <true/false>]
[--failure-retry-count <count>]
[--filesplit-size <file-split-size>]
[--log-filecount <log-file-count>]
[--log-filesize <log-file-size>]
[--logging]
[--logging-level <logging-level>]
[--logging-path <logging-path>]
[--max-backup-io-speed <speed>]
[--max-memory-usage <maximum-number-of-bytes>]
[--metadata-file-wait-time <wait-time>]
[--master-ip <x.x.x.x/x:x:x:x:x:x:x:x>]
[--obs-server-ip <address>]
[--parallel-process <process-count>]
[--password <pass-word>]
[--pre-disk-space]
[--prefetch-block <block num>]
[--prior-backup-key <prior-full-backup-key>]
[--prior-cluster-unique-id <prior-cluster-unique-id>]
[--reader-thread-count <number-of-threads>]
[--reader-thread-file-count <number-of-thread-files>]
[--reader-thread-file-size <filesize>]
[--resource-retry-count <count>]
[--resume-backup]
[--retry-wait-time <time>]
[--single]
[--snapshot-action] <action>
[--split-local-disk-storage]
[--standby-mode]
[--username <user-name>]
[-W <password>]
[--verbose]
[--disable-ssl]
[--remote-ip]
[--remote-port]
[--remote-config]
[--gbr-table-list]
[--pdb-list <pdb-list>]
[--prior-dn-list <dn-list>]
[--node-list <node-list>]
```

 版权所有 © 华为云计算技术有限公司

参数说明

表 4-1 备份命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。指定该参数为backup，表示调用备份功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t backup
--master-port	不涉及	整数	Roach主代理进程的执行端口。	1024-65535	不涉及	--master-port *6812*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-type	不涉及	字符串	备份所需的介质类型。 ● DISK（磁盘） ● OBS ● NAS ● REMOTE	● DISK（磁盘） ● OBS ● NAS ● REMOTE	不涉及	● 对于DISK：--media-type *Disk* ● 对于OBS：--media-type OBS ● 对于NAS：--media-type NAS

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-destination	不涉及	字符串	指定介质的目的备份路径。 ● DISK（磁盘） ● REMOTE：样例策略 ● OBS：不生效 ● NAS：挂载的NAS共享盘路径 **说明** 使用备份数据库到NAS时，确保数据库实例上所有节点的指定路径挂载的是同一个NAS共享盘	不涉及	不涉及	● 对于DISK：--media-destination */home/cam/backup* ● 对于NAS：--media-destination */home/cam/backup*
--metadata-destination	不涉及	字符串	元数据文件位置。	不涉及	不涉及	--metadata-destination */home/username*
--block-dw-flush-timeout	不涉及	整数	快照备份期间阻塞双写页面刷盘时长。 **说明** 当前版本该参数不生效。	0-3600（秒）	30	--block-dw-flush-timeout 60

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--bucket-name	不涉及	字符串	指定OBS桶名称。	不涉及	不涉及	--bucket-name *rdsbucket.resource.user4444*
--buffer-block-size	不涉及	整数	单个缓冲块大小，需要是8192的整数倍。	524288–268435456（字节）	非OBS，非REMOTE：2097152（字节） OBS：67108864（字节） REMOTE：16777216（字节）	--buffer-block-size 524288
--buffer-size	不涉及	整数	缓存大小。	256–16384（MB）	非OBS：256（MB） OBS：2048（MB）	--buffer-size *10000*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--cbm-recycle-level	不涉及	整数	cbm文件回收的等级。	0~2 0：代表不删除cbm文件。 1：代表只在全量备份的时候删除全备之前的cbm文件。 2：代表增备时候删除当前增备之前的cbm文件。	0	--cbm-recycle-level 2
--check-balance-retry-times	不涉及	整数	断点续备或者双数据库备份时检查数据库为balance状态的次数。	整数	4	--check-balance-retry-times 4
--cluster-unique-id	不涉及	字符串	指定数据库备份时的数据库ID。	不涉及	不涉及	--cluster-unique-id *dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02*
--compression-level	不涉及	整数	压缩级别。 0代表快速或无压缩。 9代表慢速或最大压缩。 **说明** 值越小，压缩越快。 值越大，压缩越好。	0 - 9	1	--compression-level *5*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--compression-type	不涉及	整数	压缩类型： ● 0：不压缩 ● 1：zlib ● 2：lz4	0 - 2	2	--compression-type 2
--conf-location	-r	字符串	gs_roach.conf文件的位置。	不涉及	不涉及	--conf-location /home/backup/roach
--cpu-relinquish-size	不涉及	整数	数据用量达到该参数指定的值后，Roach必须切换至低CPU使用率一段时间，该时间段由--cpu-relinquish-time指定，单位为秒。	1–10000（GB）	不涉及	--cpu-relinquish-size *10 GB*
--cpu-relinquish-time	不涉及	整数	CPU的释放时间。此期间Roach必须切换为低CPU使用率。 **说明** ● 将CPU释放时间设置为60秒或60秒以上以减小CPU使用率。 ● 如果该参数未设置值，执行中会忽略该参数。	0–3600（秒）	不涉及	--cpu-relinquish-time *1000*
--enable-logging	不涉及	字符串	是否开启记录到日志文件功能。 建议使用此参数而不是--logging。	● True ● False	False	--enable-logging *True*
--failure-retry-count	不涉及	整数	配置Roach工具失败后最大重试次数。如果设为n，Roach会进行n-1次重试。	0–256	32	--failure-retry-count *3*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--filesplit-size	不涉及	整数	被拆分的备份文件的大小。 0：文件不会被拆分。 1024：被拆分文件最大为1024GB。 **说明** 用户必须指定被拆分文件的大小（文件大小应是4的倍数）。	0–1024（GB）	4（GB）	--filesplit-size 40
--log-filecount	不涉及	整数	日志文件最大个数。	5–1024	10	--log-filecount *5*
--log-filesize	不涉及	整数	最大日志文件的大小。	5–20（MB）	5（MB）	--log-filesize *5*
--logging	不涉及	布尔型	是否开启日志文件记录功能。 如果命令中指定了该参数，则参数值为true。否则值为false。 此参数将来会退化，建议使用--enable-logging	不涉及	False	--logging

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--logging-level	不涉及	字符串	日志级别。 • FATAL（致命）：导致系统停止工作的、无法恢复的故障。该级别为最严重级别。 • ERROR（错误）：重大错误。 • WARNING（警告）：异常情况。系统在该情况下可能会继续处理任务。 • INFO（提示）：日志记录过程中的提示性日志。 • DEBUG（调试）：用于调试的详细信息。 • DEBUG2（调试2）：最详细的日志信息，通常会过滤不显示。该级别严重程度最轻。	• FATAL • ERROR • WARNING • INFO • DEBUG • DEBUG2	INFO	--logging-level *info*
--logging-path	不涉及	字符串	日志存放路径。	不涉及	*$GAUSSLOG/roach/controller*	--logging-path *$GAUSSLOG/roach/controller*
--max-backup-io-speed	不涉及	整数	备份时，对存储介质的写入速度进行限制 **说明** 开启限速时，可以限制的最大写入速度为2GB/s	0-2147483647 (byte/s)	0（不进行限速）	--max-backup-io-speed 1073741824

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--max-memory-usage	不涉及	无符号整数	Roach可用的内存最大字节数。 **说明** Roach默认可以使用系统提供的最大内存。	2–256（单位为GB）	不涉及	--max-memory-usage 2
--metadata-file-wait-time	不涉及	整数	读取元数据文件的等待时长。	60–3600（秒）	60（秒）	--metadata-file-wait-time 100
--master-ip	不涉及	IP地址	输入当前节点所在数据库的IP地址，即当前登录节点在数据库配置文件中指定的IP地址。	• x.x.x.x • x:x:x:x:x:x:x:x	不涉及	--master-ip x.x.x.x
--obs-server-ip	不涉及	IP地址	指定OBS的服务器地址。	• x.x.x.x • x:x:x:x:x:x:x:x	不涉及	--obs-server-ip x.x.x.x
--parallel-process	不涉及	整数	Roach可以使用的子进程个数。这个参数只在混合部署下生效，如果一个节点上只有一个DN副本，则该参数会被内部调整成1。	1–32	1	--parallel-process *10*
--pre-disk-space	不涉及	布尔型	检查是否有足够的磁盘空间完成即将进行的操作。 如果命令中指定了该参数，则参数值为true。否则值为false。	不涉及	False	--pre-disk-space
--prefetch-block	不涉及	整数	增量备份时读取磁盘上表文件增量修改页面的预取页面个数。	1-8192	不涉及	--prefetch-block 64

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--prior-backup-key	不涉及	字符串	表示与前备份使用的**backup key**，增量备份基于该备份执行。	不涉及	不涉及	--prior-backup-key *20170117_041504*
--prior-cluster-unique-id	不涉及	字符串	节点替换后做增量备份时，表示前一次备份使用的cluster-unique-id。该参数只有在备份到OBS介质时才会设置。	不涉及	不涉及	--prior-cluster-unique-id DB
--prior-dn-list	不涉及	整数	快照备份时指定上一次备份的DN实例ID，本次快照备份会优先选择指定的DN进行快照备份操作。	不涉及	不涉及	--prior-dn-list 6002
--reader-thread-count	不涉及	整数	指定要处理的阅读器线程数。	0 - 256	不涉及	--reader-thread-count *8*
--reader-thread-file-count	不涉及	整数	指定每个阅读器线程的缓冲区文件数量。	0 - 256	不涉及	--reader-thread-file-count *2*
--reader-thread-file-size	不涉及	整数	指定阅读器线程的每个缓冲区文件的缓冲区文件大小（以字节为单位）。	1 - 65535（字节）	不涉及	--reader-thread-file-size 16384

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--resource-retry-count	不涉及	整数	Roach工具在资源分配失败后的重试次数。 如果设为n，Roach会进行n-1次重试。 **说明** 该参数可用于在高斯工具中系统资源分配失败时进行的任何操作。系统资源包括堆内存、线程、数据库API。对其他失败操作的重试由--failure-retry-count参数控制。	0–256	0	--resource-retry-count 3
--resume-backup	不涉及	布尔型	该参数用于指定本次全量备份是否为断点续做。 当数据较多时，全量备份失败如果不想重头再做，可以在全量备份开始时加上此参数，可在断掉的位置继续做全量备份。 **说明** 此参数只用于全量备份，备份介质适用于DISK、OBS和NAS。	不涉及	False	--resume-backup
--retry-wait-time	不涉及	整数	配置Roach工具在失败后重试时需要等待的时间。	1–3600（秒）	10（秒）	--retry-wait-time *1200*
--single	不涉及	布尔型	指定为单数据库模式。	不涉及	False	-single
--snapshot-action	不涉及	字符串	快照备份操作。	backup_prepare|backup_xlog|flash_backup_prepare|flash_backup_xlog	不涉及	--snapshot-action backup_prepare

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--split-local-disk-storage	不涉及	布尔型	备份的不同的DN数据分开存储。	不涉及	False	--split-local-disk-storage
--standby-mode	不涉及	布尔型	是否打开备机备份。	不涉及	False	--standby-mode
--username	-U	字符串	数据库用户名。 **说明** 只支持数据库实例初始化用户、系统管理员用户、Operatoradmin权限用户且operation_mode开启。	不涉及	不涉及	--username *cluster_user*
-W	不涉及	字符串	数据库用户的密码。出于安全考虑，不建议使用该接口传递密码。建议使用echo和pipe方式来传递密码。	不涉及	不涉及	-W foo_pswd
--verbose	不涉及	布尔型	是否启用verbose选项来显示详细信息。 如果命令中指定了该参数，则参数值为true。否则值为false。	不涉及	False	--verbose
-X	-X	字符串	备份的数据库的XML路径。 **说明** 数据库内核为Paxos时且备份到本地盘时需要指定此参数。	绝对路径	不涉及	-X /home/*.xml
--disable-ssl	不涉及	布尔型	gs_roach组件的SSL认证默认开启，若想关闭SSL认证，则在命令中指定该参数。	不涉及	False	--disable-ssl

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--remote-ip	不涉及	字符串	指定远端第三方存储介质所在服务器IP。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-ip *x.x.x.x*
--remote-port	不涉及	整数	指定远端第三方存储介质所在服务器侦听端口。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-port “8009”
--remote-config	不涉及	字符串	第三方厂商配置参数内容。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-config “null”

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例

--gbr-table-list	不涉及	字符串	物理备份的表信息对应的json文件路径。json文件示例：
			{
			"TableList": [{
			"DBName": "db1",
			"SchemaName": "schema1",
			"TableName": "t1"
			},{
			"DBName": "db2",
			"SchemaName": "schema2",
			"TableName": "t2"
			}]}
			说明
			该版本不支持对M-compatibility的表进行细粒度备份。
			db_name、schema_name、table_name分别是备份恢复的表的库名、模式名、表名。
				不涉及	不涉及	--gbr-table-list /data/backup_table.json

--pdb-list	不涉及	字符串	指定当前进行备份的PDB列表。 **说明** 该版本不支持对M-compatibility的PDB进行细粒度备份。	仅可包含大小写字母、数字及下划线字符，最多包含128个PDB。	不涉及	--pdb-list "mypdb1,mypdb2"
--node-list	不涉及	字符串	指定闪回快照备份需要备份xLog的DN节点列表。	仅包含DN实例ID。	不涉及	--node-list "6001,6002"

使用指南

- 对于数据库实例级备份，必须提供介质类型、目标介质和主代理端口，否则Roach工具会报错。
- 对于物理表级备份，必须提供介质类型、目标介质、主代理端口和表信息对应json文件路径，否则Roach工具会报错。

版权所有 © 华为云计算技术有限公司

- 对于物理PDB级备份，必须提供介质类型、目标介质、主代理端口和PDB列表信息，否则Roach工具会报错。
- 数据库实例级备份前，请执行如下命令检查数据库运行状态，cluster_state为Normal时表示数据库正常运行，可以备份数据库实例，使用NAS或OBS或REMOTE介质时，处于Degraded时也可以备份数据。

```
gs_om -t status

-----------------------------------------------------------------------
cluster_state : Normal
redistributing : No
balanced : Yes
-----------------------------------------------------------------------
```

命令示例

- 备份数据库到OBS

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/userA/media --media-type OBS --metadata-destination /home/userA/metadata --obs-server-ip x.x.x.x --bucket-name rdsbucket.resource.user4444 --cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
```

说明

- 备份到OBS，支持数据库实例级备份和表级备份。
- 执行备份到OBS前，请确定已申请OBS账户，并配置账户权限。

- 备份数据库到REMOTE

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/gauss/clusterid/xxx --media-type REMOTE --metadata-destination /home/userA/metadata --remote-ip x.x.x.x --remote-port "8009" --remote-config "null"
```

说明

- 执行备份数据库实例到REMOTE前，请确定REMOTE第三方环境中已开启备份程序backUpAgent。
- 增量备份时间间隔应小于参数max_cbm_retention_time参数所设定时间。

- 备份数据库到DISK

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/userA/media --media-type DISK --metadata-destination /home/userA/metadata
```

- 备份数据库到NAS

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/userA/media --media-type NAS --metadata-destination /home/userA/metadata --cluster-unique-id gauss_backup
```

- 增量备份数据库

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/userA/media --media-type Disk --metadata-destination /home/userA/metadata --prior-backup-key 20170313_131629
```

- 增量备份数据库到REMOTE

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/gauss/clusterid/xxx --media-type REMOTE --metadata-destination /home/userA/metadata --remote-ip x.x.x.x --remote-port "8009" --remote-config "null" --prior-backup-key 20210119_161749
```

- 实例级备份表到REMOTE

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/gauss/clusterid/xxx --media-type REMOTE --metadata-destination /home/userA/metadata --remote-ip x.x.x.x --remote-port "8009" --remote-config "null" --gbr-table-list /xxx/table.json
```

- 物理备份PDB到NAS

```
python3 GaussRoach.py -t backup --master-port 6000 --media-destination /home/userA/media --media-type NAS --metadata-destination /home/userA/metadata --cluster-unique-id gauss_backup --pdb-list "mypdb1,mypdb2"
```


版权所有 © 华为云计算技术有限公司

- 物理备份PDB到OBS
  python3 GaussRoach.py **-t backup** **--master-port** 6000 **--media-destination** /home/userA/media **--media-type** OBS **--metadata-destination** /home/omm/metadata **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 **--pdb-list** "mypdb1,mypdb2"
- 物理备份表到OBS
  python3 GaussRoach.py **-t backup** **--master-port** 6000 **--media-destination** /home/userA/media **--media-type** OBS **--metadata-destination** /home/omm/metadata **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 **--gbr-table-list** /data/table.json
- 物理备份表到NAS
  python3 GaussRoach.py **-t backup** **--master-port** 6000 **--media-destination** /home/userA/media **--media-type** NAS **--metadata-destination** /home/omm/metadata **--cluster-unique-id** gauss_backup **--gbr-table-list** /data/table.json

相关命令

恢复命令

4.1.2.3 停止备份命令

功能

stop命令用于停止运行中的备份操作。

命令格式

```
python3 $GPHOME/script/GaussRoach.py
-t stop
[-F]
[--logging-path <logging-path>]
[--stop-object <object-type>]
[--stop-timeout <timeout-value>]
[--pause-resume-backup]
[--logging-path <filepath>]
[--metadata-destination <path-name>]
```


版权所有 © 华为云计算技术有限公司

参数说明

表 4-2 停止备份命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach界面支持多种功能。 指定该参数为stop表示调用停止功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t stop

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-F	-	字符串	可选参数。 为了兼容老版本功能保留该参数，当前版本携带或不携带该参数均会强制停止备份操作。	不涉及	不涉及	stop -F
--logging-path	不涉及	字符串	日志路径。	不涉及	*$GAUSSLOG/roach/controller*	--logging-path *$GAUSSLOG/roach/controller*
--stop-object	不涉及	字符串	停止操作类型。	backup|restore|cluster	backup	--stop-object backup
--stop-timeout	不涉及	字符串	停止操作的超时时间。超过该时间段后会退出停止操作。	30-57600（单位：秒）	600	--stop-timeout *1500*
--pause-resume-backup	不涉及	布尔型	暂停断点续做。	不涉及	False	--pause-resume-backup
--metadata-destination	不涉及	字符串	备份时的metadata的地址。	不涉及	不涉及	--metadata-destination /data/metadata/

使用指南

- stop命令用于停止当前备份操作。该命令停止所有进程并清理所有备份文件。
- stop命令可从数据库中的任意节点上执行。
- 如果执行该命令时未进行备份操作，则会提示没有正在执行的备份命令，同时也会尝试清理数据库中可能存在的备份残留状态。
- Roach也兼容使用**python3 GaussRoach.py –t stop –F**命令停止备份，有-F和无-F参数的执行结果相同。
- 如果有一个全量备份和一个增量备份同时执行，那么stop操作会一起停止这两个备份任务。

版权所有 © 华为云计算技术有限公司

命令示例

停止备份。

```
python3 $GPHOME/script/GaussRoach.py -t stop
```

系统回显

```
sparrow@BLR1000014692:~/db/bin/script> python3 GaussRoach.py -t stop
Stopping backup.
Stopping gs_roach process
Successfully stopped gs_roach process
Stopping the backup process in database
Successfully stopped the backup process in database
Waiting python backup process, stop timeout is 600 seconds.
Successfully stopped backup.
Roach operation stop completed.
```

相关命令

备份命令

4.1.2.4 备份删除命令

功能

delete命令用于删除DISK、OBS或NAS上的数据库级备份，不支持删除REMOTE上的备份。当指定了级联选项时，且此时最新一次备份集依赖的backup key链中包含了指定的备份集时，可以级联删除指定备份集以及所有直接或间接依赖它的增备备份集。

命令格式

```
python3 GaussRoach.py
-t delete
--backup-key <backup-key>
--master-port <master-port>
--media-type <media-type>
--media-destination <media-destination-path>
--metadata-destination <metadata-path>
--bucket-name <bucket-name>
--cluster-unique-id <cluster-unique-id>
--obs-server-ip <obs-server-ip>
[--cascade]
[--disable-ssl]
```


版权所有 © 华为云计算技术有限公司

参数说明

表 4-3 备份删除命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。删除期间指定该参数为delete，表示调用删除功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t delete
--backup-key	-c	字符串	从备份集中恢复使用的backup key。	不涉及	不涉及	--backup-key *20150315-16103*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--master-port	不涉及	整数	Roach主代理进程的执行端口。	1024–65535	不涉及	--master-port *6812*
--media-type	不涉及	字符串	介质类型。 ● DISK（磁盘） ● OBS ● NAS	● DISK ● OBS ● NAS	不涉及	--media-type DISK
--media-destination	不涉及	字符串	所选介质的目的备份路径。	不涉及	不涉及	--media-destination */home/cam/backup*
--metadata-destination	不涉及	字符串	指定元数据文件位置	不涉及	不涉及	*--metadata-destination/home/username*
--bucket-name	不涉及	字符串	指定OBS桶名称。	不涉及	不涉及	--bucket-name rdsbucket.resource.user4444
--cluster-unique-id	不涉及	字符串	指定数据库备份时的数据库ID。	不涉及	不涉及	--cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
--obs-server-ip	不涉及	IP地址	指定OBS的服务器地址。	x.x.x.x x:x:x:x:x:x:x:x	不涉及	--obs-server-ip x.x.x.x

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--cascade	不涉及	不涉及	指定备份密钥的所有相关备份也将被删除。 如果指定了参数，则该值为true，否则为false。	不涉及	不涉及	--cascade
--disable-ssl	不涉及	布尔型	gs_roach组件的SSL认证默认开启，若想关闭SSL认证，则在命令中指定该参数。	不涉及	False	--disable-ssl

使用指南

- 要删除备份，必须提供介质类型，目标介质、主代理端口和backup key。如果用户没有设置这些参数，Roach工具会显示错误信息。
- 如果全量备份含有相关增量备份，则只能使用级联选项将其删除。

命令示例

- 从DISK删除备份文件。
  ```
  python3 GaussRoach.py -t delete --media-type Disk --media-destination $HOME/back1 --master-port 6000  --metadata-destination $HOME/back2 --backup-key 20180404_113522
  ```
- 从OBS删除备份文件。
  ```
  python3 GaussRoach.py -t delete --media-type OBS --media-destination $HOME/back1 --master-port 6000  --metadata-destination $HOME/back2  --obs-server-ip x.x.x.x --bucket-name rdsbucket.resource.user4444 --cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 --backup-key 20180404_113522
  ```
- 从NAS删除备份文件。
  ```
  python3 GaussRoach.py -t delete --media-type NAS --media-destination $HOME/back1 --master-port 6000  --metadata-destination $HOME/back2 --cluster-unique-id gaussdb_backup --backup-key 20180404_113522
  ```

相关命令

无。

4.1.2.5 恢复命令

功能

restore命令用于备份数据的恢复，其中数据库恢复支持从DISK、OBS、NAS、REMOTE第三方介质恢复数据，细粒度物理恢复支持从OBS、NAS、REMOTE的实例级

版权所有 © 华为云计算技术有限公司

备份数据中恢复库/表，从OBS、NAS、REMOTE的表级备份数据中恢复表。在多租实例中，支持从OBS、NAS的实例备份数据中恢复PDB，以及从OBS、NAS的PDB级备份数据中恢复PDB。

命令格式

```
python3 GaussRoach.py
-t restore
--backup-key <backup-key>
--master-port <master-port>
--media-type <media-type>
--media-destination <media-destination-path>
--metadata-destination <metadata-path>
[--bucket-name <bucket name>]
[--buffer-size <buffer-size>]
[--buffer-block-size <block-size>]
[--clean]
[--cluster-unique-id <cluster unique id>]
[--continue]
[--cpu-relinquish-size <cpu-relinquish-size-value>]
[--cpu-relinquish-time <cpu-relinquish-time-value>]
[--enable-logging <true/false>]
[--failure-retry-count <count>]
[--logging]
[--logging-level <logging-level>]
[--logging-path <logging-path>]
[--log-filecount <maximum-log-file-count>]
[--log-filesize <maximum-log-file-size>]
[--max-memory-usage <maximum-number-of-bytes>]
[--metadata-file-wait-time <wait time>]
[--obs-server-ip <address>]
[--old-cluster-hostname]
[--parallel-process <process-count>]
[--parallel-restore]
[--pitr]
[--pitr-time <UnixTimeStamp>]
[--pre-disk-space]
[--resource-retry-count <count>]
[--master-ip <x.x.x.x/x:x:x:x:x:x:x:x>]
[--master-port <master port>]
[--media-type <media type>]
[--media-destination <media destination>]
[--metadata-destination <path-name>]
[--restore-buffer-threshold <buffer-percentage>]
[--resume-restore]
[--task-id <task-id>]
[--restore-configs]
[--restore-new-cluster]
[--restore-target-time <YYYY-MM-DD hh:mm:ss>]
[--timezone <timezone>]
[--retry-wait-time <time>]
[--single]
[--snapshot-action] <action>
[--skip-dependent-backup]
[--username <user-name>]/[-U <user-name>]
[--verbose]
[--disable-ssl]
[--remote-ip]
[--remote-port]
[--remote-config]
[--gbr-db-list]
[--gbr-db-remap]
[--gbr-table-list]
[--aux-db-path]
[--aux-db-data-path]
[--on-demand-tsp-location]
[--origin-cluster]
[--gbr-owner]
```

 版权所有 © 华为云计算技术有限公司

```
[--expdp-throttle-level]
[--impdp-copy-control]
[--reserve-auxdb]
[--pdb-list <pdb-list>]
[--pdb-remap <pdb-remap-list>]
[--prior-backup-key]
```

参数说明

表 4-4 恢复命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。恢复期间指定该参数为restore，表示调用恢复功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t restore
--backup-key	不涉及	字符串	执行恢复操作使用的backup key。	不涉及	不涉及	--backup-key *20150315-16103*
--master-port	不涉及	整数	Roach主代理进程的执行端口。	1024–65535	6812	--master-port 8808

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-type	不涉及	字符串	介质类型。 • OBS • DISK（磁盘） • NAS • REMOTE	• DISK • OBS • NAS • REMOTE	不涉及	• 对于DISK：--media-type *Disk* • 对于OBS：--media-type OBS • 对于NAS：--media-type NAS • 对于REMOTE：--media-type REMOTE

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-destination	不涉及	字符串	指定介质的目的路径。 ● OBS：样例策略 ● Disk（磁盘）：绝对路径 ● NAS：绝对路径 ● REMOTE：样例策略	● Disk ● OBS ● NAS ● REMOTE	不涉及	● 对于OBS：--media-destination *Samplepolicy* ● 对于磁盘：--media-destination */home/cam/backup* ● 对于NAS：--media-destination */home/cam/backup* ● 对于REMOTE：--media-destination *Samplepolicy*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--metadata-destination	不涉及	字符串	元数据文件位置。	不涉及	不涉及	--metadata-destination */home/username*
--bucket-name	不涉及	字符串	指定OBS桶名称。	不涉及	不涉及	--bucket-name rdsbucket.resource.user4444

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--buffer-block-size	不涉及	整数	单个缓冲块大小。	8192–268435456（字节）	非OBS，非REMOTE：2097152（字节） OBS：67108864（字节） REMOTE：16777216（字节）	--buffer-block-size *9058*
--buffer-size	不涉及	整数	缓存大小。	256–16384（MB）	非OBS：256（MB） OBS：2048（MB）	--buffer-size *10000*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--clean	不涉及	布尔型	恢复前是否清除数据。 如果命令中指定了该参数，则参数值为true。否则值为false。	不涉及	False	--clean
--cluster-unique-id	不涉及	字符串	指定数据库备份时的数据库ID。	不涉及	不涉及	--cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
--cpu-relinquish-size	不涉及	整数	处理释放大小后CPU的释放大小值。	1-10000 GB	不涉及	--cpu-relinquish-size *10 GB*
--cpu-relinquish-time	不涉及	整数	CPU的释放时间。 **说明** 将CPU释放时间设置为60秒或60秒以上以减小CPU使用率。	0到3600秒	不涉及	--cpu-relinquish-time *1000*
--enable-logging	不涉及	字符串	启用/禁用记录到日志文件功能。 建议使用此参数而不是--logging。	• True • False	False	--enable-logging *True*
--failure-retry-count	不涉及	整数	配置Roach工具失败后重试次数。如果设为n，Roach会进行n-1次重试。	0–256	32	--failure-retry-count *3*
--getdata-waittime-afterthreshold	不涉及	整数	达到缓冲极限后的等待时间，单位为微秒。	0–1800000000（微秒）	0	--getdata-waittime-afterthreshold *2000*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--logging	不涉及	布尔型	启用或禁用日志文件记录功能。 如果命令中指定了该参数，则参数值为true。否则值为false。 此参数将来会退化，建议使用--enable-logging。	不涉及	False	--logging
--logging-level	不涉及	字符串	日志级别。 ● FATAL（致命）：导致系统停止工作的、无法恢复的故障。该级别为最严重级别。 ● ERROR（错误）：重大错误。 ● WARNING（警告）：异常情况。系统在该情况下可能会继续处理任务。 ● INFO（提示）：日志记录过程中的提示性日志。 ● DEBUG（调试）：用于调试的详细信息。 ● DEBUG2（调试2）：最详细的日志信息，通常会过滤不显示。该级别严重程度最轻。	● FATAL ● ERROR ● WARNING ● INFO ● DEBUG ● DEBUG2	INFO	--logging-level *info*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--logging-path	不涉及	字符串	日志在磁盘中的存放路径。	不涉及	*$GAUSSLOG/roach/controller*	--logging-path $GAUSSLOG/roach/controller
--log-filecount	不涉及	整数	日志文件最大个数。	5~1024	10	--log-filecount *5*
--log-filesize	不涉及	整数	最大日志文件的大小。	5~20 MB	5（MB）	--log-filesize *5*
--max-memory-usage	不涉及	无符号整数	Roach可用的最大字节数。 默认值0表示无指定配额，即Roach可使用任意内存。	• 0 • 2–256（GB）	2–256	--max-memory-usage 2
--metadata-file-wait-time	不涉及	整数	读取元数据文件的等待时长。	60–3600（秒）	60（秒）	--metadata-file-wait-time 100
--obs-server-ip	不涉及	IP地址	指定OBS的服务器地址。	• x.x.x.x • x:x:x:x:x:x:x:x	不涉及	--obs-server-ip x.x.x.x
--old-cluster-hostname	不涉及	字符串	旧数据库的hostname。	不涉及	不涉及	--old-cluster-hostname roach001
--parallel-process	不涉及	整数	应使用的进程个数。	1–32	1	--parallel-process *10*

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--parallel-restore	不涉及	布尔型	副本间并行恢复，如果指定此参数，则恢复时，分片内的DN的数据均从存储介质上恢复，否则，备机的数据将在启动阶段通过全量build来恢复。 **说明** 此参数只用于从OBS/NAS/REMOTE介质恢复的场景。	不涉及	False	--parallel-restore
--pitr	不涉及	布尔型	PITR恢复。	不涉及	False	--pitr
--pitr-time	不涉及	字符串	PITR恢复时间，格式为Unix时间戳。	有效Unix时间戳	不涉及	--pitr-time 1621487640
--pre-disk-space	不涉及	布尔型	是否有足够的磁盘空间完成即将进行的操作。 如果命令中指定了该参数，则参数值为true。否则值为false。	不涉及	False	--pre-disk-space
--resource-retry-count	NA	整数	Roach工具在资源分配失败后的重试次数。如果设为n，Roach会进行n-1次重试。 **说明** 该参数可用于Roach工具任何操作中系统资源分配失败时进行的任何操作。系统资源包括堆内存、线程、数据库API。对其他失败操作的重试由--failure-retry-count参数控制。	0–256	0	--resource-retry-count 3

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--master-ip	不涉及	IP地址	输入当前节点所在数据库的IP地址，即当前登录节点在数据库配置文件中指定的IP地址。	• x.x.x.x • x:x:x:x:x:x:x:x	不涉及	--master-ip x.x.x.x
--restore-buffer-threshold	不涉及	整数	恢复操作前的可用缓存百分比。	1–100（%）	100	--restore-buffer-threshold 10
--restore-configs	不涉及	字符串	将旧数据库的默认配置项恢复到目的数据库。将备份恢复到新数据库时可使用该参数。该参数仅能与--restore-new-cluster一起使用。	-	-	--restore-new-cluster --restore-configs
--restore-new-cluster	不涉及	布尔型	将数据库恢复到另一主机上。 如果命令中指定了该参数，则参数值为true。否则值为false。 **说明** 在使用--restore-new-cluster时，restore操作后的start命令还必须包含--restore-new-cluster -X选项，如下所示：GaussRoach.py **-t start --restore-new-cluster-X** *<xml path of the new cluster>*	不涉及	不涉及	--restore-new-cluster

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--restore-target-time	不涉及	日期	指定目标日期/时间格式为YYYY-MM-DD hh:mm:ss。指定该参数，恢复时会根据指定的时间找到最近的全备或者增备的备份集进行恢复。 **说明** 当前版本，只有DISK介质恢复支持该入参，且只支持用于恢复最后一条备份线上的备份集。	有效日期和时间。	不涉及	--restore-target-time *2018-03-31 08:00:00*
--resume-restore	不涉及	布尔型	该参数用于指定本次恢复是否为断点续做。 当数据较多时，恢复失败如果不想重头再做，可以在恢复开始时加上此参数，可在断掉的位置继续做恢复。 **说明** 此参数只用于恢复，备份介质只适用于DISK、OBS和NAS。	不涉及	False	--resume-restore
--retry-wait-time	不涉及	整数	配置Roach工具在失败后重试时需要等待的时间。	1–3600（秒）	10（秒）	--retry-wait-time *1200*
--snapshot-action	不涉及	字符串	快照备份恢复操作。	restore_xlog|flash_restore_xlog	不涉及	--snapshot-action restore_xlog
--skip-dependent-backup	不涉及	布尔型	跳过指定备份集依赖的备份。	不涉及	False	--skip-dependent-backup

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--timezone	不涉及	整数	当地时区（未启动夏令时）距离格林威治的偏移秒数。 **说明** 仅PITR恢复时使用，可以通过python中time.timezone获取当前环境的参数。	不涉及	不涉及	--timezone 0
--username	-U	字符串	数据库用户名。 **说明** 只支持数据库实例初始化用户、系统管理员用户。	不涉及	不涉及	--username *cluster_user*
--task-id	不涉及	字符串	PITR恢复时归档日志的路径。	不涉及	不涉及	--task-id gaussdb
--verbose	不涉及	布尔型	是否启用verbose选项来显示详细信息。 如果命令中指定了该参数，则参数值为true。否则值为false。	不涉及	False	--verbose
--disable-ssl	不涉及	布尔型	gs_roach组件的SSL认证默认开启，若想关闭SSL认证，则在命令中指定该参数。	不涉及	False	--disable-ssl
--remote-ip	不涉及	字符串	指定远端第三方存储介质所在服务器IP。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-ip x.x.x.x

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--remote-port	不涉及	整数	指定远端第三方存储介质所在服务器侦听端口。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-port "8009"
--remote-config	不涉及	字符串	第三方厂商配置参数内容。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-config “null”
--gbr-db-list	不涉及	字符串	细粒度恢复时要恢复的数据库名的列表，库级恢复必选。如果恢复多个库，需要使用逗号隔开。 **说明** 该版本不支持对M-compatibility的数据库进行细粒度恢复。	不涉及	不涉及	--gbr-db-list 'db1，db2'
--gbr-db-remap	不涉及	字符串	细粒度恢复时要恢复的数据库恢复后重命名的列表，库级恢复可选。如果恢复多个库，需要使用逗号隔开，且数量与--gbr-db-list的数量保持一致。	不涉及	不涉及	--gbr-db-remap 'db1_new,db2_new'

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例

--gbr-table-list	不涉及	字符串	细粒度恢复时要恢复的表信息对应的json文件路径，表级恢复必选。
			json文件示例：
			{
			"TableList": [{
			"DBName": "db1",
			"SchemaName": "schema1",
			"TableName": "t1",
			"NewDBName": "new_db1",
			"NewSchemaName":"new_schema1",
			"NewTableName":"new_t1"
			},{
			"DBName": "db2",
			"SchemaName": "schema2",
			"TableName": "t2"
			"NewDBName": "new_db2",
			"NewSchemaName":"new_schema2",
			"NewTableName":"new_t2"
			}]}
				不涉及	不涉及	--gbr-table-list /data/table.json

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
			说明 该版本不支持对M-compatibility的表进行细粒度恢复。 db_name、schema_name、table_name分别是备份恢复的表的库名、模式名、表名。new_db_name、new_schema_name、new_table_name分别是恢复的表重命名后的的库名、模式名、表名。			
--aux-db-path	不涉及	字符串	细粒度恢复时的辅助数据库路径，细粒度恢复必选。	不涉及	不涉及	--aux-db-path /data/aux_db
--aux-db-data-path	不涉及	字符串	集中式部署，该参数不生效。该参数当前版本暂不支持。	不涉及	不涉及	--aux-db-data-path /data/gbr_aux_db_data_path.json
--on-demand-tsp-location	不涉及	字符串	细粒度恢复代码调试参数（正式版本中无效）。	不涉及	不涉及	--on-demand-tsp-location /data/tsp_location
--origin-cluster	不涉及	布尔型	细粒度恢复时是否恢复到原数据库实例，默认恢复到原数据库实例，该参数只对细粒度恢复生效，为细粒度恢复可选参数。	不涉及	不涉及	--origin-cluster

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--gbr-owner	不涉及	字符串	细粒度恢复时恢复后的库或表的owner，必须指定数据库实例内的已有用户，细粒度恢复必选。	不涉及	不涉及	--gbr-owner 'user1'
--expdp-throttle-level	不涉及	整型	该参数用于指定库级恢复时，将导出的文件发送给远端节点时的流控力度。仅当辅助数据库数据目录与数据库数据目录在同一硬盘时生效。取值value表示在复制过程中，每向硬盘写入一定次数后，sleep这段写入操作耗时的value%的时间。	大于等于0	0	--expdp-throttle-level 50
--impdp-copy-control	不涉及	字符串	该参数用于指定库级恢复时，数据导入是否进行磁盘I/O控制。	• Full • Normal	Full	--impdp-copy-control Full
--reserve-auxdb	不涉及	布尔型	该参数用于将恢复失败的辅助库保留下来。	不涉及	不涉及	--reserve-auxdb
--pdb-list	不涉及	字符串	指定当前进行恢复的PDB列表。 **说明** 该版本不支持对M-compatibility的PDB进行细粒度恢复。	仅可包含大小写字母、数字及下划线字符，最多包含128个PDB。	不涉及	--pdb-list "mypdb1,mypdb2"
--pdb-remap	不涉及	字符串	指定当前进行恢复的PDB列表在完成恢复后的新名称。	仅可包含大小写字母、数字及下划线字符，最多包含128个PDB。	不涉及	--pdb-remap "newpdb1,newpdb2"

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--prior-backup-key	不涉及	字符串	表示闪回快照PITR恢复时少数派节点所用的备份集。	不涉及	不涉及	--prior-backup-key *20170117_041504*

使用指南

对于数据库级恢复，必须提供介质类型、目标介质、备份密钥和主代理端口。如果用户没有指定这些参数，Roach工具会报错。

支持物理表级备份和恢复，包括单表和多表物理备份恢复，以及从实例级备份数据中恢复表/数据库的功能。

对于细粒度表级恢复，必须提供介质类型、目标介质、备份密钥、主代理端口、辅助数据库路径、表信息对应json文件路径、数据库安装用户和指定的恢复后owner。如果用户没有指定这些参数，Roach工具会报错。

对于细粒度库级恢复，必须提供介质类型、目标介质、备份密钥、主代理端口、辅助数据库路径、恢复的库列表、数据库安装用户和指定的恢复后owner。如果用户没有指定这些参数，Roach工具会报错。

对于细粒度PDB级恢复，必须提供介质类型、目标介质、备份密钥、主代理端口、辅助数据库路径、恢复的PDB列表、数据库安装用户和指定的恢复后owner。如果用户没有指定这些参数，Roach工具会报错。

命令示例

- 从OBS恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** *samplepolicy* **--media-type** *OBS*--**backup-key** *20160121_190923* **--metadata-destination** */home/userA/metadata* **--obs-server-ip** x.x.x.x **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
- 从REMOTE恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** /home/gauss/clusterid/xxx **--media-type** *REMOTE* **--backup-key** *20160121_190923* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"null"*
- 从DISK恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** *Disk* **--backup-key** *20160121_190548* **--metadata-destination** */home/userA/metadata*
- 从NAS恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** *NAS* **--backup-key** *20160121_190548* **--metadata-destination** */home/userA/metadata* **--cluster-unique-id** gaussdb_backup
- 增量恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** DISK **--backup-key** *20170630_220442* **--metadata-destination** */home/userA/metadata*

版权所有 © 华为云计算技术有限公司

- 从REMOTE增量恢复实例
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/gauss/clusterid/xxx* **--media-type** *REMOTE* **--backup-key** *20170630_220442* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"null"*
- 从REMOTE将实例恢复到新实例
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/gauss/clusterid/xxx* **--media-type** *REMOTE* **--backup-key** *20230202_092726* **--restore-new-cluster --metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"null"*
- 从REMOTE将实例恢复到PITR指定的时间点
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/gauss/clusterid/xxx* **--media-type** *REMOTE* **--backup-key** *20170630_220442* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"null"* **--pitr-time** *1684976963* **--parallel-restore**
- 从REMOTE将实例级备份恢复指定表
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-type** *REMOTE* **--task-id** *xxx_xxx* **--backup-key** *xxx* **--media-destination** */home/gauss/clusterid/xxx* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"xxx.config"* **--aux-db-path** */xxx/aux_db* **--gbr-table-lis**t */xxx/table.json* **--gbr-owner** *'user1'* **-U** *user*
- 从REMOTE将实例级备份恢复指定库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-type** *REMOTE* **--task-id** *xxx_xxx* **--backup-key** *xxx* **--media-destination** */home/gauss/clusterid/xxx* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"xxx.config"* **--aux-db-path** */xxx/aux_db* **--gbr-db-list** *'db1'* **--gbr-owner** 'user1' **-U** user
- 从REMOTE将实例级备份指定表恢复到PITR指的时间点
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-type** *REMOTE* **--task-id** *xxx_xxx* **--backup-key** *xxx* **--media-destination** */home/gauss/clusterid/xxx* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"xxx.config"* **--pitr-time** *1710471600* **--aux-db-path** */xxx/aux_db* **--gbr-table-list** */xxx/table.json* **--gbr-owner** *'user1'* **-U** *user*
- 从REMOTE将实例级备份指定库恢复到PITR指的时间点
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-type** *REMOTE* **--task-id** *xxx_xxx* **--backup-key** *xxx* **--media-destination** */home/gauss/clusterid/xxx* **--metadata-destination** */home/userA/metadata* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"xxx.config"* **--pitr-time** *1710577800* **--aux-db-path** */xxx/aux_db* **--gbr-db-list** *'db1'* **--gbr-owner** *'user1'* **-U** *user*
- 从REMOTE将表级备份指定表恢复到PITR指的时间点
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/gauss/clusterid/xxx* **--metadata-destination** */home/userA/metadata* **--media-type** *REMOTE* **--task-id** *xxx_xxx* **--backup-key** *xxx* **--remote-ip** *x.x.x.x* **--remote-port** *"8009"* **--remote-config** *"xxx.config"* **--gbr-table-list** */xxx/table_restore.json* **--aux-db-path** */xxx/aux_db* **--origin-cluster --pitr-time** *1712663512* **--gbr-owner** *'user1'* **-U** *user*
- 从OBS数据库级备份中物理恢复表
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** *OBS* **--backup-key** *20160121_190548* **--metadata-destination** */home/userA/metadata* **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 **--gbr-table-list** /data/table.json **--aux-db-path** /data/aux_db **--origin-cluster --gbr-owner** 'user1' **-U** user
- 从NAS数据库级备份中物理恢复表
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** *NAS* **--backup-key** *20160121_190548* **--metadata-destination** */home/userA/metadata* **--cluster-unique-id** gaussdb_backup **--restore-new-cluster --gbr-table-list** /data/table.json **--aux-db-path** /data/aux_db **--origin-cluster --gbr-owner** 'user1' **-U** user
- 从OBS数据库级备份中物理恢复数据库
  python3 GaussRoach.py **-t restore --clean --master-port** *6000* **--media-destination** */home/userA/media* **--media-type** *OBS* **--backup-key** *20160121_190548* **--metadata-destination** */home/userA/metadata* **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 **--restore-new-cluster --gbr-db-list** 'db1, db2' **--gbr-db-remap** 'newdb1, newdb2' **--aux-db-path** /data/aux_db **--origin-cluster --gbr-owner** 'user1' **-U** user
- 从NAS数据库级备份中物理恢复数据库

 版权所有 © 华为云计算技术有限公司

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type NAS --backup-key 20160121_190548 --metadata-destination /home/userA/metadata --cluster-unique-id gaussdb_backup --restore-new-cluster --gbr-db-list 'db1, db2' --gbr-db-remap 'newdb1, newdb2' --aux-db-path /data/aux_db --origin-cluster --gbr-owner 'user1' -U user
```

- 从OBS表级备份中物理恢复表

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type OBS --backup-key 20160121_190548_table --metadata-destination /home/userA/metadata --obs-server-ip x.x.x.x --bucket-name rdsbucket.resource.user4444 --cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 --restore-new-cluster --gbr-table-list /data/table.json --aux-db-path /data/aux_db --origin-cluster --gbr-owner 'user1' -U user
```

- 从NAS表级备份中物理恢复表

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type NAS --backup-key 20160121_190548_table --metadata-destination /home/userA/metadata --cluster-unique-id gaussdb_backup --restore-new-cluster --gbr-table-list /data/table.json --aux-db-path /data/aux_db --origin-cluster --gbr-owner 'user1' -U user
```

- 从NAS实例级备份集中恢复PDB

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type NAS --backup-key 20160121_190548 --metadata-destination /home/userA/metadata --cluster-unique-id gaussdb_backup --pdb-list "mypdb1,mypdb2" --pdb-remap "newpdb1,newpdb2"
```

- 从NAS PDB级备份集中恢复PDB

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type NAS --backup-key 20160121_190548_pdb --metadata-destination /home/userA/metadata --cluster-unique-id gaussdb_backup --pdb-list "mypdb1,mypdb2" --pdb-remap "newpdb1,newpdb2"
```

- 从OBS实例级备份集中恢复PDB

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type OBS --backup-key 20160121_190548 --metadata-destination /home/userA/metadata --obs-server-ip x.x.x.x --bucket-name rdsbucket.resource.user4444 --cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 --restore-new-cluster --pdb-list 'mypdb1,mypdb2' --pdb-remap 'newpdb1,newpdb2' --aux-db-path /data/aux_db --origin-cluster --gbr-owner 'user1' -U user
```

- 从OBS PDB级备份集中恢复PDB

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type OBS --backup-key 20160121_190548_pdb --metadata-destination /home/userA/metadata --obs-server-ip x.x.x.x --bucket-name rdsbucket.resource.user4444 --cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02 --restore-new-cluster --pdb-list 'mypdb1,mypdb2' --pdb-remap 'newpdb1,newpdb2' --aux-db-path /data/aux_db --origin-cluster --gbr-owner 'user1' -U user
```

- 将数据库恢复到新数据库

```
python3 GaussRoach.py -t restore --clean --master-port 6000 --media-destination /home/userA/media --media-type Disk --backup-key 20160621_190548 --restore-new-cluster --restore-configs --metadata-destination /home/userA/metadata
```


版权所有 © 华为云计算技术有限公司

说明

将数据库恢复到新数据库的前提条件和注意事项：

- 新数据库与备份数据库的拓扑相似。
- 新数据库与备份数据库的平台、操作系统相似。
- 确保新数据库节点ID与备份文件所在原节点ID相同。
- 确保运行Roach的主代理ID与备份数据库的Roach主代理ID相同。
- 如果备份数据库中存在表空间文件夹，必须在新数据库中新建表空间文件夹，且确保：
  - 新建表空间文件夹的路径与备份数据库中的表空间文件夹路径匹配。
  - 用户有权限访问这些文件夹。
- 用户可对将数据备份到DISK、OBS、REMOTE或NAS的数据库执行恢复操作。
- 对数据备份到DISK的数据库执行此操作时，必须将备份内容（包括backup.in文件）从源数据库复制到目标数据库。
- 源数据库的节点必须1对1映射到目标数据库的节点，并且备份内容必须复制到其各自的节点。
- restore和start两个操作必须在一个节点上执行, 否则会恢复失败。
- 在恢复数据库之前用户需要执行文件清理操作，建议用户在执行恢复命令时增加--clean参数来完成该操作。这样系统会先检测是否满足恢复条件，只有满足时才会清理数据库数据。

  用户也可选择在恢复之前单独执行clean命令清理数据库数据，但是在数据库恢复时如果系统检测到不满足恢复条件，那么数据库将无法回到正常状态。
- 恢复时不支持数据库中存在级联备只读节点。

相关命令

备份命令

4.1.2.6 genstack 命令

功能

genstack命令用于生成调用堆栈信息并将其打印到日志文件。

命令格式

```
python3 GaussRoach.py
-t genstack
```


版权所有 © 华为云计算技术有限公司

参数说明

表 4-5 genstack 命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	-t	字符串	Roach接口支持多种功能。使用genstack值指定此选项将调用生成调用堆栈功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t genstack

版权所有 © 华为云计算技术有限公司

使用指南

无

命令示例

生成调用堆栈跟踪。

```
python3 GaussRoach.py -t genstack
```

系统回显

```
Generating stack trace for all gs_roach processes into logfile
List of nodes for stack trace generation: ['hghproach00004', 'hghproach00005', 'hghproach00008']
Successfully generated stack trace
Roach operation genstack completed.
```

4.1.2.7 启动命令

功能

start命令用于在数据库恢复操作后启动数据库。

命令格式

```
python3 GaussRoach.py
-t start
[--cluster-timeout] <timeout>
[--logging-path] <logging-path>
[--parallel-jobs] <jobs-count>
[--parallel-restore]
[--restore-new-cluster]
[--standby-build-timeout] <timeout>
[-X <XML path of the new cluster>]
[--obs-server-ip <address>]
[--bucket-name <bucketname>]
[--obs-ak <ak>]
[--task-id <task id>]
[--restore-target-time <UNIX TIMESTAMP>]
[--pitr]
[--media-type <OBS|NAS|REMOTE>]
[--media-destination <media-destination>]
[--username <user-name>]
[-W <password>]
```

 版权所有 © 华为云计算技术有限公司

参数说明

表 4-6 启动命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。指定该参数为start，表示启动数据库。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t start
--cluster-timeout	不涉及	整数	数据库实例启停超时时间。	整数	1800(s)	--cluster-timeout 1800

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--logging-path	不涉及	字符串	日志存放路径。	不涉及	*$GAUSSLOG/roach/controller*	--logging-path *$GAUSSLOG/roach/controller*
--parallel-jobs	不涉及	整数	备份或恢复操作的并行进程数量。	1–32	5	--parallel-jobs 6
--parallel-restore	不涉及	布尔型	并行恢复对应的启动数据库的参数，此时备DN是通过start而非全量build来启动。 **说明** 此参数只用于从OBS/NAS/REMOTE介质的场景，且恢复时也需要指定为备机并行恢复的模式。	不涉及	False	--parallel-restore
--restore-new-cluster	不涉及	布尔型	将数据库恢复到另一主机上。 如果命令中指定了该参数，则参数值为true。否则值为false。 **说明** 该参数仅在media type为disk/OBS/NAS时有效。	不涉及	False	--restore-new-cluster

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--standby-build-timeout	不涉及	整数	主机build备机超时时间。	整数	600000(s)	--standby-build-timeout 600000
-X	-X	字符串	新数据库的XML路径。 **说明** • 该参数区分大小写，参数值必须为大写。 • 数据库内核为Paxos模式时，该参数为必选参数。	绝对路径	不涉及	-X /home/db/restore
--obs-server-ip	不涉及	字符串	OBS服务器地址。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--obs-server-ip obs.test.huawei.com
--bucket-name	不涉及	字符串	OBS桶名。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--bucket-name dyk
--obs-ak	不涉及	字符串	OBS用户名。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--obs-ak Test72JBCACXX2222222

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--task-id	不涉及	字符串	归档时使用的task_id。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--task-id test_uuid
--restore-target-time	不涉及	字符串	pitr恢复指定恢复的时间。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--restore-target-time 1624419494
--pitr	不涉及	布尔值	是否是pitr恢复启动。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--pitr
--media-type	不涉及	字符串	pitr启动时的存储介质。 **说明** 该参数仅在PITR恢复启动时使用。	• OBS • NAS • REMOTE	不涉及	--media-type OBS
--media-destination	不涉及	字符串	pitr归档日志存放的位置，只有NAS介质下pitr启动时需要指定。 **说明** 该参数仅在PITR恢复启动时使用。	不涉及	不涉及	--media-destination /data/backup/media/

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--username	-U	字符串	数据库用户名。 **说明** 只支持数据库实例初始化用户、系统管理员用户。	不涉及	不涉及	--username *cluster_user*
-W	不涉及	字符串	数据库用户的密码。出于安全考虑，不建议使用该接口传递密码。建议使用echo和pipe方式来传递密码。	不涉及	不涉及	-W foo_pswd

使用指南

在恢复数据库之后，使用start命令启动数据库。并行参数（可选）可以同start命令一同提供，用以同步运行流程。

命令示例

恢复操作后启动数据库。

```
python3 GaussRoach.py -t start
```

将数据库恢复到不同节点组后启动数据库。

```
python3 GaussRoach.py -t start --restore-new-cluster -X <XML path of the new cluster>
```

将实例恢复到pitr指定的时间后启动实例。(仅针对REMOTE介质)。

```
python3 GaussRoach.py -t start --restore-new-cluster -X <xml path of the new cluster> --task-id GaussKernel_archive --restore-target-time 1684976963 --pitr --parallel-restore --remote-ip x.x.x.x --remote-port "8009" --remote-config "null" --metadata-destination <metadata_destination_path> --media-type REMOTE
```

系统回显

```
sparrow@node200:~/db/bin> python3 GaussRoach.py -t start
Stopping the cluster.
Successfully stopped the cluster.
Pre-cleaning ETCD before 'restore-new-cluster' startup...
Successfully pre-cleaned ETCD.
Starting the cluster.
Starting ETCD.
```


版权所有 © 华为云计算技术有限公司

```
Successfully started ETCD.
Starting DN master.
Successfully started DN master.
Starting build DN standby.
Successfully start DN standby.
Starting the cluster.
Successfully started primary instance. Please wait for standby instances.
.
Successfully started standby instances.
Waiting for the cluster become normal...
Successfully started the cluster.
Roach operation start completed.
```

相关命令

恢复命令

4.1.2.8 清除命令

功能

clean命令用于在恢复操作之前清除该数据库。

命令格式

```
python3 GaussRoach.py
-t clean
[--logging-path <logging-path>]
[--restore-new-cluster]
```


版权所有 © 华为云计算技术有限公司

参数说明

表 4-7 清除命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。指定该参数为clean，表示调用清除功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t clean

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--logging-path	不涉及	字符串	日志存放路径。	不涉及	*$GAUSSLOG/roach/controller*	--logging-path *$GAUSSLOG/roach/controller*
--restore-new-cluster	不涉及	布尔型	将数据库恢复到另一主机上。 如果命令中指定了该参数，则参数值为true。否则值为false。 **说明** 该参数仅在media type为disk、OBS、REMOTE和NAS时有效。	不涉及	False	--restore-new-cluster

使用指南

- 在恢复操作之前，用户使用**清除**命令删除数据库。
- 在执行清理操作之前，请确保存在有效的备份。
- 执行清理和恢复时，请确保备份集有效。

说明

用户可以通过使用clean脚本或者作为恢复命令的参数来执行清除操作。

命令示例

恢复操作前清除数据库。

```
python3 GaussRoach.py -t clean
```

将数据库恢复到不同节点组前清除数据库。

```
python3 GaussRoach.py -t clean --restore-new-cluster
```

系统回显

```
Cleaning cluster.
Parsing the configuration file.
Stopping cluster.
Successfully stopped cluster.
Cleaning data before restoration.
Successfully cleaned data before restoration.
```


版权所有 © 华为云计算技术有限公司

```
Successfully cleaned cluster.
Roach operation clean completed.
```

相关命令

备份命令

4.1.2.9 配置命令

功能

config命令用于更新Roach的配置设置。

命令格式

```
python3 GaussRoach.py
-t config
[--archive <true-false> -p] [-f] [-c]
```

参数说明

表 4-8 配置命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	-t	字符串	Roach接口支持多种功能。指定该参数为config，表示调用配置功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t config

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--archive	不涉及	不涉及	该接口当前版本实际操作不会有任何作用。 启用和禁用数据库归档模式。 如果指定该参数为true，表示启用归档模式。 如果指定该参数为false，表示禁用归档模式。 添加-p加载更新后的参数。	• True • False	不涉及	--archive *true* -p
-f	不涉及	不涉及	gs_roach配置文件的绝对路径。将gs_roach配置文件分发到各个节点。	-	不涉及	-f
-c	不涉及	不涉及	gs_roach临时配置文件的绝对路径。当添加此参数，会在此路径下生成临时配置文件。	-	不涉及	-c
--max-backup-io-speed	不涉及	非负整数	备份操作的磁盘I/O最大速率，该配置实时生效，可以影响已经开始的备份操作。	0-2147483647	0	--max-backup-io-speed=100000000

使用指南

- 用户可使用config命令启用归档模式进行数据库备份和恢复操作。
- 用户可使用config命令禁用归档模式，防止在备份后永久启用归档模式。

版权所有 © 华为云计算技术有限公司

命令示例

启用归档模式。

```
python3 GaussRoach.py -t config --archive true -p
```

禁用归档模式。

```
python3  GaussRoach.py -t config --archive false -p
```

相关命令

无

4.1.2.10 归档命令

功能

archive命令用于开启关闭归档的功能，配合备份功能可以实现PITR功能，目前只支持OBS、REMOTE和NAS介质。

命令格式

```
echo passwords | python3 GaussRoach.py
-t archive
--archive-mode <on|off|clean|show|manually>
--bucket-name <bucketname>
--media-type <OBS|NAS|REMOTE>
--obs-ak <ak>
--obs-server-ip <address>
--task-id <task id>
--username <username>
--delete-all-archived-xlog
--need-recycle-xlog
--is-auto-archive False
--backup-key-list <backupkey1,backupkey2,...>
--remote-ip <remote ip>
--remote-port <remote port>
--remote-config <remote config>
[--disable-ssl]
```

 版权所有 © 华为云计算技术有限公司

参数说明

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	-t	字符串	Roach接口支持多种功能。指定该参数为archive，表示调用归档功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t archive
--archive-mode	不涉及	字符串	归档的模式。	• on：开启归档 • off：关闭归档 • clean：清理归档 • show：查询PITR可恢复时间 • manually：手动归档	不涉及	--archive-mode on
--bucket-name	不涉及	字符串	OBS桶名。	不涉及	不涉及	--bucket-name dyk
--media-type	不涉及	字符串	存储介质。	• OBS • NAS	不涉及	--media-type OBS

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-destination	不涉及	字符串	指定介质的目的归档路径，仅对NAS有效。	不涉及	不涉及	--media-destination /home/backup/media
--obs-ak	不涉及	字符串	OBS用户名。	不涉及	不涉及	--obs-ak Test72JBCACXX2222222
--obs-server-ip	不涉及	字符串	OBS服务器IP地址。	不涉及	不涉及	--obs-server-ip obs.test.huawei.com
--task-id	不涉及	字符串	归档id。	不涉及	不涉及	--task-id gaussdb_archive
--username	-U	字符串	数据库用户名。 **说明** 只支持数据库实例初始化用户、系统管理员用户、Operatoradmin权限用户且operation_mode开启。	不涉及	不涉及	--username rdsAdmin
password	-W	字符串	数据库用户的密码。	不涉及	不涉及	******
--backup-key	不涉及	字符串	当archive mode为show时，表示最近一次备份集对应的backupkey，用来查询此次备份集中备份的cn是否在归档最大时间内发生了切换。当archive mode为clean时表示需删除此次backupkey之前的日志，该参数支持分布式场景查询CN时使用，集中式场景可忽略。	不涉及	不涉及	--backup-key 20220101_201206

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--cluster-unique-id	不涉及	字符串	备份集所在的路径用来查询备份集对应的backupkey。	不涉及	不涉及	--cluster-unique-id GaussKernel_roach
--start-time	不涉及	字符串	所需查询可恢复时间的起点。	格式为Unix时间戳。	不涉及	--start-time 1555323441
--delete-all-archived-xlog	不涉及	字符串	是否删除所有已归档日志，archive mode为clean时使用。	不涉及	不涉及	--delete-all-archived-xlog
--need-recycle-xlog	不涉及	布尔型	开启归档时是否删除当前归档位置之后的归档日志，archive mode为on时使用。	不涉及	不涉及	--need-recycle-xlog
--disable-ssl	不涉及	布尔型	gs_roach组件的SSL认证默认开启，若想关闭SSL认证，则在命令中指定该参数。	不涉及	False	--disable-ssl
--is-auto-archive	不涉及	布尔型	仅当archive mode为on时有效，标识为手动或者自动归档（仅针对REMOTE介质，当前版本仅支持false取值）。	不涉及	False	--is-auto-archive False
--backup-key-list	不涉及	字符串	backupkey列表，backupkey个数最大值为24，仅当archive mode为manually时有效（仅针对REMOTE介质）。全部填写实例级备份的backupkey，或全部填写表级备份的backupkey，不能混用。	不涉及	不涉及	--backup-key-list 20230424_151220,20230424_170146

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--remote-ip	不涉及	字符串	远端第三方存储介质所在服务器IP。	不涉及	x.x.x.x	--remote-ip x.x.x.x
--remote-port	不涉及	字符串	远端第三方存储介质所在服务器的侦听端口。	不涉及	8009	--remote-port 8009
--remote-config	不涉及	字符串	远端第三方厂商配置参数内容。	不涉及	null	--remote-config null

使用指南

开启归档时，要保证每次归档日志的路径是唯一的，对应的OBS/NAS归档路径中没有已有的归档日志，否则在PITR恢复中会导致恢复时使用了已有的归档日志导致恢复失败。

说明

- 在REMOTE介质进行归档时，单次归档命令12小时超时，如果未归档完成，不会继续归档。
- 在REMOTE关闭归档后再打开归档，上次可恢复点到关闭归档时间段无法进行恢复，建议在关闭归档前进行手动归档后再关闭。
- 利用REMOTE介质归档，打开归档后，需要立即执行一次全量备份。
- 通过REMOTE介质日志归档成功之后，数据库内核会把可恢复时间区间段以basic_recovery_time_{unix_timestamp}文件的形式上传到第三方存储，第三方需要解析此文件的内容以获取准确的可恢复时间区间。在进行PITR恢复时，选取的时间点要为可恢复时间区间内的时间，如果选取不在区间内的时间用于恢复，会导致恢复失败。一个文件内可能包含多个可恢复时间区间，最后还包含一行里面有all cn build flag的字符串，取值是1或者0。如果是1，表示所有节点发生过build，第三方解析之后需要立刻执行一次全量备份/增量备份；如果是0，则忽略。

REMOTE 介质归档日志空洞场景约束

表 4-9 REMOTE 介质归档日志空洞场景表

产生可恢复区间空洞情况	产生空洞原因	空洞区间	减小空洞区间措施
手动关闭归档，然后打开归档。	关闭归档后不会barrier打点，日志里没有时间信息。	从关闭归档前日志归档成功的时间点到开启归档后第一次全量备份结束的时间段。	操作前手动执行日志归档，使得归档进度往前推进靠近关闭归档时间点。

版权所有 © 华为云计算技术有限公司

产生可恢复区间空洞情况	产生空洞原因	空洞区间	减小空洞区间措施
恢复到原实例。	恢复完成后相当于一个新实例，没有归档槽相关信息，需要重新打开归档，执行一次全量备份，之后一直做手动归档日志才有可恢复时间。	从恢复开始前日志归档成功的时间点到第一次全量备份结束的时间段。	操作前手动执行日志归档。
大版本升级（内核）。	不同大版本之间产生的日志不兼容，所以和之前的日志不是连续的，需要先关闭归档，再打开归档，再全量备份，然后再日志归档。	从上一次日志归档完成的时间点到升级提交后第一次全量备份结束的时间段。	操作前手动执行日志归档，使得归档进度往前推进，减小空洞区间。
日志还未归档就被强制回收。	日志被回收后，这段日志是缺失的，和之后的日志都是不连续的，无法用于恢复，必须做一次全量备份，然后继续做归档日志，才能让这次全量备份之后的时间可以PITR。	从日志被强制回收的开始时间点到下一次全量备份结束的时间段。	1. 降低归档时间间隔。 2. 根据业务量和留存期设置充足的日志保留量。 3. 如果1天1归档，需要确保每天的归档成功，如果失败，建议手动重试。
主DN故障发生主备切换，PITR可恢复时间无法正常推进。	主DN故障，无法barrier打点，可恢复时间区间会出现空洞。	主DN故障到恢复正常的时间段。	节点部署在不同的可用区，最大程度避免同时发生故障。
产品约束：5天以上不手动归档日志，PITR可恢复时间无法正常推进。	5天不归档barrier缓冲区满，无法barrier打点，可恢复时间区间会出现空洞。	从barrier缓冲区满的时间点到执行手动归档日志成功前的这段时间。	请根据实际的业务确定归档频率，两次归档时间间隔不应该超过5天，推荐每天至少归档一次。

版权所有 © 华为云计算技术有限公司

产生可恢复区间空洞情况	产生空洞原因	空洞区间	减小空洞区间措施
Quorum模式和DCF模式相互切换	模式切换过程中，数据库内核涉及到线程关闭和拉起，可能会影响barrier打点间距扩大超过5s，模式切换次数超过固定次数（3次），需要重启数据库实例，对barrier打点的时间影响不超过1分钟	从切换开始的1分钟内	/

命令示例

- 开启OBS归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id gaussdb_archive --obs-server-ip obs.test.huawei.com --bucket-name dyk --obs-ak Test72JBCACXX2222222 -U rdsAdmin --archive-mode on --media-type OBS
  ```

  注：其中passwords由数据库密码和obs-sk base64转码后拼接而成：base64.b64encode(password.encode()).decode() + ":" + base64.b64encode(sk.encode()).decode()

- 开启NAS归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id gaussdb_archive -U rdsAdmin --archive-mode on --media-destination /home/backup/media --media-type NAS
  ```

  注：其中passwords仅为数据库用户密码，即上述列表中的-W参数。

- 开启REMOTE归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive -U rdsAdmin --archive-mode on --media-type REMOTE --is-auto-archive False --remote-ip x.x.x.x --remote-port "8009" --remote-config "null"
  ```

  注：其中passwords仅为数据库用户密码，即上述列表中的-W参数。

- 触发REMOTE手动归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive -U rdsAdmin --archive-mode manually --media-destination /home/backup/media --metadata-destination /home/omm/gs_roach/backup/metadata --media-type REMOTE --backup-key-list "20230424_151220,20230424_170146" --remote-ip x.x.x.x --remote-port "8009" --remote-config "null"
  ```

  注：其中passwords仅为数据库用户密码，即上述列表中的-W参数。

- 部分清理OBS归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive --obs-server-ip obs.test.huawei.com --bucket-name dyk --obs-ak Test72JBCACXX2222222 -U rdsAdmin --archive-mode clean --media-type OBS --backup-key 20231212_113106
  ```

  注：删除--backup-key指定的备份集最后一条xlog（不删除本条）之前的所有在归档路径--task-id下的归档xlog

- 全部清理OBS归档。

  ```
  echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive --obs-server-ip obs.test.huawei.com --bucket-name dyk --obs-ak Test72JBCACXX2222222 -U rdsAdmin --archive-mode clean --media-type OBS --delete-all-archived-xlog
  ```

  注：删除在归档路径--task-id下的所有归档xlog

- 部分清理NAS归档。

版权所有 © 华为云计算技术有限公司

```
echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive -U rdsAdmin --archive-mode clean --media-destination /home/backup/media --media-type NAS --backup-key 20231123_101426
```

- 全部清理NAS归档。

```
echo passwords | python3 GaussRoach.py -t archive --task-id GaussKernel_archive -U rdsAdmin --archive-mode clean --media-destination /home/backup/media --media-type NAS --delete-all-archived-xlog
```

4.1.2.11 快照恢复命令

功能

snapshot_restore命令用于快照备份的恢复。

命令格式

```
python3 GaussRoach.py
-t snapshot_restore
--snapshot-action snapshot_action
--backup-config-path path
[--media-type <media-type>]
[--media-destination <media-destination-path>]
[--metadata-destination <metadata-path>]
[--bucket-name <bucketname>]
[--cluster-unique-id <value>]
[--obs-server-ip <address>]
[--backup-key <backup-key>]
```


版权所有 © 华为云计算技术有限公司

参数说明

表 4-10 快照恢复命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	-t	字符串	快照恢复功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t snapshot_restore
--snapshot-action	不涉及	字符串	快照恢复的具体操作。	backup_config|restore_config|start	不涉及	--snapshot-action backup_config
--backup-config-path	不涉及	字符串	配置文件备份路径。	不涉及	不涉及	--backup-config-path /home/Ruby/config
--backup-key	不涉及	字符串	执行恢复操作使用的backup key。	不涉及	不涉及	--backup-key *20150315-16103*
--media-type	不涉及	字符串	介质类型，当前快照恢复只支持OBS、NAS、REMOTE介质	• OBS • NAS • REMOTE	不涉及	对于OBS：--media-type OBS

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-destination	不涉及	字符串	指定介质的目的路径。 • OBS：样例策略 • NAS：绝对路径 • REMOTE：样例策略	• OBS • NAS • REMOTE	不涉及	• 对于OBS：--media-destination *Samplepolicy* • 对于NAS：--media-destination */home/cam/backup* • 对于REMOTE：--media-destination *Samplepolicy*
--metadata-destination	不涉及	字符串	元数据文件位置。	不涉及	不涉及	--metadata-destination */home/username*
--bucket-name	不涉及	字符串	指定OBS桶名称。	不涉及	不涉及	--bucket-name rdsbucket.resource.user4444
--cluster-unique-id	不涉及	字符串	指定数据库实例备份时的数据库实例ID。	不涉及	不涉及	--cluster-unique-id dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
--obs-server-ip	不涉及	IP地址	指定OBS的服务器地址	x.x.x.x x:x:x:x:x:x:x:x	不涉及	--obs-server-ip x.x.x.x

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--remote-ip	不涉及	字符串	指定远端第三方存储介质所在服务器IP。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-ip x.x.x.x
--remote-port	不涉及	整数	指定远端第三方存储介质所在服务器侦听端口。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-port "8009"
--remote-config	不涉及	字符串	第三方厂商配置参数内容。 **说明** 该值仅在media type参数取值为REMOTE时需要设置。	不涉及	不涉及	--remote-config “null”

使用指南

无。

命令示例

备份配置文件。

```
python3 GaussRoach.py -t snapshot_restore --snapshot-aciton backup_config --backup-config-path /home/Ruby/config
```

系统回显

```
champ@BLR1000014692:~/db/bin/script> python3 GaussRoach.py -t snapshot_restore --snapshot-action backup_config --backup-config-path /home/Ruby/config
Roach operation snapshot_restore completed.
```

相关命令

无

版权所有 © 华为云计算技术有限公司

4.1.2.12 版本命令

功能

version命令用于查询Roach工具的版本号。

命令格式

```
python3 GaussRoach.py
--version | -v
```

参数说明

表 4-11 版本命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	-t	字符串	Roach接口支持多种功能。使用version值指定此选项可显示Roach工具版本信息。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t snapshot_restore • -t version • -t help	不涉及	-t version

使用指南

无

命令示例

查询Roach工具的版本号。

```
python3 GaussRoach.py --version
```


版权所有 © 华为云计算技术有限公司

系统回显

```
champ@BLR1000014692:~/db/bin/script> python3 GaussRoach.py --version
Roach V100R002C10SPC330
```

说明

图中的版本号仅用于举例。最终版本号以实际为准。

相关命令

无

4.1.2.13 帮助命令

功能

help命令用于提供Roach工具支持的命令的帮助信息。

命令格式

```
python3 GaussRoach.py
--help | -h
```

参数说明

无

使用指南

无

命令示例

查看Roach工具的帮助信息。

```
python3 GaussRoach.py --help
```

系统回显

```
python3 GaussRoach.py --help

GaussRoach.py is a utility to backup GaussDB to DISK, OBS, NAS or REMOTE and restore GaussDB from
DISK, OBS, NAS or REMOTE.

Usage:
  python3 GaussRoach.py -h | --help
  python3 GaussRoach.py -v | --version
  python3 GaussRoach.py -t config [-f confFile] [-c confFile] [-p] [--archive ARCHIVEMODE] [--logging-path
filepath] [--cluster-timeout timeout] [--max-backup-io-speed speed]
  python3 GaussRoach.py -t clean [-D] [--logging-path filepath][--restore-new-cluster] [--cluster-timeout
timeout]
  python3 GaussRoach.py -t start [--logging-path filepath] [--parallel-jobs] [--restore-new-cluster -X
CONFIGXMLPATH ] [--parallel-restore] [--cluster-timeout timeout] [--standby-build-timeout timeout] [--
single] [--pitr]
  python3 GaussRoach.py -t backup [options]
  python3 GaussRoach.py -t restore [options]
  python3 GaussRoach.py -t stop [--stop-object] [-F] [--stop-timeout] [--logging-path filepath] [--pitr] [--
single] [--metadata-destination path-name]
  python3 GaussRoach.py -t delete [options]
  python3 GaussRoach.py -t backup [options]
```


版权所有 © 华为云计算技术有限公司

```
  python3 GaussRoach.py -t restore [options]
  python3 GaussRoach.py -t show [options]
  python3 GaussRoach.py -t validate [options]
  python3 GaussRoach.py -t genstack
  python3 GaussRoach.py -t getobject [options]
  python3 GaussRoach.py -t retrieve [options]
  python3 GaussRoach.py -t clean_auxdb [options]
  python3 GaussRoach.py -t archive --archive-mode archive_mode [--task-id taskid] [--start-time starttime]
[--delete-all-archived-xlog] [--need-recycle-xlog] [--is-auto-archive] [--backup-key-list backupkey_list]
  python3 GaussRoach.py -t snapshot_restore --snapshot-action snapshot_action --backup-config-path path

  Common options:
  -t                                    The operation for GaussRoach followed by backup, clean, config, delete,
genstack, restore, show, start, stop, validate, getobject.
  -h  --help                            Show this help, then exit.
  -v  --version                         Display the current roach version.
  --master-port <master port>               The port in which master roach should be started.
  --media-type <media type>                 Mention type of media either DISK, OBS, NAS or REMOTE.
  --media-destination <media destination>       The path or policy that should be used for backup/restore as
per media-type.
  --buffer-size <size>                  The buffer size that used in one process. In MB, range is [256~16384].
  --buffer-block-size<size>                 The buffer size that used in one thread. In byte, range is
[524288~268435456].
  --parallel-process <count>                The number of process that should be used. Range is [1~32].
  --prefetch-block <block-num>              The number of prefetch pages for reading incremental
modification pages of table files on disk during incremental backup. Range is [1~8192]
  --logging-level <logging level>               [FATAL|ERROR|WARNING|INFO|DEBUG|DEBUG2]. Default logging
level is warning.
  --logging                             Enable logging into log file, will be disabled by default. Use --enable-
logging instead.
  --enable-logging <True/False>             Enable logging into log file. Default is True.
  --log-filesize <size>                 The maximum log file size (in MB). Range is [5~20].
  --log-filecount <count>                   The maximum number of log files that should be created. Range is
[5~1024].
  --logging-path  <path-name>               The location where log files must be created.
  --verbose                             Gives details about the command being executed.
  --failure-retry-count <count>             The number of times it should retry in case of failure. Range is
[0~256].
  --retry-wait-time     <time>              Time it should wait before retrying after failure. In seconds. Range
is [1~3600].
  --cpu-relinquish-time <time>              CPU is relinquished for the specified relinquish time. Relinquish
time in seconds. Range is [0~3600].
  --cpu-relinquish-size <size>              CPU is relinquished after processing the relinquish size. Range is
[1~10000].
  --metadata-destination <path-name>            The path where the metadata file is to be kept in backup.
  --metadata-file-wait-time <time>          Wait time for reading metadata file. Range is [60~3600].
  --max-memory-usage <value>                Maximum allowed memory usage (in GB) for Roach process.
Range is [2~256].
  --pre-disk-space                      Checks availability of Disk Space before Backup and Restore.
  --resource-retry-count                    The number of times it should retry for library APIs like fopen,
pthread.
  --master-ip                           The ip on which master roach should be started.
  --obs-server-ip <address>                 OBS server ip address or domain name address.
  --bucket-name <bucketname>                OBS bucket name.
  --cluster-unique-id <value>               Used to indicate the unique identifier of the cluster.
  --pitr                                Used to indicate the pitr backup/restore mode
  --single                              Used to indicate the single mode
  --slotname <slotname>                     Backup slot name

Options for config
  -f                                    The path of gs_roach config file that will be distributed to every host.
  -c                                    The path of gs_roach template config file that will be created.
  -p                                    Config environment parameter and archive GUC parameter.
  --archive  <true/false>                   Decide whether start archive or stop archive. Default is true.
  --cluster-timeout                     The timeout value for cm_ctl start/stop.
  --max-backup-io-speed                     Io limit of backup operation, the value will be saved to
$GAUSSHOME/bin/io_limit.conf and will effect roach backup speed immediately, runing roach progress are
also effected.
```

 版权所有 © 华为云计算技术有限公司

```
Options for start
  -X                                The path of the Config XML file used for cluster setup.
  -U  --username <username>                 User name for the cluster during backup.
  -W <password>                     Database password for the user name passed in during backup.
Deprecated parameter. Use echo and pipe method instead.
  --logging-path                    The location where log files must be created.
  --parallel-jobs                   The number of parallel jobs to be performed for complete sync of
standby data nodes from primary data nodes.
  --restore-new-cluster             To perform start for restore data to new cluster.
  --parallel-restore                To perform start for parallel restore.
  --cluster-timeout                 The timeout value for cm_ctl start/stop.
  --standby-build-timeout             The timeout value for standby instance build.
  --pitr                            Used to indicate the pitr backup/restore mode
  --single                          Used to indicate the single mode
  --restore-target-time             Time(<=) to be used for pitr restore, format Unix Timestamp in
seconds. It should be smaller than the timestamp in hadr_barrier_id

Options for clean
  --logging-path                    The location where log files must be created.
  --restore-new-cluster             To perform clean for restore data to new cluster.
  --cluster-timeout                 The timeout value for cm_ctl start/stop.

Options for backup
  -r  --conf-location <path-name>         Path for the gs_roach.conf file.
  -U  --username <username>                 User name for the cluster.
  -W <password>                     Database password for the user name passed in. Deprecated
parameter. Use echo and pipe method instead.
  --max-backup-io-speed <speed>           The maximum IO speed that backup can write to media
destination per parallel process. Unit: B/s. 0 means no limit.
  --split-local-disk-storage        Store media data on different disk partitions of each DN seperately,
the default is stored on one disk partition specified by media destination.
  --compression-level <level>           The compression level that should be used for backup.
  --compression-type <type>             The compression type that should be used for backup.
  --cbm-recycle-level <level>           The cbmRecycle level that should be used to recycle cbm files
during backup.
  --filesplit-size <size>           Approximate file size after which the file will be split.
  --validation-type <value>             Validate the contents of backup. Values "partial" or "full".
  --prior-backup-key                Previous full/incremental backup key to be considered as the base
for incremental backup.
  --prior-cluster-unique-id         Previous full/incremental backup cluster unique id. Used to
download metadata info from OBS before backup operation.
  --validate-prior-backups <value>        Validate all prior backups (if any). One of the two values i.e.
"Force" (mandatorily run validate), "Optimistic" (if not validated, it will run validate).
  --reader-thread-count             Number of reader threads for processing. Range is [1~65535].
  --reader-thread-file-count          Number of buffer files for each reader thread. Range is [1~65535].
  --reader-thread-file-size          Buffer file size for each buffer file of reader thread. Range is
[1~65535].
  --resume-backup                   Start resume backup mode.
  --resume-restore                  Start resume restore mode.
  --check-balance-retry-times         The retry times of checking cluster balance for resume backup
  --single                          Used to indicate the single mode
  --pitr                            Used to indicate the pitr backup/restore mode
  --master-port <master port>           The port in which master roach should be started.
  --media-type <media type>             Mention type of media either DISK, OBS, NAS or REMOTE.
  --media-destination <media destination>       The path or policy that should be used for backup/restore as
per media-type.
  --metadata-destination <path-name>        The path where the metadata file is to be kept in backup.
  --logging                         Enable logging into log file, will be disabled by default. Use --enable-
logging instead.
  --logging-level <logging level>         [FATAL|ERROR|WARNING|INFO|DEBUG|DEBUG2]. Default logging
level is warning.
  --skip-delay-ddl                  Skip delay ddl.
  --single-group-nodeid               Single cluster group node id.
  --single-group-datapath             Single cluster group data path.
  --standby-mode                    In standby mode, the backup is executed on standby datanodes.
  --disable-ssl                     Disable ssl of roach
  --gbr-table-list <file name>          Absolute path of input file name which contains list of tables(in
json format) to be backed up.
                                    The contentx example "{"TableList":
```

 版权所有 © 华为云计算技术有限公司

```
[{"DBName":"db1","SchemaName":"schema1","TableName":"table1"},
{"DBName":"db2","SchemaName":"schema2","TableName":"table2"}]}"
  --remote-ip <remote ip>                    BackupAgent ip address.
  --remote-port <remote port>                  BackupAgent port.
  --remote-config <remote config>              The config file used for backupAgent.
  --pdb-list <pdb list>                    List of pdbs for pdb backup/restore, separate each pdb with a
comma, format 'pdb1,pdb2,pdb3'.

Options for restore
  --restore-buffer-threshold <buffer-threshold> The percentage of buffer that should be available before
restoring the next file.
  --backup-key <backup key>                    The key used for restore.
  --clean                                  For cluster restore, performs clean up activity internally.
  --restore-new-cluster                      Restore data to new cluster.
  --restore-configs                        Few config files are restored from backup with changed values for IP,
Port, Hostname and Path from current cluster.
  --parallel-restore                       Parallel restore DN, to avoid full build when start cluster, only support
OBS.
  --validation-type <value>                  Validate the backup before doing restore. Values "partial" or "full".
  --restore-target-time                    Time(<=) to be used for restore, format '<YYYY-MM-DD HH:MM:SS>'.
  --old-cluster-hostname                     Using anyone of hostname in old cluster when restoring to new
cluster.
  --skip-dependent-backup                    Skip the dependent backups, only the specified backup will be
restored.
  --single                                 Used to indicate the single mode
  --pitr                                   Used to indicate the pitr backup/restore mode
  --master-port <master port>                  The port in which master roach should be started.
  --media-type <media type>                    Mention type of media either DISK, OBS, NAS or REMOTE.
  --media-destination <media destination>       The path or policy that should be used for backup/restore as
per media-type.
  --metadata-destination <path-name>           The path where the metadata file is to be kept in backup.
  --logging                                Enable logging into log file, will be disabled by default. Use --enable-
logging instead.
  --logging-level <logging level>              [FATAL|ERROR|WARNING|INFO|DEBUG|DEBUG2]. Default logging
level is warning.
  --pitr-time                              Pitr restore target time in Unix Timestamp Format.
  --disable-ssl                            Disable ssl of roach
  --gbr-db-list <db list>                    Db list for fine-grained backup and restoration, separate each db
with a comma, format 'db1,db2,db3'
  --gbr-db-remap <new db list>                 New Db list for fine-grained backup and restoration, separate
each db with a comma, format 'newdb1,newdb2,newdb3'
  --gbr-table-list <file name>                 Absolute path of input file name which contains list of tables(in
json format) to be restored.
                                         The contentx example "{"TableList":
[{"DBName":"db1","SchemaName":"schema1","TableName":"table1"},
{"DBName":"db2","SchemaName":"schema2","TableName":"table2"}]}"
  --aux-db-path <aux db path>                  Auxiliary database path for fine-grained backup and restoration.
  --aux-db-data-path <aux db data path>         Auxiliary database path for fine-grained backup and
restoration, if this parameter is specified, the aux db path will link to it.
  --on-demand-tsp-location                   The fine-grained backup and restoration table space.
  --origin-cluster                         If restore to origin cluster
  --gbr-owner                              The owner of database/table in fine-grained backup and restoration.
  --remote-ip <remote ip>                    BackupAgent ip address.
  --remote-port <remote port>                  BackupAgent port.
  --remote-config <remote config>              The config file used for backupAgent.
  --expdp-throttle-level                   Throttle level for distribute expdp files by build
  --pdb-list <pdb list>                    List of pdbs for pdb backup/restore, separate each pdb with a
comma, format 'pdb1,pdb2,pdb3'.
  --pdb-remap <new pdb list>                   New pdb list for pdb backup/restore, separate each pdb with a
comma, format 'pdb1,pdb2,pdb3'.
  --impdp-copy-control                     IO control for impdp database
  --reserve-auxdb                          Reserve auxdb when gbr restore failed

Options for show
  --all-backups                            Show complete backup catalogue information. Default action if no
option given.
  --backup-key <backup_key>                    Show restore path up to a given backup key.
  --output-file <file path>                  Absolute file path with filename, where output of the show
operation will be stored.
```

 版权所有 © 华为云计算技术有限公司

```
  --no-display                          To suppress the output on console, and redirect to the output file
silently.
  --overwrite                           To overwrite the existing file.
  --related-backup-keys                      To show all related backups of the provided backup key.
  --restore-target-time                 To show the latest cluster level full backup and incremental backup
before specified time that can be used to restore, format '<YYYY-MM-DD HH:MM:SS>'.

Options for stop
  --stop-object                         Stop object, include cluster
  -F                                   Forcefully stop the object or backup operation.
  --stop-timeout                        Time after which the stop operation will exit.
  --pause-resume-backup                      Pause resume backup.
  --logging-path <filepath>                 The location where log files must be created.
  --pitr                               Used to indicate the pitr backup/restore mode
  --single                              Used to indicate the single mode
  --metadata-destination <path-name>          The path where the metadata file is to be kept in backup.

Options for validate
  [options]                             Similar options as 'restore'...

Options for delete
  --cascade                             To delete all descendant backups of the provided backup key.
  --disable-ssl                         Disable ssl of roach

Options for archive
  --obs-ak                              The OBS ak.
  --task-id                            The archive task id.
  --archive-mode                           The archive mode. [on|off|show|clean|manually]
  --delete-all-archived-xlog                To delete all archived xlog
  --start-time                          The start time of archive:show.
  --need-recycle-xlog                      Delete archived xlog whose lsn is larger than current lsn when start
archive.
  --timezone                            The timezone out of /var/chroot in PUBLIC_CLOUD scene.
  --is-auto-archive                      Optional parameter, manual or automatic archiving.
  --backup-key-list                      Optional parameter, the backup key list for manual archiving.
  --media-type <media type>                   Optional parameter, only supported for backupAgent, the value
must be "REMOTE".
  --remote-ip <remote ip>                Optional parameter, backupAgent ip address.
  --remote-port <remote port>               Optional parameter, backupAgent port.
  --remote-config <remote config>            Optional parameter, the config file used for backupAgent.

Option for Getobjct
  --restore-count                       To Count the number of times recovery is required.

Option for snapshot backup/restore
  --snapshot-action                      The specific operations for snapshot backup/restore, can be
backup_config, restore_config, start
  --backup-config-path                      The path for backup instance configs
  --block-dw-flush-timeout                   The timeout of block dw flush in snapshot backup prepare
  --prior-dn-list                      The prior backup dn for snapshot backup, the specified DN is
preferentially selected for snapshot backup.
  --node-list                          The node list which includes dn instances that need to backup xlog
  --prior-backup-key                       Specify the backup key which includes minority DNs' xlog when
flash snapshot restore

Options for retrieve
  --media-type                           retrieve type, now only support OBS/NAS
  --obs-server-ip                       OBS server ip address or domain name address.
  --bucket-name                          OBS bucket name.
  --media-destination                     The path or policy that should be used for retrieve log in NAS
  --metadata-destination                   The path where the metadata file is to be kept in backup.
  --cluster-unique-id                    Used to indicate the unique identifier of the cluster.
  --task-id                            The archive task id when retrieve xlog
  --backup-key                           Show retrieve path up to a given backup key.
  --retrieve-inst                       retrieve inst, such as "cn5001" or "dn6002"
  --retrieve-local-path                    the local filepath in the node which contain retrieve inst
  --retrieve-type                       the log type to retrieve, now only support "csnlog,clog,xlog"
  --retrieve-start-lsn                   the start point of lsn range when retrieve several xlog
  --retrieve-end-lsn                     the end point of lsn range when retrieve several xlog
```

 版权所有 © 华为云计算技术有限公司

```
  --retrieve-replace                    the bool mark indicate whether replace the retrieve logs with those in
current working cluster
  --retrieve-replace-filelist           the list record the log to be replaced
  --retrieve-replace-force              the mark force to replace regardless of version change
  --retrieve-clean                      the mark clean all the file in the --retrieve-local-path/roachRetrieve

Options for clean_auxdb
  --aux-db-path <aux db path>           Auxiliary database path to clean

### Refer user manual for more details ###
```

相关命令

无

4.1.2.14 日志找回命令

功能

retrieve命令用于从指定的备份集中恢复clog/csnlog到本地，或从归档路径中恢复指定的lsn范围内的xlog文件到本地，并可由用户选择是否将恢复出的文件覆盖至当前数据库实例的数据目录。

命令格式

```
python3 GaussRoach.py
-t retrieve
[--media-type <media-type>]
[--retrieve-type <retrieve_type>]
--retrieve-inst <retrieve_inst>
--retrieve-local-path <local_path>
[--media-destination <media-destination-path>]
[--metadata-destination <metadata-path>]
[--bucket-name <bucketname>]
[--cluster-unique-id <value>]
[--obs-server-ip <address>]
[--retrieve-start-lsn <start_lsn>]
[--retrieve-end-lsn <end_lsn>]
[--task-id <task-id>]
[--retrieve-replace]
[--retrieve-replace-filelist <json_path>]
[--retrieve-replace-force]
[--retrieve-clean]
```

 版权所有 © 华为云计算技术有限公司

参数说明

表 4-12 日志找回命令的命令行参数列表

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
-t	不涉及	字符串	Roach接口支持多种功能。指定该参数为retrieve，表示调用日志找回功能。	• -t backup • -t stop • -t delete • -t restore • -t genstack • -t getobject • -t start • -t clean • -t config • -t archive • -t retrieve	不涉及	-t backup
--media-type	不涉及	字符串	待找回日志的备份集或归档集的介质类型。 • OBS • NAS	• OBS • NAS	不涉及	对于OBS：--media-type OBS 对于NAS：--media-type NAS

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--media-destination	不涉及	字符串	指定待找回日志的备份集的数据存储路径。 ● OBS：不生效 ● NAS **说明** 使用备份数据库到NAS时，确保集群上所有节点的指定路径挂载的是同一个NAS共享盘	不涉及	不涉及	对于NAS：--media-destination */home/cam/backup*
--metadata-destination	不涉及	字符串	元数据文件位置。	不涉及	不涉及	--metadata-destination */home/username*
--bucket-name	不涉及	字符串	指定OBS桶名称。	不涉及	不涉及	--bucket-name *rdsbucket.resource.user4444*
--cluster-unique-id	不涉及	字符串	指定待找回日志的备份集的集群ID。	不涉及	不涉及	--cluster-unique-id *dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02*
--obs-server-ip	不涉及	IP地址	指定OBS的服务器地址。	x.x.x.x x:x:x:x:x:x:x:x	不涉及	--obs-server-ip x.x.x.x

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--retrieve-type	不涉及	字符串	该参数用于指定要找回的日志类型。	• clog • csnlog • clog,csnlog • xlog	不涉及	--retrieve-type clog
--retrieve-inst	不涉及	字符串	该参数用于指定要找回日志的节点id，仅支持dn主备级联备。	• dnxxxx	不涉及	--retrieve-inst dn6001
--retrieve-local-path	不涉及	字符串	该参数用于指定找回的日志在指定节点的本地暂存路径，必须是非数据库实例数据目录，否则会报错。	不涉及	不涉及	--retrieve-local-path /data/retrieve
--retrieve-start-lsn	不涉及	字符串	该参数用于指定要找回的xlog的起始lsn。	xxxxxx/xxxxxx	不涉及	--retrieve-start-lsn 0/40472D8
--retrieve-end-lsn	不涉及	字符串	该参数用于指定要找回的xlog的终止lsn。	xxxxxx/xxxxxx	不涉及	--retrieve-end-lsn 0/41531D8
--retrieve-replace	不涉及	布尔型	该参数用于指定本次操作为覆盖操作。不能与--retrieve-clean同时传入。	不涉及	False	--retrieve-replace
--retrieve-replace-force	不涉及	布尔型	该参数用于在覆盖时跳过版本校验。若未指定该参数，在执行覆盖clog/csnlog时会校验备份集与当前集群的版本号，无法进行跨版本覆盖。	不涉及	False	--retrieve-replace-force

版权所有 © 华为云计算技术有限公司

完整选项	缩写选项	数据类型	说明	取值范围	缺省值	应用实例
--retrieve-replace-filelist	不涉及	字符串	该参数用于指定包含了本次覆盖操作想要覆盖的文件列表的json文件所在的绝对路径。 json文件示例： [{"pg_clog": ["000000000096","000000000097",{"255": ["00000000000000","000000000000001"]}]}, {"pg_csnlog": ["000000000000"]}, {"pg_xlog": [{"00000002": ["00000001000000000000000027", "00000001000000000000000028", "00000001000000000000000029"]},{"00000009": ["000000010000000000000000D3"]}]}] 其中xlog文件名中包含一级子目录用于指示term值，clog文件名中包含一级子目录用于指示创建hashbucket表后拆分的pg_clog下的目录名，也兼容没有bucket目录的clog，需要注意不同目录层级下的clog文件名长度不同。以上子目录及文件名均在第一步找回后由命令行返回，用户可自行选择覆盖的文件列表。	不涉及	不涉及	--retrieve-replace-filelist /data/filelist.json
--retrieve-clean	不涉及	布尔型	该参数用于指定本次操作为清理操作。不能与--retrieve-replace同时传入。	不涉及	False	--retrieve-clean

使用指南

- 对于clog/csnlog日志找回，必须指定--backup-key、--media-destination、--metadata-destination，否则Roach工具会报错。

版权所有 © 华为云计算技术有限公司

- 对于xlog日志找回，必须指定--task id、--retrieve-start-lsn、--retrieve-end-lsn，否则Roach工具会报错。
- 如果进行了xlog替换，需要进行一次全备以避免后续由于cbm文件损坏而导致的增备失败。

命令示例

- 从OBS找回clog/csnlog
  python3 GaussRoach.py **-t retrieve --media-type** *OBS* **--media-destination** */home/userA/media* **--backup-key** *20160126_165029* **--metadata-destination** */home/userA/metadata* **--retrieve-type** *clog,csnlog* **--retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
- 从NAS找回clog/csnlog
  python3 GaussRoach.py **-t retrieve --media-type** *NAS* **--media-destination** */home/userA/media* **--backup-key** *20160126_165029* **--metadata-destination** */home/userA/metadata* **--retrieve-type** *clog,csnlog* **--retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--cluster-unique-id** dws-xlnobs_9732b696-4b6d-4844-a40c-d1f6355aca02
- 从OBS找回xlog
  python3 GaussRoach.py **-t retrieve --media-type** *OBS* **--retrieve-type** *xlog* **--retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--obs-server-ip** *x.x.x.x* **--bucket-name** rdsbucket.resource.user4444 **--task-id** *GaussKernel_archive* **--retrieve-start-lsn** *0/40472D8* **--retrieve-end-lsn** *0/41531D8*
- 从NAS找回xlog
  python3 GaussRoach.py **-t retrieve --media-type** *NAS* **--media-destination** */home/userA/media* **--retrieve-type** *xlog* **--retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--task-id** *GaussKernel_archive* **--retrieve-start-lsn** *0/40472D8* **--retrieve-end-lsn** *0/41531D8*
- 将找回的日志覆盖至数据库实例数据目录
  python3 GaussRoach.py **-t retrieve --retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--retrieve-replace --retrieve-replace-force --retrieve-replace-filelist** */data/filelist.json*
- 清理日志找回的本地暂存路径
  python3 GaussRoach.py **-t retrieve --retrieve-inst** *dn6001* **--retrieve-local-path** */data/retrieve* **--retrieve-clean**

4.1.2.15 Roach 操作参数

表 4-13 Roach 功能和参数对应表

参数	备份	恢复	删除	Genstack	停止	找回clog/csnlog	找回xlog
backup-key	-	√	√	-	-	√	-
buffer-block-size	√	√	-	-	-	√	-
buffer-size	√	√	-	-	-	√	-
clean	-	√	-	-	-	-	-
compression-level	√	-	-	-	-	-	-
cpu-relinquish-size	√	√	-	-	-	-	-
cpu-relinquish-time	√	√	-	-	-	-	-

版权所有 © 华为云计算技术有限公司

参数	备份	恢复	删除	Genstack	停止	找回clog/csnlog	找回xlog
failure-retry-count	√	√	-	-	-	-	-
filesplit-size	√	-	-	-	-	-	-
log-filecount	√	√	-	-	-	-	-
log-filesize	√	√	-	-	-	-	-
logging	√	√	-	-	-	√	√
logging-level	√	√	-	-	-	√	√
logging-path	√	√	-	-	√	√	√
master-ip	√	√	-	-	-	-	-
master-port	√	√	√	-	-	-	-
max-memory-usage	√	√	-	-	-	-	-
media-destination	√	√	√	-	-	√	-
media-type	√	√	√	-	-	√	√
metadata-destination	√	√	√	-	-	√	√
parallel-process	√	√	-	-	-	√	√
parallel-restore	-	√	-	-	-	-	-
pre-disk-space	√	√	-	-	-	-	-
resource-retry-count	√	√	-	-	-	-	-
restore-buffer-threshold	-	√	-	-	-	-	-
restore-configs	-	√	-	-	-	-	-
restore-new-cluster	-	√	-	-	-	-	-
retry-wait-time	√	√	-	-	-	-	-
stop-timeout	-	-	-	-	√	-	-
username	√	-	-	-	-	-	-
verbose	√	√	-	-	-	-	-
disable-ssl	√	√	√	-	-	-	-

版权所有 © 华为云计算技术有限公司

参数	备份	恢复	删除	Genstack	停止	找回clog/csnlog	找回xlog
remote-ip	√	√	-	-	-	-	-
remote-port	√	√	-	-	-	-	-
remote-config	√	√	-	-	-	-	-
--gbr-db-list	-	√	-	-	-	-	-
--gbr-db-remap	-	√	-	-	-	-	-
--gbr-table-list	√	√	-	-	-	-	-
--aux-db-path	-	√	-	-	-	-	-
--on-demand-tsp-location	-	√	-	-	-	-	-
--origin-cluster	-	√	-	-	-	-	-
--gbr-owner	-	√	-	-	-	-	-
--retrieve-type	-	-	-	-	-	√	√
--retrieve-inst	-	-	-	-	-	√	√
--retrieve-local-path	-	-	-	-	-	√	√
--retrieve-replace	-	-	-	-	-	√	√
--retrieve-replace-filelist	-	-	-	-	-	√	√
--retrieve-replace-force	-	-	-	-	-	√	√
--retrieve-clean	-	-	-	-	-	√	√
--retrieve-start-lsn	-	-	-	-	-	-	√
--retrieve-end-lsn	-	-	-	-	-	-	√

4.2 gs_blackbox

背景信息

gs_blackbox是用于处理黑匣子文件的工具，支持dump黑匣子文件以及解析黑匣子文件，帮助诊断关键信息。

版权所有 © 华为云计算技术有限公司

说明

- 数据库生成的黑匣子文件的目录在“/dev/shm/node_port/node_name/gs_blackbox”下，其中node_port是数据库节点的端口号，node_name是数据库节点名称。
- 通过dump操作生成的数据库日志文件的目录在“$GAUSSLOG/gs_blackbox/node_name”下。

语法

```
gs_blackbox [dump|show] -f file
gs_blackbox --help
```

参数说明

表 4-14 gs_blackbox 参数说明

参数	说明
dump	将指定的黑匣子文件dump到数据库日志目录下。
show	解析黑匣子文件，并将内容展示到终端。
-f file	指定需要dump或者解析的黑匣子文件路径。
--help	查看黑匣子工具帮助信息。

示例

示例一：dump黑匣子文件

```
gs_blackbox dump -f gs_blackbox-2024-04-30_153026.box
[SUCCESS] Dump completed.
```

示例二：解析黑匣子文件

```
gs_blackbox show -f gs_blackbox-2024-04-30_153026.box
errnum          139721504749312     2024-04-30 15:30:26 CST.970    54047
errnum          139721504749312     2024-04-30 15:30:26 CST.970    54047
errnum          139721504749312     2024-04-30 15:30:26 CST.970    54047
errnum          139721504749312     2024-04-30 15:30:26 CST.970    13398
errnum          139721504749312     2024-04-30 15:30:26 CST.970    25965
errnum          139721504749312     2024-04-30 15:30:26 CST.013    13398
delete_memcxt   139721504749312     2024-04-30 15:30:26 CST.135    TZParserMe
```

4.3 gs_cgroup

背景信息

数据库实例环境下做批量任务处理时，因为多任务复杂性，会导致不同机器间的负载差距很大。为了充分利用数据库实例资源，负载管理变得尤为重要。gs_cgroup是GaussDB提供的负载管理工具。负责创建默认控制组、创建用户自定义控制组、删除用户自定义控制组、更新用户自定义组的资源配额和资源限额、显示控制组配置文件内容、显示控制组树形结构和删除用户的所有控制组。

gs_cgroup工具为使用数据库的操作系统用户创建Cgroups配置文件，并且在操作系统中生成用户设定的Cgroups。同时为用户提供增加、删除Cgroups，更新Cgroups资源

版权所有 © 华为云计算技术有限公司

配额，设定Cgroups的CPU或IO限额，设定异常处理阈值及相应操作等服务。此工具只负责当前操作系统节点的Cgroups操作，使用时需在各个节点上调用相同命令语句进行统一配置。

这里假设读者已经了解了负载管理的相关原理。

说明

docker模式下不支持gs_cgroup进行资源管理。docker模式配置详情查看《安装指南》中“环境准备”章节。

使用示例

- 使用普通用户或数据库管理员执行命令。
  a. 前置条件：需设置GAUSSHOME环境变量为数据库安装目录；且root用户已创建普通用户默认的控制组。
  b. 创建控制组及设置对应的资源配额，以便在数据库中运行作业时，指定到此控制组，使用此控制组管理的资源；通常数据库管理员为每个数据库用户创建Class组。
     i. 创建Class控制组和Workload控制组。
     gs_cgroup -c -S *class1* **-s** *40*
     创建当前用户新的Class Cgroups命名为“class1”，资源配额为总Class的40%。
     gs_cgroup -c -S *class1* **-G** *grp1* **-g** *20*
     创建当前用户新的class1 Cgroups下属的Workload控制组，命名为“grp1”，资源配额为class1 Cgroups的20%。
     ii. 删除grp1控制组和Class控制组。
     gs_cgroup -d -S *class1* **-G** *grp1*
     删除当前用户已建的“grp1”Cgroups。
     gs_cgroup -d -S *class1*
     删除当前用户已建的“class1”Cgroups。

     须知

     如果删除Class控制组，则Workload控制组也被删除。
  c. 更新已创建控制组的资源配额。
     i. 更新动态资源配额。
     gs_cgroup -u -S *class1* **-G** *grp1* **-g** *30*
     更新当前用户的class1 Cgroups下属grp1 Cgroups资源配额为class1 Cgroups的30%。
     ii. 更新限制资源配额。
     gs_cgroup --fixed -u -S *class1* **-G** *grp1* **-g** *30*
     更新当前用户的class1 Cgroups下属grp1 Cgroups限制使用CPU核数范围占上一级class1可用核数的30%。
  d. 更新Gaussdb的CPU核数范围。
     gs_cgroup -u -T *Gaussdb* **-f** *0-20*
     更新Gaussdb进程使用的CPU核数为0-20。

版权所有 © 华为云计算技术有限公司

说明

-f参数只适用于对Gaussdb设置核数范围。对于其他各控制组，如需设置核数范围，需要使用--fixed参数

e. 设置异常处理信息（class:wg组需存在）。

i. 设置组class:wg下的作业阻塞到1200秒或执行2400秒时执行终止动作：
```
gs_cgroup -S class -G wg -E "blocktime=1200,elapsedtime=2400" -a
```

ii. 设置组class:wg下的作业下盘数据量达到256MB或大表广播数据量达到100MB时执行终止动作：
```
gs_cgroup -S class -G wg -E "spillsize=256,broadcastsize=100" -a
```

iii. 设置组Class下的作业在所有DN上CPU总时间到达100s时执行降级动作：
```
gs_cgroup -S class -E "allcputime=100" --penalty
```

iv. 设置组Class下的作业在所有DN上执行时间到达2400秒，倾斜率达到90时执行降级动作：
```
gs_cgroup -S class -E "qualificationtime=2400,cpuskewpercent=90"
```

须知

给控制组设置异常处理信息，需要确保对应的控制组已经创建。指定多个阈值时用“，”分隔，不指定任何动作时默认为“penalty”操作。

f. 设置控制组使用的核数范围。

设置组class:wg的核数范围占Class控制组的20%。
```
gs_cgroup -S class -G wg -g 20 --fixed -u
```

须知

Class或Workload核数范围必须通过指定--fixed参数设置。

g. 回退上一个步骤。
```
gs_cgroup --recover
```

说明

--recover只支持对Class控制组和Workload控制组的增删改操作进行回退，且只支持回退一次操作。

h. 查看已创建的控制组信息。

i. 查看配置文件中控制组信息。
```
gs_cgroup -p
```

控制组配置信息

```
gs_cgroup -p

Top Group information is listed:
GID:   0 Type: Top    Percent(%): 1000( 50) Name: Root                Cores: 0-47
GID:   1 Type: Top    Percent(%):  833( 83) Name: Gaussdb:omm          Cores: 0-20
GID:   2 Type: Top    Percent(%):  333( 40) Name: Backend             Cores: 0-20
GID:   3 Type: Top    Percent(%):  499( 60) Name: Class               Cores: 0-20

Backend Group information is listed:
GID:   4 Type: BAKWD  Name: DefaultBackend   TopGID:   2 Percent(%): 266(80) Cores:
0-20
GID:   5 Type: BAKWD  Name: Vacuum          TopGID:   2 Percent(%):  66(20) Cores: 0-20
```


版权所有 © 华为云计算技术有限公司

```
Class Group information is listed:
GID:  20 Type: CLASS  Name: DefaultClass     TopGID:   3 Percent(%): 166(20) MaxLevel: 1
RemPCT: 100 Cores: 0-20
GID:  21 Type: CLASS  Name: class1           TopGID:   3 Percent(%): 332(40) MaxLevel: 2
RemPCT:  70 Cores: 0-20

Workload Group information is listed:
GID:  86 Type: DEFWD  Name: grp1:2          ClsGID:  21 Percent(%):  99(30) WDLevel:  2
Quota(%): 30 Cores: 0-5

Timeshare Group information is listed:
GID: 724 Type: TSWD   Name: Low             Rate: 1
GID: 725 Type: TSWD   Name: Medium            Rate: 2
GID: 726 Type: TSWD   Name: High            Rate: 4
GID: 727 Type: TSWD   Name: Rush            Rate: 8

Group Exception information is listed:
GID:  20 Type: EXCEPTION Class: DefaultClass
PENALTY: QualificationTime=1800 CPUSkewPercent=30

GID:  21 Type: EXCEPTION Class: class1
PENALTY: AllCpuTime=100 QualificationTime=2400 CPUSkewPercent=90

GID:  86 Type: EXCEPTION Group: class1:grp1:2
ABORT: BlockTime=1200 ElapsedTime=2400
```

上述示例查看到的控制组配置信息如**表4-15**所示。

表 4-15 控制组配置信息

GID	类型	名称	Percent（%）信息	特定信息
0	Top控制组	Root	1000代表总的系统资源为1000份。 括号中的50代表IO资源的50%。 GaussDB不通过控制组对IO资源做控制，因此下面其他控制组信息中仅涉及CPU配额情况。	-
1		Gaussdb:omm	系统中只运行一套数据库程序，Gaussdb:omm控制组默认配额为833，数据库程序和非数据库程序的比值为（833:167=5:1）。	-
2		Backend	Backend和Class括号中的40和60，代表Backend占用Gaussdb:dbuser控制组40%的资源，Class占用Gaussdb:dbuser控制组60%的资源。	-
3		Class		-

版权所有 © 华为云计算技术有限公司

GID	类型	名称	Percent（%）信息	特定信息
4	Backend控制组	Default Backend	括号中的80和20代表DefaultBackend和Vacuum占用Backend控制组80%和20%的资源。	TopGID：代表Top类型控制组中Backend组的GID，即2。
5		Vacuum		
20	Class控制组	Default Class	DefaultClass和class1的20和40代表占Class控制组20%和40%的资源。因为当前只有两个Class组，所有它们按照20:40的比例分配Class控制组499的系统配额，则分别为166和332。	• TopGID：代表DefaultClass和class1所属的上层控制（Top控制组中的Class组）的GID，即3。 • MaxLevel：Class组当前含有的Workload组的最大层次，DefaultClass没有Workload Cgroup，其数值为1。 • RemPCT:代表Class组分配Workload组后剩余的资源百分比。如class1中剩余的百分比为70。
21		class1		

版权所有 © 华为云计算技术有限公司

GID	类型	名称	Percent（%）信息	特定信息
86	Workload控制组	grp1:2（该名称由Workload Cgroup Name和其在class中的层级组成，它是class1的第一个Workload组，层级为2，每个Class组最多10层Workload Cgroup。）	根据设置，其占class1的百分比为30，则为332*30%=99。	• ClsGID：代表Workload控制组所属的上层控制组（class1控制组）的GID。 • WDLevel：代表当前Workload Cgroup在对应的Class组所在的层次。
724	Timeshare控制组	Low	-	Rate：代表Timeshare中的分配比例，Low最少为1，Rush最高为8。这四个Timeshare组的资源配比为Rush:High:Medium:Low=8:4:2:1
725		Medium	-	
726		High	-	
727		Rush	-	

ii. 查看操作系统中树形结构的控制组信息。

gs_cgroup -P显示控制组树形结构信息，其中shares代表操作系统中CPU资源的动态资源配额“cpu.shares”的数值，cpus代表操作系统中CPUSET资源的动态资源限额“cpuset.cpus”的数值，指的是该控制组能够使用的核数范围。

```
gs_cgroup -P
Mount Information:
cpu:/dev/cgroup/cpu
cpuset:/dev/cgroup/cpuset
cpuacct:/dev/cgroup/cpuacct

Group Tree Information:
- Gaussdb:wangrui (shares: 5120, cpus: 0-20, weight: 1000)
        - Backend (shares: 4096, cpus: 0-20, weight: 400)
                - Vacuum (shares: 2048, cpus: 0-20, weight: 200)
                - DefaultBackend (shares: 8192, cpus: 0-20, weight: 800)
```


版权所有 © 华为云计算技术有限公司

```
- Class (shares: 6144, cpus: 0-20, weight: 600)
        - class1 (shares: 4096, cpus: 0-20, weight: 400)
                - RemainWD:1 (shares: 1000, cpus: 0-20, weight: 100)
                        - RemainWD:2 (shares: 7000, cpus: 0-20, weight: 700)
                                - Timeshare (shares: 1024, cpus: 0-20, weight: 500)
                                        - Rush (shares: 8192, cpus: 0-20, weight: 800)
                                        - High (shares: 4096, cpus: 0-20, weight: 400)
                                        - Medium (shares: 2048, cpus: 0-20, weight: 200)
                                        - Low (shares: 1024, cpus: 0-20, weight: 100)
                        - grp1:2 (shares: 3000, cpus: 0-5, weight: 300)
                - TopWD:1 (shares: 9000, cpus: 0-20, weight: 900)
        - DefaultClass (shares: 2048, cpus: 0-20, weight: 200)
                - RemainWD:1 (shares: 1000, cpus: 0-20, weight: 100)
                        - Timeshare (shares: 1024, cpus: 0-20, weight: 500)
                                - Rush (shares: 8192, cpus: 0-20, weight: 800)
                                - High (shares: 4096, cpus: 0-20, weight: 400)
                                - Medium (shares: 2048, cpus: 0-20, weight: 200)
                                - Low (shares: 1024, cpus: 0-20, weight: 100)
                - TopWD:1 (shares: 9000, cpus: 0-20, weight: 900)
```

参数说明

- -a [--abort]

  对满足设定的异常阈值的作业执行终止动作。

- -b pct

  Backend Cgroups占用Top Backend资源的百分比，需同时指定“-B backendname”参数。

  取值范围为1 ~ 99。在不进行设置的情况下，默认CPU配额设置为Vacuum控制组占20%，DefaultBackend控制组占%80。Vacuum和DefaultBackend控制组配额之和设置应小于100%。

- -B name

  Backend Cgroups名称，仅可指定“-u”参数来更新此Cgroups的资源配额。

  -b percent和-B backendname参数共同作用来控制数据库后备线程的资源比例。

  取值范围：字符串，最长为64个字节。

- -c

  创建Cgroups并指定其标识。

  普通用户指定“-c”和“-S classname”可以创建新的Class Cgroups，如果再指定“-G groupname”则创建属于Class Cgroups的Workload Cgroups，新的Workload Cgroups位于Class Cgroups的最底层（4层是最底层）。

- -d

  删除Cgroups及标识。

  普通用户指定“-d”和“-S classname”可以删除已有的Class Cgroups，如果再指定“-G groupname”则删除属于Class Cgroups的Workload Cgroups，并将附属的线程重置到“DefaultClass:DefaultWD:1”Cgroups中；若删除的Workload Cgroups位于Class Cgroups的高层（1层是最高层），则需调整低层的Cgroups的层级，重建新的Cgroups附属的线程加载到新的Cgroups。

- -E data

  设定异常阈值，目前阈值包括：blocktime、elapsedtime、allcputime、spillsize、broadcastsize以及qualificationtime和cpuskewpercent，指定多个阈值时用“，”分隔。参数值0表示取消设置，设置不合法的值时会提示。

 版权所有 © 华为云计算技术有限公司

表 4-16 异常阈值类型

异常阈值类型	说明	取值范围（0表示取消设置）	支持的异常操作
blocktime	作业的阻塞时间，单位秒。包括全局并发排队以及局部并发排队的总时间。	0~UINT_MAX	abort
elapsedtime	作业的已被执行时间，单位秒。从开始执行到当前所消耗的时间。	0~UINT_MAX	abort
allcputime	作业在所有DN上执行时所耗费的CPU总时间，单位秒。	0~UINT_MAX	abort，penalty
cpuskewpercent	作业在DN上执行时的CPU时间的倾斜率，依赖于qualificationtime的设置。	0~100	abort，penalty
qualificationtime	检查作业执行cpu倾斜率的间隔时间，单位秒，需同cpuskewpercent一起设置。	0~UINT_MAX	none
spillsize	作业在DN上下盘的数据量，单位MB。	0~UINT_MAX	abort
broadcastsize	作业在DN上算子大表广播数据量，单位MB。	0~UINT_MAX	abort

- -h [--help]

  显示命令帮助信息。

- -H

  用户环境中$GAUSSHOME信息。

  取值范围：字符串，最长为1023个字符。

- -f

  设置GaussDB控制组使用的核数范围，范围必须是a-b或a的形式。其他控制组可以使用--fixed进行设置核数范围。

- --fixed

  设置控制组使用的核数范围比例占上一层级的百分比或者设置IO资源。

  设置核数范围比例时--fixed设置核数范围与'-s' '-g' '-t' '-b' 一起使用。

  核数比例范围0-100，同一层级的核数比例总和小于或者等于100，0代表核数与上一层级相同，对于所有的控制组，CPU限额默认设置为0。-f和--fixed不能同时设置。设置--fixed之后，-f设置的范围自动失效。设置的比例以quota值在-p中进行显示。

  设置I/O资源配额时，与'-R' '-r' '-W' '-w'一起使用。

- -g pct

  指定Workload Cgroups的资源占用“Class”Cgroups资源的百分比，需同时指定“-G groupname”参数；用于创建“-c”或更新“-u”Workload Cgroups。

  取值范围为1 ~ 99，默认Workload控制组CPU配额设置为20%，各Workload控制组配额之和应小于99%。

版权所有 © 华为云计算技术有限公司

- -G name

  指定Workload Cgroups的名称，需同时指定“-S classname”参数来表示该group属于哪一个Class Cgroups；可以连同“-c”参数创建新的Cgroups、“-d”参数删除此Cgroups及“-u”更新此Cgroups的资源配额；需要注意，此名称不可是Timeshare Cgroups的默认名称，如“Low”、“Medium”、“High”或“Rush”。

  如果用户自己创建Workload控制组，指定名称时不能指定带 ':' 的名称。不能创建同名控制组。

  取值范围：字符串，最长为28个字节。

- -N [--group] name

  可以将组名简写成class:wg。

- -p

  显示Cgroups配置文件的信息。

- -P

  显示Cgroups树形结构信息。

- --penalty

  对满足设定的异常阈值的作业执行降级动作，如果没有设定任何操作，则该操作将为默认操作。

- --recover

  仅用于回退Class控制组和Workload控制组的增删改操作，且只能回退一步。

- --revert

  恢复控制组为默认状态。

- -D mpoint

  指定的挂载点，默认的挂载点"/dev/cgroup/subsystem"。

- -m

  挂载cgroup。

- -M

  取消挂载cgroup。

- -U

  数据库用户名称。

- --refresh

  刷新控制组状态。

- -s pct

  指定Class Cgroups的资源占用“Top Class”Cgroups资源的百分比，需同时指定“-S classname”参数；用于创建“-c”或更新“-u”Class Cgroups。

  取值范围为1-99。默认Class控制组的CPU配额设置为20%。新创建的Class控制组的CPU配额以及默认的DefaultClass之和应小于100%。

- -S name

  指定Class group的名称；可以连同“-c”参数创建新的Cgroups、“-d”参数删除此Cgroups及“-u”更新此Cgroups的资源配额。创建子Class名称不能带“：”。

  取值范围：字符串，最长为31个字节。

 版权所有 © 华为云计算技术有限公司

- -t percent

  指定Top Cgroups（Root、Gaussdb:omm、Backend和Class Cgroups）占用资源的百分比，需同时指定“-T name”参数。若指定“-T Root”所在的Cgroups，其在Cgroups配置文件中显示的名称为“Root”，此数值代表的含义是blkio.weight值的百分比，最小值为10%，CPU资源配额如cpu.shares的数值不可修改；若指定“Gaussdb:omm” Cgroups，则表示占用整个系统CPU资源的百分比（可根据Root Cgroups的默认CPU配额1024得出该组的cpu.shares数值，此数值默认当前系统仅含有1套数据库环境），对于IO资源配额为1000，不再变化;若指定“Class”或“Backend”Cgroups，则表示资源占用“Gaussdb”Cgroups资源的百分比。

  取值范围为1 ~ 99。默认Class控制组配额为60%， Backend控制组配额为40%。修改Class控制组配额，同时会自动更新Backend控制组配额，使两者之和保持100%。

- -T name

  指定Top Cgroups的名称。

  取值范围：字符串，最长为64个字节。

- -u

  更新Cgroups。

- -V [--version]

  显示gs_cgroup工具的版本信息。

4.4 gs_check

说明

系统内部工具，不支持单独调用。

背景信息

gs_check用来检查数据库状态，统一当前系统中存在的各种检查工具，帮助用户在数据库运行过程中，全量的检查数据库运行环境，操作系统环境，网络环境及数据库执行环境，也有助于在数据库重大操作之前对各类环境进行全面检查，有效保证操作执行成功。

注意事项

- 必须指定-i或-e参数，-i会检查指定的单项，-e会检查对应场景配置中的多项。
- 黑名单检查项仅适用于R5升级到R6。
- MTU值不一致时可能导致检查缓慢或进程停止响应，当巡检工具出现提示时请修改各节点MTU值一致后再进行巡检。
- 交换机不支持当前设置的MTU值时，即使MTU值一致也会出现通信问题引起进程停止响应，需要根据交换机调整MTU大小。

语法

- 单项检查

```
gs_check -i ITEM [...] [-U USER] [-L] [-l LOGFILE] [-o OUTPUTDIR] [--skip-root-items] [--set] [--routing] [--cloud-mode] [--time-out=<SECS>] [--max-thread=NUMBERS] [--server-ip=IP] [--client-ip=IP] [--fio-ip=IP] [--fio-disk-path=PATH]
```

 版权所有 © 华为云计算技术有限公司

- 场景检查
  ```
  gs_check -e SCENE_NAME [-U USER] [-L] [-l LOGFILE] [-o OUTPUTDIR] [--skip-root-items] [--set] [--time-out=<SECS>] [--routing] [--skip-items] [--cloud-mode]
  ```
- 显示帮助信息
  ```
  gs_check -? | --help
  ```
- 显示版本号信息
  ```
  gs_check -V | --version
  ```
- 接收数据库管理员密码
  ```
  echo "password" | gs_check -i ITEM [...] [-U USER] [-L] [-l LOGFILE] [-o OUTPUTDIR] [--skip-root-items][--set][--routing][--cloud-mode][--time-out=<SECS>][--max-thread=NUMBERS][--server-ip=IP] [--client-ip=IP][--fio-ip=IP][--fio-disk-path=PATH]
  echo "password" | gs_check -e SCENE_NAME [-U USER] [-L] [-l LOGFILE] [-o OUTPUTDIR] [--skip-root-items] [--time-out=<SECS>][--set][--routing][--skip-items][--cloud-mode]
  ```

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统， 可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

- -U

  运行数据库的用户名称。

  取值范围：运行数据库的用户名称。
- -L

  本地执行
- -i

  指定检查项。格式-i CheckXX详细的检查项请参见**表4-17**。
- -e

  场景检查项。默认列表包括：inspect（例行巡检）、upgrade（升级前巡检）、expand(扩容前巡检)、binary_upgrade（就地升级前巡检）、health（健康检查巡检）、cloud_inspect（云上环境日常巡检）、replace（节点替换）、longtime（耗时长巡检）、install（安装），用户可以根据需求自己编写场景。
- -l

  指定日志文件路径，文件路径仅能以“.log”结尾。
- -o

  指定检查结果输出文件夹路径。
- --skip-items

  跳过指定的检查项。
- --format

  设置结果报告的格式，仅支持default和json。
- --set

  修复支持设置的Abnormal项。
- --cid

  检查ID仅仅被内部check进程所使用。
- --skip-root-items

  跳过具有特殊权限执行的检查项。

 版权所有 © 华为云计算技术有限公司

- --time-out

  设置超时时间。单位为秒，默认为1500s，若用户自定义超时时间不得少于1500s。

- --routing

  指定业务IP的网段，格式为IP地址：子网掩码。

- --disk-threshold="PERCENT"

  检查磁盘占用时可选指定告警阈值，可指定1-99之间的整数，不输入则默认为90。检查其他项时不需要该参数。

- --nodegroup-name

  指定缩容磁盘空间检查时执行缩容的逻辑数据库名称，仅在-i CheckSpaceForShrink时有效。

- --ShrinkNodes

  指定缩容磁盘空间检查时缩容节点名称，仅在-i CheckSpaceForShrink时有效。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

- --max-thread

  输入为数字，设置单次并发最大线程数，默认为50，仅在-i CheckFioBenchmark时有效。

- --server-ip

  输入为节点ip，指定ip来运行iperf server端线程，需与--client-ip一起使用，仅在-i CheckIperfOnce时生效。

- --client-ip

  输入为节点ip，指定ip来运行iperf client端线程，与iperf sever端进行交互，需与--server-ip一起使用，仅在-i CheckIperfOnce时生效。

- --fio-ip

  输入为节点ip，指定ip来进行基于fio的I/O性能测试，需与--fio-disk-path一起使用，仅在-i CheckFioOnce时生效。

- --fio-disk-path

  输入为磁盘路径，指定磁盘来进行基于fio的I/O性能测试，需与--fio-ip一起使用，仅在-i CheckFioOnce时生效。

表 4-17 数据库状态检查表

状态	巡检项	检查内容	是否支持--set
os	CheckCPU	检查主机CPU占用率，如果idle大于30%并且iowait小于30%。则检查项通过，否则检查项不通过。	否
	CheckTimeZone	检查数据库内各节点时区，如果时区一致则检查通过，否则检查项不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckSysParams	检查各节点操作系统参数，判断是否等于预期值。检查项不等于建议值则报warning，不等于指定值则检查项不通过，并打印不满足项。麒麟环境不支持此检查项。欧拉2.10环境不支持此检查项。 详见**操作系统参数**。	否
	CheckOSVer	检查数据库内各个节点的操作系统版本信息，如果满足版本兼容列表且全数据库在同一混搭列表中则检查通过，否则检查不通过。	否
	CheckNTPD	检查系统NTPD服务，如果服务开启且各节点时间误差在1分钟以内则检查项通过，否则检查项不通过。麒麟环境不支持此检查项。	否
	CheckTHP	检查系统THP服务，如果服务开启则检查项通过，否则检查项不通过。	是
	CheckSshdService	检查系统是否存在sshd服务，若存在则检查项通过，否则检查项不通过。	否
	CheckCrondService	检查系统是否存在crontab服务，若存在则检查项通过，否则检查项不通过。	是
	CheckCrontabLeft	检查crontab是否残留Gauss相关信息，若无该信息则检查项通过，否则检查项不通过。	是
	CheckDirLeft	检查文件目录（ /opt/huawei/Bigdata/ ,/var/log/Bigdata/, /home/omm ）是否存在，（ 若mount目录包含此目录则忽略 ）若不存在则查项通过，否则检查项不通过。	是
	CheckProcessLeft	检查是否残留gaussdb和omm进程，若未残留则检查项通过，否则检查项不通过。	是
	CheckStack	检查栈深度，若各个节点不一致则报warning，若大于等于3072则检查项通过，否则不通过。	否
	CheckSysPortRange	检查ip_local_port_range系统参数范围，若范围在26000~65535则检查项通过，否则检查项不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckEtcHosts	检查/etc/hosts没有配置localhost检查项不通过，存在带有#GaussDB注释的映射则检查项不通过，相同IP不同hostname则检查项不通过，否则通过，若hostname相同，但ip不同检查项不通过。	否
	CheckCpuCount	检查CPU核心与可用CPU，如果不符检查项则不通过，相符但存在不可用信息则报Warning， 所有节点CPU信息不相同则检查项不通过。	否
	CheckHyperThread	检查超线程，若打开则检查项通过，否则检查项不通过（ARM平台不需要做超线程检查）。	否
	CheckMemInfo	检查各节点总内存大小是否一致，若检查结果一致，则检查项通过，否则报warning。	否
	CheckBalanceState	查询Balance状态，如果查询结果为Yes则检查项通过，如果不是Yes则报Warning，如果查询失败则检查项不通过。	否
	CheckCmserverStandby	如果数据库cm_server属性为Standby，则检查项通过，否则报出Warning。	否
	CheckMaxHandle	检查操作系统最大句柄值，如果该值大于等于1000000则检查项通过，否则检查项不通过。	否
	CheckKernelVer	检查各节点系统内核版本信息，如果版本信息一致则检查项通过，否则报Warning。	否
	CheckEncoding	检查数据库内各个节点的系统编码，如果编码一致则检查项通过，否则检查项不通过。	否
	CheckFilehandle	此检查项检查以下两项，两项都通过为通过，否则为不通过： ● 检查每个gaussdb进程打开的进程数是否超过80万，不超过则检查通过，否则检查不通过。 ● 检查是否有slave进程使用的句柄数超过master进程，如果没有则检查通过，否则检查不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckKeyProAdj	检查所有关键进程，如果所有关键进程的omm_adj值为0，则通过，否则不通过。	否
	CheckCompMem	检查各个组件内存使用是否正常。	否
device	CheckSwapMemory	检查交换内存和总内存大小，若检查结果为0则检查项通过，否则检查项报Warning大于总内存时检查项不通过。	否
	CheckLogicalBlock	检查磁盘逻辑块大小，若为512则检查项通过，否则检查项不通过（SSD场景不支持磁盘逻辑块检查）。	否
	CheckMaxAsyIOrequests	获取当前异步请求值，如果异步IO值大于（dn）*1048576并且大于104857600则检查项通过，否则检查项不通过。	否
	CheckDiskFormat	检查磁盘XFS格式信息，如果配置为'inode64'则检查项通过，否则报warning；如果节点上无XFS格式硬盘，检查时提示相应信息。	否
	CheckInodeUsage	新节点：所有磁盘 老节点：数据库路径（GAUSSHOME/PGHOST/GAUSSHOME/GAUSSLOG/tmp及实例目录） 检查以上指定目录使用率，如果使用率超过warning阈值（默认为60%）报warning超过NG阈值（默认为80%）则检查项不通过，否则通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckSpaceUsage	新节点：所有磁盘 老节点：数据库路径（GAUSSHOME/PGHOST/GAUSSHOME/GAUSSLOG/tmp及实例目录） 检查磁盘以上指定目录（目录列表）使用率，如果使用率超过warning阈值（默认为70%） 报warning超过NG阈值（默认为90%）则检查项不通过。数据库路径下检查GAUSSHOME/PGHOST/GPHOME/GAUSSLOG/tmp/data路径的剩余空间，不满足阈值则检查项不通过，否则通过。	否
	CheckSpaceForShrink	检查逻辑数据库将指定主机缩容掉后，剩余主机完成重分布后磁盘空间使用率，如果使用率超过阈值（默认70%）则检查项不通过，不能进行缩容。 **说明** • gs_shrink内部调用，不建议人工调用。 • 该检查需添加参数：[--nodegroup-name=LCGROUPNAME] [--ShrinkNodes= HOSTNAME]，host之间以逗号隔开。	否
	CheckDiskConfig	检查磁盘名大小挂载点是否一致，若一致则检查项通过，否则报warning。	否
	CheckXid	查询xid的数值，如果大于10亿报Warning，大于18亿则检查项不通过。	否
	CheckSysTabSize	如果每一块磁盘的剩余容量大于该磁盘上所有实例的系统表容量总和则检查项通过，否则检查项不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckDiskUsage	检查各DN/CMS/ETCD数据目录大小以及磁盘使用率, 如果磁盘使用率超过阈值则检查不通过。 ● DN NG=70; ● DN WARNING=60; ● CMS NG=50; ● CMS WARNING=40; ● ETCD NG=70; ● ETCD WARNING=60	否
	CheckProcMem	检查节点整体内存使用率是否超过percent_total，若超过则告警，否则继续检查各个gaussdb进程是否超过max_process_memory的percent_max，超过则检查不通过。	否
	CheckSlowDisk	如果不存在磁盘的IO平均服务时间低于普遍水平则检查项通过，否则检查项不通过。	否
	CheckInstallDiskUsage	检查磁盘系统安装目录空间和索引使用率，如果使用率低于90%则检查项通过，否则检查项不通过。	否
	CheckTmpDiskUsage	检查磁盘临时目录磁盘和索引使用率，如果使用率低于90%则检查项通过，否则检查项不通过。	否
	CheckIperfBenchmark	检查数据库中两两节点间的网络性能，若执行成功会输出iperf done，性能结果会保存在gs_check日志目录下的iperf文件夹中。 **说明** ● 该检查需添加参数[-L]。 ● 如果数据库节点较多，需要增加[--time-out=]参数，增加超时时间。	否
	CheckFioBenchmark	检查数据库中节点上数据盘的I/O性能，若执行成功会输出fio done，性能结果会保存在gs_check日志目录下的fio文件夹中。 **说明** ● 该检查需添加参数[-L]。 ● 如果数据库节点较多，需要增加[--time-out=]参数，增加超时时间。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckIperfClean	清理数据库上残留的iperf线程，若执行成功会输出iperf_clean done。 **说明** 该检查需添加参数[-L]。	否
	CheckFioClean	清理数据库上残留的fio线程以及测试产生的中间文件，若执行成功会输出fio_clean done。 **说明** 该检查需添加参数[-L]。	否
	CheckIperfOnce	基于iperf检查固定两节点间的网络性能，若执行成功会输出iperf done，性能结果会保存在gs_check日志目录下的iperf文件夹中，以clientip_serverip_once的方式存储。 **说明** • 该检查需添加参数[-L]。 • 该检查需要添加[--server-ip=]以及[--client-ip=]来指定server端ip和client端ip。	否
	CheckFioOnce	基于fio单次检查固定ip上固定磁盘的io性能，若执行成功会输出fio done，性能结果会保存在gs_check日志目录下的fio文件夹中，以"ip_rw”，"ip_randread”，"ip"_randwrite”的方式存储。 **说明** • 该检查需添加参数[-L]。 • 该检查需要添加[--fio-ip=]以及[--fio-disk-path=]来指定ip和磁盘。	否
cluster	CheckClusterState	检查fencedUDF状态，如果fencedUDF状态为down则报warning；检查数据库状态，如果数据库状态为Normal则检查项通过，否则检查项不通过。	否
	CheckDBParams	主DN实例检查共享缓冲区大小和最大连接数。 共享缓冲区需要大于128KB且大于shmmax且大于shmall*PAGESIZE 以上项完全满足则检查项通过，否则检查项不通过。	是

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckDebugSwitch	在各节点检查各实例的配置文件中log_min_messages参数的值，为空则认为是Warning。判断日志级别是否是warning，不是warning则报warning。	是
	CheckUpVer	检查数据库各个节点上升级包的版本，如果一致则检查项通过，否则检查项不通过。	否
	CheckDirPermissions	检查节点目录（实例Xlog路径、GAUSSHOME、GPHOME、PGHOST、GAUSSLOG）权限，如果目录有写入权限且不大于750则检查项通过，否则检查项不通过。	否
	CheckEnvProfile	检查节点环境变量（$GAUSSHOME、$LD_LIBRARY_PATH、$PATH），检查CMS/CMA/DN进程的环境变量。如果环境变量存在并配置正确，进程的环境变量存在则检查项通过，否则检查项不通过。	否
	CheckGaussVer	检查各个节点gaussdb版本是否一致，如果版本一致则检查项通过，否则检查项不通过。	否
	CheckPortRange	若ip_local_port_range的范围在阈值范围内（默认是26000-65535），并且实例端口不在ip_local_port_range范围内则检查项通过，否则检查项不通过。	否
	CheckCatchup	检查gaussdb进程堆栈是否能搜索到CatchupMain函数，若搜索不到则检查项通过，否则检查项不通过。	否
	CheckProcessStatus	检查'om_monitor', 'cm_agent', 'cm_server', 'gaussdb'进程属主，若不存在omm以外的属主则检查项通过，否则检查项不通过。	否
	CheckSpecialFile	检查tmp目录（PGHOST）、日志目录（GAUSSLOG）、data目录、程序目录（GAUSSHOME）下文件是否存在特殊字符以及非omm用户的文件，检查OM目录（GPHOME）是否存在特殊字符以及非omm用户的文件，非root:root用户的文件。若不存在则检查项通过，否则检查项不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckCollector	在output目录下查看信息收集是否成功，若收集成功则检查项通过，否则检查项不通过。	否
	CheckOMMonitor	检查各个节点om_monitor进程是否存在，若各节点都存在则检查项通过，否则检查项不通过。	是
	CheckLargeFile	检查各个DN目录是否存在超过4G的文件。任一DN目录及其子目录有超过4G的单个文件，则检查不通过，否则检查通过。	否
	CheckProStartTime	检查关键进程启动时间是否间隔超过5分钟，超过则检查不通过，否则检查通过。	否
	CheckDilateSysTab	检查系统表是否膨胀，若膨胀则不通过，否则检查通过。	是
	CheckErrorInLog	检查各个主机$GAUSSLOG是否有最近阈值(5天)内的ERROR日志，并提供最多10条ERROR日志信息，若有则检查不通过。	否
	CheckBackupResult	检查近两天的全量备份和增量备份是否正常，如果都正常则通过，否则不通过。	否
	CheckFDLeak	数据库各进程持有句柄数超过系统限制百分之六十，则报错；文件句柄泄漏存在但是少于10个报warning，超过10个报ng。	否
	CheckProStartTime	检查关键进程启动时间是否间隔超过一分钟，超过则检查不通过，否则检查通过。	否
	CheckClusterTopo	检查数据库拓扑，仅查询，供离线分析。	否
	CheckMpprcFile	检查是否存在对环境变量分离文件的改动，若存在则检查不通过，否则检查通过。	否
database	CheckLockNum	检查数据库锁数量，查询成功检查项通过。	否
	CheckBlackList	检查数据库黑名单语法，如果黑名单不存在则检查项通过，否则检查项不通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckCurConnCount	检查数据库连接数，如果连接数小于最大连接数的90%则检查项通过，否则检查项不通过。	否
	CheckCursorNum	检查数据库的游标数，检查成功则检查项通过，否则检查项不通过。	否
	CheckMaxDatanode	检查最大DN数，若最大DN数小于xml配置的节点数*DN数（默认值为90*5）报warning，否则检查项通过。	是
	CheckLockState	检查数据库是否被锁，若数据库被锁则不通过，否则检查项通过。	否
	CheckIdleSession	检查非空闲会话数，如果数量为0则检查项通过，否则检查项不通过。	否
	CheckDBConnection	检查能否连接数据库，如果连接成功则检查项通过，否则检查项不通过。	否
	CheckGUCConsistent	• 对于逻辑数据库，下面两项都一致时，检查通过，否则检查不通过。 – 检查逻辑数据库相同实例上不可修改的GUC参数的值是否一致。 – 检查逻辑数据库相同实例上可修改的GUC参数的值在逻辑数据库内部是否一致。 • 对于非逻辑数据库，数据库相同实例上的GUC参数的值一致则检查通过，否则检查不通过。	否
	CheckGUCValue	检查(max_connections + max_prepared_transactions) * max_locks_per_transaction的值，若该值大于等于1000000则检查项通过，否则检查项不通过。	是
	CheckCgroupTable	检查Cgroups系统表，若不存在用户新建的Cgroups则检查项通过，否则报warning。	是
	CheckPMKData	检查数据库PMK schema是否包含有异常数据，如果不存在异常数据则检查项通过，否则检查项不通过。	是

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckSysTable	检查系统表，检查成功则检查项通过。	否
	CheckSysTabSize	如果每一块磁盘的剩余容量大于该磁盘上所有实例的系统表容量总和则检查项通过，否则检查项不通过。	否
	CheckTableSpace	表空间路径和数据库路径之间不能存在嵌套且表空间路径相互不能存在嵌套，则检查项通过，否则检查项不通过。	否
	CheckTableSkew	若存在表在数据库各DN上的数据分布不均衡，且分布数据最多的DN比最少的DN所分布的数据多100000条以上，则检查不通过，否则检查通过。	否
	CheckDNSkew	检查DN级别的表倾斜数据，若分布数据最多的DN节点比分布数据最少的DN节点数据量高于5%，则检查不通过，否则检查通过。	否
	CheckUnAnalyzeTable	若存在未做analyze的表，并且表中至少包含一条数据，则检查不通过，否则检查通过。	是
	CheckCreateView	创建视图时，如果查询语句中含有子查询，并且子查询结果查询解析和重写之后存在别名重复，检查不通过，否则检查通过。	否
	CheckHashIndex	如果存在hash index则检查不通过，否则检查通过。	否
	CheckNextvalInDefault	检查Default表达式中是否包含nextval(sequence)，若包含则不通过，否则通过。	否
	CheckPoolerNum	检查Pooler使用量，若超过33000则检查不通过，超过28000报Warning，否则检查通过。	否
	CheckReturnType	检查用户自定义函数是否包含非法返回类型，若包含则检查不通过，否则检查通过。	否
	CheckSysadminUser	检查除数据库属主外是否存在数据库管理员用户，若存在则不通过，否则检查通过。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckTDDate	检查TD模式数据库下的orc表中是否包含date类型的列，若包含检查不通过，否则检查通过。	否
	CheckDropColumn	如果存在drop column的表，则检查不通过，否则检查通过。	否
	CheckDiskFailure	对数据库中的所有数据做全量查询，若存在查询错误则检查不通过，否则检查通过。	否
	CheckXlogPrune	若xlog的数量超过（checkpoint_segments * 4 + wal_keep_segments * 2 + 2），则不通过。	否
	CheckQuerySpillSize	若临时下盘目录空间查过阈值（1024MB），则不通过。	否
	CheckKeyDBTableSize	检查最大磁盘占有量的数据库（TOP 10），然后检查其中最大的表（TOP 5）。	否
	CheckSQLResponeTime	检查80%/95%的响应时间是否超过阈值（p80 - 800ms; p95 - 1000ms）。	否
	CheckThreadPool	检查线程池的使用情况。	否
	CheckBPHitRatio	检查buffer pool命中率是否超过阈值(95%)。	否
	CheckWorkloadTrx	检查workload是否有过高的事务回滚率(1%)。	否
	CheckKernelMemoryUsage	获取DN内存使用分布。	否
	CheckDBStat	收集数据库统计信息。	否
	CheckClusterStmt	收集归一化sql的信息。	否
	CheckSQLTypeDist	收集sql语句分布信息。	否
	CheckSlowQueryLog	检查最近阈值条(1000)的慢SQL日志。	否
	CheckVacuumStatus	若日志中近30分钟内未打印autovacuum相关日志，则检查不通过。	否
	CheckBiggestTable	对数据库中所有的表进行检查，收集各数据库中最大的表。	否

版权所有 © 华为云计算技术有限公司

状态	巡检项	检查内容	是否支持--set
	CheckIntegrity	检查各个节点gaussdb的sha256值，若各节点一致则检查项通过，否则检查项不通过。	否
	CheckLogDiskUsage	检查磁盘日志目录空间和索引使用率，如果使用率低于90%则检查项通过，否则检查项不通过。	否
	CheckDnWait	检查系统表中处于wait状态的实例，若存在处于wait状态的实例，将对应的节点以及实例写入结果对象。	否
network	CheckNetWorkDrop	检查各IP1分钟内网络掉包率，如果不超过1%则检查项通过，否则检查项不通过。	否
	CheckUsedPort	检查net.ipv4.ip_local_port_range，范围大于等于OS默认值通过（32768-61000）。 检查TCP协议随机端口数，小于总随机端口数的80%通过。	否
	CheckRouting	检查各节点在业务IP网段的IP个数，超过1个则报warning，否则检查通过。	否
	CheckNetSpeed	网络满载时，检查网卡平均接收带宽大于600MB通过; 网络满载时，检查网络ping值，小于1秒通过。 网络满载时，检查网卡丢包率，小于1%通过。	否
other	CheckDataDiskUsage	检查磁盘DN目录使用率，如果使用率低于90%则检查项通过，否则检查项不通过。	否
	CheckLogDiskUsage	检查磁盘日志目录空间和索引使用率，如果使用率低于90%则检查项通过，否则检查项不通过。	否
	CheckPstree	检查是否支持pstree。	否

用户自定义场景

步骤1 以操作系统用户omm登录数据库主节点。

步骤2 在$GPHOME/script/gspylib/inspection/config路径下新建场景配置文件scene_XXX.xml。

 版权所有 © 华为云计算技术有限公司

步骤3 将检查项写进场景配置文件中，书写格式为：

```
<?xml version="1.0" encoding="utf-8" ?>
<scene name="XXX" desc="check cluster parameters before XXX.">
<configuration/>
<allowitems>
<item name="CheckXXX"/>
<item name="CheckXXX"/>
</allowitems>
</scene>
```

item name为检查项名称

注：用户需自行保证自定义xml的正确性

步骤4 在$GPHOME/script/gspylib/inspection/config执行如下命令，将此文件分发至执行检查的各个节点

```
scp scene_upgrade.xml SIA1000068994：home/package/script/gspylib/inspection/config/
```

说明

$GPHOME/script/gspylib/inspection/config就是新建的场景配置文件的绝对路径。

步骤5 执行以下命令查看检查结果。

```
gs_check -e XXX
```

----结束

用户自定义检查项

步骤1 新增巡检项配置，修改script/gspylib/inspection/config/items.xml文件，格式如下：

```
<checkitem id="10010" name="CheckCPU">
<title>
<zh>检查CPU占用率</zh>
<en>Check CPU Idle and I/O wait</en>
</title>
<threshold>
StandardCPUIdle=30;
StandardWIO=30
</threshold>
<suggestion>
<zh>如果idle不足 CPU负载过高，请扩容节点，如果iowait过高，则磁盘为瓶颈，扩容磁盘</zh>
</suggestion>
<standard>
<zh>检查主机CPU占用率，如果idle大于30%并且iowait小于30%，则检查项通过，否则检查项不通过</zh>
</standard>
<category>os</category>
<permission>user</permission>
<scope>all</scope>
<analysis>default</analysis>
</checkitems>
```

- id：巡检项id。
- name：巡检项脚本名，和巡检项脚本文件名相同。
- title：巡检项描述名称（支持多语言）。

  <zh>：中文版检查内容。

  <en>：英文版检查内容。
- standard：巡检项标准说明（支持多语言）。
- suggestion：巡检项修复建议说明（支持多语言）。
- threshold：巡检项阈值定义，多值之间使用分号隔开，示例Key1=Value1;Key2=Value2。

 版权所有 © 华为云计算技术有限公司

- category：巡检项分类，可选参数：os、device、network、cluster、database、other。
- permission：巡检项需要的执行权限，可选参数：root，user默认为user(普通用户)。
- scope：巡检项执行的节点范围，可选参数：local-仅在当前节点执行，all-在数据库所有节点执行，默认为all。
- analysis：巡检项执行结果分析方式，default-检查每个节点的结果，所有节点检查项通过，则最终检查通过，consistent-数据库内所有节点一致性检查，单节点仅返回结果，各个节点结果一致则判定检查通过，custom-自定义结果分析方式，默认为default。

注：用户需保证自定义xml的正确性

步骤2 新建检查脚本，脚本名称格式遵循CheckXXXX.py，必须以Check开头，脚本放置在script/gspylib/inspection/items目录下，该目录下脚本安装巡检项分类组织，每个分类一个单独的文件夹，巡检项脚本放置在对应的分类文件夹中。格式如下：

```
class CheckCPU(BaseItem):
def __init__(self):
super(CheckCPU, self).__init__(self.__class__.__name__)
self.idle = None
self.wio = None
self.standard = None

def preCheck(self):
# check the threshold was set correctly
if (not self.threshold.has_key('StandardCPUIdle')
or not self.threshold.has_key('StandardWIO')):
raise Exception("threshold can not be empty")
self.idle = self.threshold['StandardCPUIdle']
self.wio = self.threshold['StandardWIO']

# format the standard by threshold
self.standard = self.standard.format(idle=self.idle, iowait=self.wio)

def doCheck(self):
cmd = "sar 1 5 2>&1"
output = SharedFuncs.runShellCmd(cmd)
self.result.raw = output
# check the result with threshold
d = next(n.split() for n in output.splitlines() if "Average" in n)
iowait = d[-3]
idle = d[-1]
rst = ResultStatus.OK
vals = []
if (iowait > self.wio):
rst = ResultStatus.NG
vals.append("The %s actual value %s is greater than expected value %s" % ("IOWait", iowait, self.wio))
if (idle < self.idle):
rst = ResultStatus.NG
vals.append("The %s actual value %s is less than expected value %s" % ("Idle", idle, self.idle))
self.result.rst = rst
if (vals):
self.result.val = "|n".join(vals)
```

所有脚本基于BaseItem基类开发，基类定义的通用的检查流程，通用的检查结果分析方法，默认的结果输出格式。可扩展方法：

- doCheck: 该方法包含该检查项具体的检查方法，检查结果格式如下：

  result.rst --- 检查结果状态，可选参数:
  - OK – 检查项完成，结果通过。
  - NA – 当前节点不涉及该检查项。

 版权所有 © 华为云计算技术有限公司

  - NG – 检查项完成，结果不通过。
  - WARNING – 检查项完成，结果警告。
  - ERROR – 检查项发生内部错误，未完成检查。
- preCheck: 检查前条件判定，localPreCheck – 用于检查当前执行节点是否指定节点。可通过巡检项配置文件中的scope参数进行配置。可重载该方法实现自定义的前置检查
- postAnalysis：检查结果分析方法，内置两种实现：default，consistent。可通过巡检项配置文件中的analysis参数进行配置。可重载该方法实现自定义的结果分析。

注：用户自定义的检查项名称不得与已有检查项名称相同，同时用户需保证自定义检查项脚本的规范性。

步骤3 将此脚本分发至所有的执行节点。

```
scp CheckXXXX.py  $GPHOME/script/gspylib/inspection/items
```

步骤4 以非root用户执行以下命令，查看结果。

本地执行：

```
gs_check -i CheckXXX  -L
```

非本地执行：

```
gs_check  -i  CheckXXX
```

----结束

操作系统参数

表 4-18 操作系统参数

参数名称	参数说明	推荐取值
net.ipv4.tcp_max_tw_buckets	表示同时保持TIME_WAIT状态的TCP/IP连接最大数量。如果超过所配置的取值，TIME_WAIT将立刻被释放并打印警告信息。	10000
net.ipv4.tcp_tw_reuse	允许将TIME-WAIT状态的sockets重新用于新的TCP连接。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv4.tcp_tw_recycle	表示开启TCP连接中TIME-WAIT状态sockets的快速回收。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv4.tcp_keepalive_time	表示当keepalive启用的时候，TCP发送keepalive消息的频度。	30

 版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
net.ipv4.tcp_keepalive_probes	在认定连接失效之前，发送TCP的keepalive探测包数量。这个值乘以tcp_keepalive_intvl之后决定了一个连接发送了keepalive之后可以有多少时间没有回应。	9
net.ipv4.tcp_keepalive_intvl	当探测没有确认时，重新发送探测的频度。	30
net.ipv4.tcp_retries1	在连接建立过程中TCP协议最大重试次数。	5
net.ipv4.tcp_syn_retries	TCP协议SYN报文最大重试次数。	5（该参数值控制数据库向已经建立连接的远程主机重新发送数据的最大次数，低值可以更早的检测到与远程主机失效的连接，因此服务器可以更快的释放该连接；如需减少网络故障场景超时时间，可设置该参数值为2）。
net.ipv4.tcp_synack_retries	TCP协议SYN应答报文最大重试次数。	5
net.ipv4.tcp_retries2	控制内核向已经建立连接的远程主机重新发送数据的次数，低值可以更早的检测到与远程主机失效的连接，因此服务器可以更快的释放该连接。 发生“connection reset by peer”时可以尝试调大该值规避问题。	12
vm.overcommit_memory	控制在做内存分配的时候，内核的检查方式。 ● 0：表示系统会尽量精确计算当前可用的内存。 ● 1：表示不作检查直接返回成功。 ● 2：内存总量× vm.overcommit_ratio/100 + SWAP的总量，如果申请空间超过此数值则返回失败。 内核默认是2过于保守，推荐设置为0，如果系统压力大可以设置为1。	0

版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
net.ipv4.tcp_rmem	TCP协议接收端缓冲区的可用内存大小。分无压力、有压力和压力大三个区间，单位为页面。	8192 250000 16777216
net.ipv4.tcp_wmem	TCP协议发送端缓冲区的可用内存大小。分无压力、有压力和压力大三个区间，单位为页面。	8192 250000 16777216
net.core.wmem_max	socket发送端缓冲区大小的最大值。	21299200
net.core.rmem_max	socket接收端缓冲区大小的最大值。	21299200
net.core.wmem_default	socket发送端缓冲区大小的默认值。	21299200
net.core.rmem_default	socket接收端缓冲区大小的默认值。	21299200
net.ipv4.ip_local_port_range	服务器可用临时端口范围。	26000-65535
kernel.sem	内核信号量参数设置大小。	250 6400000 1000 25600
vm.min_free_kbytes	保证物理内存有足够空闲空间，防止突发性换页。	系统总内存的5%
net.core.somaxconn	定义了系统中每一个端口最大的侦听队列的长度，这是个全局的参数。	65535
net.ipv4.tcp_syncookies	当出现SYN等待队列溢出时，启用cookies来处理，可防范少量SYN攻击。 ● 0表示关闭SYN Cookies。 ● 1表示开启SYN Cookies。	1
net.core.netdev_max_backlog	在每个网络接口接收数据包的速率比内核处理这些包的速率快时，允许送到队列的数据包的最大数目。	65535
net.ipv4.tcp_max_syn_backlog	记录的那些尚未收到客户端确认信息的连接请求的最大值。	65535
net.ipv4.tcp_fin_timeout	系统默认的超时时间。	60
kernel.shmall	内核可用的共享内存总量。	1152921504606846720
kernel.shmmax	内核参数定义单个共享内存段的最大值。	18446744073709551615

版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
net.ipv4.tcp_sack	启用有选择的应答，通过有选择地应答乱序接受到的报文来提高性能，让发送者只发送丢失的报文段（对于广域网来说）这个选项应该启用，但是会增加对CPU的占用。 ● 0表示关闭。 ● 1表示开启	1
net.ipv4.tcp_timestamps	TCP时间戳（会在TCP包头增加12节），以一种比重发超时更精确的方式（参考RFC 1323）来启用对RTT的计算，启用可以实现更好的性能。 ● 0表示关闭。 ● 1表示开启	1
vm.extfrag_threshold	系统内存不够用时，linux会为当前系统内存碎片情况打分，如果超过vm.extfrag_threshold的值，kswapd就会触发memory compaction。所以这个值设置的接近1000，说明系统在内存碎片的处理倾向于把旧的页换出，以符合申请的需要，而设置接近0，表示系统在内存碎片的处理倾向做memory compaction。	500
vm.overcommit_ratio	系统使用绝不过量使用内存的算法时，系统整个内存地址空间不得超过swap+RAM值的此参数百分比，当vm.overcommit_memory=2时此参数生效。	90
MTU	节点网卡最大传输单元。OS默认值为1500。	8192

文件系统参数

- soft nofile

  说明：soft nofile表示软限制，用户使用的文件句柄数量可以超过该限制，但是如果超过会有告警信息。

  推荐取值：1000000

- hard nofile

  说明：hard nofile表示硬限制，是一个严格的限制，用户使用的文件句柄数量一定不能超过该设置。

  推荐取值：1000000

- stack size

  说明：线程堆栈大小。

  推荐值：3072

版权所有 © 华为云计算技术有限公司

示例

执行单项检查结果：

```
perfadm@lfgp000700749:/opt/huawei/perfadm/tool/script> gs_check -i CheckCPU
Parsing the check items config file successfully
Distribute the context file to remote hosts successfully
Start to health check for the cluster. Total Items:1 Nodes:3

Checking...                [=========================] 1/1
Start to analysis the check result
CheckCPU.....................................OK
The item run on 3 nodes.  success: 3

Success. All check items run completed. Total:1  Success:1  Failed:0
For more information please refer to /opt/huawei/wisequery/script/gspylib/inspection/output/
CheckReport_201902193704661604.tar.gz
```

本地执行结果：

```
perfadm@lfgp000700749:/opt/huawei/perfadm/tool/script> gs_check -i CheckCPU -L

2017-12-29 17:09:29 [NAM] CheckCPU
2017-12-29 17:09:29 [STD] 检查主机CPU占用率，如果idle 大于30%并且iowait 小于 30%.则检查项通过，否则检查项不通过
2017-12-29 17:09:29 [RST] OK

2017-12-29 17:09:29 [RAW]
Linux 4.4.21-69-default (lfgp000700749)  12/29/17  _x86_64_

17:09:24       CPU     %user     %nice   %system   %iowait    %steal     %idle
17:09:25       all      0.25      0.00      0.25      0.00      0.00     99.50
17:09:26       all      0.25      0.00      0.13      0.00      0.00     99.62
17:09:27       all      0.25      0.00      0.25      0.13      0.00     99.37
17:09:28       all      0.38      0.00      0.25      0.00      0.13     99.25
17:09:29       all      1.00      0.00      0.88      0.00      0.00     98.12
Average:       all      0.43      0.00      0.35      0.03      0.03     99.17
```

4.5 gs_checkos

说明

系统内部工具，不支持单独调用。

背景信息

gs_checkos用于检查操作系统、控制参数、磁盘配置等内容，并对系统控制参数、I/O配置、网络配置和THP服务等信息进行配置。

前提条件

- 当前的硬件和网络环境正常。
- 各主机间root互信状态正常。
- 只能使用root用户执行gs_checkos命令。

说明

该工具不支持独立调用，出于安全考虑，前置完成后会自动删除。

版权所有 © 华为云计算技术有限公司

语法

- 检查操作系统信息
  gs_checkos -i *ITEM* [-f *HOSTFILE*] [-h *HOSTNAME*] [-X *XMLFILE*] [--detail] [-o *OUTPUT*] [-l *LOGFILE*]
- 显示帮助信息
  gs_checkos -? | --help
- 显示版本号信息
  gs_checkos -V | --version

参数说明

- -i

  列表编号。格式：-i A、-i B1、-i A1 -i A2或-i A1,A2。

  取值范围：A1...A14、B1...B6

  说明

  - A1...A14 表示只检查操作系统参数，并不进行参数设置。
  - B1...B6 表示将参数系统参数设置为期望值。
  - A和B不能同时输入。

  详细的检查项请参见**表4-19**。
- -f

  主机名称列表文件。

  说明

  -f和-h参数不能同时使用。

  取值范围：主机名称的列表。
- -h

  指定需要检查的主机名称，可以同时指定多个主机，主机之间使用“,”分割。

  取值范围：数据库实例的主机名称。如果不指定主机，则检查当前主机。
- -X

  数据库实例XML配置文件。
- --detail

  显示检查结果详情。
- -o

  指定OS检查报告输出到指定的文件。

  不指定则将检查结果输出到屏幕上。
- -l

  指定日志文件及存放路径。

  默认值：/tmp/gs_checkos/gs_checkos-YYYY-MM-DD_hhmmss.log。
- -?, --help

  显示帮助信息。
- -V, --version

  显示版本号信息。

 版权所有 © 华为云计算技术有限公司

表 4-19 操作系统检查项

编号	检查项	检查项详细说明
A1	OS version status	操作系统版本检查：保证主机平台是支持的SUSE11、SUSE12、RHEL和CentOS平台中的一种。同时保证数据库实例各个主机的平台是属于同一混编范围的。
A2	Kernel version status	内核版本检查。
A3	Unicode status	字符集设置检查：保证数据库实例各个主机的字符集一致，目前OM工具只支持操作系统默认utf-8的语言和编码。
A4	Time zone status	时区状态设置检查：保证数据库实例各个主机的时区状态一致。
A5	Swap memory status	Swap分区检查：各个主机的swap分区大小必须小于等于各主机内存总大小。
A6	System control parameters status	内核参数检查：保证当前内核中的参数值和默认的期望值保持一致。 详细的内核参数请参见**操作系统参数**。
A7	File system configuration status	系统文件句柄检查：系统文件句柄的参数值必须大于等于100万。 线程堆栈大小检查：线程堆栈大小大于等于3072KB。 进程最大可用虚拟内存：系统进程最大可用虚拟内存为unlimited。
A8	Disk configuration status	硬盘挂载参数检查：格式为xfs的硬盘，其挂载格式必须为："rw,noatime,inode64,allocsize=16m"。
A9	Pre-read block size status	预读参数检查：预读参数期望值为16384。
A10	IO scheduler status	IO调度策略检查：IO调度策略的方式为deadline。
A11	Network card configuration status	万兆以下网卡参数检查： • 确保网卡mtu=1500。 万兆及以上网卡参数检查： • 当网卡为非绑定模式，同时网卡类型为万兆时，确保mtu=1500，rx/tx >=4096。 • 当网卡为绑定模式时，需保证绑定的每个网卡类型均为万兆，然后再确保mtu=1500，rx/tx>=4096。 • 数据库实例周边设备（例如交换机等）的网卡也要设置成与数据库实例部署的主机一致的值。

版权所有 © 华为云计算技术有限公司

编号	检查项	检查项详细说明
A12	Time consistency status	时间一致性检查：确认ntp服务启动，数据库实例各个主机的系统时间误差不超过60s。
A13	Firewall service status	防火墙检查： 确认防火墙的状态为关闭状态。
A14	THP service status	THP服务检查：确认THP服务状态为关闭状态。
B1	Set system control parameters	内核参数设置：当实际值不满足检查条件期望时，对结果为Abnormal项参数进行修改设置；对Warning项不进行设置，由用户根据实际环境设置。
B2	Set file system configuration value	系统文件句柄设置：当系统文件句柄的参数值小于100万时，对其进行修改设置。 线程堆栈大小设置：线程堆栈大小小于3072KB时，对其进行修改设置。 进程最大可用虚拟内存设置：系统进程最大可用虚拟内存不为unlimited时，对其进行修改设置。
B3	Set pre-read block size value	硬盘预读参数设置：当系统的实际值小于16384时，对其进行修改设置。
B4	Set IO scheduler value	I/O配置项的设置：当系统的实际值不为deadline时，对其进行修改设置。
B5	Set network card configuration value	万兆网卡参数设置：对不满足万兆网卡的RX、TX参数进行设置，不对MTU参数进行设置。
B6	Set THP service	THP服务设置：如果THP服务启动，将该服务关闭。

说明

- A6选项检查标准来自配置文件check_list.conf下[/etc/sysctl.conf]、[SUGGEST:/etc/sysctl.conf]域：

  若[/etc/sysctl.conf]下参数值与系统实际参数值不同，A6检查提示Abnormal，可以使用B1参数进行设置。

  若[SUGGEST:/etc/sysctl.conf]下参数值与系统实际参数值不同，A6检查提示Warning，B1参数不会进行设置，需根据实际情况进行手动设置。
- A7选项检查系统文件句柄标准来自配置文件check_list.conf下[/etc/security/limits.conf]域open file参数，可以使用B2参数进行设置。
- A11选项检查万兆网卡参数"MTU"、"RX"、"TX"标准来自配置文件check_list.conf下[/sbin/ifconfig]域，可以使用B5参数对RX、TX进行设置，MTU需用户手动设置。
- 使用gs_checkos设置的内核参数值和文件句柄参数值，需重启新会话生效。
- CCE容器不支持进入容器后配置系统文件句柄设置以及CCE容器部分内核配置文件不存在，因此docker模式下默认不执行B2项配置系统文件句柄，不执行A6项检查容器系统控制参数。

版权所有 © 华为云计算技术有限公司

操作系统参数

表 4-20 操作系统参数

参数名称	参数说明	推荐取值
net.ipv4.tcp_max_tw_buckets	表示同时保持TIME_WAIT状态的TCP/IP连接最大数量。如果超过所配置的取值，TIME_WAIT将立刻被释放并打印警告信息。	10000
net.ipv4.tcp_tw_reuse	允许将TIME-WAIT状态的sockets重新用于新的TCP连接。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv4.tcp_tw_recycle	表示开启TCP连接中TIME-WAIT状态sockets的快速回收。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv4.tcp_keepalive_time	表示当keepalive启用的时候，TCP发送keepalive消息的频度。	30
net.ipv4.tcp_keepalive_probes	在认定连接失效之前，发送TCP的keepalive探测包数量。这个值乘以tcp_keepalive_intvl之后决定了一个连接发送了keepalive之后可以有多少时间没有回应。	9
net.ipv4.tcp_keepalive_intvl	当探测没有确认时，重新发送探测的频度。	30
net.ipv4.tcp_retries1	在连接建立过程中TCP协议最大重试次数。	5
net.ipv4.tcp_syn_retries	TCP协议SYN报文最大重试次数。	5（该参数值控制数据库向已经建立连接的远程主机重新发送数据的最大次数，低值可以更早的检测到与远程主机失效的连接，因此服务器可以更快的释放该连接；如需减少网络故障场景超时时间，可设置该参数值为2）。
net.ipv4.tcp_synack_retries	TCP协议SYN应答报文最大重试次数。	5

版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
net.ipv4.tcp_retries2	控制内核向已经建立连接的远程主机重新发送数据的次数，低值可以更早的检测到与远程主机失效的连接，因此服务器可以更快的释放该连接。 发生“connection reset by peer”时可以尝试调大该值规避问题。	12
vm.overcommit_memory	控制在做内存分配的时候，内核的检查方式。 ● 0：表示系统会尽量精确计算当前可用的内存。 ● 1：表示不作检查直接返回成功。 ● 2：内存总量× vm.overcommit_ratio/100＋SWAP的总量，如果申请空间超过此数值则返回失败。 内核默认是2过于保守，推荐设置为0，如果系统压力大可以设置为1。	0
net.ipv4.tcp_rmem	TCP协议接收端缓冲区的可用内存大小。分无压力、有压力和压力大三个区间，单位为页面。	8192，250000，16777216
net.ipv4.tcp_wmem	TCP协议发送端缓冲区的可用内存大小。分无压力、有压力和压力大三个区间，单位为页面。	8192，250000，16777216
net.core.wmem_max	socket发送端缓冲区大小的最大值。	21299200
net.core.rmem_max	socket接收端缓冲区大小的最大值。	21299200
net.core.wmem_default	socket发送端缓冲区大小的默认值。	21299200
net.core.rmem_default	socket接收端缓冲区大小的默认值。	21299200
net.ipv4.ip_local_port_range	服务器可用临时端口范围。	26000-65535
kernel.sem	内核信号量参数设置大小。	250，6400000，1000，25600
vm.min_free_kbytes	保证物理内存有足够空闲空间，防止突发性换页。	系统总内存的5%
net.core.somaxconn	定义了系统中每一个端口最大的侦听队列的长度，这是个全局的参数。	65535

版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
net.ipv4.tcp_syncookies	当出现SYN等待队列溢出时，启用cookies来处理，可防范少量SYN攻击。 ● 0表示关闭SYN Cookies。 ● 1表示开启SYN Cookies。	1
net.core.netdev_max_backlog	在每个网络接口接收数据包的速率比内核处理这些包的速率快时，允许送到队列的数据包的最大数目。	65535
net.ipv4.tcp_max_syn_backlog	记录的那些尚未收到客户端确认信息的连接请求的最大值。	65535
net.ipv4.tcp_fin_timeout	系统默认的超时时间。	60
kernel.shmall	内核可用的共享内存总量。	1152921504606846720
kernel.shmmax	内核参数定义单个共享内存段的最大值。	18446744073709551615
net.ipv4.tcp_sack	是否启用有选择的应答，通过有选择地应答乱序接受到的报文来提高性能，让发送者只发送丢失的报文段（对于广域网来说）。启用该功能可以提高性能，但是会增加对CPU的占用，建议启用该功能。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv4.tcp_timestamps	是否启用TCP时间戳（会在TCP包头增加12节），以一种比重发超时更精确的方式（参考RFC 1323）来启用对RTT的计算，启用该功能可以提高性能。 ● 0表示关闭。 ● 1表示开启。	1
net.ipv6.conf.all.ndisc_notify	定义地址和设备更改的通知模式。 ● 0表示不开启通知。 ● 1表示当设备启动或硬件地址更改时，生成未经请求的邻居通告。	1
vm.extfrag_threshold	系统内存不够用时，linux会为当前系统内存碎片情况打分，如果超过vm.extfrag_threshold的值，kswapd就会触发memory compaction。如果该参数的值接近1000，说明系统在内存碎片的处理上倾向于把旧的页换出，以符合申请的需要；而接近0，表示系统在内存碎片的处理倾向于做memory compaction。	500

版权所有 © 华为云计算技术有限公司

参数名称	参数说明	推荐取值
vm.overcommit_ratio	系统使用绝不过量使用内存的算法时，系统整个内存地址空间不得超过内存总量×vm.overcommit_ratio/100 + SWAP的总量，当vm.overcommit_memory=2时此参数生效。	90
MTU	节点网卡最大传输单元。OS默认值为1500。	8192

文件系统参数

- soft nofile

  说明：soft nofile表示软限制，用户使用的文件句柄数量可以超过该限制，但是如果超过会有告警信息。

  推荐取值：1000000

- hard nofile

  说明：hard nofile表示硬限制，是一个严格的限制，用户使用的文件句柄数量一定不能超过该设置。

  推荐取值：1000000

- stack size

  说明：线程堆栈大小。

  推荐值：3072

示例

使用如下命令检查操作系统参数。

```
gs_checkos -i A -h plat1 -X /opt/software/GaussDB_Kernel/clusterconfig.xml --detail -o /var/log/checkos
Performing operation system check/set. Output the result to the file /var/log/checkos.
Operation system check/set is completed.
Total numbers:14. Abnormal numbers:0. Warning number:1.
```

查看操作系统参数检查结果。

```
vim /var/log/checkos
Checking items:
   A1. [ OS version status ]                    : Normal
   A2. [ Kernel version status ]                : Normal
   A3. [ Unicode status ]                       : Normal
   A4. [ Time zone status ]                     : Normal
   A5. [ Swap memory status ]                   : Normal
   A6. [ System control parameters status ]     : Normal
   A7. [ File system configuration status ]     : Normal
   A8. [ Disk configuration status ]            : Normal
   A9. [ Pre-read block size status ]           : Normal
   A10.[ IO scheduler status ]                  : Normal
   A11.[ Network card configuration status ]    : Normal
   A12.[ Time consistency status ]              : Warning
   A13.[ Firewall service status ]              : Normal
   A14.[ THP service status ]                   : Normal
```

 版权所有 © 华为云计算技术有限公司

4.6 gs_collector

> **说明**
>
> 集中式单节点不支持该工具。

背景信息

当GaussDB数据库发生故障时，使用此工具收集数据库的OS信息、数据库的日志信息以及数据库配置文件等信息，来定位问题。可以使用-C参数，指定收集不同的信息内容，具体支持收集的内容信息如表4-21所示。

表 4-21 gs_collector 内容收集对照表

TypeName	Content	描述	是否默认收集
System	HardWareInfo RunTimeInfo	收集现在操作系统的状态信息： • HardWareInfo • Cpu（cat /proc/cpuinfo） • 内存（cat /proc/meminfo） • 磁盘（df -h） • RunTimeInfo： • topStatus（ps ux） • ioStat（iostat -xm 2 3） • 网络状态（cat /proc/net/dev） • 内存使用情况（free -m）	是
Database	具体列表见表8-1。	收集系统视图或者系统表的内容。	是，默认收集pg_locks，pg_stat_activity和pg_thread_wait_status三个视图的信息。

版权所有 © 华为云计算技术有限公司

TypeName	Content	描述	是否默认收集
Log	DataNode ClusterManager Roach Perf ffic dcf gns FeatureCollect Flamegraph	• Datanode收集： – gs_log – asp_data – pg_perf • ClusterManager收集： cm目录： – cm_agent – cm_server – om_monitor – cm_ctl bin目录： – gs_ctl – gs_guc – gs_initdb – gs_encrypt – gs_obs om目录 • ffic收集 ffic_log目录。 • dcf收集 dcf_log目录。 • gns收集 gns目录。 • Flamegraph收集： gs_flamegraph目录 • Memory收集： mem_log目录 • Feature收集： feature_collect目录	是

 版权所有 © 华为云计算技术有限公司

TypeName	Content	描述	是否默认收集
Config	DataNode	● 收集Datanode的配置信息: – gaussdb.conf – gaussdb.state – gs_hba.conf – pg_control – pg_replslot – gs_ident.conf ● 当开启多租特性功能后，会额外收集配置信息: – pdb.conf – pg_pdb_control	是
Gstack	DataNode	利用gstack命令得到DN的当前堆栈信息。	否
CoreDump	gaussdb cm_server cm_agent gs_ctl gs_rewind gaussdb_stack gs_rewind_stack cm_server_stack cm_agent_stack gs_ctl_stack	通过时间筛选gaussdb进程的core文件。 **注意** 关于Core的配置和方式请按照如下要求：在/proc/sys/kernel/core_pattern文件中添加core文件的格式：core-%e-%p-%t。	否
XLog	DataNode	通过时间筛选收集xLog。	否
ForceFailover	DataNode	收集DN的强起日志，以及截断的xLog日志。	否
Plan	*	收集计划复现信息： ● 用户输入*表示收集所有数据库上的计划复现信息。 ● 用户输入具体的数据库名称表示收集指定数据库上的计划复现信息。	否

版权所有 © 华为云计算技术有限公司

前提条件

日志收集工具依赖操作系统工具，例如gstack为其中一种，如果未安装该工具，则提示错误后，跳过该收集项。

语法

- 日志收集（非root用户）
  ```
  gs_collector --begin-time="BEGINTIME" --end-time="ENDTIME" [-h HOSTNAME | -f HOSTFILE] [--keyword=KEYWORD] [--speed-limit=SPEED] [-o OUTPUT] [-l LOGFILE] [-C CONFIGFILE]
  ```
- 显示帮助信息
  ```
  gs_collector -? | --help
  ```
- 显示版本号信息
  ```
  gs_collector -V | --version
  ```
- 接收数据库管理员密码
  ```
  echo "password" |  gs_collector --begin-time="BEGINTIME" --end-time="ENDTIME" [-h HOSTNAME | -f HOSTFILE] [--keyword=KEYWORD] [--speed-limit=SPEED] [-o OUTPUT] [-l LOGFILE] [-C CONFIGFILE]
  ```

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

- -h

  收集信息的主机名称。

  取值范围：主机名称，只能指定一个主机名称。

  如果不指定则默认收集所有主机的信息。

- -f

  主机名称列表文件。该文件为纯文本格式。

  说明

  -f和-h参数不能同时使用。

  取值范围：主机名称列表。

- -o

  将收集日志以压缩包形式输出到指定的文件夹。

  不指定则将检查结果以压缩包形式输出到数据库的数据库日志目录。

  即检查结果以压缩包形式输出到$GAUSSLOG目录中。

- -l

  指定的日志文件以及存放路径。

- -C

  指定待收集内容的配置文件，利用TypeName指定需要收集的信息类型，利用Content指定每一类信息的具体内容，利用Count指定此类信息收集的次数，利用Interval指定收集间隔，单位为秒。

  TypeName和Content不允许缺失或者内容为空。

  Interval和Count可以不指定，如果没有指定Count，则默认收集一次。如果没有指定Interval则表示间隔为0秒，Interval和Count的值不能小于0。

版权所有 © 华为云计算技术有限公司

如果不指定则使用默认的配置文件。

配置文件格式采用json格式，模板如下：

```
{
   "Collect":
   [
      {"TypeName": "name", "Content":"value","Interval":"seconds", "Count":"count"}
   ]
}
```

说明

默认配置文件内容如下：

```
{
   "Collect":
   [
      {"TypeName": "System", "Content":"RunTimeInfo, HardWareInfo","Interval":"0",
"Count":"1"},
      {"TypeName": "Log", "Content" : "DataNode,ClusterManager,Roach,Perf,ffic,dcf,gns",
"Interval":"0", "Count":"1"},
      {"TypeName": "Database", "Content":
"pg_locks,pg_stat_activity,pg_thread_wait_status","Interval":"0", "Count":"1"},
      {"TypeName": "Config", "Content": "DataNode", "Interval":"0", "Count":"1"}
   ]
}
```

TypeName和对应的Content取值范围见表1gs_collector内容收集对照表。

对于Log，CoreDump，Config，XLog四种类型，Interval和Count参数不生效。

- --keyword=KEYWORD

  包含关键字KEYWORD的日志文件。

  若关键字KEYWORD中含有空格，需要使用双引号包围。

- --begin-time

  日志的开始时间。输入格式为“yyyymmdd hh:mm”。

- --end-time

  日志的结束时间。输入格式为“yyyymmdd hh:mm”。

- --speed-limit

  日志收集时的收集速率，输入格式为非负整数，单位为MB/s。

  该参数主要是为了防止日志收集过程中产生过高的磁盘或网络I/O，导致ETCD或者DN故障（如果它们与$GAUSSLOG/$PGHOST部署在同一个磁盘上）。该值应当不超过数据库内上述磁盘I/O与网络I/O速率的最小值的1/3。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

4.7 gs_encrypt

说明

集中式单节点不支持该工具。

背景信息

GaussDB提供了gs_encrypt工具为输入的明文字符串进行加密操作。

版权所有 © 华为云计算技术有限公司

语法

```
gs_encrypt [OPTION]... PLAINTEXT
```

说明

- “PLAINTEXT”前面不需要加短选项或长选项。
- “PLAINTEXT”是需要加密的明文字符串。

参数说明

OPTION的取值如下所示：

- -?, --help
  显示帮助信息。
- -V, --version
  显示版本号信息。
- -k *PASSWORD*， --key=*PASSWORD*
  用户输入的密码，用于派生出加密操作时使用的密钥。密码长度有效范围为8~128个字符，并且至少包含大写字母、小写字母、数字和特殊字符中的三种字符。
- -v *VectorValue*，--vector=*VectorValue*
  用户输入的盐值，计算密码单向哈希时用于对密码进行加扰。字符串长度必须等于16。
- -f *FilePrefix*, --file-prefix=*FilePrefix*
  加密操作时使用的cipher和rand文件前缀字符串。cipher和rand文件必须存放在$GAUSSHOME/bin目录下。
- -B *Value*，--key-base64=*Value*
  加密时使用的明文密钥，以base64编码。
- -D *Value*， --vector-base64=*Value*
  加密时使用的随机数，以base64编码。base64编码前的字符串长度必须为16。

  说明

  出于安全考虑，建议用户在使用gs_encrypt工具前设置HISTCONTROL=ignorespace，强制Linux不记录敏感历史命令，并在执行gs_encrypt命令行前面加空格。

示例

示例1：通过输入key和vector明文值来对明文字符串进行加密操作。由于每次加密时使用的IV值是随机生成的，因此使用同样的key和vector以及明文字符串加密后对应的密文串不相同。单次支持加密的明文字符串最大长度为335字节。

```
gs_encrypt -k password -v 1234567890123456 test_encrypt
```

命令回显：

```
G0R7Y2oROfuiEjWtUBx9+eHP3I21A1wMOJ2onVQIiAsxMjM0NTY3ODkwMTIzNDU2
```

示例2：已经存在cipher和rand文件，通过文件前缀来对明文字符串进行加密操作。

1. 生成cipher和rand文件
   ```
   gs_guc generate -S Mppdb@123 -D $GAUSSHOME/bin -o test
   ```

 版权所有 © 华为云计算技术有限公司

命令回显：

```
gs_guc encrypt -S ***
```

2. 执行加密操作

```
gs_encrypt -f test test_encrypt
```

命令回显：

```
tSWpR8aFyF4GHVIVO7QNXhEMNlTcqRMrk172nAf2hLgSQwhcFWyP7X0QVS0SONon
```

示例3：通过输入base64转码后的key和vector密文来对明文字符串进行加密操作。

1. 输入base64转码后的key

```
python
>>> import base64
>>> print base64.b64encode("XXXXXXXXX");
```

命令回显：

```
R2F1c3NAMTIz
```

2. 输入base64转码后vector值

```
>>> print base64.b64encode("1234567890abcdef");
```

命令回显：

```
MTIzNDU2Nzg5MGFiY2RlZg==
```

3. 执行加密操作

```
gs_encrypt -B R2F1c3NAMTIz -D MTIzNDU2Nzg5MGFiY2RlZg==  test_encrypt
```

命令回显：

```
V7mpnsfIHTjVTYGw1YHI0c6aY0N4J5L7BtU4CqDoeIkxMjM0NTY3ODkwYWJjZGVm
```

4.8 gs_guc

背景信息

目前GaussDB配置文件（“gaussdb.conf”、“gs_hba.conf”，“pdb.conf”）中的参数默认值都是单机的配置模式。应用程序可以通过调用gs_guc来设置适合自己的参数。

- “gs_guc-current.log”

  这个文件记录了gs_guc产生的日志。

  默认路径：$GAUSSLOG/bin/gs_guc。

- “gs_guc-年-月-日_HHMMSS.log”

  当“gs_guc-current.log”文件大小超过16MB后，会根据当前时刻生成一个备份文件。

- “server.key.cipher”，“server.key.rand”

  使用gs_guc encrypt命令加密用户密码时，如果指定的-M的值为server，则会生成这两个文件。其中“server.key.cipher”存储用户密码的密文，“server.key.rand”存储的是加密因子。

- “client.key.cipher”，“client.key.rand”

  使用gs_guc encrypt命令加密用户密码时，如果指定的-M的值为client，则会生成这两个文件。其中“client.key.cipher”存储用户密码的密文，“client.key.rand”存储的是加密因子。

 版权所有 © 华为云计算技术有限公司

说明

“client.key.cipher”和“client.key.rand”是不指定-U选项时生成的密文文件和加密因子文件。如果通过-U指定了用户名，则会生成以用户名开头的文件。以指定-U test为例，生成的文件名为:“test.key.cipher、test.key.rand”。

- “obsserver.key.cipher”，“obsserver.key.rand”

  使用gs_guc generate命令加密输入的密码时，则会生成这两个文件。其中“obsserver.key.cipher”存储输入密码的密文，“obsserver.key.rand”存储的是加密因子。

  说明

  “obsserver.key.cipher”和“obsserver.key.rand”是创建OBS服务对象时调用的密钥文件。

语法

- -Z为datanode或gnsserver
  - 检查配置文件中参数（仅支持DN）

    ```
    gs_guc check [-Z datanode] [-N NODE-NAME] [-I INSTANCE-NAME | -D <DATA_DIR>] -c "parameter"
    ```

  - 修改配置文件(或PDB配置文件)中参数

    ```
    gs_guc set [-Z datanode | -Z gnsserver][--pdbname pdbname] [-N NODE-NAME] [-I INSTANCE-NAME | -D <DATA_DIR>] -c "parameter = value"
    ```

  - 将已设置的参数值修改为内核(或PDB)默认值

    ```
    gs_guc [ set | reload ] [-Z datanode | -Z gnsserver][--pdbname pdbname] [-N NODE-NAME] [-I INSTANCE-NAME | -D <DATA_DIR>] -c "parameter"
    ```

  - 修改配置文件中参数，同时发送信号量到gaussdb.conf

    ```
    gs_guc reload [-Z datanode | -Z gnsserver] [-N NODE-NAME] [-I INSTANCE-NAME | -D <DATA_DIR>] -c parameter=value
    ```

  - 修改客户端认证策略。若选择reload会同时发送信号量到gs_hba.conf，即无需重启即可生效。

    ```
    gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME IPADDR IPMASK AUTHMEHOD authentication-options"
    gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME IPADDR-WITH-IPMASK AUTHMEHOD authentication-options"
    gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME HOSTNAME AUTHMEHOD authentication-options"
    ```


版权所有 © 华为云计算技术有限公司

说明

- AUTHMEHOD后的authentication-options为可选参数，AUTHMEHOD支持以下选项：
  - trust：不验密，禁止远程主机使用trust方式访问数据库。
  - reject：拒绝访问。
  - md5：md5认证，默认不支持（MD5加密算法安全性低，存在安全风险，不建议使用）。
  - sha256：sha256认证（推荐使用）。
  - sm3：sm3认证（国密SM3）。
  - cert：客户端证书认证。
  - gss：kerberos认证，仅用于内部节点间认证。
  - peer：获取客户端所在操作系统用户名，并检查与数据库初始用户名是否一致，若一致则通过认证。此方式只支持数据库初始用户通过local模式本地连接，并支持通过配置gs_ident.conf建立操作系统用户与数据库初始用户映射关系。
- gs_hba.conf中的认证策略越靠前优先级越高，使用gs_guc工具配置时会按一定规则排序将新策略插入到原有认证策略中。配置字段比较顺序为：IPADDR/HOSTNAME > HOSTTYPE > DATABASE > USERNAME，即优先比较IPADDR或HOSTNAME，如果无法区分优先级则继续比较HOSTTYPE，以此类推。对于每个配置字段，通常越严格的配置参数优先级越高、排序越靠前，越宽松的配置参数优先级越低、排序越靠后，具体如下：
  - IPADDR：当配置为全0时表示不限制IP，会放在指定具体某个IP地址的策略后面。
  - DATABASE：当配置为all时表示不限制数据库，会放在指定具体某个数据库的策略后面；当数据库配置为replication时会放在其他策略后面。
  - USERNAME：当配置为all时表示不限制用户，会放在指定具体某个用户的策略后面。
  - HOSTTYPE：local > hostssl > hostnossl > host。

- 注释已经设置的客户端认证策略。 若选择reload会同时发送信号量到gs_hba.conf，即无需重启即可生效。

```
gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME IPADDR IPMASK"
gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME IPADDR-WITH-IPMASK"
gs_guc {set | reload} -Z NODE-TYPE [-N NODE-NAME] {-I INSTANCE-NAME | -D <DATA_DIR>} -h "HOSTTYPE DATABASE USERNAME HOSTNAME"
```

- 显示帮助信息

```
gs_guc -? | --help
```

- 显示版本号信息

```
gs_guc -V | --version
```

- 生成加密密文文件和加密因子文件
  - gs_guc encrypt生成加密密码和加密因子文件

```
gs_guc encrypt [-M keymode] [-K password] [-U username] -D <DATA_DIR>
或
gs_guc encrypt [-M keymode] [-K password] [-U username] -C [cipher_dir] -R [rand_dir]
```

说明

- -K是用户指定的密码，gs_guc会对该密码进行长度（详情见**命令参考**）和密码复杂度校验，如果不满足，将会报错。如果不带-K参数则会提示交互式输入密码。为了系统安全，推荐使用交互式输入密码方式。此密码用于保证生成密码文件的安全性和唯一性，用户无需保存或记忆。
- -M是加密类型，当前仅支持server、client和source。默认值为server。

版权所有 © 华为云计算技术有限公司

- gs_guc generate生成obsserver或其他前缀的加密密码和加密因子文件
  ```
  gs_guc generate [-o prefix] [-S cipherkey] -D <DATA_DIR>
  ```

  说明

  - -o是输出cipher和rand文件前缀名称，默认输出文件名前缀为obsserver。其内容仅支持数字、字母和下划线。
  - -S是用户指定的密码，密码需要满足长度要求（8<=len<16）和复杂度要求，如不满足将会报错。当其值为default时，会随机生成一段字符串作为密码，该密码长度为13。如果不带-S参数则会提示交互式输入密码。为了系统安全，推荐使用交互式输入密码方式。此密码用于保证生成密码文件的安全性和唯一性，用户无需保存或记忆。
  - 使用gs_guc encrypt或generate命令生成加密密码和加密因子文件时只是参数不同，本质上是一样的。生成过程中会使用随机数作为加密密钥材料和盐值，因此是每次生成的文件都是不同的。每次生成的加密密码和加密因子文件需要成对使用，不能更换或交替使用，加密和解密时需要使用相同的加密密码和加密因子文件。

说明

- gs_guc工具不支持参数值中包含'#'，'`'的设置。可以使用vi工具通过手工修改配置文件来设置。
- 如果已经在环境变量中设置PGDATA，则可以省略-D参数。否则提示设置参数失败。
- 环境变量PGDATA设置方法：先将参数PGDATA设置为*/gaussdb/data*，然后在执行“export PGDATA”使设置生效。
- authpolicy包含一串认证参数：HOSTTYPE DATABASE USERNAME IPADDR IPMASK，或者HOSTTYPE DATABASE USERNAME IPADDR-WITH-IPMASK，或者HOSTTYPE DATABASE USERNAME HOSTNAME。
- 如果设置GUC参数时使用-c "parameter"，则会将已设置的GUC参数值设置成该参数的数据库启动时参数默认值（系统视图PG_SETTINGS中boot_val字段值）。注意log_directory和audit_directory不会被设置为数据库启动时参数默认值，而是设为$GAUSSLOG/pg_audit/instance_name。因GUC参数间存在依赖关系，因此请慎用该功能。
- 设置-c参数时，参数都可以省略双引号。
- 如果value中含有特殊字符（如$），请转义后使用。
- 如果同一个配置参数在配置文件里面出现多行，且有两行或多于两行同时生效（即没有用"#"注释掉），那么只有最后一个配置参数会被设置，而前面的都会被忽略。
- 通过reload模式设置或修改数据库节点配置文件（gaussdb.conf）的参数，生效存在短暂延迟，有可能导致配置后数据库各实例参数极短时间不一致。
- gs_guc设置浮点类型的参数时，由于浮点数在计算机中不能精确表示，所以参数值在误差范围1e-9内都可以设置成功。
- gs_guc设置整型类型的参数时，可以接受十进制、十六进制、八进制的数据，以0x开始表示十六进制，以0开始表示八进制，其他情况表示十进制。
- gs_guc设置参数时，不支持关联参数校验。
- gs_guc只支持设置$GAUSSHOME/bin/cluster_guc.conf文件中的参数。
- 推荐使用gs_guc设置参数。不建议手动配置gaussdb.conf文件，因为过程中如果出现错误，没有校验，可能导致数据库重启失败。
- gs_guc设置string参数，参数中的单引号会计算字符数。

命令参考

- set

  表示只修改配置文件中的参数。
- check

  表示只检查配置文件中的参数。

 版权所有 © 华为云计算技术有限公司

说明

check结果中出现*parameter*=NULL，表示配置文件中没有配置此参数。

- reload

  表示修改配置文件中的参数，同时发送信号量给数据库进程，使其重新加载配置文件。

- encrypt/generate

  生成加密密码文件。

- -N

  需要设置的主机名称。

  取值范围：已有主机名称。

  当参数取值为ALL时，表示设置数据库中所有的主机。

  暂不支持一次设置中指定多个-N参数，指定多个-N会报错。

- -I INSTANCE-NAME

  需要设置的实例名称。

  取值范围：已有实例名称。

  当参数取值为ALL时，表示设置主机中所有的实例。

  暂不支持一次设置中指定多个-I参数，指定多个-I会报错。

- -D

  需要执行命令的数据库实例路径。使用encrypt命令时，此参数表示指定的密码文件生成的路径。

  说明

  - 数据库实例路径-D与实例名称"-I" 不能同时使用

- -C

  使用encrypt命令时，明文密钥文件生成路径。

- -R

  使用encrypt命令时，随机数密钥文件生成路径。

- -c *parameter*=*value*

  要设定的数据库配置参数的名称和参数值。

  说明

  - 如果参数是一个字符串变量，则使用-c *parameter*="'*value*'"或者使用-c "*parameter* = '*value*'"。
  - 当使用gs_guc set/reload为"log_directory" 恢复默认值时，其默认值会被置为具体的data目录。
  - 当参数是一个字符串变量，gs_guc工具不会进行合法性校验，如果配置参数不合法导致数据库运行异常，可以查看gs_log日志进行问题排查。

  取值范围：gaussdb.conf中的所有参数。

- --ignore-node=NODES

  需要忽略的主机名称。

 版权所有 © 华为云计算技术有限公司

说明

- 该参数必须与set/reload一起使用，且-Z只支持datanode。
- 该参数不支持与-D一起使用。
- 在与reload一起使用时，如果--ignore-node没有指定主节点，则数据库实例中所有节点的参数依然会全部同步修改。
- --ignore-node必须在-N all时才可生效。

- -c parameter

  当进行check操作时，表示需要检查的参数名称；当进行set/reload操作时，表示需要恢复为数据库默认值的参数名称。

- -h host-auth-policy

  指定需要在“gs_hba.conf”增加的客户端认证策略。

  取值范围：

  – HOSTTYPE DATABASE USERNAME IPADDR IPMASK [authmehod-options]
  – HOSTTYPE DATABASE USERNAME IPADDR-WITH-IPMASK [authmehod-options]
  – HOSTTYPE DATABASE USERNAME HOSTNAME [authmehod-options]

  HOSTTYPE是必输参数，取值：

  – local
  – host
  – hostssl
  – hostnossl

  其中“local”是一个Unix域套接字，“host”是普通或SSL加密的TCP/IP套接字，“hostssl”是SSL加密的TCP/IP套接字，“hostnossl”是一个纯TCP/IP套接字。

  authmehod-options是可选输入，取值：

  – trust
  – reject
  – md5（MD5加密算法安全性低，存在安全风险，不建议使用。）
  – sha256
  – cert
  – gss

  详细的参数说明请参见“gs_hba.conf”配置文件中的描述。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本信息。

- -Z NODE_TYPE

  实例类型。

  取值范围：

  – datanode
  – cmserver
  – cmagent

 版权所有 © 华为云计算技术有限公司

  - gnsserver

- -M, --keymode=MODE

  设定该密码在数据库运行过程中，用于服务端(server)或客户端(client)调用。

  说明

  在使用ssl功能时，会涉及到服务端证书和私钥文件的加密密码和客户端证书和私钥文件的加密密码，在加密存储加密密码时，需要指定模式，否则默认是生成服务端的密码文件。

  取值范围：

  - server（缺省）
  - client
  - source
  - tde：仅用于透明数据加密特性场景。该模式加密-K中的参数并保存在$GAUSSHOME/gs_security/tde.key.cipher与$GAUSSHOME/gs_security/tde.key.rand文件中。

- -K PASSWORD

  gs_guc encrypt指定需要加密的密码字符串。如果不加此参数则会提示交互式输入，为了系统安全，推荐使用交互式输入密码方式。

  不同加密类型（MODE）取值范围如表4-22所示：

  表 4-22 -K 参数取值范围

-M 取值	密码长度要求	密码复杂度要求
client, server	[8, 128]字节	符合复杂度要求
source	[8, 15]字节	符合复杂度要求
tde	[8, 256]字节	无复杂度要求

  须知

  密码长度兼容问题：

  旧版gs_guc（GaussDB Kernel 505.2.0之前的版本）支持加密[8, 15]字节的密码，新版gs_guc（GaussDB Kernel 505.2.0及之后的版本）支持加密[8, 128]字节的密码，新版和旧版加密后部分场景文件格式不一致。

  兼容场景如下：

  1. 旧版gs_guc加密的密码，新版数据库可解密。
  2. 新版gs_guc加密[8, 15]字节的密码，旧版数据库可解密。
  3. 新版gs_guc加密[16, 128]字节的密码，旧版数据库无法解密。该场景下，建议重新使用新版或旧版gs_guc加密[8, 15]字节的密码，供旧版数据库使用。

- -U, --keyuser=USER

  指定要加密的用户，隶属于OS用户。

版权所有 © 华为云计算技术有限公司

说明

比如，GaussDB可以为每个用户配置不同的ssl证书和私钥文件，指定该选项，可以生成按用户名区分的密码文件。

- -S CIPHERKEY

  gs_guc generate指定需要加密的字符串。如果不加此参数则会提示交互式输入，为了系统安全，推荐使用交互式输入密码方式。

  取值范围：字符串，符合密码长度（8~15）和复杂度的要求。当其值为default时，会随机生成一段字符串作为密码，该密码长度为13。

- -o PREFIX

  指定需要输出的OBS类型的cipher和rand文件前缀名称字符串，默认输出文件名前缀为obsserver。

  取值范围：仅支持数字、字母和下划线

- --base64

  参数值为base64格式。

- --pdbname

  参数值为需要设置参数的PDB的名称，仅支持set/check命令、集中式DN节点。

示例

示例1：修改数据库允许的最大连接数为800。修改后需要重启数据库才能生效。

```
gs_guc set -Z datanode -D /gaussdb/data/datanode -c "max_connections = 800" -N plat1

Total instances: 1. Failed instances: 0.
Success to perform gs_guc!
```

示例2：将已设置的数据库允许的最大连接数值修改为默认值。修改后需要重启数据库才能生效。

```
gs_guc set -Z datanode -D /gaussdb/data/datanode -c "max_connections" -N plat1

Total instances: 1. Failed instances: 0.
Success to perform gs_guc!
```

示例3：设置客户端认证策略。

```
gs_guc set -Z datanode -N all -I all -h "host replication testuser 10.252.95.191/32 sha256"

Total instances: 2. Failed instances: 0.
Success to perform gs_guc!
```

示例4：注释清理已经设置的客户端认证策略。

```
gs_guc set -Z datanode -N all -I all -h "host replication testuser 10.252.95.191/32"

Total instances: 2. Failed instances: 0.
Success to perform gs_guc!
```

示例5：加密服务器端存储用户密码的明文，避免密码泄露。

```
gs_guc encrypt -K XXXXXXXXX -D /data/xzg/wisequery/gaussdb
gs_guc encrypt -K ***
```

示例6：检查中各个实例的参数配置情况。

```
gs_guc check -Z datanode -D /gaussdb/data/datanode -c "max_connections" -N plat1

Total GUC values: 1. Failed GUC values: 0.
```


版权所有 © 华为云计算技术有限公司

```
The value of parameter max_connections is same on all instances.
   max_connections=800
```

示例7：修改DN的参数，例如修改数据库允许的最大连接数为800。修改后需要重启数据库才能生效。

```
gs_guc set -Z datanode -N all -I all -c "max_connections = 800"
Begin to perform gs_guc for all datanodes.

Total instances: 3. Failed instances: 0.
Success to perform gs_guc!

Begin to perform gs_guc for all datanodes.

Total instances: 9. Failed instances: 0.
Success to perform gs_guc!
```

示例8：修改DN的参数，修改archive_command归档命令，由用户自己保证归档路径的有效性。修改后需要重启数据库才能生效。推荐使用此方法。

```
gs_guc set -Z datanode -N all -D /data/datanode  -c "archive_command= 'cp %p /data/datanode/pg_xlog/
archive_status/%f' "
Total instances: 3. Failed instances: 0.
Success to perform gs_guc!
```

示例9：修改DN的参数，修改archive_command归档命令，由用户自己保证归档路径的有效性。修改后需要重启数据库才能生效。不推荐使用此方法。

```
gs_guc set -Z datanode -N all -I all -c "archive_command= 'cp %p pg_xlog/archive_status/%f' "
Total instances: 3. Failed instances: 0.
Success to perform gs_guc!
```

示例10：通过用户输入的密码，生自定义的test.key.cipher和test.key.rand文件。

```
gs_guc generate -S Mppdb@123 -D /data/xzg/wisequery/gaussdb -o test
gs_guc encrypt -S ***
```

示例11：设置PDB的参数，详情请见《 特性描述 》多租数据库一章。

```
gs_guc set --pdbname pdbname -Z datanode -D /gaussdb/data/datanode -c
"global_syscache_threshold=70MB"

Total instances: 1. Failed instances: 0.
Success to perform gs_guc!
```

4.9 gs_om

说明

系统内部工具，不支持单独调用。

背景信息

GaussDB提供了gs_om工具帮助对数据库进行维护，包括启动数据库、停止数据库、查询数据库状态、切换主备实例、修改IP、生成静态配置文件、SSL证书替换、执行日常巡检并上报告警、显示帮助信息和显示版本号信息等功能。

说明

使用root执行的功能需调用数据库工具临时存放的包路径下的gs_om执行。

版权所有 © 华为云计算技术有限公司

语法

- 启动数据库实例
  gs_om -t start [-h *HOSTNAME*] [-D *INSTANCEDIR*] [--az=*AZNAME*] [--time-out=*SECS*] [-l *LOGFILE*]

  说明

  集中式不支持--az参数。

- 停止数据库实例
  gs_om -t stop [-h *HOSTNAME*] [-D *INSTANCEDIR*] [--az=*AZNAME*] [-m *MODE*] [-l *LOGFILE*] [--enable_alt]

  说明

  集中式不支持--az参数。

- 重启数据库实例
  gs_om -t restart [--nodegroup-name=*NODEGROUPNAME*] [-l *LOGFILE*]
- 查询数据库实例状态
  gs_om -t status [-h *HOSTNAME*] [-o *OUTPUT*] [--detail] [-l *LOGFILE*]
- 切换主备数据库实例

  说明

  1. switch为维护操作：确保数据库实例状态正常，所有业务结束，并使用pg_get_senders_catchup_time视图确认无主备追赶后，再进行switch操作。
  2. 集中式不支持--az参数。
  3. dcf模式不支持--az，--reset参数。

  gs_om -t switch [-h *HOSTNAME*] [--az=*AZNAME*] [-D *INSTANCEDIR*] [--reset] [--time-out=*SECS*] [-l *LOGFILE*] [--enable_alt]
- 管理CN
  gs_om -t managecn -m *MODE* -X *XMLFILE* [-l *LOGFILE*]
- 修改IP
  gs_om -t changeip -X *XMLFILE* [--rollback] [-l *LOGFILE*] [--auto-stop]
- 生成静态配置文件
  gs_om -t generateconf -X *XMLFILE* [--distribute] [-l *LOGFILE*]
- 升级、安装、扩容、节点替换、温备、增加CN时间评估

  说明

  - 耗时评估建议在服务器上使用。
  - 安装耗时与节点数、DN数成正比。节点数与DN数越多，耗时越长。
  - 小数据库实例状态规模下，安装耗时较短，评估意义不大，建议在10节点及以上数据库实例状态规模下使用。
  - 为了评估的准确性，建议在每节点4到8个DN下进行评估。

  gs_om -t estimate -m upgrade [-l *LOGFILE*]
  gs_om -t estimate -m install -X *XMLFILE* [-l *LOGFILE*]
  gs_om -t estimate -m expand -X *XMLFILE* [-l *LOGFILE*]
  gs_om -t estimate -m replace -h *HOSTNAME* [-l *LOGFILE*]
  gs_om -t estimate -m warm-standby -h *HOSTNAME* [-l *LOGFILE*]
  gs_om -t estimate -m addCN [-l *LOGFILE*]
- SSL证书替换
  gs_om -t cert --cert-file=*CERTFILE* [-l *LOGFILE*] [-m *COMPONENT*]
  gs_om -t cert --rollback [-m *COMPONENT*]

 版权所有 © 华为云计算技术有限公司

说明

- 证书替换：
  - 不加-m参数时，默认替换DN数据目录下的根证书和服务端证书及私钥，同时替换$GAUSSHOME/share/sslcert/gsql目录下的证书，并将替换前的证书打包备份在当前目录。
  - 加-m gsql参数时，同不加-m参数一样。
  - 加-m gds参数时，替换$GAUSSHOME/share/sslcert/gds目录下的证书，并将替换前的证书打包备份在当前目录。
  - 加-m xbsa参数时，替换$GAUSSHOME/share/sslcert/xbsa目录下的证书，并将替换前的证书打包备份在当前目录。
  - 加-m roach参数时，替换$GAUSSHOME/share/sslcert/roach目录下的证书，并将替换前的证书打包备份在当前目录。
  - 加-m etcd参数时，必须增加-L参数，替换ETCD数据目录下的证书，同时替换$GAUSSHOME/share/sslcert/etcd目录下的证书，并将替换前的证书打包备份在当前目录。
  - 通过ADAPTOR模式安装的集群，不需要指定-X参数；通过XML模式安装的集群，需要指定-X参数（如果不指定，只会替换dn的证书，不会替换gns的证书，保持前向兼容）。
- 证书回滚和证书替换的-m参数的使用同理。

- 输出相对表空间路径
  ```
  gs_om -t tablespace [-h HOSTNAME] [-l LOGFILE]
  ```
- 开启、关闭数据库实例内kerberos认证
  ```
  gs_om -t kerberos -m install -U USER --krb-server IP1 --krb-standby IP2
  gs_om -t kerberos -m install -U USER --krb-client
  gs_om -t kerberos -m uninstall -U USER
  ```

  说明

  - 部署Kerberos需要保证数据库实例所有节点时间同步，时间误差不超过3min。
  - USER为数据库实例初始化用户。
  - 安装、卸载kerberos，需保证数据库实例处于停止状态，部署完成之后启动数据库实例。
  - 安装服务端，IP1表示主server所在的IP节点，IP2表示备server所在的IP节点。服务端主备IP指定数据库实例内任意IP即可。安装client，在数据库实例内任一节点执行一次命令即可。
  - 命令行执行成功后重启数据库服务，数据库实例内kerberos认证开启。
  - 关闭数据库实例内kerberos认证命令在任意节点上执行一次即可。
- 统计磁盘信息
  ```
  gs_om -t show_disk
  ```
- 统计数据库实例磁盘信息
  ```
  gs_om -t show_cluster_disk
  ```
- 统计磁盘和日志信息
  ```
  gs_om -t show_disk_and_log
  ```
- 统计报警日志信息
  ```
  gs_om -t show_alarm_log
  ```
- 统计数据库实例主磁盘信息
  ```
  gs_om -t get_all_primary_disk
  ```
- 数据库实例拓扑结构比较
  ```
  gs_om -t compare-cluster --remote-host=HOSTNAME --remote-env-file=MPPRCFILE [-l LOGFILE]
  ```
- 终止会话
  ```
  gs_om -t killsession [--close-hba|--restore-hba] [-l LOGFILE]
  ```


版权所有 © 华为云计算技术有限公司

说明

- gs_om -t killsession可以终止数据库的所有活跃用户会话。
- gs_om -t killsession --close-hba关闭客户端访问的白名单，只保留数据库实例内部认证的配置。
- gs_om -t killsession --restore-hba恢复白名单访问。

- 热补丁命令

```
gs_om -t hotpatch -n NAME -c [load | unload | active | deactive | info | recover | cleanLock]
gs_om -t hotpatch -n NODETYPE -c list
```

说明

- 热补丁操作需要数据库实例中待打补丁的进程处于可响应信号、网络可达的正常状态，如果数据库实例中进程处于非正常状态，则不保证热补丁操作能够成功，当进程恢复正常状态后，可通过重试来达到最终成功。
  - load：加载补丁，将指定目录中的补丁分发到数据库实例中各节点中的补丁目录，并将补丁目录中补丁加载到内存。
  - unload：卸载补丁，将加载到内存中的补丁卸载。
  - active：激活补丁，将加载到内存中补丁激活，补丁只有在激活后才生效。
  - deactive：去激活补丁，将激活的补丁去激活，此时补丁在内存中，但是不生效。
  - info：查看补丁信息。
  - recover：OM补丁专有命令，用于恢复由于preinstall等操作被损坏的OM补丁。
  - cleanLock：OM补丁专有命令，用于删除故障操作残留的热补丁锁文件。
  - list：查看补丁列表。
  - 当-c后面的命令为load时，-n的内容为路径+补丁名称。

    例如：gs_om -t hotpatch -n /home/hotpatch/GaussDB-Kernel_503.0.1.HP0001.pat -c load
  - 当-c后面的命令为list时，-n的内容表示nodetype，nodetype可选内容为KERNEL、CMSERVER、ALL，分别代表内核、CMSERVER、全部类型。例如：gs_om -t hotpatch -n KERNEL -c list。
  - 当-c后面的命令为list和load之外的命令时，-n的内容为补丁名称。为了使用方便，-n后面也可以加上补丁路径。例如：gs_om -t hotpatch -n GaussDB-Kernel_503.0.1.HP0001.pat -c info。
- 如果补丁操作成功，会返回包含“[PATCH-SUCCESS]”的字符串，如果失败，则会返回包含“[PATCH-ERROR]”的字符串。常见的错误及处理方法如下：
  - “this patch has been loaded before”，补丁已经被加载，重复执行load命令导致，不需要处理。
  - “No patch loaded now”， 补丁未加载而执行加载外的命令，需要先加载补丁。
  - “The patch and bin are not at same version”，补丁文件和可执行文件不匹配，请确定并更换补丁文件。
  - “patch lib return error: 0xXXXXXXXX”，热补丁库内部错误，需要将错误码（0xXXXXXXXX）及当时使用的补丁文件发送给华为工程师进行进一步分析处理。

- 性能诊断采集任务

```
gs_om -t perf_job -m [enable_all|disable_all|is_enable]
```

 版权所有 © 华为云计算技术有限公司

说明

- 性能诊断采集任务会以一定频率访问视图并存储为json文件，性能诊断采集任务运行后，会在$GAUSSLOG/pg_perf下定期生成json文件。
- enable_all：安装内置的性能诊断采集任务。安装成功后，可以查询pg_job以及pg_job_proc表，获取性能诊断采集任务的详细信息。

  例如：性能诊断采集任务在pg_job_proc表内what字段类似为：select capture_view_to_json('dbe_perf.statement', 0);
- disable_all：卸载内置的性能诊断采集任务。
- is_enable：查询是否安装内置性能诊断采集任务。

- 设置虚拟ip

  ```
  gs_om -t virtualip -m MODE [-X XMLFILE | --virtual-ip=Ips] [-L] [-l LOGFILE]
  ```

- quorum模式到dcf模式切换

  ```
  gs_om -t change_mode -m [dcf|quorum] [-L] [-l LOGFILE] [--time-out=SECS]
  ```

  说明

  - 切换操作需要数据库实例中实例的进程处于可响应信号、网络可达的正常状态，如果数据库实例中进程处于非Normal状态，则不保证切换操作能够成功，当进程恢复正常状态后，可通过重试来达到最终成功。容灾和升级待观察阶段不支持切换。
  - 线下模式切换完成之后需要手动更改xml文件以用于执行扩容，替换，升副本，升级，PITR恢复等需要使用xml文件的操作。
  - 初始是安装dcf模式的数据库实例不建议用户切换到quorum模式。
  - DN副本数小于3不支持模式切换。
  - 主备节点差异较大时会导致切换失败，需要等待备机同步后重入切换动作。
  - 级联备只读实例升备后不支持切换。
  - DCF模式下带logger部署形态不支持切换为Quorum模式。若容灾中的主集群或者备集群带logger部署，也不支持切换为Quorum。
  - 主集群已完成模式切换且灾备未执行切换时，主备集群复制协议不一致期间灾备集群无法执行switchover动作，需将主备集群进行模式切换成一致后才能执行。
  - 流式容灾模式模式切换必须先进行主集群的模式切换，再进行备集群的模式切换。

- ETCD多数派故障一键修复

  ```
  gs_om -t repair_etcd
  ```

- dorado双数据库实例相关命令

  ```
  gs_om -t dorado_switch -m [primary|standby] [--logic]
  gs_om -t dorado_failover --dorado_status=[normal|abnormal] [--logic]
  gs_om -t dorado_restore
  gs_om -t dorado_disk_mount
  gs_om -t dorado_query
  gs_om -t dorado_refresh_dr_relations -m [replacement|reduce_replica|add_replica|switch_etcd|switch_dcc] --data_ips=hostnames
  gs_om -t dorado_dr_manual_recovery -c [switchover|query] [-m primary] [--time-out=SECS] [--logic]
  gs_om -t dr_post_check -c [switchover|failover|query|restore] -m [primary|disaster_standby] --dr_type dorado_cluster
  gs_om -t dr_cluster_manual_recovery -m [primary|disaster_standby] --dr_type dorado_cluster [--logic]
  ```

- 流式容灾相关命令

  ```
  gs_om -t streaming_disaster_recovery_start -m [primary|disaster_standby] [--time-out=SECS] [--json=string][--primary_cluster_name=string] [--add_other_dr_cluster]
  gs_om -t streaming_disaster_recovery_stop [--json=string]
  gs_om -t streaming_disaster_recovery_failover [--delete-dr-info=0|1] [--primary_cluster_name=string] [--logic]
  gs_om -t streaming_disaster_recovery_switchover -m [primary|disaster_standby] [--time-out=SECS] [--simple-mode][--primary_cluster_name=string] [--appoint_cluster_name=string] [--logic]
  gs_om -t streaming_disaster_recovery_query[--appoint_cluster_name=string]
  gs_om -t prepare_disaster_info -m [write|read] -U [USER] [--time-out=SECS] [--json=string][--
  ```


版权所有 © 华为云计算技术有限公司

```
add_other_dr_cluster]
gs_om -t dr_post_check -c [switchover|failover|query] -m [primary|disaster_standby] --dr_type streaming_cluster[--primary_cluster_name=string][--appoint_cluster_name=string]
gs_om -t dr_cluster_manual_recovery -m [primary|disaster_standby] --dr_type streaming_cluster[--primary_cluster_name=string]
gs_om -t dr_cluster_serial_change -m [reduce_dr_cluster|add_dr_cluster|change_primary_cluster][--json=string][--primary_cluster_name=string]
gs_om -t break_dr_action -m [dorado_cluster|streaming_cluster] --break-action=[start|multi_start]
gs_om -t streaming_disaster_recovery_restore [--primary_cluster_name=string]
gs_om -t streaming_disaster_recovery_simulation -m [primary|disaster_standby|disaster_simulation] --xlog-keep=INT --simulation-state=[start|end]
gs_om -t dr_change_hadr_user_info -U [USER]
gs_om -t streaming_disaster_recovery_keep --xlog-keep=INT [--primary_cluster_name=string]
```

- 创建templatem模板数据库

```
gs_om -t create_templatem
```

说明

- DCF模式下不支持创建templatem模板数据库。
- 在使用该命令前需要先设置环境变量PGUSER，设置方法：执行“export PGUSER=xxx”命令。
- 已经搭建的容灾数据库实例不支持templatem特性开启，除非解除容灾关系，并在主数据库实例上创建templatem之后再重新搭建，再在主和灾备数据库实例上开启GUC参数。
- 如果要开启templatem特性，需要做好特性开启前的人员值守准备计划，若出现需要运维人员的场景，需要及时修复。
- 如果用户创建templatem模板数据库失败，可以重新尝试创建templatem。
- 创建templatem模板库时，需要确保数据库实例所有节点的状态为normal。
- 在创建templatem模板库期间，不允许客户有启动或停止节点的操作。
- 在创建templatem过程中，会停止数据库实例，直到创建完成后才会启动数据库实例，此过程中数据库处于离线状态。
- 在创建过程中，如果主DN节点故障，可能会导致数据库实例异常，此时需要运维人员介入，常见的错误及处理方法如下：
  - 修复主DN节点，在主DN节点上手动执行drop命令进行修复，停止集群，命令如下：

    ```
    gs_initdb -D <DATA_DIR> --nodename=<NODENAME> --templatemdrop
    ```

  - 主DN节点异常无法修复时，如果同时存在备DN节点build失败的情况，需要依托于CM的多数派进行选主。对选择的主节点可以通过如下SQL语句查询当前节点是否有templatem模板库：

    ```
    select count(1) from pg_database where datname='templatem' and datcompatibility='M';
    ```

    - 如果选取的升主节点上面没有templatem，向增量build过程失败的备DN节点执行全量bulid修复数据，其余备DN节点执行增量build修复数据。
    - 如果选取的升主节点上面有templatem，手动执行drop命令，向增量build过程失败的备DN节点执行全量build修复数据，其余备DN节点执行增量build修复数据。

    增量build的命令如下：

    ```
    gs_ctl build -D <DATA_DIR> -b auto -M standby -q
    ```

    全量build的命令如下：

    ```
    gs_ctl build -D <DATA_DIR> -b full -M standby -q
    ```

- 删除templatem模板数据库

```
gs_om -t drop_templatem
```

 版权所有 © 华为云计算技术有限公司

说明

- DCF模式下不支持删除templatem模板数据库。
- 在使用该命令前需要先设置环境变量PGUSER，设置方法：执行“export PGUSER=xxx”命令。
- 执行该命令时，需要确保数据库实例所有节点的状态为normal。
- 执行该命令时，不允许客户有启动或停止节点的操作。
- 这个命令通常只用于在创建templatem失败时进行清理以便后续重试创建。
- templatem一经创建后不要删除。
- 存在M-compatibility数据库时，不可删除templatem模板库，若要删除templatem，需先删除M-compatibility数据库。
- 该命令属于高风险操作，请谨慎执行。
- 在创建templatem过程中，会停止数据库实例，直到创建完成后才会启动数据库实例，此过程中数据库处于离线状态。

- 客户使用特性自动采集
```
gs_om -t feature_collect [--time-out=SECS] [--collect-key=collect_key]
```

说明

- --time-out： 一个节点上采集的超时时间。若节点上采集超过该时间，该节点会退出采集操作。单位：s，默认取值3600，取值范围：60~604800。
- --collect-key：唯一标识一次采集任务。格式：YYYY-MM-DD_HHMMSS。默认取采集开始的时间，示例：2024-01-01_121212。

- 生效PDB对应的参数配置文件
```
gs_om -t reload_pdb_guc [--pdbname==PDBNAME]
```

说明

--pdbname：需要重载PDB配置文件的PDB的名称。

- 显示帮助信息
```
gs_om -? | --help
```
- 显示版本号信息
```
gs_om -V | --version
```
- 接收数据库管理员密码
```
echo "password" | gs_om -t [start | stop | restart | status ...]
```

说明

云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

gs_om参数可以分为如下几类：

- 通用参数：
  - -t

    gs_om命令的类型。

    取值范围：

    start、stop、restart、status、switch、managecn、changeip、generateconf、estimate、virtualip、cert、tablespace、kerberos、isolate_stop、isolate_tostandby、isolate_restore、killsession、compare-cluster，show_disk、show_cluster_disk、show_disk_and_log、

 版权所有 © 华为云计算技术有限公司

show_alarm_log、ec、get_all_primary_disk、perf_job、hotpatch、dorado_switch、dorado_restore、dorado_failover、dorado_disk_mount、streaming_disaster_recovery_start、streaming_disaster_recovery_stop、streaming_disaster_recovery_failover、streaming_disaster_recovery_switchover、streaming_disaster_recovery_query、prepare_disaster_info、dorado_query、change_mode、break_dr_options、streaming_disaster_recovery_keep、streaming_disaster_recovery_restore、streaming_disaster_recovery_simulation、dr_change_hadr_user_info、dr_post_check 、dr_cluster_manual_recovery、dorado_dr_manual_recovery、repair_etcd、create_templatem、drop_templatem、reload_pdb_guc。

说明

集中式仅支持start、stop、status、switch、generateconf、cert、killsession、show_alarm_log、get_all_primary_disk、hotpatch、changeip --auto-stop、dorado_switch、dorado_restore、dorado_failover，dorado_disk_mountdorado_disk_mount、streaming_disaster_recovery_start、streaming_disaster_recovery_stop、streaming_disaster_recovery_failover、streaming_disaster_recovery_switchover、streaming_disaster_recovery_query、prepare_disaster_info、dorado_query、streaming_dr_manual_recovery 、dorado_dr_manual_recovery、create_templatem、drop_templatem、reload_pdb_guc。

- -l

  指定日志文件及存放路径。

  默认值：$GAUSSLOG/om/gs_om-YYYY-MM-DD_hhmmss.log（virtualip的默认值：/tmp/gs_virtualip/gs_om-YYYY-MM-DD_hhmmss.log）

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

- 启动数据库参数：
  - -h

    指定需要启动的服务器名称。一次只能启动一个服务器。

    取值范围：服务器名称。

    不指定服务器名称时，表示启动整个数据库。

  - -D

    实例指定的目录。

  - --az

    指定需要启动的AZ名称。一次只能启动一个AZ。

    取值范围：AZ名称。

    不指定AZ名称时，表示启动所有AZ。

  - --time-out=SECS

    指定超时时间，如果超时，om脚本自动退出。单位：s。

    取值范围：正整数，建议值600。

    默认值：600。

版权所有 © 华为云计算技术有限公司

  - -l

    预留接口，GaussDB不支持。

- 停止数据库参数：
  - --az

    指定需要启动的AZ名称。一次只能启动一个AZ。

    取值范围：AZ名称。

    不指定AZ名称时，表示启动所有AZ。
  - -D

    实例指定的目录。
  - -h

    指定需要停止实例所在的服务器名称。一次只能停止一个服务器。

    取值范围：实例所在的服务器名称。

    不指定服务器名称时，表示停止整个数据库。
  - -m, --mode=MODE

    停止模式设置。

    取值范围：支持三种停止模式。

    - fast（f）方式：保证有主备关系的实例数据是一致的。
    - immediate（i）方式：不保证有主备关系的实例数据是一致的。
    - smart（s）方式：保证有主备关系的实例数据是一致的，且备机本地所有日志都已redo。

    默认值：fast方式。
  - --enable_alt

    指定该参数时，开启计划内应用无损透明。
- 重启数据库实例参数：
  - --time-out=SECS

    指定超时时间，如果超时，om脚本自动退出。单位：s。

    取值范围：正整数，建议值600。

    默认值：600。
- 查询状态参数：

  数据库实例状态显示结果的参数说明请参见**表4-23**。
  - -h

    指定需要待查询的服务器名称。

    取值范围：服务器名称。

    不指定服务器时，表示查询整个数据库实例。
  - -o

    输出到指定的output文件中。

    如果不指定，直接显示在屏幕上。
  - --detail

    显示详细信息。如果不指定，只提示该服务器是否正常。

 版权所有 © 华为云计算技术有限公司

- 切换主备实例参数：
  - -h
    指定需要切换实例的服务器名称。
    取值范围：实例所在服务器名称。
  - --az
    指定需要切换实例所在的AZ名称。
    取值范围：实例所在AZ名称。
  - -D
    实例指定的目录。
  - --reset
    重置主备状态。
  - --time-out=SECS
    切换节点最大等待时间。
    如果超时，om脚本自动退出。单位：s。
    取值范围：正整型。
  - -I
    预留接口，GaussDB不支持。
  - --enable_alt
    指定该参数时，开启计划内应用无损透明。
- 管理CN参数：
  - -X
    新的数据库实例配置文件路径。
    取值范围：clusterconfig.xml的路径。
  - -m, --mode=MODE
    指定所要进行的操作。
    取值范围：
    - add：增加数据库实例CN。
    - delete：删除数据库实例CN。
- 修改IP参数：
  - -X
    新的数据库实例配置文件路径。
    取值范围：clusterconfig.xml的路径。
  - --rollback
    当自动回滚失败时，用户排除错误后，输入该参数执行手动回滚操作。
  - --auto-stop
    修改IP前自动停止数据库实例。
- 生成配置文件参数：
  - -X
    选择数据库配置文件路径。
    取值范围：clusterconfig.xml的路径。

 版权所有 © 华为云计算技术有限公司

  - --distribute

    将静态配置文件发布到数据库实例的安装目录。

- 升级、安装、扩容、节点替换、温备、增加CN时间评估参数：
  - -m, --mode=MODE

    指定所要进行的评估操作

    取值范围：upgrade、install、expand、replace、warm-standby、addCN。
  - -X

    执行安装或扩容所配置的xml文件路径。当评估模式为安装或扩容时，该参数是必选项。

    取值范围：clusterconfig.xml的路径。
  - -h

    指定故障主机的名称。可以指定多个主机名称，以逗号分隔。当评估模式为节点替换或温备时，该参数是必选项。

    取值范围：主机名称。
  - -L

    指定该参数时，仅对当前节点进行操作。
- SSL证书替换：
  - -L

    指定该参数时，仅对当前节点进行操作。
  - --cert-file

    指定本地SSL证书zip压缩包所在路径，如果启用国密认证，指定本地国密SSL双证书zip压缩包所在路径。

    --rollback

    证书回退到上一次成功备份的状态。
  - -m

    指定改参数时，仅对该组件进行操作。
  - -X

    新的集群配置xml文件路径。

    取值范围：clusterconfig.xml的路径。
- 输出相对表空间路径
  - -h

    指定需要待输出的服务器名称。
- 开启、关闭数据库实例内kerberos认证：
  - -m

    指定所要进行的操作。

    取值范围：

    - install：开启数据库实例内kerberos认证。
    - uninstall：关闭数据库实例内kerberos认证。
  - -U

    指定数据库实例部署用户。

    取值范围：数据库实例部署用户。

 版权所有 © 华为云计算技术有限公司

  - --krb-server

    为服务端执行安装，此参数只对install起作用，和--krb-standby一同使用。
  - --krb-client

    为客户端执行安装，仅适用于安装。
  - --krb-standby

    创建备用服务器。仅适用于安装。
- 数据库实例拓扑结构比较
  - --remote-host

    指定远程数据库实例的一个正常节点名。

    取值范围：和本地数据库实例用户已建立互信的远程数据库实例节点。
  - --remote-env-file

    指定远程节点的用户环境变量文件。

    取值范围：指定节点数据库实例用户的环境变量文件，若未分离环境变量则为用户目录下的.bashrc文件。
- 热补丁
  - -n

    热补丁名称。
  - -c

    热补丁支持的操作，支持[load | unload | active | deactive | info | list | recover | cleanLock]。
- 性能诊断采集任务
  - -m，--mode=MODE

    指定所要进行的操作。

    取值范围：

    - enable_all：安装内置的性能诊断采集任务。
    - disable_all：卸载内置的性能诊断采集任务。
    - is_enable：查询是否安装内置性能诊断采集任务。
- 设置虚拟ip
  - -m, --mode=MODE

    指定所要进行的操作。

    取值范围：

    - add：增加虚拟ip。
    - delete：删除虚拟IP。
  - -X

    XML配置文件的路径，该参数为必选项。
  - --virtual-ip

    指定要删除的虚拟IP。
  - -L

    指定该参数时，仅对当前节点进行操作。

 版权所有 © 华为云计算技术有限公司

- quorum模式切换dcf模式：
  - -m, --mode=MODE
    指定所要进行的操作。
    取值范围：
    - quorum：dcf模式切换quorum模式。
    - dcf：quorum模式切换dcf模式。
  - --time-out=SECS
    指定DN超时时间，如果超时，om脚本报错回滚。单位：s。
    取值范围：正整数（60~2147483646），建议值180。
    默认值：180。
  - -L
    指定该参数时，仅对以当前节点为主的分片进行操作。
- 调用链接口参数
  - --trace-id
    每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。
    取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。

表 4-23 状态说明

字段	字段含义	字段值
cluster_state	数据库状态。显示整个数据库是否运行正常。	• Normal：表示数据库可用，且数据有冗余备份。所有进程都在运行，主备关系正常。 • Unavailable：表示数据库不可用。 • Degraded：表示数据库可用，但数据没有冗余备份。
redistributing	数据重分布状态	• Yes：表示数据库处于数据重分布状态。 • No：表示数据库未处于数据重分步状态。
balanced	平衡状态。显示是否有数据库实例发生过主备切换而导致主机负载不均衡。	• Yes：表示数据库处于负载均衡状态。 • No：表示数据库未处于负载均衡状态。
node	主机名称	表示该实例所在的主机名称。多AZ时会显示AZ编号。
node_ip	主机IP	表示该实例所在的主机IP。
instance	实例ID	表示该实例的ID。

版权所有 © 华为云计算技术有限公司

字段	字段含义	字段值
state	实例状态	• Primary：表示实例为主实例。 • Standby：表示实例为备实例。 • Pending：表示该实例在仲裁阶段。 • Unknown：表示实例状态未知。 • Down：表示实例处于宕机状态。 • StateLeader：表示该ETCD为主实例状态。 • StateFollower：表示该ETCD为备实例状态。 • StateCandidate：表示该ETCD为仲裁阶段。 **说明** 由于ETCD的主备关系切换时间很短，所以系统很难获取到StateCandidate这种状态。

4.10 gs_replace

说明

集中式单节点不支持该工具。

背景信息

数据库是由多台主机组成的，当数据库中某些主机故障或者主机上的某些实例发生故障后，为了使GaussDB快速地恢复正常，用户可以使用gs_replace工具将发生故障的主机替换为正常主机，将发生故障的实例替换为正常实例。

其中替换主机根据使用场景不同，可分为如下两种：

- 主机替换前后IP一致：将新主机的名称、IP与故障主机的名称、IP修改一致，然后做替换。适用于用户保证替换主机的名称、IP和被替换主机的名称、IP一致的场景，使用此方案时要求的前置条件少，能快速实现替换。
- 主机替换前后IP不一致：无需修改新主机的名称、IP，直接做替换, 不支持仅将新主机名称修改为数据库中已有节点主机名称。适用于替换主机的名称、IP和被替换主机的名称、IP不一致的场景，也可用于将一个高性能的机器替换一个性能较差但其上实例状态正常的场景，使用此方案时要求的前置条件较多。

“不换IP”和“换IP”是从数据库实例配置角度的差异，“不换IP”指不修改数据库实例配置文件的IP和hostname信息，“换IP”是指需要修改数据库实例配置文件的IP和hostname信息。

如果替换前需要修改以下GUC参数：log_dir，listen_addresses，local_bind_address，port，pgxc_node_name，pooler_port， log_directory和audit_directory，由于替换后服务器或实例无法同步这些参数在原来设置的值，会设置为默认值，所以在替换成功后，请使用gs_guc工具以reload方式修改替换后主机或实例中的GUC参数。

版权所有 © 华为云计算技术有限公司

前提条件和注意事项

修复故障实例

前提条件

- 数据库实例处于启动状态，且处于没有加锁。
- 修复操作需要在一个正常主机上执行。
- 该工具支持对故障的ETCD/DN/CM进行节点替换。
- 如果参数enable_synclist_single_inst值为off， 则一主多备场景下，一组DN中，至少保证多数派实例正常工作，只能少数派故障，使用gs_replace修复故障DN不影响业务。
- 使用gs_replace替换DN实例时，被替换的DN不能是主DN，否则替换不成功。
- 数据库实例中至少存在一个正常运行的CM Server。
- 修复时数据库实例内所有节点都不可宕机，如果某台机器宕机会导致cm agent实例不可用，修复会失败。
- 如果数据库实例中部署有ETCD，则正常的ETCD个数必须大于ETCD总个数的一半。
- 强制修复多个节点时，由于会停止需要修复节点上的所有DN主、备实例，所以指定修复的节点的DN主、备均不能在同一个DN环内。
- 一主多备部署下，修复DN实例时，为保证数据正确，DN环中必须有CM可监控的主存活。
- 忽略故障节点升级后，由于故障节点和正常节点的版本不一致，不能直接进行修复。
- 开启GNS的情况下，修复GNS需要数据库实例中至少存在一个正常运行的GNS Server。

注意事项

- 如果故障实例所在主机的安装目录下（$GAUSSHOME/bin/）的二进制文件损坏或丢失，则不能通过替换实例进行修复。需要复制其他正常主机对应的二进制文件到该主机，或者将该主机卸载后，再通过替换主机修复。
- 在前一次修复结束后才能再次执行修复。因此请不要同时在多个主机上执行修复操作。
- 实例修复操作会修复故障节点下的全部故障实例。
- 在修复实例的config阶段，先将指定节点的CM Agent组件修复好，这样才能获取到节点实例的状态。如果主机上的某些实例被人为停止，在CM Agent组件修复好之后，这些原来正常的实例会被正常拉起，而不会被修复。如果在一定时间内拉起失败，这些实例将会被修复。
- 实例修复前用户手动在故障实例上配置的guc参数、gs_hba.conf配置的白名单会丢失，需要重新设置。
- 实例修复过程中可能由于某些步骤出错导致失败，失败后需要使用相同的命令来重入修复过程。一些实例可能已经执行完修复并被成功拉起，但实际上状态还没有完全恢复正常，需要继续修复。

替换主机（不换IP）

前提条件

- 数据库实例处于启动状态，且处于没有加锁。

 版权所有 © 华为云计算技术有限公司

- 替换操作需要在一个正常主机上执行。
- 云环境部署中，如果参数enable_synclist_single_inst值为off， 则一组DN的主实例、备实例，其中只能有一个损坏。
- 线下部署中，如果参数enable_synclist_single_inst值为off， 则一主多备场景下，一组DN中，至少保证多数派实例正常工作，只能少数派故障，使用gs_replace修复故障DN不影响业务。
- 使用gs_replace替换DN实例时，被替换的DN不能是主DN，否则替换不成功。
- 数据库实例内如下实例分别至少存在一个正常运行的：CM Server、CM Agent。
- 如果数据库实例中部署有ETCD，则正常的ETCD个数必须大于ETCD总个数的一半。
- 新主机的IP、主机名等信息要和损坏的主机一致。
- 在备份恢复过程中，暂不支持节点替换操作。
- 不支持备DN节点替换为日志节点。

注意事项

- 在前一次替换结束后才能再次执行替换。因此请不要同时在多个主机上执行替换操作。
- 用户手动在故障主机上配置的guc参数、gs_hba.conf配置的白名单会丢失，需要重新设置。

替换主机（换IP）

替换主机-换IP的方案相当于数据库实例温备方案，即无需修改新机IP、hostname的情况下，将处于运行态的主机直接替换为新机。

前提条件

- 数据库实例处于启动状态，且没有加锁。
- 替换操作需要在一个正常主机上执行。
- 如果参数enable_synclist_single_inst值为off， 则一主多备场景下，一组DN中，至少保证多数派实例正常工作，只能少数派故障，使用gs_replace修复故障DN不影响业务。
- 使用gs_replace替换DN实例时，被替换的DN不能是主DN，否则替换不成功。
- 数据库实例内如下实例分别至少存在一个正常运行的：CM Server、CM Agent。
- 如果数据库实例中部署有ETCD，则正常的ETCD个数必须大于ETCD总个数的一半。
- 如果数据库实例中未部署ETCD，则正常的CM Agent个数必须大于CM Agent总个数的一半。
- 用户可以执行单节点卸载操作，卸载掉需要被替换的节点或者手动清理损坏节点上的数据库相关目录和进程。
- 基于新主机IP的数据库实例配置文件已准备好，与现有的数据库实例配置文件相比，仅修改了IP和hostname。
- 温备前请先检查所有故障节点的cm_agent进程。如果不存在cm_agent进程，但是存在DN进程，需要手动启动cm_agent进程或者停止DN进程。
- 在备份恢复过程中，暂不支持节点替换操作。

注意事项

 版权所有 © 华为云计算技术有限公司

- 在前一次替换结束后才能再次执行替换。因此请不要同时在多个主机上执行替换操作。
- 用户手动在故障主机上配置的guc参数、gs_hba.conf配置的白名单会丢失，需要重新设置。
- 如果一个节点故障，还有别的实例故障，在满足温备条件的情况下温备成功后，故障的节点需要进行一次实例替换来修复。
- 如果故障节点处于断连情况下，执行替换主机后，该节点上数据需要运维人员进行手工清理，同时残留的数据库进程，实例目录，环境变量也需要运维人员手动清理（该节点上的实例目录，环境变量文件与新节点的路径一致，可类比查看新节点的这些信息对该节点进行清理）。

说明

节点故障，在其他节点上实例故障，修复原则：优先修复故障节点。

例如：某3节点（A,B,C）数据库实例,其中A节点故障，B节点上某个实例故障，C节点正常，在满足修复条件的前提下，修复原则为：

方案一（推荐）：

先对A节点进行节点替换的install步骤，然后再对A，B节点一起进行 config操作。

方案二（不推荐）：

1. 优先对节点A进行节点替换操作，将节点A完全修复。
2. 再对节点B进行实例修复。

语法

- 新增主机
  ```
  gs_replace -t install -h HOSTNAME  [-l LOGFILE]
  ```
- 配置实例、主机
  ```
  gs_replace -t config -h HOSTNAME [--force] [-l LOGFILE]
  ```
- 启动新增主机或实例
  ```
  gs_replace -t start -h HOSTNAME [--force] [--time-out=<SECS>] [-l LOGFILE]
  ```
- 替换主机（换IP）
  ```
  gs_replace -t warm-standby -X XMLFILE [-l LOGFILE]
  ```
- 显示帮助信息
  ```
  gs_replace -? | --help
  ```
- 显示版本号信息
  ```
  gs_replace -V | --version
  ```
- 接收数据库管理员密码
  ```
  echo "password" |  gs_replace -t [ install | config | start | warm-standby ]
  ```

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

gs_replace参数可以分为如下几类：

- 通用参数：
  - -t

    om命令的类型。

    取值范围：warm-standby、install、config和start.

 版权所有 © 华为云计算技术有限公司

  - -l

    指定日志文件及存放路径。

    默认值：$GAUSSLOG/om/gs_replace-YYYY-MM-DD_hhmmss.log

- 安装新增实例、主机参数：

  -h

  指定新增主机或新增实例所在主机的名称。可以指定多个主机名称，以逗号分隔。

  取值范围：主机名称。

- 配置新增实例、主机参数：
  - -h

    指定替换主机或替换实例所在主机的名称。可以指定多个主机名称，以逗号分隔。

    取值范围：主机名称。

  - --esc

    esc参数，"True"|other

  - --force

    如果指定主机的表空间所在磁盘出现故障，从而导致表空间中的数据损坏，更换新磁盘后指定该参数进行实例修复时，将对该主机强制进行表空间数据的恢复。如果在config阶段指定该参数，则在start阶段也必须指定该参数。

- 启动新增实例、主机参数：
  - -h

    指定需要启动主机或实例所在主机的名称。可以指定多个主机名称，以逗号分隔。

    取值范围：主机名称。

  - --force

    如果指定主机的表空间所在磁盘出现故障，从而导致表空间中的数据损坏，更换新磁盘后指定该参数进行实例修复时，将对该主机强制进行表空间数据的恢复。如果在config阶段指定该参数，则在start阶段也必须指定该参数。

  - --time-out=<SECS>

    启动超时等待时间。

    若--time-out取值过小时，有可能会因为需要启动的实例数过多而引起超时。若在超时时间内启动不成功，则会报错启动超时，但是数据库实例会在后台继续启动。可以等待一段时间后，查询数据库实例状态，检查数据库实例是否启动成功。

    取值范围：正整数，默认值300。

    单位：秒（s）

- 替换节点（换IP）参数：
  - -X

    数据库实例配置文件路径。

  - -l

    指定日志文件及存放路径。

    在内部会自动给日志名添加一个时间戳。

 版权所有 © 华为云计算技术有限公司

  - 当既不明确指定-l，又不在XML文件中配置gaussdbLogPath时，默认值为“$GAUSSLOG/om/gs_replace-YYYY-MM-DD_hhmmss.log”。
  - 当不明确指定-l，但在XML文件中配置了gaussdbLogPath时，默认值为“gaussdbLogPath/用户名/om/gs_replace-YYYY-MM-DD_hhmmss.log”。
- 调用链接口参数
  - --trace-id

    每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

    取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。
- 其他参数：
  - -?, --help

    显示帮助信息。
  - -V, --version

    显示版本号信息。

示例（替换故障实例）

以替换主机plat1、plat2上的实例为例。

步骤1 以操作系统用户omm登录GaussDB数据库任一主机。

说明

操作系统用户omm登录的主机为非故障主机。

步骤2 （可选）使用如下命令在需要替换实例的主机上清理可能存在的残留文件。此命令仅在上次修复故障实例执行失败的情况下需要执行。

```
(if [ -f $PGHOST/GaussReplace.dat ];then rm $PGHOST/GaussReplace.dat;fi)
```

说明

该文件为替换故障实例、替换主机中产生的用于记录执行步骤的临时文件，如果在上次执行过程中出现宕机或网卡中断等，可能会导致该文件残留。在替换故障实例前检查该文件是否存在，且生成时间非本次替换故障实例的时间，则可判断为上次执行的残留文件，删除该文件后，继续执行替换故障实例。

步骤3 使用如下命令对需要替换实例的主机进行配置操作。

```
gs_replace -t config -h plat1, plat2
```

配置操作会清理替换实例的空间，初始化替换实例，配置替换实例。

如果收到提示：“GAUSS_50201: The XXX does not exist.”，则请检查对应的实例数据目录是否存在。如果不存在，请重新创建目录后再次执行上述命令。

说明

如果指定主机的表空间所在磁盘出现故障，从而导致表空间中的数据损坏，更换新磁盘后，需要指定--force参数对该主机强制进行表空间数据的恢复。如果在config阶段指定--force参数，则在start阶段也必须指定--force参数。

步骤4 使用如下命令对需要修复实例的主机进行启动操作。

```
gs_replace -t start -h plat1 , plat2
```

 版权所有 © 华为云计算技术有限公司

启动操作会启动数据库实例替换实例的主机。

步骤5 使用如下命令重置实例状态。

说明

switch为维护操作：确保数据库实例状态正常，所有业务结束，并使用pg_get_senders_catchup_time视图查询无主备追赶后，再进行switch操作。

```
gs_om -t switch --reset
```

重置过程会恢复数据库实例初始状态，以保证各主机的负载都是均衡的。

步骤6 执行如下命令查询数据库实例状态。

```
gs_om -t status
```

----结束

示例（替换主机-不换 IP）

以需要替换新主机名称为plat3为例。

步骤1 以root用户登录到新主机plat3，检查字符集。

```
echo $LANG
en_US.UTF-8
```

新主机上的字符集、编码方式等信息，需与其它主机保持一致。

如果不一致请使用如下方式设置字符集。

```
export LANG="en_US.UTF-8"
```

步骤2 以root用户登录数据库实例的任意一台没有完全损坏的主机。

步骤3 进入安装包解压后的script文件夹。例如，安装包存放路径为*/opt/software/gaussdb*。

```
cd /opt/software/gaussdb/script
```

步骤4 准备数据库实例环境。

```
./gs_preinstall -U omm -G dbgrp -X ../clusterconfig.xml --alarm-type=5
```

omm为运行数据库实例的操作系统用户，dbgrp为操作系统用户的群组名称，clusterconfig.xml为数据库实例配置文件，此示例中假设其存储在安装包存放路径下。

步骤5 切换为运行数据库实例的操作系统用户omm。

```
su - omm
```

步骤6 执行如下命令完成新增主机的安装操作。

```
gs_replace -t install -h plat3
```

步骤7 执行如下命令完成新增主机的配置操作。

```
gs_replace -t config -h plat3
```

步骤8 执行如下命令完成新增主机的启动操作。

```
gs_replace -t start -h plat3
```

步骤9 执行如下命令恢复数据库实例初始状态，以保证负载均衡。

```
gs_om -t switch --reset
```

步骤10 执行gs_om -t status --detail重新查看数据库实例状态。确认数据库实例balanced状态是否已恢复为“Yes”。

版权所有 © 华为云计算技术有限公司

```
[   Cluster State   ]

cluster_state   : Normal
redistributing  : No
balanced        : Yes
```

----结束

示例（替换主机-换 IP）

步骤1 以root用户登录新主机，并检查新主机上的字符集。

```
echo $LANG
en_US.UTF-8
```

新主机上的字符集、编码方式等信息，使与其它主机保持一致。

如果不一致请使用如下方式设置字符集。

```
export LANG="en_US.UTF-8"
```

步骤2 以root身份登录数据库实例的任意一台正常主机。

步骤3 进入安装包解压后的script文件夹。例如，安装包存放路径为*/opt/software/gaussdb*。

```
cd /opt/software/gaussdb/script
```

步骤4 准备数据库实例环境。

```
./gs_preinstall -U omm -G dbgrp -X ../clusterconfig.xml --alarm-type=5
```

omm为运行数据库实例的操作系统用户，dbgrp为操作系统用户的群组名称，clusterconfig.xml为数据库实例配置文件，此示例中假设其存储在安装包存放路径下。

步骤5 切换为运行数据库实例的操作系统用户omm。

```
su - omm
```

步骤6 执行如下命令进行替换主机操作。

```
gs_replace -t warm-standby -X clusterconfig.xml
```

如果执行过程中出现异常，请根据提示信息排除异常后重新执行该命令。

说明

在温备的过程中，如果出现报错导致温备失败，用户不选择重入，需要使用进行其他工具(升级等)操作数据库，则需进入到数据库实例内各节点$PGHOST下，确认是否存在warm_standby_bak文件夹，如果存在需要删除此文件夹。

步骤7 执行如下命令恢复数据库实例初始状态，以保证负载均衡。

```
gs_om -t switch --reset
```

步骤8 执行gs_om -t status --detail重新查看数据库实例状态。确认数据库实例balanced状态是否已恢复为“Yes”。

```
[   Cluster State   ]

cluster_state   : Normal
redistributing  : No
balanced        : Yes
```

----结束

版权所有 © 华为云计算技术有限公司

4.11 gs_roach

背景信息

数据库备份/恢复工具。

gs_roach分为master进程和agent进程两种。其中，master进程只有在备份主节点上才有，agent进程在其它参与备份的节点上存在。master进程接受和建立来自所有agent进程的连接（长连接），并且通过这些连接，来推动整个分布式备份流程的步骤和状态，具体原理如下：

1. master进程采用ip白名单认证，只允许来自数据库实例内的ip连接。
2. master进程只有在启动之后的固定超时时间内（通常为10分钟）接受来自agent进程的tcp连接请求。如果在该时间段内，数据库实例内所有节点的agent进程没有都连接上，那么master进程会报错退出，该次备份失败；如果在该时间段内，数据库实例内所有节点的agent进程都和master进程成功建连，那么在最后一个agent进程建连成功之后，master进程就会停止对master port的侦听，并开始正式的备份操作。后续整个备份过程中master进程不再接受其他任何agent进程的连接请求。

语法

```
gs_roach OPTIONS ACTION [backup | restore | update | delete | uploadmeta | downloadmeta | obs-upload | obs-download | obs-delete | query | retrieve]
```

参数说明

OPTION:

- --master
  设置该实例为master。默认为agent。
- --with-serverlog
  包括备份服务器日志文件时的备份。
- --enable-crc
  开启crc校验。
- --master-ip <master ip>
  设置主“roach”IP地址。
- --master-port <master port>
  主roach应该启动的端口。
- --media-type <media-type>
  备份目的地。
- --media-destination <media-destination>
  指定备份的路径。
- --metadata-destination <path>
  备份元数据文件的路径。
- --restore-buffer-threshold <buffer-threshold>
  在还原下一个文件之前应该可用的缓冲区的百分比

 版权所有 © 华为云计算技术有限公司

- --backup-key <backup key>
  还原restore操作的键。
- --prior-backup-key <backup key>
  增量备份的backup key。
- --backup-type <backup type>
  备份的类型。
- --exec-delay-ddl
  备份时延迟DDL语句。
- --single
  指定为单机模式。
- --single-group-nodeid
  单机模式下的nodeid。
- --single-group-datapath
  单机模式下的数据路径。
- --snapshot-action
  快照备份操作
- --split-local-disk-storage
  备份的不同的DN数据分开存储。
- --standby-mode
  开启备机备份。
- --no-display
  指定不显示回显信息。
- --resume-backup
  标记本次备份开启断点续做。
- --resume-restore
  标记本次恢复开启断点续做。
- --target-backup-key <backup key>
  恢复到指定的backup key。
- --incremental
  增量恢复。
- --restore-new-cluster
  恢复数据到新数据库。
- --restore-configs
  恢复配置文件到新数据库。
- --parallel-restore
  指定使用备DN并行恢复。
- --skip-restore-gtmcm
  恢复时不恢复cm的数据。
- --compression-level <level>
  备份时应该使用的压缩级别。

 版权所有 © 华为云计算技术有限公司

- --compression-type <type>
  备份时应该使用的压缩类型。
- --cbm-recycle-level <level>
  备份时cbm文件回收应该使用cbmRecycle级别。
- --reader-thread-count <count>
  备份时读线程的数量。
- --reader-thread-file-count <count>
  备份时线程文件的数量。
- --reader-thread-file-size <size>
  备份时线程文件的大小。
- --buffer-size <size>
  应该使用的缓冲区大小。
- --buffer-block-size <block-size>
  每个缓冲区应该使用的块大小。单位：B
- --max-backup-io-speed <speed>
  每个并行进程备份写入媒体目标的最大速度。单位：B/s。必须大于buffer-block-size。0表示不限制
- --parallel-process <count>
  应该使用的进程数。
- --failure-retry-count <count>
  在失败的情况下，应该重试的次数。
- --resource-retry-count <count>
  在失败的情况下，重试的次数。
- --retry-wait-time <time>
  失败后重试前等待的时间。
- --logging-level <logging level>
  [FATAL|ERROR|WARNING|INFO|DEBUG|DEBUG2]. 设置日志的等级，默认的等级是warning.
- --logging
  启用记录日志文件。在以后的版本中将不推荐使用此版本。请使用--enable-logging代替。
- --enable-logging <True/False>
  启用记录日志文件。默认值为False。
- --log-filesize
  最大日志文件大小。
- --filesplit-size
  将要分割的文件的近似文件大小。
- --log-filecount
  创建日志文件的最大数量。
- --logging-path
  需要创建日志文件的位置。

 版权所有 © 华为云计算技术有限公司

- --dbname <dbname>
  设置数据库名。
- --username <username>
  数据库的用户名。
- -W <password>
  数据库用户的密码。
- --pipeline
  使用管道传输密码，禁止在终端使用。
- --cpu-relinquish-time
  放弃CPU的时间。放弃时间在秒级的范围内。
- --cpu-relinquish-size
  处理完释放大小后，释放CPU。
- --max-memory-usage <value>
  Roach进程允许使用的最大内存，单位为GB。范围2-256。
- --pre-disk-space
  备份恢复前检查磁盘空间可用性。
- --mark-failed
  标记备份状态失败。
- --mark-transparent
  标记备份状态为透明。
- --mark-corrupted
  标记备份状态为已损坏。
- --mark-restored
  标记该备份集已经被恢复过。
- --obs-server-ip <address>
  OBS服务器ip地址或域名地址。
- --bucket-name <bucketname>
  OBS桶名。
- --cluster-unique-id <value>
  数据库唯一标识。
- --content <content>
  上传到obs对象的文件内容。
- --obs-ak <obs ak>
  连接obs服务的ak。
- --obs-sk <obs sk>
  连接obs服务的sk。
- --prefetch-block <block-bum>
  增量备份时读取磁盘上表文件增量修改页面的预取页面个数。
- --verbose
  当前正在执行的命令的详细信息。

 版权所有 © 华为云计算技术有限公司

- --miniaturized-deployment
  小型化部署数据库标识。
- --old-cluster-hostname <hostname>
  旧数据库的hostname。
- --all-backups
  显示所有备份信息。
- --backup-key <backup key>
  显示给定备份的详细信息。
- --overwrite
  覆盖已经存在的文件。
- --related-backup-keys
  显示给定备份的相关的backup key。
- --start-time
  查询PITR可恢复时间时的开始时间。
- --version
  显示版本信息。
- -V
  显示版本信息。
- --help
  显示帮助信息。
- --disable-ssl
  关闭SSL认证，gs_roach默认开启SSL认证，若指定该参数则关闭SSL认证。
- --remote-ip
  指定远端第三方存储介质所在服务器IP，用于与REMOTE第三方通信进行备份恢复。
- --remote-port
  指定远端第三方存储介质所在服务器侦听端口，用于与REMOTE第三方通信进行备份恢复。
- --remote-config
  REMOTE第三方厂商配置参数内容。
- gbr-sys-data-recovery
  指示本次恢复的目标为template0、template1和global中的数据。
- gbr-db-list
  要恢复的数据库列表。

  说明

  该版本不支持对M-compatibility的数据库进行细粒度恢复。
- gbr-table-list
  备份或恢复的表列表文件的绝对路径。

  说明

  该版本不支持对M-compatibility的表进行细粒度恢复。

 版权所有 © 华为云计算技术有限公司

- aux-db-path
  辅助数据库路径。
- retrieve-type
  找回的日志类型。
- retrieve-shard
  找回日志的分片。
- retrieve-local-path
  找回日志的本地暂存路径。
- retrieve-start-lsn
  找回xlog的起始lsn。
- retrieve-end-lsn
  找回xlog的终止lsn。
- archive-clean-timeout
  清理归档的超时时间。
- archive-clean-shard
  清理参与该次备份实例所在分片的分片名。
- archive-clean-stoplsn
  清理参与该次备份实例所在分片的归档xLog slice的指定LSN前的所有xLog slice。
- last-backup-endtime
  清理基于该次备份集归档所对应的备份结束时间。
- reverse
  清理指定归档后的所有归档日志。
- skip-restore-dn
  恢复时跳过DN的恢复步骤。
- download-target-path
  obs-download时指定下载目标文件的本地路径。
- archive-backup-version
  清理归档时传入的备份集备份时的版本信息。
- pdb-list
  执行PDB级备份时，指定当前进行备份的PDB列表。执行恢复时，指定当前进行恢复的PDB列表。

  说明

  该版本不支持对M-compatibility的PDB进行细粒度备份恢复。
- gbr-pdb-sys-data-recovery
  执行PDB级恢复时，指定当前恢复是否为PDB级第一轮恢复。
- flash-restore-node-count
  执行闪回快照恢复时，指定本次恢复的节点数量。

ACTION:

- backup

 版权所有 © 华为云计算技术有限公司

备份操作。

- restore
  恢复操作。
- update
  更新操作。
- delete
  删除操作。
- uploadmeta
  上传元数据操作。
- downloadmeta
  下载元数据操作。
- obs-upload
  上传内容到obs对象操作。
- obs-download
  从obs上读取对象内容操作。
- obs-delete
  删除obs上指定前缀的对象操作。
- query
  查询PITR可恢复时间。
- retrieve
  关键日志找回操作。

4.12 gs_ssh

说明

系统内部工具，不支持单独调用。

背景信息

GaussDB提供了gs_ssh工具帮助用户在数据库各节点上执行相同的命令。

前提条件

- 数据库已经正确安装部署。
- 调用命令可用which查询到且在当前用户下有执行权限。
- 执行的命令返回的内容不可以太多或者无限制。

语法

- 同步执行命令
  ```
  gs_ssh -c cmd
  ```
- 显示帮助信息
  ```
  gs_ssh -? | --help
  ```


版权所有 © 华为云计算技术有限公司

- 显示版本号信息
  ```
  gs_ssh -V | --version
  ```

> **须知**
>
> gs_ssh工具支持以管道符形式传入被执行的cmd，如：echo 'gsql -p xxx -d xxx -U xxx -W xxx -c "xxx;"' | gs_ssh，使用该方式执行命令时，进程中不显示具体的cmd。当使用gs_ssh工具执行包含敏感信息的命令时，请使用管道符形式传入cmd，避免敏感信息在进程中可查。

参数说明

- -c
  指定需要在数据库各主机上执行的linux shell命令名。
- -?, --help
  显示帮助信息。
- -V, --version
  显示版本号信息。

示例

在数据库各主机上执行相同命令。以执行"hostname"命令为例。

```
gs_ssh -c "hostname"
Successfully execute command on all nodes.
Output:
[SUCCESS] plat1:
plat1
[SUCCESS] plat2:
plat2
[SUCCESS] plat3:
plat3
[SUCCESS] plat4:
plat4
```


版权所有 © 华为云计算技术有限公司

5 系统内部调用的工具

本节所涉及的工具为GaussDB进程间运行使用的工具，或者说模块间调用的工具。这些工具只用于满足内部调用场景，未对其他场景下的使用正确性做过验证。因此，不建议用户直接使用这些工具执行业务，以防对系统运行带来影响。

说明

沙箱部署场景，所有工具需要进入到沙箱内部才能正确执行。

进入沙箱命令参考：chroot /var/chroot

5.1 etcd

说明

集中式单节点不支持该工具。

功能介绍

etcd是GaussDB引入的第三方软件，能够通过其启动一个键值存储系统，管理来自存储系统的数据。

一个etcd进程管理来自同一个存储系统的数据。一个系统上可以同时运行多个etcd进程，只要使用不同的数据目录和不同的端口号。etcd启动时需要知道数据目录的位置，该位置必须通过--data-dir选项指定。

缺省时etcd在前台启动并将日志信息输出到标准错误。但在实际应用中，etcd应当作为后台进程启动，而且多数是在系统启动时自动启动。

参数说明

公共参数：

- --version
  显示etcd的版本。
- -h,--help
  显示etcd的帮助信息。
- --config-file

 版权所有 © 华为云计算技术有限公司

服务器配置文件路径。

- gateway
  运行无状态直通etcd TCP连接转发代理。
- grpc-proxy
  运行无状态etcd v3 gRPC L7反向代理。

表5-1中的参数主要是etcd数据库中单个成员相关的参数。

表 5-1 成员参数

参数	参数说明	取值范围
--name NAME	指定成员名。	缺省是default。
--data-dir <DATA_DIR>	指定数据目录。	缺省是${name}.etcd。
--wal-dir WALDIR	指定WAL目录。	-
--snapshot-count SNAPSHOTCOUNT	指定触发做一次snapshot的事务数量。	缺省是100000。
--heartbeat-interval INTERVAL	指定心跳周期。	缺省是100。 单位为ms。
--election-timeout TIMEOUT	指定心跳超时重新选举时间。	缺省是1000。 单位为ms。
--initial-election-tick-advance ENABLE	是否提前初始化选举时钟启动，以便更快的选举。	ENABLE可以取以下值： true：提前初始化选举时钟。 false：按正常流程初始化。 缺省是true。
--listen-peer-urls URLS	指定侦听其他成员连接的URL。	缺省是https://localhost:2380。
--listen-client-urls URLS	指定侦听客户端连接的URL。	缺省是https://localhost:2379。
--max-snapshots COUNT	要保留的最大快照文件数（0是无限的）。	缺省是5。
--max-wals COUNT	要保留的最大wal文件数（0是无限的）。	缺省是5。
--quota-backend-bytes SIZE	当后端大小超过给定限额时发出警报（0默认为低空间限额）。	缺省是0，2G空间。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--backend-bbolt-freelist-type	后端（bboltdb）使用的自由列表类型（支持数组和映射的类型）。	缺省是map。
--backend-batch-interval	提交后端事务之前的最长时间	缺省是0。
--backend-batch-limit	提交后端事务之前的最大数量的操作。	缺省是0。
--max-txn-ops COUNT	事务中允许的最大操作数。	缺省是128。
--max-request-bytes SIZE	服务器将接受的最大客户端请求大小（以字节为单位）。	缺省是1572864字节。
--grpc-keepalive-min-time TIME	客户端在ping服务器之前应该等待的最小持续时间间隔。	缺省是5秒。
--grpc-keepalive-interval TIME	检查连接是否激活的服务器到客户机ping的频率持续时间（0表示禁用）。	缺省是2小时。
--grpc-keepalive-timeout TIME	关闭无响应连接之前的额外等待时间（0表示禁用）。	缺省是20秒。
--socket-reuse-port ENABLE	启用此选项可在侦听器上设置套接字选项SO_REUSEPORT，允许重新绑定已在使用的端口。	ENABLE可以取以下值： ● true：启用套接字选项SO_REUSEPORT。 ● false：不启用套接字选项SO_REUSEPORT。 缺省是false。
--socket-reuse-address ENABLE	启用此选项可在侦听器上设置套接字选项SO_REUSEADDR，允许绑定到处于TIME_WAIT的地址。	ENABLE可以取以下值： ● true：启用套接字选项SO_REUSEADDR。 ● false：不启用套接字选项SO_REUSEADDR 缺省是false。

表5-2中的参数主要是etcd数据库相关的参数。

 版权所有 © 华为云计算技术有限公司

表 5-2 数据库参数

参数	参数说明	取值范围
--initial-advertise-peer-urls URLS	指定对其他成员公开的侦听对方连接的URL。	缺省是https://localhost:2380。
--initial-cluster CLUSTER	指定etcd数据库配置。	缺省是default=https://localhost:2380。
--initial-cluster-state CLUSTERSTATE	指定etcd数据库模式。	CLUSTERSTATE可以取以下值： ● new：新的数据库。 ● existing：数据库已存在。 缺省是new。
--initial-cluster-token CLUSTERTOKEN	指定etcd数据库标识。	缺省是etcd-cluster。
--advertise-client-urls URLS	指定对外公开的侦听客户端连接的URL。	缺省是https://localhost:2379。
--discovery DISCOVERY	指定自发现模式启动的URL。	-
--discovery-fallback FALLBACK	指定自发现服务不可用时数据库的处理方式。	FALLBACK可以取以下值： ● exit：etcd节点退出。 ● proxy：etcd节点以代理模式运行。 缺省是proxy。
--discovery-proxy PROXY	指定自发现服务的代理。	-
--discovery-srv SRV	指定DNS自发现模式启动的DNS域名。	-
--discovery-srv-name	使用DNS引导时查询的DNS srv名称的后缀。	-
--strict-reconfig-check ENABLE	拒绝可能导致仲裁丢失的重新配置请求。	ENABLE可以取以下值： ● true：拒绝。 ● false：允许。 缺省是true。
--pre-vote	启用运行Raft选举。	ENABLE可以取以下值： ● true：拒绝。 ● false：允许。 缺省是true。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--auto-compaction-retention RETENTION	存储的自动压缩保留时间（小时）。 0表示禁用自动压缩。	缺省是0。
--auto-compaction-mode MODE	自动压缩模式。	MODE可以取以下值： • periodic：时间周期 • revision：修改版本数量 缺省是'periodic'。
--enable-v2 ENABLE	是否接受响应etcd V2版本的客户端请求。	ENABLE可以取以下值： true：接受响应etcd V2版本请求。 false：拒绝响应etcd V2版本请求。 缺省是true。
--v2-deprecation MODE	允许选择加入更高的兼容性模式弃用v2。	MODE可以取以下值： • not-yet：如果v2包含有意义的内容，则发出警告。 • write-only：不允许自定义v2状态。 • write-only-drop-data：自定义v2状态将被删除。 • gone：不再维护v2。 缺省是not-yet。

表5-3中的参数主要是etcd代理相关的参数。

表 5-3 代理参数

参数	参数说明	取值范围
--proxy PROXYMODE	指定代理模式。	PROXYMODE可以取以下三个值： • off：关闭代理模式。 • readonly：打开代理模式，但只能读。 • on：打开代理模式，读写皆可。 缺省是off。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--proxy-failure-wait FAILUREWAIT	指定代理故障状态下有效时间。	缺省是5000。 单位为ms。
--proxy-refresh-interval INTERVAL	指定更新周期。	缺省是30000。 单位为ms。
--proxy-dial-timeout TIMEOUT	指定拨号超时时间。	缺省是1000。 单位为ms。
--proxy-write-timeout MEOUT	指定写超时时间。	缺省是5000。 单位为ms。
--proxy-read-timeout TIMEOUT	指定读超时时间。	缺省是0。 单位为ms。

表5-4中的参数主要是etcd安全相关的参数。

表 5-4 安全参数

参数	参数说明	取值范围
--cert-file PATH	指定客户端TLS cert文件路径。	-
--key-file PATH	指定客户端TLS密钥文件路径。	-
--client-cert-auth AUTH	指定是否支持客户端cert认证。	AUTH可以取以下值： ● true：支持客户端cert认证。 ● false：不支持客户端cert认证。 缺省是false。
--client-crl-file PATH	指定客户端证书撤销列表文件路径。	-
--client-cert-allowed-hostname	允许客户端证书身份验证的TLS名称。	-
--trusted-ca-file PATH	指定客户端TLS信任的CA密钥文件路径。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--auto-tls ENABLE	客户端TLS使用自动生成的证书。	ENABLE可以取以下值： true：使用自动生成的证书。 false：不适用自动生成证书。 缺省是false。
--peer-ca-file PATH	指定其他成员TLS CA文件路径（不推荐使用）。	-
--peer-cert-file PATH	指定其他成员TLS cert文件路径。	-
--peer-key-file PATH	指定其他成员TLS密钥文件路径。	-
--peer-client-cert-auth AUTH	指定是否支持其他成员客户端cert认证。	AUTH可以取以下值： ● true：支持其他成员cert认证。 ● false：不支持其他成员cert认证。 缺省是false。
--peer-trusted-ca-file PATH	指定其他成员TLS信任的CA密钥文件路径。	-
--peer-cert-allowed-cn COMMONNAME	允许进行对等身份验证的通用名字。	缺省是''。
--peer-auto-tls ENABLE	peer间是否TLS使用自动生成的证书。	ENABLE可以取以下值： true：使用自动生成的证书。 false：不使用自动生成证书。 缺省是false。
--self-signed-cert-validity	etcd自动生成的客户端证书和对端证书的有效期，单位为年，默认值为1。	缺省是1。
--peer-crl-file PATH	指定peer间证书撤销列表文件路径。	-
--cipher-suites CIPHERS	以逗号分隔的服务器/客户端和对等方之间受支持的TLS密码套件列表。	-
--cors	以逗号分隔的CORS来源白名单。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--host-whitelist	如果服务器不安全，则HTTP客户端请求中的可接受主机名。	缺省是*。

表5-5中的参数主要是便于开发者调试的参数。

表 5-5 调试参数

参数	参数说明	取值范围
--logger	指定“zap”用于结构化日志记录。	缺省是zap。
--log-output DEST	指定日志的输出位置。	DEST可以取以下值： ● stdout：标准输出 ● stderr：标准错误 缺省default，日志输出在系统日志中。
--log-level LEVEL	指定日志级别。	LEVEL可以取以下值： ● debug ● info ● warn ● error ● panic ● fatal 缺省是info。
--enable-log-rotation ENABLE	启用单个日志输出文件。	ENABLE可以取以下值： ● true：启用单个日志输出文件。 ● false：不启用单个日志输出文件。 缺省是false。

表5-6中的参数是在etcd数据库不可用时强制转换到单机模式运行的参数。

版权所有 © 华为云计算技术有限公司

表 5-6 维护参数

参数	参数说明	取值范围
--force-new-cluster FORCENEWCLUSTER	指定是否强制创建一个新的数据库。	FORCENEWCLUSTER可以取以下值： • true：强制创建一个新的数据库。 • false：不强制创建一个新的数据库。 缺省是false。

表5-7中的参数是便于使用者获取etcd内部各项指标参数，便于分析现象和定位问题。

表 5-7 分析参数

参数	参数说明	取值范围
--enable-pprof ENABLE	是否通过HTTPS服务器启用运行时分析数据。	ENABLE可以取以下值： true：启用HTTPS运行数据分析服务。 false：禁用HTTPS运行数据分析服务。 缺省是false。
--metrics LEVEL	设置分析数据的详细程度。	缺省是basic。指定extensive将展示服务端grpc直方图等分析指标。
--listen-metrics-urls URL	额外响应/metrics和/health类请求的URL列表。	缺省是空。

表5-8中主要是etcd鉴权的相关参数。

表 5-8 权限参数

参数	参数说明	取值范围
--auth-token TOKEN	指定v3认证令牌类型及其选项('simple' 或 'jwt')。JWT的示例选项：-auth-token jwt,pub-key=app.rsa.pub,privkey=app.rsasign-method = RS512,ttl = 10m	缺省是simple。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--auth-token-ttl	指定超时时间	缺省值是300，单位为s。

表5-9中的参数是便于使用者试用etcd第三个版本原型功能的参数。

表 5-9 预研参数

参数	参数说明	取值范围
--experimental-initial-corrupt-check ENABLE	响应客户端和peer端请求前是否进行数据损坏检查。	ENABLE可以取以下值： true：进行数据损坏检查。 false：跳过数据损坏检查。 缺省是false。
--experimental-corrupt-check-time INTERVAL	群集数据损坏检查的时间间隔。	缺省是0s。
--experimental-enable-v2v3 ''	指定V2请求转换V3请求的前缀。	缺省为''。

5.2 etcdctl

说明

集中式单节点不支持该工具。

功能介绍

etcdctl是GaussDB引入的etcd键值存储系统控制工具，可以用来创建和查询键的值。

etcdctl的主要功能有：

- 解除、展示告警。
- 认证。
- 查询内存及性能数据。
- 压缩历史记录日志。
- 碎片整理。
- 删除键。
- 观察并参与leader选举。
- 查看节点状态、健康状态及KV历史记录。
- 获取键值。
- 显示帮助信息。
- 租约管理。
- 获取锁。

版权所有 © 华为云计算技术有限公司

- 创建镜像。
- 添加、更新、删除和查看实例成员信息。
- 转换leader角色。
- 设置键值。
- 添加、查看、授权、注销角色。
- 快照管理。
- 添加、查看、授权、注销用户、更改用户密码。
- 监控键值变化。

参数说明

etcdctl参数可分为如下几类：

- command参数，详细请参见**表5-10**
- option参数，详细请参见**表5-11**

表 5-10 command 参数

参数	参数说明
alarm	解除、展示告警。
auth	认证。
check	查询内存及性能数据。
compaction	压缩历史记录日志。
defrag	碎片整理。
del	删除键。
elect	观察并参与leader选举。
endpoint	查看节点状态、健康状态及KV历史记录。
get	获取键值。
help	显示帮助信息。
lease	租约管理。
lock	获取锁。
make-mirror	创建镜像。
member	添加、更新、删除和查看实例成员信息。
move-leader	转换leader角色。
put	设置键值。
role	添加、查看、授权、注销角色。
snapshot	快照管理。

版权所有 © 华为云计算技术有限公司

参数	参数说明
txn	处理事务中所有请求。
user	添加、查看、授权、注销用户、更改用户密码。
version	查看版本信息。
watch	监控键值变化。

表 5-11 option 参数

参数	参数说明
--cacert	指定客户端CA文件路径。
--cert	指定客户端SSL cert文件路径。
--command-timeout	执行命令超时时间。
--debug	打印调试日志。
--dial-timeout	建连超时时间。
--discovery-srv	指定实例域名。
--discovery-srv-name	指定实例DNS域名。
--endpoints	逗号分隔指定一系列etcd服务端URL。 缺省是https://127.0.0.1:2379。
--help, -h	显示帮助信息。
--hex	以十六进制打印字符串。
--insecure-discovery	接受来自不安全SRV记录的实例信息。
--insecure-skip-tls-verify	跳过服务器证书认证。
--insecure-transport	禁用不安全客户端连接。
--keepalive-time	客户端连接保持时间。
--keepalive-timeout	客户端连接keepalive超时时间。
--key	指定客户端SSL密钥文件路径。
--password	使用密码进行身份验证。
--user	使用用户进行身份验证。
--write-out，-w	指定指定输出格式。

 版权所有 © 华为云计算技术有限公司

5.3 gaussdb

背景信息

gaussdb是GaussDB数据库的主进程，也是一个可执行的命令，能够通过其启动一个数据库进程。

语法

```
gaussdb [OPTION]...
```

描述

客户端应用程序为了访问数据库，将连接（通过网络或本地）到一个正在运行的gaussdb进程。然后该进程实例会启动一个独立的线程来处理这个连接。

须知

通过gaussdb启动数据库时，需要再开一个窗口来连接到数据库，也可以使用&符号使程序在后台执行。

一个gaussdb进程总是管理来自同一个数据库的数据。一个系统上可以同时运行多个gaussdb进程，只要使用不同的数据目录和不同的端口号。gaussdb启动时需要知道数据目录的位置，该位置必须通过-D指定。通常，-D直接指向由gs_initdb创建的数据库目录。

缺省时gaussdb在前台启动并将日志信息输出到标准错误。但在实际应用中，gaussdb应当作为后台进程启动，而且多数是在系统启动时自动启动。

gaussdb还能以单用户模式运行，这种用法主要用于gs_initdb的初始化过程中，有时候也被用于调试灾难性恢复。不过，单用户模式运行的服务器并不适合于调试，因为没有实际的进程间通讯和锁动作发生。当从shell上以单用户模式调用时，用户可以输入查询，然后结果会在屏幕上以一种更适合开发者阅读（不适合普通用户）的格式显示出来。在单用户模式下，将把会话用户ID设为1并赋予系统管理员权限。该用户不必实际存在，因此单用户模式运行的服务器可以用于对某些意外损坏的系统表进行手工恢复。

选项

gaussdb接受表5-12中所示的命令行参数。

表 5-12 参数说明

参数	参数说明	取值范围
-B NBUFFERS	设置服务器进程使用的共享内存缓冲区的数量。	-
-b BINARY UPGRADES	binary upgrade标志。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-c NAME=VALUE	给一个正在运行的参数命名。-c可以出现多次从而设置多个参数。	-
-C NAME	打印一个指定的运行时的参数的值然后退出。可以用在正在运行的服务器上，从gaussdb.conf返回值。	取值为gaussdb.conf中的所有参数。
-d 1-5	设置调试级别，1-5记录对应debug级别信息。级别越高，写到服务器日志的调试输出越多。	取值范围：1-5。 **说明** 如果取值小于等于0，则只会记录notice级别的信息。
-D <DATA_DIR>	声明数据目录或者配置文件的文件系统路径。	用户自定义。
-e	把缺省日期风格设置为“European”，也就是说用DMY规则解释日期输入，并且在一些日期输出格式里日在月份前面打印。	-
-F	关闭fsync调用以提高性能，但是要冒系统崩溃时数据毁坏的风险。声明这个选项等效关闭了fsync参数。	-
-h HOSTNAME	指定gaussdb侦听来自前端应用TCP/IP连接的主机名或IP地址。	前端存在的主机名或IP地址。
-i	该选项允许远程客户通过TCP/IP（网际域套接字）与服务器通讯。	-
-k DIRECTORY	指定gaussdb侦听来自前端应用连接的Unix域套接字的目录。	缺省通常是/tmp，但是可以在编译的时候修改。
-l	该选项允许远程客户通过SSL（ 安全套接层）与服务器通讯。	-
-N MAX-CONNECT	设置服务器接受的客户端连接的最大数。缺省时由gs_initdb自动选择。声明这个选项等价于声明max_connections配置参数。	取值范围：正整数。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-M SERVERMODE	在启动时指定数据库的启动模式。	SERVERMODE可以取下面三个值： ● primary：本端以主机模式启动。 ● standby：本端以备机模式启动。 ● pending：本端处于等待状态，等待提升为主机或者降为备机。 ● cascade_standby：本端以级联备模式启动。 ● hadr_main_standby：本端以流式容灾中的首备启动。
-o OPTIONS	向每个服务器进程传递“OPTIONS”。保留语法，已没有实际功能。	-
-p PORT	指定gaussdb侦听客户端连接的TCP/IP端口或本地Unix domain socket文件的扩展。指定-p参数会按照给定的参数设置连接端口号，不指定时按照配置文件中的port值连接。	正整数，在操作系统支持的端口范围内。
-s	在每条命令结束时打印时间信息和其他统计信息。	-
-S WORK-MEM	声明内部排序和散列在求助于临时磁盘文件之前可以使用的内存数量。	单位为KB。
--single_node	拉起单DN数据库。	-
-V, --version	打印gaussdb的版本信息然后退出。	-
--NAME=VALUE	给一个正在运行的参数命名。	-
--describe-config	描述配置参数然后退出。	-
--securitymode	以安全模式运行。 **须知** 在安全模式下，不允许执行COPY TO/FROM filename，不允许对public角色授权，不允许创建和修改表空间，用户需要具有USEFT属性，才能创建、修改、删除和使用外表。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-?, --help	显示关于gaussdb命令行参数的帮助信息，然后退出。	-

表5-13中的参数主要是便于开发人员调试使用，有时也用于帮助恢复严重损坏的数据库。在应用程序使用数据库提供服务时，请不要使用这些参数进行调试。

表 5-13 开发者选项

参数	参数说明	取值范围
-f s|i|n|m|h	禁止某种扫描和连接方法的使用。	取值范围： ● s：关闭顺序 ● i：索引扫描 ● n：关闭嵌套循环 ● m：融合（merge）连接 ● h：Hash连接
-n	主要用于调试导致服务器进程异常崩溃的问题。一般策略是通知所有其他服务器进程终止并重新初始化共享内存和信号灯。该选项指定gaussdb不重新初始化共享内存。	-
-O	允许修改系统表的内部。	-
-P	读系统表时忽略系统索引，但在修改表时仍然更新索引。	-
-t pa|pl|ex	打印与每个主要系统模块相关的查询记时统计。	-
-T	主要用于调试导致服务器进程异常崩溃的问题。该选项指定gaussdb通过发送SIGSTOP信号停止其他所有服务器进程，但是并不让它们退出。这样就允许系统程序员手动从所有服务器进程搜集内核转储。	-
-W NUM	指定一个新的服务器进程开始需要等待的秒数。	单位：秒
--localxid	使用本地事务ID，而不是全局事务ID。 **须知** 此选项仅用于gs_initdb。使用此选项可能会导致数据库不一致。	已存在的本地事务ID

下面的选项表5-14仅在单用户模式下可用。

版权所有 © 华为云计算技术有限公司

表 5-14 单用户模式选项

参数	参数说明	取值范围
--single	启动单用户模式。必须是命令行中的第一个选项。	-
DBNAME	要访问的数据库的名称。	字符串。默认为用户名。
-d 0-5	重新指定调试级别。	0-5
-E	回显所有命令。	-
-j	禁止使用新行作为语句分隔符。	-
-r <FILE_NAME>	将所有服务器标准输出和标准错误保存到文件filename中。	-
-L	服务器启动时加载共享库，共享库名称是参数shared_preload_libraries指定。	-

表 5-15 自启动模式选项

参数	参数说明	取值范围
--boot	启动自启动模式。必须是命令行中的第一个选项。	-
-r <FILE_NAME>	将所有服务器标准输出和标准错误保存到文件filename中。	-
-x NUM	指定一个新的服务器线程的类型。在初始化数据库时，会用到自启动模式，通过设置这个参数，启动不同线程来执行一些逻辑，正常情况下，不会用到，因为自启动模式很少会用到。	-

表 5-16 升级模式选项

参数	参数说明	取值范围
-u NUM	指定升级前数据库内核版本号。	-

除以上参数外，出现在postgres.conf中的所有参数，均可以用--guc_name=guc_value的方式在gaussdb启动时传给gaussdb。例如，--enable_fast_query_shipping=off

 版权所有 © 华为云计算技术有限公司

环境变量

- PGCLIENTENCODING：客户端使用的缺省字符编码。客户端可以独立地覆盖它。这个值也可以在配置文件里设置。
- PGDATESTYLE：运行时参数DateStyle的缺省值。现在反对使用该环境变量。
- TZ：服务器的时区。

错误处理

一个提到了semget或shmget的错误信息可能意味着需要重新配置内核，提供足够的共享内存和信号灯。可以通过降低shared_buffers值以减少GaussDB的共享内存的消耗，或者降低max_connections值减少GaussDB的信号灯的消耗。

如果发现类似“另外一个服务器正在运行”的错误信息，可以根据系统使用不同的命令：

```
ps ax | grep gaussdb
```

或

```
ps -ef | grep gaussdb
```

如果确信没有冲突的服务器正在运行，可以删除消息里提到的锁文件然后再次运行。

无法绑定端口的错误信息可能表明该端口已经被其他非GaussDB进程使用。如果终止gaussdb后又马上用同一端口号运行它，也可能得到错误信息。这时，必须多等几秒，等操作系统关闭了该端口再试。最后，如果使用了一个操作系统认为是保留的端口，也可能导致这个错误信息。例如：Unix版本认为低于1024的端口号是“可信任的”，因而只有Unix系统管理员可以使用它们。

须知

- 使用SIGKILL终止主进程会阻止gaussdb在退出前释放它持有的系统资源（例如共享内存和信号灯），会影响新的进程。
- 可以使用SIGTERM，SIGINT，SIGQUIT信号正常结束服务器进程。第一个信号将等待所有的客户端退出后才退出。第二个将强制断开所有客户端，而第三个将不停止立刻退出，导致在重启时的恢复运行。
- 信号SIGHUP将会重新加载服务器配置文件。它也可能给单个服务器进程发送SIGHUP信号，但是这通常是不明显的。
- 要取消一个正在执行的查询，向正在运行的进程发送SIGINT信号。

用法

启动一个单用户模式的服务器：

```
gaussdb --single -D /usr/local/pgsql/data other-options my_database
```

用-D给服务器提供正确的数据库目录的路径。同时还要声名已存在的特定数据库名称。

版权所有 © 华为云计算技术有限公司

须知

- 通常，独立运行的服务器把换行符当做命令输入完成字符；要想把一行分成多行写，必需在除最后一个换行符以外的每个换行符前面敲一个反斜杠。
- 如果使用了-j命令行选项，新行将不被当作命令结束符。此时服务器将从标准输入一直读取到EOF标志为止，然后把所有读到的内容当作一个完整的命令字符串看待，并且反斜杠与换行符也被当作普通字符来看待。
- 输入EOF（Control+D）即可退出会话。如果已经使用了-j则必须连续使用两个EOF才行。
- 单用户模式运行的服务器不会提供复杂的行编辑功能（比如没有命令历史）。单用户模式也不会做任何后台处理，像自动检查点。

示例

1. 用缺省值在后台启动gaussdb：
   ```
   nohup gaussdb >logfile 2>&1 </dev/null  &
   ```
2. 在指定端口(如1234)启动gaussdb，-D参数为gs_initdb初始化后的数据目录：
   ```
   gaussdb --coordinator -p 1234 -D /usr/local/pgsql/data
   ```
   使用gsql与服务器连接，用-p指定端口号：
   ```
   gsql -d gaussdb -p 1234
   ```

5.4 gs_backup

说明

系统内部工具，不支持单独调用。

背景信息

数据库实例部署成功后，在数据库运行的过程中，会遇到各种问题及异常状态。GaussDB提供了gs_backup工具帮助数据库实例备份、恢复重要数据、显示帮助信息和版本号信息。

前提条件

- 可以正常连接GaussDB数据库。
- 在进行还原时，需要保证各节点备份目录中存在备份文件，若备份文件丢失，则需要从其他节点进行复制（二进制文件需修改文件名中的节点名）。

语法

- 备份数据库主机
  ```
  gs_backup -t backup --backup-dir=BACKUPDIR [-h HOSTNAME] [--parameter] [--binary] [--all] [-l LOGFILE] [--ignore-miss YES|NO]
  ```
- 恢复数据库主机
  ```
  gs_backup -t restore --backup-dir=BACKUPDIR [-h HOSTNAME] [--parameter] [--binary] [--all] [-l LOGFILE][--ignore-miss YES|NO]
  ```
- 显示帮助信息
  ```
  gs_backup -? | --help
  ```


版权所有 © 华为云计算技术有限公司

- 显示版本号信息
  ```
  gs_backup -V | --version
  ```

参数说明

gs_backup参数可以分为如下几类：

- 备份数据库主机参数：
  - -h

    指定存储备份文件的主机名称。

    取值范围：主机名称。如果不指定主机名称，则分发到数据库实例。
  - --backup-dir=BACKUPDIR

    备份文件保存路径。
  - --parameter

    备份参数文件，不指定--parameter、--binary、--all参数时默认只备份参数文件。
  - --binary

    备份二进制文件。
  - --all

    备份二进制和参数文件。
  - --ignore-miss

    备份文件保存路径或备份文件不存在时忽略异常。
  - -t

    指定gs_backup的行为，备份时为backup。
  - -l

    指定日志文件及存放路径。

    默认值：$GAUSSLOG/om/gs_backup-YYYY-MM-DD_hhmmss.log
- 恢复数据库主机参数：
  - -h

    指定需要恢复主机的名称。

    取值范围：主机名称。如果不指定主机，则恢复数据库实例。
  - --backup-dir=BACKUPDIR

    恢复文件提取路径。
  - --parameter

    恢复参数文件，不指定--parameter、--binary、--all参数时默认只恢复参数文件。
  - --binary

    恢复二进制文件。
  - --all

    恢复二进制和参数文件。
  - --ignore-miss

    备份文件保存路径或备份文件不存在时忽略异常。
  - -t

 版权所有 © 华为云计算技术有限公司

指定gs_backup的行为，恢复时为restore。

– -l

指定日志文件及存放路径。

默认值：$GAUSSLOG/om/gs_backup-YYYY-MM-DD_hhmmss.log

- 调用链接口参数
  – --trace-id

  每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

  取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。
- 其他参数：
  – -?, --help

  显示帮助信息。
  – -V, --version

  显示版本号信息。

示例

- 使用gs_backup脚本备份数据库主机。
```
gs_backup -t backup --backup-dir=/opt/software/gaussdb/backup_dir -h plat1 --parameter
Parsing configuration files.
Successfully parsed the configuration file.
Performing remote backup.
Remote backup succeeded.
Successfully backed up cluster files.
```
- 使用gs_backup脚本恢复数据库主机。
```
gs_backup -t restore --backup-dir=/opt/software/gaussdb/backup_dir -h plat1 --parameter
Parsing configuration files.
Successfully parsed the configuration file.
Performing remote restoration.
Successfully restored cluster files.
```

5.5 gs_ctl

说明

集中式单节点不支持主备间操作。

功能介绍

gs_ctl是GaussDB提供的数据库服务控制工具，可以用来启停数据库服务和查询数据库状态。主要供数据库管理模块调用。

- 启动、停止、重启GaussDB节点。
- 在不停止数据库的情况下，重新加载配置文件（gaussdb.conf，gs_hba.conf）。
- 主备切换、主备状态查询、DN重建和重建状态查询。

参数说明

gs_ctl参数可分为如下几类：

 版权所有 © 华为云计算技术有限公司

- option参数，详细请参见表5-17。
- 公共参数，详细请参见表5-18。
- start和restart模式的参数，详细请参见表5-19。

须知

通常不建议用户直接通过gs_ctl命令来启动数据库实例，严重情况下可能导致集群双主、数据丢失等问题。请使用cm_ctl集群管理工具来进行实例的启动和主备切换，cm_ctl工具内部会进行多数派副本一致性校验，从而避免双主、数据丢失等问题的发生。

- stop和restart模式的参数，详细请参见表5-20。
- failover参数，详细请参见表5-18和表5-32。
- switchover模式的参数，详细请参见表5-21。
- build参数，详细请参见表5-22。
- member参数，详细请参见表5-24。
- changerole参数，详细请参见表5-25。
- setrunmode参数，详细请参见表5-26。
- fixleader参数，详细请参见表5-33。
- switchreplication参数，详细请参见表5-31。
- copy参数，详细请参见表5-27。

表 5-17 option 参数

参数	参数说明
init[db]	初始化DN。
start	启动DN。
restart	重启DN。
build	在需要重建的实例节点重建实例。
stop	停止DN。
reload	重载配置文件（gaussdb.conf, gs_hba.conf）。
status	显示DN运行状态。
promote	在指定的数据目录运行的备用服务器被命令退出故障恢复，并开始进行读写操作。
failover	在主DN异常时，将备DN切换为主DN。 **注意** 该参数属于高风险操作，请谨慎执行。如果集群有部署CM，请使用cm_ctl命令，若执行此命令可能会影响CM的仲裁，存在数据丢失风险。

版权所有 © 华为云计算技术有限公司

参数	参数说明
switchover	在主备DN正常时，出于维护的需要，将备DN切换为主DN，可保证切换过程中数据不丢失。
query	查询主备DN之间的状态信息。
notify	启动后再指定主备DN。
kill	给指定的进程发送信号量。
querybuild	查看DN的重建进度。
hotpatch	对DN执行热补丁操作。
finishredo	强制停止redo。 **注意** 该参数属于高风险操作，请谨慎执行。
copy	基于共享存储的双中心容灾模式，进行xlog日志复制。
stack	获取gaussdb的调用栈。
getlocalmaxlsn	获取指定目录里xlog可读最大位置。
decode	指定xlog目录进行解码。
startwalrcv	通过对日志预解析找到有效日志尾部，启动接收物理日志的walreceiver线程。
member	DCF模式下，增删单个DN副本。建议使用gs_om命令执行增删副本操作。
changerole	DCF模式下，变更单个DN节点角色。建议使用gs_om命令执行角色变更操作。
setrunmode	DCF模式下，设置集群运行模式为多数派或者少数派。建议使用cm_ctl命令设置运行模式。
fixleader	DCF模式下，设置定主功能。该操作仅支持多数派模式下生效。 **警告** 该参数属于高风险操作，请谨慎执行。定主成功后，集群无法执行switchover和failover操作。如果定主成功的主节点手工无法修复，可使用少数派命令强制升主，此操作会清除定主标记，且存在丢数据风险。
switchreplication	切换复制模式，DCF模式与quorum模式相互切换。建议使用gs_om命令执行切换。 **注意** 该参数属于高风险操作，请谨慎执行。
primary_demote	用于dorado主备数据库实例容灾倒换流程，将主DN降为备DN。

版权所有 © 华为云计算技术有限公司

表 5-18 公共参数

参数	参数说明	取值范围
-D, --pgdata=<DATA_DIR>	指定数据目录的位置。	<DATA_DIR>的取值必须为有效的数据目录。
-s, --silent	对promote，reload，restart，stop命令仅打印部分提示信息，对其他命令不生效。 打印信息如“gstl stopped ,xxx”等信息。不打印如“waiting for server to shut down，server stopped”等提示信息。	-
-t, --timeout=<SECS>	这个参数同时控制如下两个功能： 等待数据库启动或者关闭完成的最大秒数。	取值范围：整型（秒）。 默认值： 等待数据库启动或者停止完成：60秒。
-V , --version	打印gs_ctl的版本信息然后退出。	-
-w	用户执行启动或者关闭数据库的操作后，会等待事务操作的完成后再执行。	默认值：省略此参数时，默认会等待事务操作完成。
-W	用户执行启动或者关闭数据库的操作后，会强制完成启动或者关闭，不等待事务操作完成。	-
-M	-M后面需要跟SERVERMODE参数，表示在启动时指定数据库的启动模式。	SERVERMODE的取值范围： ● primary：本端以主机模式启动。 ● standby：本端以备机模式启动。 ● pending：本端处于等待状态，等待提升为主机或者备机。 **须知** 通常不建议用户直接通过gs_ctl命令来启动数据库实例，严重情况下可能导致数据库实例双主、数据丢失等问题。请使用cm_ctl数据库实例管理工具来进行实例的启动和主备切换，cm_ctl工具内部会进行多数派副本一致性校验，从而避免双主、数据丢失等问题的发生。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-T	-T 后面跟term，升主命令时会用此term作为主机term，build时会连接大于等于此term的主机进行build。	term的取值范围，无符号整型。
-q	build完成后不自动启动，需要调用者启动。	-
-d	打印更多的debug信息。	-
-L	查询lsn:XX/XX有效性并显示max_lsn。	-
-P PASSWORD	与-U参数配合使用，指定连接本地数据库的用户密码。 **说明** 当省略-U参数，且认证方法为trust时，-P参数无效。	-
-U USERNAME	指定连接数据库的用户。此参数只能与notify、query和querybuild参数配合使用。 对于域用户，使用DOMAIN\username格式。	取值范围：GaussDB中存在的用户。 默认值：省略此参数则使用与当前操作系统用户同名的用户。
-Z NODE-TYPE	指定当前连接数据库的运行模式。	NODE-TYPE的取值为single_node。
-?, -h, --help	显示关于gs_ctl命令行参数的帮助信息。	-

表 5-19 start 或 restart 共有参数

参数	参数说明	取值范围
-c, --core-file	允许gaussdb进程产生内核文件。	-
-l, --log=<FILE_NAME>	把服务器日志输出附加在<FILE_NAME>文件上。	<FILE_NAME>的取值为启动数据库服务的用户有权限的文件。 例如，data目录下的某文件。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-o OPTIONS	声明要直接传递给由GaussDB执行的gaussdb的命令行选项。 参数通常都用单或者双引号包围以保证它们作为一个整体传递。	gaussdb支持的参数。
-p PATH-TO-POSTGRES	声明gaussdb可执行文件的位置。	缺省位于gs_ctl自身所在目录，一般不需要此选项。

表 5-20 stop 和 restart 共有参数

参数	参数说明	取值范围
-m, --mode=MODE	声明关闭模式。	mode的取值： ● fast：不等待客户端中断连接，所有活跃事务都被回滚并且客户端都被强制断开，然后服务器将被关闭。 ● immediate：强行关闭，在下次重新启动的时候将导致故障恢复。 默认值：fast

表 5-21 switchover 参数

参数	参数说明	取值范围
-f	不等待客户端中断连接，所有活跃事务都被回滚并且客户端都被强制断开，然后服务器将被切换，且不做checkpoint。	无参数

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-m SWITCHOVER-MODE	声明双机主备实例正常时切换的切换模式。	mode的取值： ● smart：等待所有客户端中断连接后才切换。如果服务器处于紧迫等待，一旦所有的客户断开连接，恢复和复制流将会中断。备注：由于在数据库实例环境中，数据库实例管理连接、客户业务连接会导致smart模式几乎很难执行成功，因此在本版本中，该模式已经被废弃。 ● fast：不等待客户端中断连接，所有活跃事务都被回滚并且客户端都被强制断开，然后服务器将被切换。 默认值：fast
-T	-T 后面跟term，升主命令时会用此term作为主机term，switchover须指定term>当前term。	term的取值范围，无符号整型。

版权所有 © 华为云计算技术有限公司

表 5-22 build 参数

参数	参数说明	取值范围
-b, --mode=MODE	指定重建DN备机的模式。	mode的取值： ● full：通过全量镜像的方式重新同步DN主机的数据目录。 ● incremental：通过解析Xlog日志获取主备DN差异的数据进行增量修复备DN。 **警告** ● 增量重建适用于主备双主等因日志造成的不一致场景。 ● 备机数据文件损坏、数据目录丢失等故障通过增量重建的方式无法修复，此时可通过全量重建的方式重新修复备机。 ● 增量重建会全量复制ustore引擎的undo文件，当undo文件较多时，增量重建耗时会变长。 ● auto(不指定)：先增量，根据失败后是否可以再增量选择继续增量或者全量，三次增量失败后进行全量。 ● standby_full：使用备机重建备机或容灾场景下的首备重建灾备数据库实例内故障的备机。需要数据源备机或首备状态为Normal。 ● standby_incremental：使用备机增量重建备机或容灾场景下的首备增量重建灾备数据库实例内故障的备机。需要数据源备机或首备状态为Normal。 ● standby_auto：先进行备机增量重建备机，根据失败后是否可以再增量选择继续增量或者全量，三次备机增量失败后进行备机全量重建。需要数据源备机或首备状态为Normal。 ● copy_secure_files：在流式容灾场景下，获取指定节点数据目录下gs_secure_files目录对应内容。 ● copy_upgrade_file：在dorado容灾场景下，获取指定节点下升级所需指定文件。 ● cross_cluster_full：基于共享存储的同城双中心场景下的跨数据库实例全量build

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
		• cross_cluster_incremental：基于共享存储的同城双中心场景下的跨数据库实例增量build • copy_expdp_files：库级恢复场景下，获取指定节点辅助数据库目录下的导出文件用于导入。 默认值：auto
-r, --recvtimeout=INTERVAL	指定重建过程中备机日志接收等待主机响应的超时时间。	取值范围：整型（秒）。 默认值：120秒
-d	打印更多build过程信息	-
-C CONNECTOR	到源端DN的连接串，通过连接源端DN进行本节点上DN的重建。只支持备机到备机的连接。	连接串包含本节点待重建DN的IP和端口和源端DN的IP和端口，格式如下： 'localhost=Local DN IP localport=Local DN Port+1 remotehost=DN IP remoteport=DN Port+1'
--expdppath	指定copy_expdp_files模式下要复制文件的所在路径。	字符串。
--expdp_throttle	指定copy_expdp_files模式下的流控力度。	取值范围：整型（大于等于0）。 默认值：-1 未传入该参数，则流控不生效，取值value表示在复制过程中，每向硬盘写入一定次数后，sleep这段写入操作耗时的value%的时间。

build过程中的文件传输采用黑名单形式，控制部分文件（如配置文件、白名单文件等）不传，具体如下：

- 全量build时主实例不会发送给备实例的文件：
  - 以pgsql_tmp开头的文件
  - 以standby_read开头的文件
  - 参数log_directory配置的日志文件夹
  - backup_label
  - disable_conn_file
  - postmaster.pid
  - postmaster.opts
  - gs_gazelle.conf
  - pg_ctl.lock
  - gaussdb.conf

版权所有 © 华为云计算技术有限公司

  - gaussdb.cong.lock
  - gaussdb.conf.bak
  - gaussdb.conf.bak.old
  - gaussdb.state
  - gs_build.pid
  - disc_readonly_test
  - pg_rewind_bak
  - pg_rewind_filemap
  - full_upgrade_bak
  - pg_location
  - delay_xlog_recycle
  - delay_ddl_recycle
  - standby_inc_filemap
  - page_verifylogs
  - 参数ssl_cert_file配置的文件
  - 参数ssl_key_file配置的文件
  - 参数ssl_enc_cert_file配置的文件
  - 参数ssl_enc_key_file配置的文件
  - 参数ssl_ca_file配置的文件
  - 参数ssl_crl_file配置的文件
  - server.key.cipher
  - server.key.rand
  - client.crt
  - client.key
  - global/pg_dw开头的文件
  - global/pg_dw_single
  - global/pg_dw.build
  - global/config_exec_params
  - gs_hba.conf（仅跨集群build或从CN重建时不传输）
  - gs_ident.conf（仅跨集群build时不传输）
  - cn_drop_backup（仅从CN重建时不传输）
  - gs_local_barriers（仅跨集群build时不传输）
  - pg_replslot（dorado跨集群build时全部不传输）
  - pg_replslot文件夹下的物理复制槽
  - pg_replslot文件夹下的归档槽和备份槽（仅跨集群build时不传输）
  - core文件
- 增量build时不会修改的文件：
  - postmaster.pid
  - postmaster.opts
  - gaussdb.conf

 版权所有 © 华为云计算技术有限公司

- gaussdb.conf.lock
- gaussdb.conf.bak
- gaussdb.conf.old
- pg_ctl.lock
- build_completed.start
- build_completed.done
- backup_label
- client.crt
- client.key
- gs_build.pid
- gaussdb.state
- pg_errorinfo
- pg_location
- pg_copydir
- gs_local_barriers
- standby_inc_filemap
- config_exec_params
- delay_xlog_recycle
- delay_ddl_recycle
- pg_rewind_bak
- pg_rewind_filemap
- barrier_lsn
- disc_readonly_test
- pg_dw开头的文件
- cacert.pem
- server.crt
- server.key
- server.key.cipher
- server.key.rand
- pg_replslot文件夹下除逻辑槽以外的槽（dorado跨集群build时全部不传输）
- page_verifylogs
- pgsql_tmp

 版权所有 © 华为云计算技术有限公司

表 5-23 hotpatch 参数

参数	参数说明	取值范围
-a ACTION	指定需要执行的热补丁操作。	• load： 加载热补丁 • unload：卸载热补丁 • active：激活热补丁 • deactive：去激活热补丁 • info：显示热补丁信息 • list：显示热补丁列表
-n NAME	指定需要执行操作的热补丁名称，需要完整的路径名。	字符串

表 5-24 member 参数

参数	参数说明	取值范围
--operation	集中式DN数据库实例，DCF模式下，增删节点的操作。	• add： 向数据库实例的DCF节点配置中增加一个节点。 • remove：从数据库实例DCF节点配置中删除一个节点。
-u, --nodeid	新增节点的DCF NODE ID (DCF模式下，CM仲裁failover指定备机需要同步的LOGGER节点ID复用此字段， 0-无需同步)。	无符号整型
-i, --ip	新增节点与数据库实例DCF模块通信的IP。	字符串。
-e, --port	新增节点与数据库实例DCF模块通信的port。	整型。
[-D, --pgdata=]<DATA_DIR>	指定数据目录的位置。	<DATA_DIR>的取值：用户自定义。
-G	DN实例，DCF模式下节点角色的group值。	int范围 0 ~ MAX_INT
--priority	DN实例，DCF模式下节点角色的priority值。	int范围 0 ~ MAX_INT

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-R, --role	新增节点的角色（参数可选）。	• passive： passive角色。若不指定此参数默认为passive角色。 • logger：logger角色。 • cascade_follower: 级联备角色。

表 5-25 changerole 参数

参数	参数说明	取值范围
-R, --role	集中式DN数据库实例，DCF模式下节点角色。	• passive： passive角色。 • follower：follower角色。
-u, --nodeid	新增节点的DCF NODE ID。	无符号整型。
[-D, --pgdata=]<DATA_DIR>	指定数据目录的位置。	<DATA_DIR>的取值：用户自定义。

表 5-26 setrunmode 参数

参数	参数说明	取值范围
-v, --votenum	投票个数。	无符号整型。
-x, --xmode	运行模式（在少数派的DN节点上使用）。	• minority：少数派强起。 • normal：加回。
[-D, --pgdata=]<DATA_DIR>	指定数据目录的位置。	<DATA_DIR>的取值：用户自定义。

版权所有 © 华为云计算技术有限公司

表 5-27 copy 参数

参数	参数说明	取值范围
-Q	指定xLog日志复制方向。	• copy_from_local：从本地往共享存储上复制。 • copy_from_share：从共享存储往本地存储上复制。 • force_copy_from_local：强制从本地往共享存储上复制。

表 5-28 stack 参数

参数	参数说明	取值范围
-I	用于指定需要获取调用栈的线程的lwtid。	正整数

gs_ctl stack方式获取调用栈使用示例：

- 执行以下命令获取指定线程的调用栈。
  ```
  gs_ctl stack -D data_dir -I lwtid
  ```
  上述命令中-D data_dir用于指定需要获取调用栈的gaussdb进程的数据目录，-I lwtid用于指定目标线程的lwtid，lwtid可以ls /proc/pid/task/获取。具体步骤如下所示。
  a. 获取gaussdb进程号和数据目录。
     ```
     ps -ux | more
     USER        PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
     perfadm     308  9.3 10.1 8719348 1649108 ?     Sl   May20  58:58 /xxx/bin/gaussdb -u 92617 -D /xxx/datafile1/cluster/data1/dn1 -M pending
     ```
  b. 通过进程号获取lwtid，task目录下的目录名就是lwtid。
     ```
     ls /proc/308/task/
     1096  505  522  525  529  532  536  539  542  546  549  552  555  558  561  565  569  575  584
     833  923  926  929  932  935  938
     ```
  c. 获取指定lwtid的调用栈。
     ```
     gs_ctl stack -D /xxx/datafile1/cluster/data1/dn1 -I 1096
     [2022-05-21 10:52:51.354][24520][][gs_ctl]: gs_stack start:
     tid<140409677575616> lwtid<1096>
     __poll + 0x2d
     CommWaitPollParam::caller(int (*)(pollfd*, unsigned long, int), unsigned long) + 0x34
     int comm_socket_call<CommWaitPollParam, int (*)(pollfd*, unsigned long,
     int)>(CommWaitPollParam*, int (*)(pollfd*, unsigned long, int)) + 0x28
     comm_poll(pollfd*, unsigned long, int) + 0xb1
     ServerLoop() + 0x72b
     PostmasterMain(int, char**) + 0x329a
     main + 0x617
     __libc_start_main + 0xf5
     0x55cf616e7647
     [2022-05-21 10:52:51.354][24520][][gs_ctl]: gs_stack finished!
     ```
- 执行以下命令获取所有线程的调用栈。

版权所有 © 华为云计算技术有限公司

```
gs_ctl stack -D data_dir
```

上述命令中，-D data_dir用于指定需要获取调用栈的gaussdb进程的数据目录。具体步骤如下所示。

a. 获取gaussdb进程号和数据目录。

```
ps -ux | more
USER        PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
perfadm     308  9.3 10.1 8719348 1649108 ?     Sl   May20  58:58 /xxx/bin/gaussdb -u 92617 -
D /xxx/datafile1/cluster/data1/dn1 -M pending
```

b. 获取所有线程的调用栈。

```
$ gs_ctl stack -D /xxx/datafile1/cluster/data1/dn1
[2022-05-21 10:59:44.063][34511][][gs_ctl]: gs_stack start:
Thread 0 tid<140409677575616> lwtid<21045>
__poll + 0x2d
CommWaitPollParam::caller(int (*)(pollfd*, unsigned long, int), unsigned long) + 0x34
int comm_socket_call<CommWaitPollParam, int (*)(pollfd*, unsigned long,
int)>(CommWaitPollParam*, int (*)(pollfd*, unsigned long, int)) + 0x28
comm_poll(pollfd*, unsigned long, int) + 0xb1
ServerLoop() + 0x72b
PostmasterMain(int, char**) + 0x329a
main + 0x617
__libc_start_main + 0xf5
0x55cf616e7647

Thread 1 tid<140405343516416> lwtid<21060>
__poll + 0x2d
WaitLatchOrSocket(Latch volatile*, int, int, long) + 0x29f
SysLoggerMain(int) + 0xc86
int GaussDbThreadMain<(knl_thread_role)17>(knl_thread_arg*) + 0x45d
InternalThreadFunc(void*) + 0x2d
ThreadStarterFunc(void*) + 0xa4
start_thread + 0xc5
clone + 0x6d
```

此处省略剩余调用栈。

表 5-29 getlocalmaxlsn 参数

参数	参数说明	取值范围
-D	指定对应目录下lsn最小的xlog文件。	绝对路径

表 5-30 decode 参数

参数	参数说明	取值范围
-D	指定主DN的实例路径。	绝对路径
--startpos	指定解码的开始位置。	LSN值
--endpos	指定解码的结束位置。	LSN值
--xlogfolder	指定待解码xlog所在的文件夹。	绝对路径

 版权所有 © 华为云计算技术有限公司

表 5-31 switchreplication 参数

参数	参数说明	取值范围
-A	用于触发stream和DCF复制模式切换的起始和结束阶段的操作。	• dcf_start：触发切换到dcf复制模式起始操作。 • stream_start：触发切换到stream复制模式起始操作。 • dcf_finish：触发切换到dcf复制模式结束操作。 • stream_finish：触发切换到stream复制模式结束操作。

表 5-32 failover 参数

参数	参数说明	取值范围
-u, --nodeid	DCF模式下，指定备机需要同步的LOGGER节点ID。 仅支持DCF模式下一主一备一LOGGER组网下使用。	无符号整型。

表 5-33 fixleader 参数

参数	参数说明	取值范围
--fixval	DCF模式下，开启或者关闭定主功能，默认关闭。 仅支持DCF模式集群在多数派模式下使用，少数派不支持。	• 1：开启定主功能，开启生效后其他节点无法当选为主。 • 0：关闭定主功能。

5.6 gs_finishredo_retrieve

背景信息

强切会停止未redo的xlog的回放。在强切发生后，可从未回放的xlog里解码出对应的内容，进行参考。当cm的GUC参数enable_finishredo_retrieve打开的时候，可以通过此脚本实现自动化数据找回，并将结果保存在本地。

版权所有 © 华为云计算技术有限公司

语法

- 强切找回
  ```
  gs_finishredo_retrieve -t retrieve
  ```
- 显示帮助信息
  ```
  gs_finishredo_retrieve -? | --help
  ```
- 显示版本号信息
  ```
  gs_finishredo_retrieve -t cleanup
  ```

gs_finishredo_retrieve参数可以分为如下几类：

- 通用参数：

  -t

  gs_finishredo_retrieve命令参数类型。

  取值范围：retrieve，cleanup
- 调用链接口参数：

  --trace-id

  每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

  取值范围：1~128位的字符串，允许的字符为数字、字母和'-'。

说明

由于om和gs_ctl都支持免密，所以不需要额外权限。

相关功能

强切，gs_ctl finishredo, gs_ctl decode, gs_ctl getlocalmaxlsn

5.7 gs_expand

说明

集中式仅支持升副本操作，且不支持单独调用。

集中式单节点不支持该工具。

背景信息

GaussDB提供了gs_expand工具来帮助用户在业务发展和数据量增加的情况下，数据库实例性能和存储等方面出现瓶颈后，可以使用gs_expand工具增加主机来提升数据库实例的性能及存储。

扩容（升副本）前提条件

- 升副本操作要求在数据库实例normal状态下进行，degraded和unavailable状态不支持升副本。
- 升副本支持一主多备集群和单副本集群。
- 一主多备集群升副本操作支持从2副本集群最大扩展到5副本集群；单副本集群无此限制。
- 升副本如果是级联备，只支持集中式。

版权所有 © 华为云计算技术有限公司

- 升副本（增加ETCD）操作中需要满足以下数据库配置ETCD的节点数。数据库中配置ETCD的节点总个数因为多数派选主，应为单数。
  - 2AZ部署：AZ1中ETCD个数>=2；AZ2中ETCD个数>=1；ETCD总个数=3或5或7。
  - 3AZ部署：AZ1中ETCD个数>=2；AZ2中ETCD个数>=2；AZ3中ETCD个数>=1；ETCD总个数=5或7。
  - 指定参数增加只读节点前，需要将当前集群中的主备DN的sync_config_strategy参数设置为only_sync_node。
  - 单副本集群升副本至一主多备集群，配置节点时需要满足一主多备集群各实例配置的数量约束：

    集中式：ETCD 数量（3-7 奇数）个；CMS（2-8）个；DN >=2 个。

注意事项

- 新增主机机器至少满足数据库实例的安装要求（数据库实例安装要求，请参考《安装指南》中“软硬件环境要求”一章）。
- 新数据库实例整体性能取决于数据库实例中配置最低的机器。
- 新增主机的locale、encoding等需要和原有主机上保持一致。
- 新的主机上已经做了数据库实例安装路径和用户组规划。
- 使用环境字符集不支持中文类型（zh_CN.UTF-8），可通过locale命令查看。
- 升副本支持新增主机中只含DN，或者包含DN和ETCD，或者包含DN和CMS，或者三者都含场景；不支持不包含DN实例的场景进行升副本操作。
- 若新增主机硬件规格（内存、硬盘）低于现有数据库实例，需要逐步将性能较低的服务器替换为性能更好的服务器以满足数据库实例的性能要求。
- 升副本过程中，可能发生DN主备切换，请合理分配运维窗口期，避免对业务造成影响。

语法

- 数据库实例扩容
  ```
  gs_expand -t dilatation -X XMLFILE [-l LOGFILE] [--parallel-jobs=NUMBER] [--time-out=<SECS>] [--dilatation-mode=MODE] {--nodegroup-name=LCGROUPNAME | --addto-elastic-group}
  ```
- 数据重分布
  ```
  gs_expand -t redistribute  [--parallel-jobs=NUMBER] [--resource-level=f/h/m/l] [--failure-limit=NUMBER]
  [--build-redistb] [--redis-mode=MODE] [-l LOGFILE] [--priority-tables=FILE] [--exclude-tables=FILE] [--time-out=<SECS>] [--max-worker] [--max-dop] [--join-tables=FILE]
  ```
- 回滚
  ```
  gs_expand -t crashclean -X XMLFILE [-l LOGFILE]
  ```
- 重分布资源管控
  ```
  gs_expand -t resourcectl --resource-level=f/h/m/l [--max-worker][--max-dop]
  ```
- 扩容dilatation阶段扩容进度查询
  ```
  gs_expand -t progress -X XMLFILE
  ```
- 数据库实例升副本

 版权所有 © 华为云计算技术有限公司

说明

- 集中式仅支持此功能，且不支持单独调用。
- DCF模式下，升副本后在主DN节点执行如下命令查看dcf_replication_info中角色是否包含PASSIVE：
  ```
  gs_ctl query -D <<DATA_DIR>>
  ```
  上述命令中参数<DATA_DIR>是主DN的数据目录。
  若上述结果中包含PASSIVE，在主DN节点执行如下命令进行角色切换：
  ```
  cm_ctl changerole --role  FOLLOWER  -n <<NODE_ID>> -D <<DATA_DIR>>
  ```
  上述命令中参数<NODE_ID>是待切换角色的节点ID，<DATA_DIR>是待切换角色DN的数据目录。

```
gs_expand -t add_replication -X XMLFILE [--dn-guc]
```

- 重分布参数设置
  ```
  gs_expand -t parametercfg [--lockwait-timeout=MSEC] [--lockwait-interval=<SECS>] [--trylock-threshold=NUMBER] [--enable-cancel=Boolean] [--last-catchup-threshold=MSEC] [--catchup-times=NUMBER] [--write-error-mode=Boolean] [--catchup-query-dop=NUMBER] [--parallel-catchup-threshold=NUMBER] [--parallel-reindex-jobs=NUMBER] [--redis-retry-times=NUMBER][--redis-retry-interval=NUMBER]
  ```
- 显示帮助信息
  ```
  gs_expand -? | --help
  ```
- 显示版本信息
  ```
  gs_expand -V | --version
  ```

参数说明

gs_expand参数可以分为如下几类：

- 通用参数：
  - -t

    gs_expand命令参数类型。

    取值范围：dilatation，redistribute，crashclean，progress，resourcectl，add_replication和parametercfg。

    说明

    集中式仅支持add_replication。

  - -X

    安装配置文件路径。

    说明

    如果需要修改XML文件名称时，需要将修改后的XML文件分发到数据库实例的其他主机上。

    取值范围：clusterconfig.xml的文件路径。

    如果-t参数指定的是progress，-X参数传入的是扩容前原始数据库实例的XML配置文件 。

  - -l

    指定日志文件及存放路径。

    默认值：$GAUSSLOG/om/gs_expand-YYYY-MM-DD_hhmmss.log

  - --time-out=<SECS>

    超时等待时间。单位：s。

 版权所有 © 华为云计算技术有限公司

取值范围：≥43200，正整型。

默认值：43200

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

- 数据库实例扩容参数：
  - --parallel-jobs=NUMBER

    指定数据库实例扩容过程重建实例的并发任务个数。

    参考取值：正整数。

    数据库实例扩容默认值为20。

  - --dilatation-mode=MODE

    指定扩容执行模式。

    取值范围：

    - read-only：只读模式
    - insert：插入模式

    如果不指定，默认为read-only模式。

  - --nodegroup-name=LCGROUPNAME

    逻辑数据库实例模式时指定要扩容的逻辑数据库实例名称，指定该参数必须同时指定--dilatation-mode=insert参数。

  - --addto-elastic-group

    逻辑数据库实例模式时指定是否给弹性数据库实例扩容，指定该参数必须同时指定--dilatation-mode=insert参数。

  - --call-thread=

    脚本执行调用链ID，可以为空。

- 数据重分布参数：
  - --parallel-jobs=NUMBER

    指定数据重分布过程的并发执行任务个数。

    参考取值：正整数（1~8）。数据重分布推荐并发任务数为4。

    如果不指定--parallel-jobs参数：数据重分布默认值为1，为非并发执行方式。

  - --build-redistb

    创建pgxc_redistb表，但不进行重分布，此表中记录了本次重分布所有需要进行重分布的表。

  - --redis-mode=MODE

    指定数据重分布的执行模式。

    取值范围：

    - read-only：只读模式
    - insert：插入模式，正在进行重分布的表支持并发插入（insert）以及导入（Gds导入）。

 版权所有 © 华为云计算技术有限公司

如果不指定，默认为read-only模式。

- --priority-tables=FILE

  优先重分布模式，系统将优先执行用户在指定文件中所写表的重分布，默认不选此参数。

  取值范围：指定的文件路径，指定文件内容不能超过1000行。

  说明

  指定文件格式,每一行：database schema tablename

  例如：

  testdb public test

  testdb "pubLic" test （ schema为pubLic ）

  注意：不支持database、schema和tablename数据对象名称存在空格，schema和tablename名称中存在大写字母或者特殊字符时需要加双引号进行转义。

- --exclude-tables=FILE

  跳过重分布模式，系统将跳过用户在指定文件中所写表的重分布，默认不选此参数。

  与优先重分布模式同时使用时，不允许两个模式中指定的表有重叠。

  取值范围：指定的文件路径，指定文件内容不能超过1000行。

  说明

  指定文件格式,每一行：database schema tablename

  例如：

  testdb public test

  testdb "pubLic" test （ schema为pubLic ）

  注意：不支持database、schema、tablename数据对象名称存在空格，schema和tablename名称中存在大写字母或者特殊字符时需要加双引号进行转义。

- --translate-table=*FILE*

  数据迁移模式，系统会将用户指定的文件中的表完成数据迁移，默认不选此参数。

  取值范围：指定的文件路径，指定文件内容不能超过1000行。

  不支持从普通表或段页式表转换为hashbucket表，也不支持hashbucket表转换为其他类型的表。

  说明

  指定文件格式,每一行：database schema tablename transfertype ，例如：

  ```
  testdb public test t segment
  ```

  注意：不支持database、schema、tablename数据对象名称存在空格，schema和tablename名称中存在大写字母或者特殊字符时需要加双引号进行转义。

- --join-tables=*FILE*

  多表扩容模式，输入为具有join关系的用户表配置文件。

  取值范围：指定的文件路径，指定文件内容不能超过1000行。

 版权所有 © 华为云计算技术有限公司

说明

指定文件格式。每一行格式为：database schema1 tablename1 schema2 tablename2 ...（每行的多张表构成一个分组）。例如：

```
testdb public test1 public test2
testdb public test3 public test4 public test5
```

注意：

- 不支持database、schema、tablename数据对象名称存在空格，schema和tablename名称中存在大写字母或者特殊字符时需要加双引号进行转义。
- 每一行schema tablename必须成对出现，唯一确定一张表。
- 一张表不能出现在多个分组中，每组的表数量必须大于等于2，小于等于5。
- 如果存在一个schema里的tablename为其他schema的名字时，会正常处理这张表不会报错。

– --call-thread=

脚本执行调用链ID，可以为空。

– --failure-limit=NUMBER

跳过重分布失败表模式，系统在未超过用户指定失败数量前不会中断重分布，默认不选此参数。

取值范围：0 ~ 10

说明

指定此参数后，如果未达到失败表数量，重分布不会中断，并且在执行完成后报错，打印出失败表的列表。

– 当资源管控级别设置为f时，支持表内并发度（--max-worker）和表内扫描（--max-dop）可配置，通过调整以下参数进行：

▪ --max-worker

表内并发度。

取值范围：[1,8]

默认值：8

▪ --max-dop

单表query_dop参数值。

取值范围：[1,64]

默认值：4

- 资源管控参数：

--resource-level=RESLEVEL

指定数据重分布的资源等级。

取值范围：

– f：100%资源

– h：50%资源

– m：30%资源

– l：10%资源

默认值：l

 版权所有 © 华为云计算技术有限公司

说明

当资源管控参数设置为f时，进入快速扩容场景，会充分使用资源，通过并发操作提升数据重分布的性能。此时，如果shared_buffers参数值设置过小（小于推荐值），会影响此场景下的性能优化效果。

当资源管控级别设置为f时，支持表内并发度（--max-worker）和表内扫描（--max-dop）可配置，通过调整以下参数进行：

- --max-worker

  表内并发度。

  取值范围：[1,8]

  默认值：8

- --max-dop

  单表query_dop参数值。

  取值范围：[1,64]

  默认值：4

- 调用链接口参数

  --trace-id

  每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

  取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。

重分布参数设置接口参数：

- --lockwait-timeout

  锁超时时间。如果等锁线程在该时间内无法拿到锁，则报锁超时失败。

  取值范围：[1,1200000]

  单位：毫秒

  默认值：8000

- --lockwait-interval

  等锁步长。即使用CANCEL后等锁失败的等锁时间间隔，如果此轮等锁超时，间隔该参数指定的时长进行下一轮尝试。

  取值范围：[0,300]

  单位：秒

  默认值：0

- --trylock-threshold

  等锁尝试次数上限。如果到达该上限还因为等锁超时无法成功获取到锁，则此表此轮重分布失败。

  取值范围：[1,INT_MAX], -1

  单位：次

  默认值：-1（表示无上限）

- --enable-cancel

  是否允许用户业务快速失败。即是否允许达到等锁尝试次数上限时，快速失败用户业务从而保证重分布成功拿锁。

  取值范围：[false,true]

 版权所有 © 华为云计算技术有限公司

默认值：false

- --last-catchup-threshold

  最后一轮追增时间上限。即最后一轮完成delete+insert追增耗时。

  取值范围：[1000,60000]

  单位：毫秒

  默认值：5000

- --catchup-times

  追增轮次上限。即追增轮次的上限。

  取值范围：[1,INT_MAX], -1

  单位：次

  默认值：-1（表示无上限）

- --write-error-mode

  写报错模式。true表示写报错模式。false表示采用全阻塞模式进行最后一轮追增。

  取值范围：[false,true]

  默认值：true

- --catchup-query-dop

  追增并行扫描度。

  取值范围：[1,32]

  单位：个

  默认值：4

⚠ 注意

该参数设置过小可能导致追增性能下降，设置过大占用资源较多。

- --parallel-catchup-threshold

  多表重分布模式中开启并行追增所需的耗时阈值，一个分组中所有表追增耗时小于该值时，开启并行追增。

  取值范围：[1000,1800000]

  单位：毫秒

  默认值：300000

- --parallel-reindex-jobs

  同时临时表创建索引时的并发度，即多少个索引同时创建。

  取值范围：[1,64]

  单位：个

  默认值：2

- --redis-retry-times

  自动拉起重分布进程的次数上限，触发重分布进程主动退出时使用，拉起次数超出上限本轮重分布失败结束。

  对应参数：redis_retry_times

  取值范围：[0,100]

版权所有 © 华为云计算技术有限公司

单位：次

默认值：5

- --redis-retry-interval

  自动拉起重分布进程的时间间隔，触发重分布进程主动退出时使用。

  对应参数：redis_retry_interval

  取值范围：[0,36000]

  单位：秒

  默认值：600

增加副本参数：

- --dn-guc

  增加只读副本时，可指定--dn-guc参数来设置只读副本的默认GUC参数。

  使用方式：gs_expand -t add_replication -X /usr/local/tool/clusterConfig_rep.xml --dn-guc max_process_memory=10485760 --dn-guc wal_file_init_num=0

升副本操作：

```
gs_expand -t add_replication -X /usr/local/tool/clusterConfig.xml
Ready to restore local static config file.
Restore local static config file successfully.
Current cluster replication number is 3
Init global successfully.
Ready to create step directory: /usr/local/temp/add_replication_step
Successfully to create step directory: /usr/local/temp/add_replication_step
From 3 replication to 4 replication.
Start check new nodes in XML.
New nodes are added to the cluster ['10.30.58.219'].
Check new nodes successfully.
Stage begin: Start backup file stage.
Stage finish: Successfully backup config file stage.
Transport XML file successfully. /usr/local/tool/clusterConfig.xml
Successfully generate static config file on new instance nodes.
Stage begin: Start install app on new nodes stage.
Start clean new nodes.
Uninstalling applications on ['10.30.58.219'].
Successfully uninstalled applications on ['10.30.58.219'].
Clean new nodes [['10.30.58.219']] successfully.
Stage finish: Install stage perform successfully.
Start copying license file and so file.
Copy cert files and so files successfully.
Start compress tar ball.
Start decompress tar ball on remote new node.
Distribute tar file successfully.
Start to config new DN instances.
Initializing instances on new instances.
Init new instances successfully.
Start to build cm_agent on new nodes.
Transport cm_agent config file to new nodes successfully.
Start rebuild new datanode instance.
Rebuild new datanode instance successfully.
New instance config successfully.
Start to configure all data node instances stage.
Start configure gaussdb.conf on origin DN instance.
Stage finish: Successfully config gaussdb.conf file stage.
Cluster type is single-inst, no need to build pgxc_node.
Update dcf_config for dcf successfully.
Cm start reload configure.
Start new node begin.
New node Start successfully.
Wait for cluster to normal.
```

 版权所有 © 华为云计算技术有限公司

```
Starting to check replication information.
Cluster state is [Degraded].Please check repair cluster to Normal.
Cluster wait to build DN. Time spent 0.0886058807373s
Starting to check replication information.
Cluster state is [Degraded].Please check repair cluster to Normal.
Cluster wait to build DN. Time spent 10.1608610153s
Starting to check replication information.
Cluster state is [Degraded].Please check repair cluster to Normal.
Cluster wait to build DN. Time spent 20.2248330116s
Starting to check replication information.
Cluster state is [Degraded].Please check repair cluster to Normal.
Cluster wait to build DN. Time spent 30.2797150612s
Starting to check replication information.
Cluster state is [Normal].
Cluster state is : [Normal]
Start Clean backup files
Clean backup files successfully.
Add replication perform successfully.Current replication number is 5
```

5.8 gs_hotpatch

说明

集中式单节点不支持该工具。

功能介绍

gs_hotpatch是一个用于给节点内所有DN实例应用热补丁的工具。

前提条件

- 数据库实例已经安装完成。
- 数据库实例环境变量正确设置。
- gs_ctl工具已正确安装。

语法

```
gs_hotpatch -t [load|unload|active|deactive|info|list|recover|cleanLock|setLock] -n PATCH_NAME [-X
XMLFILE] [-S SOURCE_IP]
```


版权所有 © 华为云计算技术有限公司

参数说明

表 5-34 gs_hotpatch 参数说明

参数	参数说明	取值范围
-t	需要执行的热补丁动作。	• load：加载补丁。 • unload：卸载补丁。 • active：激活补丁。 • deactive：去激活补丁。 • info：查看补丁信息。 • list：查看补丁列表。 • recover：恢复因gs_install操作被损坏的补丁，仅om补丁使用。 • cleanLock：清除补丁操作中的锁文件，仅om补丁使用。 • setLock：设置补丁锁，仅om补丁使用。
-n	需要操作的热补丁名称。	包含热补丁路径和名称的字符串。
-X	选择数据库配置文件路径。	数据库配置文件路径。
-S	执行热补丁命令的源IP。	IP地址字符串，例如123.123.123.123。
-V --version	显示版本信息。	无参数，直接输入-V或者--version。
-? | --help	显示帮助信息。	无参数，直接输入-?或者--help。

注意

gs_hotpatch命令仅操作单个节点上实例，除特殊情况，不建议用户直接使用，操作热补丁建议使用gs_om -t hotpatch命令。

5.9 gs_install

说明

系统内部工具，不支持单独调用。

背景信息

数据库实例部署是一个复杂的过程，分布式数据库的安装和部署更是复杂。GaussDB提供了gs_install工具来帮助完成数据库实例的安装和部署。

版权所有 © 华为云计算技术有限公司

数据库实例安装部署，要求用户指定配置文件，配置文件中会指定程序安装路径、实例数据目录、主备关系、实例数和各实例的业务IP端口等信息。

前提条件

- 已成功执行前置脚本gs_preinstall。
- 用户需确保各个节点上的locale保持一致。
- 需要使用前置时设置的数据库实例用户进行安装操作。

语法

- 安装数据库实例

  gs_install -X *XMLFILE* [--gsinit-parameter="*PARAMETER*" [...]] [--gtminit-parameter="*PARAMETER*" [...]] [--cn-guc="*PARAMETER*" [...]] [--dn-guc="*PARAMETER*" [...]] [--alarm-component=*ALARMCOMPONENT*] [--time-out=*<SECS>*] [-l *LOGFILE*] [--product="*PARAMETER*"] [--call-thread="*PARAMETER*"][--cascade-guc="*PARAMETER*" [...] [--base64-guc]]

  说明

  执行gs_install脚本时，如果输入参数--autostart=no，则工具脚本在配置（config）步骤完成后退出，不会自动启动数据库实例，需要用户通过执行gs_om -t start命令手动启动。

- 显示帮助信息

  gs_install -? | --help

- 显示版本号信息

  gs_install -V | --version

参数说明

- -X

  数据库实例配置文件的路径。

  取值范围：XML文件的存储路径。

- -l

  指定安装日志文件及日志文件存放的路径。

  当既不明确指定-l，又不在XML文件中配置gaussdbLogPath时，默认值为“$GAUSSLOG/om/gs_install-YYYY-MM-DD_hhmmss.log”；

  当不明确指定-l，但在XML文件中配置了gaussdbLogPath时，默认值为“gaussdbLogPath的值、用户名和om/gs_install-YYYY-MM-DD_hhmmss.log”的组合。

- --gsinit-parameter=PARAM

  数据实例参数指定。例如：--gsinit-parameter="--encoding utf-8"，表示指定新数据库编码格式为utf-8。

  取值范围请参见**gs_initdb**的参数说明，其中对-D、-W、-C、-A、-X、-s参数的设置不生效，对-w与--pwpasswd参数不支持。

  说明

  执行gs_install脚本时，如果不指定初始用户密码，安装完成后密码为空，在执行其他操作前需要通过gsql客户端修改初始用户的密码。如果初始用户密码为空，则除修改密码外无法执行其他SQL操作以及升级、扩容和节点替换等操作。如果通过参数指定初始用户密码，为避免敏感信息泄露，请在执行完成后及时清理history中的历史操作记录。为安全起见，推荐用户在安装时不指定密码，而是在安装完成再修改初始用户的密码。

- --gtminit-parameter=PARAM

  预留接口，用于后续GTM实例参数指定。

 版权所有 © 华为云计算技术有限公司

> **说明**
>
> 集中式不支持该参数。

- --cn-guc=PARAM

  配置CN参数。

  取值范围请参见**gs_guc**的参数说明。

  > **说明**
  >
  > 集中式不支持该参数。

- --dn-guc=PARAM

  配置DN参数。

  取值范围请参见**gs_guc**的参数说明。

- --cascade-guc=PARAM

  配置级联备实例参数。

  取值范围请参见**gs_guc**的参数说明。

- --time-out=<SECS>

  启动超时等待时间。单位：s。

  若--time-out取值过小，有可能会因为需要启动的实例数过多而引起超时。若在超时时间内启动不成功，则会报错启动超时，但是数据库实例会在后台继续启动。可以等待一段时间后，查询数据库实例状态，检查数据库实例是否启动成功。

  取值范围：正整数，若无特殊需求，不需要配置该参数，系统会根据数据库实例规模自动调整超时时间。

- --alarm-component=alarm_component

  指定告警上报组件的绝对路径。

- --product

  指定所安装的数据库版本。

- --cms-guc

  配置CMS参数。

  取值范围请参见**gs_guc**的参数说明。

- --cma-guc

  配置CMA参数。

  取值范围请参见**gs_guc**的参数说明。

- --base64-guc

  使用base64转码后的GUC参数值。

- --autostart=yes|no

  用来指定数据库实例配置完成后是否启动。

  - yes：启动数据库实例。
  - no：不启动数据库实例。

- --trace-id

  每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

  取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。

 版权所有 © 华为云计算技术有限公司

- --call-thread=

  脚本执行调用链ID，可以为空。

- --init-password

  用于在安装时指定数据库密码，指定该参数后，安装过程中会提示输入密码，用户密码的复杂度及长度要求如下：

  - 包含大写字母（A-Z）的最少个数（根据GUC参数password_min_uppercase配置）。
  - 包含小写字母（a-z）的最少个数（根据GUC参数password_min_lowercase配置）。
  - 包含数字（0-9）的最少个数（根据GUC参数password_min_digital配置）。
  - 包含特殊字符的最少个数（根据GUC参数password_min_special配置，特殊字符的列表请参见**表5-35**）。
  - 密码的最小长度（根据GUC参数password_min_length配置）。
  - 密码的最大长度（根据GUC参数password_max_length配置）。

    说明

    上述GUC参数配置请联系管理员处理。
  - 至少包含上述四类字符中的三类。
  - 不能和用户名、用户名倒写相同，本要求为非大小写敏感。
  - 不能和当前密码、当前密码的倒写相同。

  表 5-35 特殊字符

编号	字符	编号	字符	编号	字符	编号	字符
1	~	2	!	3	@	4	#
5	$	6	%	7	^	8	&
9	*	10	(	11	)	12	-
13	_	14	=	14	+	16	\
17	|	18	[	19	]	20	{
21	}	22	;	23	:	24	,
25	<	26	>	27	.	28	/
29	?	-	-	-	-	-	-

  执行命令如下：

```
gs_install -X xxx.xml --init-password
```

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

 版权所有 © 华为云计算技术有限公司

示例

使用gs_install脚本进行数据库实例安装。

```
gs_install -X /opt/software/GaussDB_Kernel/clusterconfig.xml
Parsing the configuration file.
Check preinstall on every node.
Successfully checked preinstall on every node.
Creating the backup directory.
Successfully created the backup directory.
Installing the cluster.
Checking the installation environment on all nodes.
begin install Cluster..
Installing applications on all nodes.
Successfully installed APP.
begin init Instance..
Cluster installation is completed.
Configuring.
Deleting instances from all nodes.
Successfully deleted instances from all nodes.
Checking node configuration on all nodes.
Initializing instances on all nodes.
Are you sure you want to specify the password of the database (yes/no)? yes
Please enter the password
Password:
Please enter the password again
Password:
Start to init instance.
Updating instance configuration on all nodes.
Check consistence of memCheck and coresCheck on db nodes.
Successful check consistence of memCheck and coresCheck on all nodes.
Configuring gs_hba on all nodes.
Configuration is completed.
Starting cluster.
=========================================================================
Successfully started primary instance. Wait for standby instance.
=========================================================================
.
Successfully started cluster.
=========================================================================
cluster_state      : Normal
redistributing     : No
node_count         : 3
Coordinator State
    normal         : 0
    abnormal       : 0
GTM State
    primary        : 0
    standby        : 0
    abnormal       : 0
    down           : 0
Datanode State
    primary        : 1
    standby        : 2
    secondary      : 0
    building       : 0
    abnormal       : 0
    down           : 0

Successfully installed application.
```

5.10 gs_initdb

版权所有 © 华为云计算技术有限公司

5.10.1 gs_initdb 简介

背景信息

gs_initdb初始化数据库时，会创建数据库目录、生成系统表、创建默认数据库和模板数据库。

系统表

初始化数据库时会生成大量的系统表和视图，其中绝大部分都对任何数据库用户开放查看权限。

说明

pg_user_status、pg_auth_history系统表权限只对初始化数据库用户和sysadmin用户开放。

生成的数据库

- template1：是一个模板数据库，当以后再创建一个新的数据库时，template1数据库里的所有内容都会复制到新数据库中。通过gs_initdb的参数可以决定template1数据库的设置。
- template0：是GaussDB提供的最初始的备份数据库，当需要时可用template0作为模板生成“干净”的数据库。
- templatem: 是一个M-Compatibility类型模板数据库，用以创建M-Compatibility类型数据库。当需要开启M-Compatibility特性时，可以使用gs_initdb工具来创建。
- template_pdb：是一个用于内核多租特性的模板数据库，用以创建用户PDB。初始化创建需要使用--enable-mtd参数，默认不创建。
- postgres：是一个提供给用户、工具和第三方应用的缺省数据库。

5.10.2 使用指导

背景信息

在安装时，推荐使用-D参数调用gs_initdb初始化数据库。如果由于故障恢复等原因，需要重新初始化一个数据库，可以通过执行gs_initdb来完成。

- 尽管gs_initdb会尝试创建相应的数据目录，但可能没有权限执行此操作，因为要创建目录的父目录通常被root所拥有。如果要创建数据目录，首先用root用户创建一个空数据目录，然后用chown把该目录的所有权交给数据库用户。
- gs_initdb决定template1数据库的设置，而该设置将会成为其他数据库的默认设置。
- gs_initdb初始化数据库的缺省区域和字符集编码。字符集编码、字符编码排序（LC_COLLATE）和字符集类（LC_CTYPE，如大写、小写数字等）可以在创建数据库时独立设置。

操作步骤

步骤1 以操作系统用户omm登录GaussDB数据库任一主机。

步骤2 规划数据库目录。

 版权所有 © 华为云计算技术有限公司

1. 切换到root用户。
   ```
   su - root
   ```
2. 根据提示输入用户名和密码。

步骤3 进入“/opt/gaussdb”目录，并创建“data1”目录，即DN实例路径。

```
cd /opt/gaussdb
mkdir data1
```

1. 将data1目录的所有权分配给数据库用户omm，dbgrp是omm所属的用户组。
   ```
   chown omm:dbgrp data1
   ```
2. 退出root用户。
   ```
   exit
   ```

步骤4 执行gs_initdb命令初始化数据库，并根据提示设定密码。

```
gs_initdb /opt/gaussdb/data1 --nodename dn1
Enter new system admin password:[PASSWORD]
Enter it again:[PASSWORD]
```

gs_initdb的语法：

```
gs_initdb [OPTION]... [<DATA_DIR>]
```

OPTION参数请参见**命令参考**。

----结束

5.10.3 命令参考

命令参考

gs_initdb支持的常用参数和不常用参数请分别参见**表5-36**和**表5-37**。

表 5-36 常用参数说明

参数	参数说明	取值范围
-A, --auth=METHOD	指定本地用户连接数据库时的认证方法，即“gs_hba.conf”配置文件中host和local所在行的认证方法。 除非用户对本地用户都是信任的，否则不要使用默认值trust。	METHOD的取值： • trust • reject • md5（不安全的算法，为了兼容老版本而存在） • sha256 • sm3 默认值：trust

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--auth-host=METHOD	指定本地用户通过TCP/IP连接数据库时的认证方法，即：“gs_hba.conf”配置文件中host所在行的认证方法。 指定此参数则会覆盖-A参数的值。	METHOD的取值： ● trust ● reject ● md5（不安全的算法，为了兼容老版本而存在） ● sha256 ● sm3 默认值：trust
--auth-local=METHOD	指定本地用户通过Unix域套接字连接数据库时的认证方法，即“gs_hba.conf”配置文件中local所在行的认证方法。 指定此参数则会覆盖-A参数的值。	METHOD的取值： ● trust ● reject ● md5（不安全的算法，为了兼容老版本而存在） ● sha256 ● sm3 ● peer（仅用于local模式） 默认值：trust
-c, --enable-dcf	设置安装的节点为DCF模式。	-
[-D, --pgdata=]<DATA_DIR>	指定数据目录的位置。	<DATA_DIR>的取值：用户自定义。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
--nodename=NODENAME	初始化的节点名称。	不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。

-E, --encoding=ENCODING	为新数据库设置编码格式。	● 如果使用此参数，需要加上--locale选项指定支持此编码格式的区域。如果不加--locale选项，则采用系统默认的区域，如果系统默认区域的编码格式和用此参数指定的编码格式不匹配则会导致数据库初始化失败。
		● 如果不指定此参数，则使用系统默认区域的编码格式。系统默认区域和编码格式可以使用locale命令查看，如下：
		`omm@linux:~> locale|grep LC_CTYPE`
		`LC_CTYPE="en_US.UTF-8"`
		其中UTF-8表示系统默认区域的编码格式。
		不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--locale=LOCALE	为新数据库设置缺省的区域。 **须知** • 如果用户设置了数据库的编码格式，则用户选择区域的编码格式必须与用户设置的编码格式一致，否则数据库初始化会失败。 • 当指定了数据库的编码格式为GB18030_2022时，则参数locale的取值范围与GB18030保持一致。 • 当指定了数据库的编码格式为ZHS16GBK时，则参数locale的取值范围与GBK保持一致。	例如用户要将数据库编码格式初始化为GBK，可以采用如下步骤： 1. 用locale -a |grep gbk命令查看系统支持gbk编码的区域，如下： omm@linux:~> locale -a|grep gbk zh_CN.gbk zh_SG.gbk 2. 初始化数据库时加入--locale=zh_CN.gbk选项。 3. 不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
--dbcompatibility=DBCOMPATIBILITY	为新数据库设置默认db兼容性。 **说明** 不支持指定创建的默认数据库为M-compatibility数据库。	不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
--lc-collate=LOCALE --lc-ctype=LOCALE --lc-messages=LOCALE --lc-monetary=LOCALE --lc-numeric=LOCALE --lc-time=LOCALE	为新数据库设置指定范畴的区域。	各参数的取值必须是操作系统支持的值。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。 **说明** • 如果用户在数据库安装的时候没有指定--lc-collate参数，则--lc-collate参数的默认值为C。 • 当指定了数据库的编码格式为GB18030_2022时，则参数--lc-collate和参数--lc-ctype的取值范围与GB18030保持一致。 • 当指定了数据库的编码格式为ZHS16GBK时，则参数--lc-collate和参数--lc-ctype的取值范围与GBK保持一致。
--no-locale	和--locale=C等价。	-
--pwfile=FILE	gs_initdb时从文件FILE中读取数据库中系统管理员的密码。该文件的第一行将被当作密码使用。	FILE可以是“相对路径+文件”的形式，也可以是“绝对路径+文件”的形式。相对路径是相对当前路径的。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-U, --username=NAME	选择数据库系统管理员的用户名。	取值范围：正常的数据库用户。不能包括"|", ";", "&", "<", ">", "`", "\\", "!"这几个字符。 默认值：运行gs_initdb的操作系统用户。
-W, --pwprompt	gs_initdb时强制交互式输入数据库管理员的密码。	-
-w, --pwpasswd=PASSWD	gs_initdb时通过命令行指定管理员用户的密码，而不是交互式输入。	设置的密码要符合复杂度要求： ● 最少包含8个字符； ● 不能和用户名及当前密码（ALTER）相同，或和当前密码反序； ● 至少包含如下字符中的三类： – 大写字母（A-Z）。 – 小写字母（a-z）。 – 数字（0-9）。 – 特殊字符（限定为~!@#$%^&*()-_=+|[{}];:,<.>/?）。
-2, --pipeline	使用管道传输密码。	-
-C, --enpwdfiledir=DIR	gs_initdb时指定的经AES128加密算法加密过的密码文件所在目录。gs_initdb会将该目录下密码文件进行解密，同时把解密后的密码做密码复杂度校验，校验通过的话会将此密码作为用户的密码。 **说明** ● 加密密码文件需使用gs_guc工具生成：gs_guc encrypt -K *password* -D *Dir*。 ● 如果用户指定多个-w和-C参数，gs_initdb会将用户输入的最后一个-w或-C参数作为用户的需求，即输入密码的明文或经过AES128加密后的密码。	不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
-X, --xlogdir=XLOGDIR	声明事务日志存储的目录。 所设置的目录，必须满足运行GaussDB的用户有读写权限。	支持绝对路径和相对路径，相对路径是相对于数据目录的。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-S, --security	安全方式初始化数据库。	以-S方式初始化的数据库后，创建的数据库用户权限受到限制，默认不再具有public schema的使用权限。
--undostoragetype	指定UStore的回滚段使用的存储方式。 **说明** • UStore表与回滚段应统一使用页式。	取值范围： • segpage：段页式。预留参数，暂不支持。 • page：页式。 默认值：page。 例如：gs_initdb --undostoragetype='page';
--templatemcreate	创建模板数据库templatem。 **说明** • 执行该命令时数据库实例应处于关闭状态。 • gs_initdb创建templatem过程中发生故障，需手动执行drop清理后恢复集群状态，再下发重新创建，否则可能会导致集群异常。 • 不可在备机下执行该命令，可能会导致数据库实例异常。	该命令不可以和templatemdrop同时使用。
--templatemdrop	删除模板数据库templatem。 **说明** • 执行该命令时数据库实例应处于关闭状态。 • 该命令通常只用于当创建templatem失败时进行清理以便之后的再次创建。templatem一经创建，请不要删除。 • 不可在备机下执行该命令，可能会导致数据库实例异常。	该命令不可以和templatemcreate同时使用。
--enablem	初始安装时创建M-Compatibility模式数据库。 **说明** • 执行该命令时数据库实例应处于关闭状态。 • 该命令用于安装时创建M-Compatibility模式数据库。	• 不带enablem和disablem默认安装M-Compatibility模式数据库。 • 不可与disablem同时使用。

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
--disablem	初始安装时不创建M-Compatibility模式数据库。 **说明** • 执行该命令时数据库实例应处于关闭状态。 • 该命令用于安装时禁止创建M-Compatibility模式数据库。	• 不带enablem和disablem默认安装M-Compatibility模式数据库。 • 不可与enablem同时使用。

--enable-mtd	初始化时创template_pdb模板PDB。
	注意
	如果在数据库实例初始化时下发该参数，并且不为数据库实例指定密码，则在完成实例初始化后，为postgres数据库配置密码阶段，须使用初始用户额外执行以下步骤，为模板PDB配置密码：
	gaussdb=# ALTER PLUGGABLE DATABASE template_pdb OPEN;
	ALTER PLUGGABLE DATABASE
	gaussdb=# \c template_pdb
	Non-SSL connection (SSL connection is recommended when requiring high-security)
	You are now connected to database "template_pdb" as user "rdsAdmin".
	template_pdb=# ALTER role "rdsAdmin" PASSWORD '<password>';
	ALTER ROLE
	template_pdb=# \c postgres
	Non-SSL connection (SSL connection is recommended when requiring high-security)
	You are now connected to database "postgres" as user "rdsAdmin".
	gaussdb=# ALTER PLUGGABLE DATABASE template_pdb CLOSE;
	ALTER PLUGGABLE DATABASE
		不带enable-mtd参数时默认不创建template_pdb.

表 5-37 不常用参数说明

参数	参数说明	取值范围
-d, --debug	从初始化后端打印调试输出信息。初始化后端是gs_initdb用于创建系统表的程序。	-

版权所有 © 华为云计算技术有限公司

参数	参数说明	取值范围
-L DIRECTORY	gs_initdb初始化数据库所需要的输入文件的路径。通常是不必要的。如果需要明确声明的话，程序会提示输入。	初始化数据库所需的输入文件路径。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
-n, --noclean	不指定该参数时，当gs_initdb发现一些错误妨碍它完成创建数据库的工作时，它将在检测到不能结束工作之前将其创建的所有文件删除。这个选项禁止任何清理动作，因而对调试很有用。	-
-s, --show	显示内部设置。	-
-H, --host-ip	节点node_host初始化。	不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。
-V, --version	打印gs_initdb版本信息，然后退出。	-
-?, --help	显示关于gs_initdb命令行参数的帮助信息，然后退出。	-
-g	针对共享存储的跨AZ容灾数据库实例的初始化，指定共享xlog日志文件的路径。	指定的共享日志文件路径。不能包括"|", ";", "&", "$", "<", ">", "`", "\\", "!"这几个字符。

5.11 gs_persist

功能介绍

gs_persist工具用来在共享存储设备上实现注册、预留和抢占写权限等功能。该工具是系统内部工具，由DN实例自动调用，不建议用户使用。

前提条件

确保存储设备支持SCSI-3协议。

语法

- 执行gs_persist命令：。
  ```
  gs_persist [DEVICEPATH] [KEY] [ACTION]
  ```
- 显示帮助信息。
  ```
  gs_persist
  -? | --help
  ```


版权所有 © 华为云计算技术有限公司

参数说明

- DEVICEPATH
  共享存储设备的路径。
- KEY
  指定的注册字符串，不同机器需使用不同的字符串作为KEY。
- ACTION
  指定执行的操作：
  - -I：查询预留信息。
  - -R：进行预留抢占。
  - -C：清理预留。
- -?,--help
  打印帮助信息。

5.12 gs_postuninstall

背景信息

GaussDB提供了gs_postuninstall工具来帮助清理准备数据库实例环境阶段所做配置，使得卸载后的环境得到清理。

前提条件

- 数据库实例卸载执行成功。
- 在全量环境清理的时候，该工具会自行创建root互信。
- 只能使用root用户执行gs_postuninstall命令，该接口出于安全考虑，前置完成后会自动删除。
- 如果gs_preinstall命令执行时指定了“--sep-env-file=MPPRCFILE”参数，则执行本命令之前，需要export 环境变量分离文件，source环境变量分离文件的global文件，该场景下只支持-L模式。例如：

  export MPPDB_ENV_SEPARATE_PATH=/opt/huawei/env && source /etc/gauss_env_file_global && gs_postuninstall -X /var/chroot/usr/local/package/clusterConfig.xml -U omm -L

  说明

  如果使用om_agent进行安装，则：

  - 该后置功能文件需要重新解压om包获取，om包位于GaussDB-Kernel_xxx_64bit_Om.tar.gz下的script目录中。
  - 安装时默认使用了“--sep-env-file=MPPRCFILE”参数，此时默认环境变量文件为/home/User/gauss_env_file，XML配置文件位于$GPHOME下，该场景下只支持-L模式。

  以omm为集群用户举例：

  ```
  source /home/omm/gauss_env_file && source /etc/gauss_env_file_global && source /home/omm/.bashrc && gs_postuninstall -X $GPHOME/clusterConfig.xml -U omm -L
  ```


版权所有 © 华为云计算技术有限公司

语法

- 数据库实例清理用户、用户组以及虚拟IP清理
  ```
  gs_postuninstall -U USER -X XMLFILE [-L] [--delete-user] [--delete-group] [--delete-virtual-IP] [-l LOGFILE] [--sep-env-file=MPPRCFILE]
  ```
- 显示帮助信息
  ```
  gs_postuninstall -? | --help
  ```
- 显示版本号信息
  ```
  gs_postuninstall -V | --version
  ```

参数说明

- -U

  运行数据库实例的操作系统用户名。

  取值范围：字符串，要符合标识符的命名规范。

- -X

  数据库实例配置文件路径。

  取值范围：XML文件的存储路径。

- -L

  只清理本主机的环境。

  如果数据库实例内某主机做单机环境清理后，数据库实例不能再做全量环境清理。

- --delete-user

  删除-U参数指定的操作系统用户。

  如果在redhat|euleros|centos|oracle任一系统环境下，且用户名与用户组名相同，选择此项必须指定--delete-group参数。

- --delete-group

  删除操作系统用户所在的用户组（选择此选项必须指定--delete-user参数）。

- --delete-virtual-IP

  删除数据库实例配置文件中的虚拟IP。

- -l

  指定日志文件名及路径。在内部会自动给日志名添加一个时间戳。

  - 当既不明确指定-l，又不在XML文件中配置gaussdbLogPath时，默认值为“$GAUSSLOG/om/gs_postuninstall-YYYY-MMDD_hhmmss.log”。
  - 当不明确指定-l，但在XML文件中配置了gaussdbLogPath时，默认值为“gaussdbLogPath/用户名/om/gs_postuninstall-YYYY-MMDD_hhmmss.log”。

  说明

  由于在执行gs_postuninstall后，系统会自动删除数据库实例相关目录（包含$GAUSSLOG目录）。因此建议用户通过该参数指定日志文件到非GaussDB数据库相关路径。

- --sep-env-file=MPPRCFILE

  保存环境变量的文件。如果指定该参数，表示将使用环境变量分离的版本。

- -?, --help

  显示帮助信息。

 版权所有 © 华为云计算技术有限公司

- -V, --version
  显示版本号信息。

5.13 gs_preinstall

说明

系统内部工具，不支持单独调用。

背景信息

GaussDB提供了gs_preinstall工具来帮助完成数据库实例的环境配置，以保证数据库实例安装的正常进行。

注意事项

- 用户需要检查上层目录权限，保证安装用户对安装包和配置文件目录读写执行的权限。
- XML文件中各主机的名称与IP映射配置正确。
- 只能使用root用户执行gs_preinstall命令。出于安全考虑，线下场景用户需自行保证数据库实例工具临时存放的包路径下的文件属主为root，权限为700，若执行gs_preinstall之前，临时包路径下用户已创建文件，则在执行gs_preinstall后请手动修改该文件权限。该工具不支持独立调用，出于安全考虑，前置完成后会自动删除。
- 出于安全考虑，数据库实例工具临时存放的包路径与xml文件中“gaussdbToolPath”路径不可配置为一个路径。
- 当为数据库实例用户设置密码时，密码不可包含如下特殊字符：“;”、“|”、“$”。
- gs_preinstall指定的用户不支持使用-d手动选定家目录，仅支持系统默认创建的/home下和用户同名的家目录。
- 本产品支持的操作系统版本列表如下：
  x86服务器支持EulerOS V2.0SP5（intel）、EulerOS V2.0SP10（intel、hygon）、Kylin V10 SP2（intel、hygon）、UnionTech V20 (intel、hygon)、Huawei Cloud EulerOS 2.0 (intel、hygon)、Suse12.5、EulerOS V2.0SP12（intel）。
  ARM服务器支持EulerOS V2.0SP9（kunpeng）、EulerOS V2.0SP10（kunpeng、phytium）、Kylin V10 SP3（kunpeng、phytium）、Kylin V10 SP2（kunpeng、phytium）、Kylin V10 SP1（kunpeng、phytium）、UnionTech V20 (kunpeng、phytium)、Huawei Cloud EulerOS 2.0 (kunpeng、phytium)、、EulerOS V2.0SP12（kunpeng）、BCLinux-for-Euler release 21.10 (LTS-SP2)（kunpeng）。

  说明

  操作系统只支持英文操作系统。

语法

- 准备数据库实例环境（root用户进入项目script目录下，执行gs_preinstall脚本）。
  ```
  gs_preinstall -U USER -G GROUP -X XMLFILE --alarm-type=ALARMTYPE [--alarm-server-addr=HOSTADDR] [-L] [--skip-os-set] [--env-var="ENVVAR" [...]] [--sep-env-file=MPPRCFILE] [--skip-hostname-set] [-l LOGFILE] [--non-interactive] [-M MOUNTDIR] [--agent-mode] [--libnet-set] [--dorado-mode] [--trace-id] [--ddb_type]
  ```
- 显示帮助信息（root用户进入项目script目录下执行）。

 版权所有 © 华为云计算技术有限公司

```
./gs_preinstall -? | --help
```

- 显示版本号信息（root用户进入项目script目录下执行）。

```
./gs_preinstall -V | --version
```

参数说明

- -U

  运行数据库实例的操作系统用户名。

  取值范围：字符串，要符合标识符的命名规范。

  说明

  配置-U参数值时，字符串中不能包含"|"，";"，"&"，"$"，"<"，">"，"`"，"\\"，"'"，"\"，"{"，"}"，"("，")"，"["，"]"，"~"，"*"，"?"特殊字符。

- -G

  运行数据库实例的操作系统用户的群组名。

  取值范围：字符串，要符合标识符的命名规范。

- -X

  数据库实例配置文件路径。

  取值范围：xml文件的存储路径。

- --alarm-type=ALARMTYPE

  告警类型。

  取值范围：1、5

  - 取值为1表示使用华为FusionInsight的告警管理工具管理告警。
  - 取值为5表示告警将写入system_alarm日志中，此方式只支持数据库实例管理组件和gtm组件上报告警。

- --alarm-server-addr=HOSTADDR

  系统日志服务器端IP。

  不指定该参数，则不配置系统日志参数。

- -L

  指定该参数则仅对当前节点进行预安装。适用于不允许使用root互信的场景和设置安全加固参数PermitRootLogin为no的场景。需要满足如下前提条件：

  - 安装包和数据库实例配置文件已提前上传到所有主机；
  - 数据库实例安装用户已提前创建好并建立好互信关系；
  - 用户需要自己保证各台主机上所执行命令的一致性。

- -M

  指定Cgroups组件的mount路径。如果第一次进行预安装时已经挂载成功，再次进行预安装时将不会重新挂载到指定的mount路径。

- --skip-os-set

  是否设置操作系统参数。默认设置系统参数。如果指定该参数则不设置。

- --dorado-mode

  是否为dorado双数据库实例。默认不设置该参数。

- --libnet-set

  是否设置用户态网络，目前支持线下部署集中式欧拉2.9ARM机器数据库实例。默认不设置改参数。

 版权所有 © 华为云计算技术有限公司

- --env-var="ENVVAR"

  配置普通用户环境变量。可以指定多个。

  取值范围：字符串。

  说明

  不能指定数据库实例默认创建的环境变量。默认创建的环境变量见**表5-38**。

- --sep-env-file=MPPRCFILE

  保存环境变量的文件。如果指定该参数，表示将使用环境变量分离的版本。

  说明

  不能指定环境变量文件为数据库实例相关目录或者和这些目录同名，不能指定环境变量文件所在目录为“gaussdbToolPath”任何一层目录，不能和数据库实例用户的目录相同。

- --skip-hostname-set

  是否将xml配置文件中主机名与IP的映射关系写入“/etc/hosts”文件中。默认写入，如果指定该参数则不写入。

- -l

  指定日志文件名及路径。在内部会自动给日志名添加一个时间戳。

  默认值：“$GAUSSLOG/om/gs_preinstall.log”

  - 当既不明确指定-l，又不在XML文件中配置gaussdbLogPath时，默认值为“$GAUSSLOG/om/gs_preinstall-YYYY-MM-DD_hhmmss.log”。
  - 当不明确指定-l，但在XML文件中配置了gaussdbLogPath时，默认值为gaussdbLogPath的值、用户名和“om/gs_preinstall-YYYY-MM-DD_hhmmss.log”的组合。

- --non-interactive

  指定预安装模式。

  - 当不指定该参数时，则为安全交互模式，在此模式下用户需要人机交互输入密码。
  - 当指定该参数时，为非交互模式，不需要进行人机交互。

  说明

  交互式输入密码的过程中，密码不能包含";"，"'"，"$"3个特殊字符。

- --agent-mode

  是否使用代理通讯方式，默认不设置改参数。

- --trace-id

  - 每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

    取值范围：1-128位的字符串，允许的字符为数字，字母和"-"。

- --ddb_type

  目标ddb_type切换类型，取值范围ETCD|DCC，仅在ETCD/DCC切换场景中的前置使用。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

 版权所有 © 华为云计算技术有限公司

表 5-38 GaussDB 默认创建的环境变量

环境变量名称	说明
MPPDB_ENV_SEPARATE_PATH	数据库实例环境变量分离文件路径。
GPHOME	数据库实例工具目录。
PATH	数据库实例工具脚本目录。
LD_LIBRARY_PATH	数据库实例引用第三方动态库路径。
PYTHONPATH	python软件路径。
GAUSS_WARNING_TYPE	告警类型。
GAUSSHOME	数据库实例安装路径。
GAUSS_VERSION	数据库实例版本号。
PGHOST	数据库实例用户的临时目录路径。
GS_CLUSTER_NAME	数据库实例名称。
DORADO_REARRANGE	数据库dorado双实例模式专用，区分其他部署模式。
GAUSSLOG	数据库实例日志路径。
GAUSS_ENV	数据库实例环境变量标识。
UNPACKPATH	数据库实例工具临时存放的包路径。
BACKUP_USER	数据库实例备份恢复用户（记录rdsBackupUser，agent模式专有）。
PGUSER	数据库实例管理员用户（记录rdsAdminUser，agent模式专有）。

说明

该配置在数据库实例的每台机器上都需要修改。

示例

切换root用户，进入script目录下，使用非交互方式进行数据库实例环境准备：

```
./gs_preinstall -U omm -G dbgrp -X /opt/software/GaussDB_Kernel/clusterconfig.xml --alarm-type=5 --non-interactive
Parsing the configuration file.
Successfully parsed the configuration file.
Installing the tools on the local node.
Successfully installed the tools on the local node.
Distributing package.
Successfully distributed package.
Installing the tools in the cluster.
Successfully installed the tools in the cluster.
Checking hostname mapping.
Successfully checked hostname mapping.
Checking OS version.
Successfully checked OS version.
```

 版权所有 © 华为云计算技术有限公司

```
Creating cluster's path.
Successfully created cluster's path.
Set and check OS parameter.
Setting OS parameters.
Successfully set OS parameters.
Warning: Installation environment contains some warning messages.
Please get more details by "/package_r8/sudo/gs_checkos -i A -h host179174,host179175,host179176".
Set and check OS parameter completed.
Preparing CRON service.
Successfully prepared CRON service.
Preparing SSH service.
Successfully prepared SSH service.
Setting user environmental variables.
Successfully set user environmental variables.
Configuring alarms on the cluster nodes.
Successfully configured alarms on the cluster nodes.
Setting the dynamic link library.
Successfully set the dynamic link library.
Setting Cgroup.
Successfully set Cgroup.
Setting finish flag.
Successfully set finish flag.
Preinstallation succeeded.
```

使用交互模式进行数据库实例环境准备：

```
./gs_preinstall -U omm -G dbgrp -X /opt/software/GaussDB_Kernel/clusterconfig.xml --alarm-type=5
Parsing the configuration file.
Successfully parsed the configuration file.
Installing the tools on the local node.
Successfully installed the tools on the local node.
Are you sure you want to create trust for root (yes/no)? yes
Please enter password for root.
Password:
Creating SSH trust for the root permission user.
Checking network information.
All nodes in the network are Normal.
Successfully checked network information.
Creating SSH trust.
Creating the local key file.
Successfully created the local key files.
Appending local ID to authorized_keys.
Successfully appended local ID to authorized_keys.
Updating the known_hosts file.
Successfully updated the known_hosts file.
Appending authorized_key on the remote node.
Successfully appended authorized_key on all remote node.
Checking common authentication file content.
Successfully checked common authentication content.
Distributing SSH trust file to all node.
Successfully distributed SSH trust file to all node.
Verifying SSH trust on all hosts.
Successfully verified SSH trust on all hosts.
Successfully created SSH trust.
Successfully created SSH trust for the root permission user.
Setting pssh path
Successfully set core path.
Distributing package.
Begin to distribute package to tool path.
Successfully distribute package to tool path.
Begin to distribute package to package path.
Successfully distribute package to package path.
Successfully distributed package.
Are you sure you want to create the user[omm] and create trust for it (yes/no)? yes
Preparing SSH service.
Successfully prepared SSH service.
Installing the tools in the cluster.
Successfully installed the tools in the cluster.
Checking hostname mapping.
Successfully checked hostname mapping.
```

 版权所有 © 华为云计算技术有限公司

```
Creating SSH trust for [omm] user.
Please enter password for current user[omm].
Password:
Checking network information.
All nodes in the network are Normal.
Successfully checked network information.
Creating SSH trust.
Creating the local key file.
Successfully created the local key files.
Appending local ID to authorized_keys.
Successfully appended local ID to authorized_keys.
Updating the known_hosts file.
Successfully updated the known_hosts file.
Appending authorized_key on the remote node.
Successfully appended authorized_key on all remote node.
Checking common authentication file content.
Successfully checked common authentication content.
Distributing SSH trust file to all node.
Successfully distributed SSH trust file to all node.
Verifying SSH trust on all hosts.
Successfully verified SSH trust on all hosts.
Successfully created SSH trust.
Successfully created SSH trust for [omm] user.
Checking OS version.
Successfully checked OS version.
Creating cluster's path.
Successfully created cluster's path.
Set and check OS parameter.
Setting OS parameters.
Successfully set OS parameters.
Warning: Installation environment contains some warning messages.
Please get more details by "/data1/omm/ssh_test/script/gs_checkos -i A -h
pekpomdev00062,pekpomdev00061,pekpomdev00060".
Set and check OS parameter completed.
Preparing CRON service.
Successfully prepared CRON service.
Setting user environmental variables.
Successfully set user environmental variables.
Configuring alarms on the cluster nodes.
Successfully configured alarms on the cluster nodes.
Setting the dynamic link library.
Successfully set the dynamic link library.
Setting pssh path
Successfully set pssh path.
Setting Cgroup.
Successfully set Cgroup.
Set ARM Optimization.
Successfully set ARM Optimization.
Setting finish flag.
Scene: non-huaweiyun
Successfully set finish flag.
Preinstallation succeeded.
```

5.14 gs_redis

说明

集中式不支持该工具。

版权所有 © 华为云计算技术有限公司

参数说明

表 5-39 gs_redis 参数说明

参数	说明
-A password	连接数据库的默认用户密码。
-p port	执行重分布的CN的端口号。
-d database	连接的数据库名称，默认为postgres（大小写敏感）。
-u username	连接数据库的用户名。
-v	创建pgxc_redistb表，用于存储需要重分布的表。
-V	版本信息。
-r	重新执行重分布，如果pgxc_redistb没有创建则重新创建。
-t timeout	超时时间。
-j jobs	指定重分布的并发数。
-m mode	运行不同的数据分布模式。 ● read-only（默认值） ● insert
-P filepath	优先重分布模式，系统将优先执行用户在指定文件中所写表的重分布，默认不选此参数。
-W	非默认用户连接数据库时使用的密码
-x filepath	指定排他文件路径，文件中记录不需要重分布的用户表列表。文件中每行按照“database名称、schema名称和table名称”的格式指定不需要重分布的用户表信息，默认不选此参数。
-s num	所允许的重分布失败表的数量上限值，默认值是0。
-y filepath	指定需要迁移的文件，文件内记录需要转换存储类型的表，文件中每行按照“database名称、schema名称、table名称、transfer类型”的格式指定需要迁移的用户表信息，默认不选此参数。
-a	指定重分布是否支持重试。
-i	指定测试用例文件目录。
-e	用于退出、恢复测试，该参数仅用于测试。
--help	帮助信息。

版权所有 © 华为云计算技术有限公司

5.15 gs_redis_bucket

说明

集中式不支持该工具。

gs_redis_bucket工具是GaussDB提供的用来对扩容后的集群中的hashbucket表执行数据重分布的工具，主要由OM脚本来调用完成数据重分布。

参数说明

表 5-40 gs_redis_bucket 参数说明

参数	说明
-A password	连接数据库的管理员用户密码。
-p port	执行hashbucket重分布的CN的端口号。
-u username	连接数据库的用户名。
-V	版本信息。
--help	帮助信息。

5.16 gs_rescue

背景信息

gs_rescue是用于在极端场景下抢救客户数据，将磁盘数据文件输出为COPY命令可识别的数据格式，可将数据导入新数据库实例，供客户选择是否使用。

版权所有 © 华为云计算技术有限公司

⚠ 注意

- 仅限当数据库实例无法启动，而且无备份、无备机、无其它技术手段使数据库恢复正常情况下使用。
- 为防止数据泄露，用户需要合理设置抢救数据导出文件的权限，保证权限最小化。
- 仅解析落盘数据。
- 离线解析数据无法保证数据一致性。
- 在解析ustore表数据文件时，若tuple总长度为24字节(等于MIN_UHEAP_TUPLE_SIZE)，当前无法保证数据解析的准确性，会有相关提示"Warning: due to current limitation, for the tuple which size is MIN_UHEAP_TUPLE_SIZE(24), the parsed data may be inaccurate."
- 不解析xlog/undo内数据。
- 数据库应处于离线状态。
- 仅支持astore/ustore普通heap表、astore段页式、unlogged表、分区表、二级分区表、ilm表的数据解析，不支持临时表、透明加密表、全密态表、压缩表、内存表。
- 仅支持常用数据列类型。
- 数据库数据文件需可正常访问。
- 在普通机械盘且IO非瓶颈情况下，解析速度大于20MB/s。
- gs_rescue/gs_rescue_tui解析或恢复的M-Compatibility模式数据文件仅支持在M-Compatibility模式数据库下进行数据使用，其他模式下数据文件不兼容。
- M-Compatibility模式数据库下仅支持ASTORE/USTORE普通heap表、astore段页式、分区表及二级分区表的数据解析。
- M-Compatibility模式数据库下解析含有列类型为SET类型、ENUM类型的表需给出完整的类型定义，详情请参考示例五。

语法

```
gs_rescue [-abcdhiktxy] [-R startblock [endblock]] [-D attrlist] [-E segpageoid] [-T toast_relation_filenode] [-s segsize] [-n segnumber] file
```

⚠ 注意

- 使用解析数据类型参数的顺序需要和创建表时数据类型的顺序一致。
- file文件参数需要放在最后面。

参数说明

表 5-41 gs_rescue 参数说明

参数	说明
-a	打印数据时使用绝对地址（Page头信息将一直使用相对地址）。

版权所有 © 华为云计算技术有限公司

参数	说明
-b	打印数据时直接使用ASCII码形式（该选项会关闭其它打印选项）。
-c	指定具体需要解析的表所属库的兼容模式，与-D参数组合使用，当前分为两种模式： -c A：A兼容模式（解析Tuple行为与不使用-c参数时一致）。 -c M: M-Compatibility模式（支持解析M-Compatibility模式数据库的Tuple，具体支持数据类型如表5-43所示。）。
-d	使用十六进制打印块原始数据（该选项会关闭其它打印选项）。
-D attrlist	使用指定数据类型解析Tuple，类型之间使用逗号分隔，具体支持数据类型见表5-42。
-E segpageoid	解析段页式表，segpageoid为表oid。
-f	解析数据过程中打印Page内容，包含十六进制和字符。
-h	打印帮助信息。
-i	打印解析的元组详细信息，如：XMIN/XMAX/infomask等。
-k	尝试解析crc失败的Page，开启此参数无法保证抢救数据的正确性。
-o	开启此参数，将不会打印此页面的老元组，当前分为两种模式： -o不指定：块上所有数据均抢救。 -o指定：以事务均提交模式抢救块上数据。
-R	指定具体需要解析的Page范围（Page计数从0开始）： ● [startblock]：开始Page下标。 ● [endblock]：结束Page下标。
-s segsize	指定segment大小，默认为1GB，单位为字节，用于校验页面checksum值，不建议修改。
-t	是否解析toast表，对于变长字段，如:text/clob等，数据存储于toast表内，非用户指定路径，如果需要解析，则指定此选项。
-T toast_relation_filenode	通常Tuple如果涉及toast，会在Tuple内存储具体的Toast文件信息，但如果toast表被vacuum full过，会造成toast的relfilenode发生改变，这种情况下需要指定toast表的relfilenode。
-v	对于toast表输出额外的信息。
-n segnumber	强制指定segment的编号，用于校验（大于1G文件生成的xxx.1，xxx.2，xxx.3...)文件的checksum是否正确。
-S	强制指定Page大小，当前默认为8K。

版权所有 © 华为云计算技术有限公司

参数	说明
-x	强制指定解析为index元组，用于解析索引表。
-y	强制指定解析为heap元组，用于解析普通表，防止文件损坏时不能自动解析heap表。
-m	解析pg_filenode.map文件（其它选项将会被忽略）。

表 5-42 gs_rescue 支持的数据类型

类型	数据类型	解析对应类型	备注
数值类型	tinyint	tinyint	-
	smallint	smallint	-
	integer	integer	-
	binary_integer	binary_integer	-
	bigint	bigint	-
	tinyint unsigned	uint1	-
	smallint unsigned	uint2	-
	int unsigned	uint4	-
	bigint unsigned	uint8	-
	numeric[(p[,s])]	numeric	-
	decimal[(p[,s])]	decimal	-
	number[(p[,s])]	number	-
	smallserial	smallserial	-
	serial	serial	-
	bigserial	bigserial	-
	real	real	-
	float4	float4	-
	double precision	double_precision	-
	float8	float8	-

版权所有 © 华为云计算技术有限公司

类型	数据类型	解析对应类型	备注
	float(p)	float4/float8	建表时精度p选择不同，存储空间大小会不同。根据系统表pg_class、pg_attribute、pg_type结合用户表查看typlen大小，4字节使用float4，8字节使用float8。
	binary_double	float8	-
	dec[(p[,s])]	numeric	-
	integer[(p[,s])]	numeric	-
货币类型	money	money	-
布尔类型	boolean	boolean/bool	boolean或bool参数都可以解析boolean类型。
字符类型	char(n)	charn	-
	character(n)	charn	-
	nchar(n)	charn	-
	varchar(n)	varchar	-
	character varying(n)	varchar	-
	varchar2(n)	varchar2n	-
	nvarchar2(n)	nvarchar2n	-
	text	text	-
	"char"	char	-
	clob	clob	huge clob暂不支持（超过1GB的clob）。
	name	name	-
二进制类型	blob	blob	-
	raw	raw	-
	bytea	bytea	-
日期类型	date	timestamp	对于date类型，使用timestamp参数。
	time [(p)] [without time zone]	time	-
	time [(p)] [with time zone]	timetz	-

版权所有 © 华为云计算技术有限公司

类型	数据类型	解析对应类型	备注
	timestamp[(p)] [without time zone]	timestamp	-
	timestamp[(p)] [with time zone]	timestamptz	-
	smalldatetime	smalldatetime	-
	interval day (l) to second (p)	interval	-
	reltime	reltime	-
	abstime	abstime	-
	interval [fields] [(p)]	interval	-
	year	tinyint	对于year类型，解析出为非零值时需要增加1900得到实际年份。
几何类型	point	point	-
	lseg	lseg	-
	box	box	-
	path	path	-
	polygon	polygon	-
	circle	circle	-
网络地址类型	cidr	cidr	-
	inet	inet	-
	macaddr	macaddr	-
位串类型	bit(n)	bit	-
	bit varying(n)	varbit	-
uuid类型	uuid	uuid	-
范围类型	int4range	int4range	-
	int8range	int8range	-
	numrange	numrange	-
	tsrange	tsrange	-
	tstzrange	tstzrange	-
	daterange	daterange	-

版权所有 © 华为云计算技术有限公司

类型	数据类型	解析对应类型	备注
文本搜索类型	tsvector	tsvector	-
	tsquery	tsquery	-
json类型	json	json	-
	jsonb	jsonb	-
对象标识类型	oid	oid	-
	cid	cid	-
	tid	tid	-
	xid	xid	-
	regconfig	regconfig	-
	regdictionary	regdictionary	-
	regoper	regoper	-
	regoperator	regoperator	-
	regproc	regproc	-
	regprocedure	regprocedure	-
	regclass	regclass	-
	regtype	regtype	-
向量数据类型	floatvector	floatvector	-
	boolvector	boolvector	-

表 5-43 gs_rescue 支持的 M-Compatibility 模式数据类型

类型	数据类型	解析对应类型	备注
数值类型	tinyint	int1	-
	smallint	int2	-
	mediumint	int3	-
	integer/int	int4	-
	bigint	int8	-
	tinyint unsigned	uint1	-
	smallint unsigned	uint2	-

版权所有 © 华为云计算技术有限公司

类型	数据类型	解析对应类型	备注
	mediumint unsigned	uint3	-
	integer/int unsigned	uint4	-
	bigint unsigned	uint8	-
	numeric[(p[,s])]	numeric	-
	decimal[(p[,s])]	numeric	-
	fixed[(p[,s])]	numeric	-
	dec[(p[,s])]	numeric	-
	float4[(p[,s])]	float4	FLOAT只能保证6位有效
	float8[(p[,s])]	float8	FLOAT只能保证6位有效
	float(p,s)	float4	FLOAT只能保证6位有效
	double[(p[,s])]	float8	FLOAT只能保证6位有效
	dec[(p[,s])]	numeric	-
布尔类型	boolean	boolean/bool	BOOLEAN或BOOL参数都可以解析BOOLEAN类型。
字符类型	char(n)	bpchar	-
	varchar(n)	varchar	-
	text	text	-
	tinytext	tinytext	-
	mediumtext	mediumtext	-
	longtext	longtext	-
二进制类型	binary(n)	binary	-
	varbinary(n)	varbinary	-
	blob	blob	-
	tinyblob	tinyblob	-
	mediumblob	mediumblob	-
	longblob	longblob	-
日期类型	date	date	tui工具恢复带有全零值的数据0000-00-00时，需要保证文件中不存在\0（即十六进制00）。

版权所有 © 华为云计算技术有限公司

类型	数据类型	解析对应类型	备注
	time[(p)]	time	tui工具恢复带有全零值的数据00:00:00时，需要保证文件中不存在\0（即十六进制00）。
	timestamp[(p)]	timestamp	tui工具恢复带有全零值的数据0000-00-00 00:00:00时，需要保证文件中不存在\0（即十六进制00）。
	datetime[(p)]	datetime	tui工具恢复带有全零值的数据0000-00-00 00:00:00时，需要保证文件中不存在\0（即十六进制00）。
	year	year	-
位串类型	bit(n)	bit	-
集合类型	set('val1','val2',...)	set('val1','val2',...)	-
枚举类型	enum('val1','val2',...)	enum('val1','val2',...)	-
JSON类型	json	jsonm	-

说明

- 解析数组时，可使用array_i、array_d参数进行解析。不同数据类型数组的对齐性质可能不同，解析时也需要相应参数。根据系统表pg_class、pg_attribute、pg_type并结合用户表查看该数组的typalign，对齐性质为'i'使用array_i，对齐性质为'd'使用array_d。
- 针对 date 类型解析的参数，在不同环境下可能存在差异。如在集中式环境下，解析参数为timestamp；在分布式环境下，解析参数为date；在小型化环境下，解析参数为timestamp。如果数据库可用，可以查询pg_type数据表获取该date类型的存储类型作为参数；如果数据库不可用，可以参考示例三中的信息来确定该类型的参数。
- M-Compatibility模式解析数据时，类型名称需要小写。

示例

示例一：直接解析数据。

```
gs_rescue -D integer /data1/install/single_node/base/16706/16707

*******************************************************************
* GaussDB File/Block Formatted Dump Utility
*
* File: /data1/install/single_node/base/16706/16707
* Options used: -D integer
*******************************************************************
This is an Astore table!

Block    0 ********************************************************
<Header> -----
 Block Offset: 0x00000000         Offsets: Lower      48 (0x0030)
```

 版权所有 © 华为云计算技术有限公司

```
 Block: Size 8192  Version    6            Upper    8128 (0x1fc0)
 LSN:  logid      0 recoff 0x02f00158      Special  8192 (0x2000)
 Items:    2                      Free Space: 8080
 Checksum: 0x2863  Prune XID: 0x00000000  Flags: 0x0040 ()
 Length (including item array): 48

<Data> -----
 Item   1 -- Length:   28  Offset: 8160 (0x1fe0)  Flags: NORMAL
COPY: 1

*** End of File Encountered. Last Block Read: 0 ***
```

示例二：解析数据，解析后可重定向数据至文件，然后通过copy命令导入新的可用数据库实例。

```
gs_rescue -D integer /data1/install/single_node/base/16706/16707 | grep COPY | perl -lne 's/^COPY: //g;
print;'
1
```

示例三：此示例主要描述如何在数据库不启动且未知待抢救表路径、表字段等信息的情况下解析目标表数据的过程。主要流程是通过解析pg_filenode.map文件，来获取系统表的OID和relfilenode，然后继续解析各系统表（pg_database/pg_class/pg_attribute等）来获取用户表的路径、列类型等。

说明

本示例假设用户使用默认表空间。 以下示例假设数据目录为/data1/install/single_node。

1. 通过使用数据目录下global/pg_filenode.map获取pg_database的OID和relfilenode，然后通过relfilenode可获取pg_database的物理路径。

   说明

   pg_filenode.map文件内存储系统表的oid和filenode的对应关系，常见的系统表的oid如下：

   - pg_tablespace(OID:1213)
   - pg_database(OID:1262)
   - pg_namespace(OID:2615)
   - pg_class(OID:1259)
   - pg_attribute(OID:1249)
   - pg_type(OID:1247)

   ```
   gs_rescue -m ./global/pg_filenode.map

   ******************************************************************
   * GaussDB File/Block Formatted Dump Utility
   *
   * File: /data1/install/single_node/global/pg_filenode.map
   * Options used: -m
   ******************************************************************
   Magic Number: 0x592718
   Number Mappings: 68
   Mappings list details:
     NOTE: listing has been limited to the first 62 mappings
           (perhaps the file is not a valid pg_filenode.map file?)
   OID: 1262       Filenode: 16702
   OID: 2964       Filenode: 16349
   OID: 1213       Filenode: 16355
   ...
   ```

   上述示例内OID 1262对应的Filenode 16702，从而可获取pg_database的物理路径为：./global/16702。

版权所有 © 华为云计算技术有限公司

2. pg_database存储可用数据库的信息，通过解析pg_database物理文件（./global/16702），获取目标database的详细信息（主要为获取待抢救database的OID）。

```
gs_rescue -i -D name,oid,integer,name,name,bool,bool,integer,oid,xid32,oid,name,array_i,xid,xid,name ./global/16702
...
 Item   8 -- Length:  400  Offset: 6896 (0x1af0)  Flags: NORMAL
  XMIN: 2  XMAX: 0  CID|XVAC: 0  OID: 17697
  Block Id: 0  linp Index: 8   Attributes: 16   Size: 32
  infomask: 0x0909 (HASNULL|HASOID|XMAX_LOCK_ONLY|XMIN_COMMITTED)
  t_bits: [0]: 0xff [1]: 0xef

COPY: gaussdb  10     0      C      C      f      t      -1     17692   7142530613     1663    A
\N      14955   2       PRC
...
```

   在上述示例内，gaussdb数据库的OID为17697，所以可以得到gaussdb数据库的文件路径为：./base/17697，此数据库内的相关系统表（pg_class/pg_attribute/pg_type等）和用户表数据，均存储在此目录下。

3. 解析目标库内系统表信息，因为pg_class系统表存储对应数据库下所有表的信息，可通过本数据库下pg_filenode.map获取pg_class的物理路径，pg_filenode.map的路径为：./base/17697/pg_filenode.map。

```
gs_rescue -m /data1/install/single_node/base/17697/pg_filenode.map | grep OID
OID: 1249       Filenode: 17186
OID: 1259       Filenode: 17206
OID: 1247       Filenode: 17082
...
```

   pg_class在pg_filenode.map内的OID为1259，所以目标数据库的pg_class路径为：./base/17697/17206。

4. 获取本库下的用户表名称和路径（relfilenode）信息。

```
gs_rescue -i -D
name,oid,oid,oid,oid,oid,oid,oid,float8,float8,integer,oid,oid,oid,oid,oid,oid,bool,bool,char,char,smallint,smallint,bool,bool,bool,bool,bool,tinyint,bool,bool,char,xid32,array_i,array_i
,char,xid,oid,int2vector,xid ./base/17697/17206  | grep COPY
...
 Item  22 -- Length:  280  Offset: 2560 (0x0a00)  Flags: NORMAL
  XMIN: 16675  XMAX: 16675  CID|XVAC: 6  OID: 34097
  Block Id: 2  linp Index: 22   Attributes: 40   Size: 32
  infomask: 0x252b (HASNULL|HASVARWIDTH|HASOID|COMBOCID|XMIN_COMMITTED|UPDATED|KEYS_UPDATED)
  t_bits: [0]: 0xff [1]: 0xff [2]: 0xff [3]: 0xff
          [4]: 0x9d
COPY: t_user    2200    34093   0       10      0       34091   0       0       0       0       0       0       0       0       0       0       f       f       pr      3       0       f       f       f       f       f       1       f       f       n       1627792621      \N      {orientation=row,compression=no}    d16632
...
```

   在上述示例内，可以获取待抢救表t_user的：

   - 路径 - ./base/17697/34097
   - OID - 34091

5. 获取待抢救表t_user的列类型列表。首先需要解析pg_attribute表获取t_user的所有列类型的oid， 然后再通过解析pg_type根据类型oid获取所有列对应的类型名称。

   解析表字段信息，pg_attribute系统表存储关于表字段的信息，pg_attribute表在pg_filenode.map内的OID为1249，获得物理路径为：./base/17697/17186，通过目标表t_user的OID 34091，获取在pg_attribute表中关于t_user表的所有列信息。

```
gs_rescue -D
oid,name,oid,integer,smallint,smallint,integer,integer,integer,bool,char,char,bool,bool,bool,bool,tinyint,integer,oid,array_i,array_i,array_i,bytea,tinyint ./base/17697/17186 | grep 34091
COPY: 34091     id      23      -1      4       1       0       -1      -1      t       p       i       f       f       f
```

 版权所有 © 华为云计算技术有限公司

```
t       127     0       0   \N      \N      \N      \N      0
COPY: 34091     name    1042    -1      -1      2       0       -1      24      f       x       i       f       f       f
t       127     0   100     \N      \N      \N      \N      0
COPY: 34091     age     23      -1      4       3       0       -1      -1      t       p       i       f       f       f
t       127     0       0   \N      \N      \N      \N      0
COPY: 34091     ctid    27      0       6       -1      0       -1      -1      f       p       s       t       f       f
t       0       0       0   \N      \N      \N      \N      0
COPY: 34091     xmin    28      0       8       -3      0       -1      -1      t       p       d       t       f       f
t       0       0       0   \N      \N      \N      \N      0
...
```

在上述示例内，第6列attnum大于0为用户所创建的字段，也是字段编号，可以获取表t_user字段的：

- 列名称 - id name age
- 列类型OID - 23 1042 23

解析字段的数据类型信息，pg_type系统表存储数据类型的相关信息，pg_type表在pg_filenode.map内的OID为1247，获得物理路径为：./base/17697/17082，通过目标字段age的OID 23，获取在pg_type表中关于目标字段的数据类型信息。

```
gs_rescue -i -D
name,oid,oid,smallint,bool,char,char,bool,bool,char,oid,oid,oid,regproc,regproc,regproc,regproc,regproc,r
egproc,regproc,char,char,bool,oid,integer,integer,oid,charn,text,array_i ./base/17697/17082
...
Item   9 -- Length:  172  Offset: 6608 (0x19d0)  Flags: NORMAL
  XMIN: 1  XMAX: 0  CID|XVAC: 0  OID: 23
  Block Id: 0  linp Index: 9   Attributes: 30   Size: 32
  infomask: 0x0909 (HASNULL|HASOID|XMAX_LOCK_ONLY|XMIN_COMMITTED)
  t_bits: [0]: 0xff [1]: 0xff [2]: 0xff [3]: 0x07

COPY: int4  11  10  4   t   b   N   f   t   ,   0   0   1007    42  43  2406    2407    0   0   0   i   p   f   0
-1  0   0   \N  \N  \N
...
```

在上述示例内，在解析pg_type表的详情中，通过目标列age的OID查询到数据类型为int4。

6. 通过上述步骤获取的t_user表的路径以及列类型列表，可解析用户表内数据。

   说明

   -D 后跟随列的解析参数可以是pg_type中的数据类型，也可以是**表5-42**所对应的解析参数。

```
gs_rescue -D int4,charn,int4 ./base/17697/34097

****************************************************************
* GaussDB File/Block Formatted Dump Utility
*
* File: 34091
* Options used: -D int,charn,int
****************************************************************
This is an Astore table!

Block    0 ********************************************************
<Header> -----
 Block Offset: 0x00000000         Offsets: Lower      44 (0x002c)
 Block: Size 8192  Version    6           Upper    8136 (0x1fc8)
 LSN:  logid      1 recoff 0x610668c0      Special  8192 (0x2000)
 Items:    1                      Free Space: 8092
 Checksum: 0xca7f  Prune XID: 0x00000000  Flags: 0x0040 ()
 Length (including item array): 44

<Data> -----
 Item   1 -- Length:   56  Offset: 8136 (0x1fc8)  Flags: NORMAL
COPY: 0 Jack                    21

*** End of File Encountered. Last Block Read: 0 ***
```

示例四：对数据表的列有变动的情况可参考此示例。

版权所有 © 华为云计算技术有限公司

在按照创建表时的数据类型顺序参数解析时，修改列后的数据解析可能会失败，因此对修改列后的数据可以用相应的数据类型顺序参数解析。

如数据表新增了一列，可按照如下步骤解析恢复：

1. 按照创建表时的数据类型顺序参数解析数据重定向至文件file01，按照新增列后数据表的数据类型顺序参数解析数据至文件file02。
2. 在可用数据库中创建original表和amend表，copy file01文件数据至original表，copy file02文件数据至amend表。
3. 将original表数据插入到amend表得到全部数据。

```
gs_rescue -D integer /data1/install/single_node/base/16706/16708 | grep COPY | perl -lne 's/^COPY: //g;
print;' > /data2/file01
gs_rescue -D integer,integer /data1/install/single_node/base/16706/16708 | grep COPY | perl -lne 's/
^COPY: //g; print;' > /data2/file02

gaussdb=# create table original(col1 integer);
CREATE TABLE
gaussdb=# copy original from '/data2/file01';
COPY 1
gaussdb=# create table amend(col1 integer,col2 integer default 0);
CREATE TABLE
gaussdb=# copy amend from '/data2/file02';
COPY 1
gaussdb=# insert into amend(col1) select * from original;
INSERT 0 1
```

示例五：对M-Compatibility模式下的含有SET类型、ENUM类型的数据表进行解析。

1. 对于SET类型定义和ENUM类型，需给出其在表格定义时使用的完整定义。
2. 在shell中输入-D参数时，需要合理使用单引号、双引号对参数进行包裹，如遇无法正确包裹的情况下需使用反斜杠来转义单引号或双引号，可参考enum类型的示例。

SET类型：

```
CREATE TABLE tt1(c1 set('aa', 'bb'));
INSERT INTO tt1 VALUES('aa');
INSERT INTO tt1 VALUES('bb');
gs_rescue -c M -D "set('aa', 'bb')" /data/single_data/base/16723/24583

*****************************************************************
* GaussDB File/Block Formatted Dump Utility
*
* File: /data/single_data/base/16723/24583
* Options used: -c M -D set('aa', 'bb')
*****************************************************************
This is a Ustore table!

Block    0 ********************************************************
<Header> -----
 Block Offset: 0x00000000         Offsets: Lower      168 (0x00a8)
 Block: Size 8192  Version    9            Upper     8140 (0x1fcc)
 LSN:  logid      0 recoff 0x09dd47d8      Special  8192 (0x2000)
 Items:    2      td_count : 0x0004(4)     Free Space :  7972
 potrntial_freespace:0x1f24               pagesize_version:0x2009
 Checksum: 0x8a29  XID_base : 0x57213      Flags: 0x0040
 Length (including item array): 56

        ******************************************************
UHeap Page TD information, nTDSlots = 4
        TD Slot #1, xid:57216, urp:0
        TD Slot #2, xid:35184372089024, urp:3749773313
        TD Slot #3, xid:0, urp:2305843009229029376
        TD Slot #4, xid:65536, urp:0
<Data> -----
```

 版权所有 © 华为云计算技术有限公司

```
 Item   1 -- Length: 25  Offset: 8166 (0x1fe6)  Flags: NORMAL
COPY: aa
 Item   2 -- Length: 25  Offset: 8140 (0x1fcc)  Flags: NORMAL
COPY: bb


*** End of File Encountered. Last Block Read: 0 ***
```

ENUM类型：

```
CREATE TABLE ttt1(c1 enum('aa', 'bb','{"key1": "value1", "key2": "value2"}'));
INSERT INTO ttt1 VALUES('aa');
INSERT INTO ttt1 VALUES('{"key1": "value1", "key2": "value2"}');
gs_rescue -c M -D "enum('aa', 'bb','{\"key1\": \"value1\", \"key2\": \"value2\"}')" /data/single_data/base/
16723/16743
*******************************************************************
* GaussDB File/Block Formatted Dump Utility
*
* File: /data/single_data/base/16723/16743
* Options used: -c M -D enum('aa', 'bb','{"key1": "value1", "k...
*******************************************************************
This is a Ustore table!

Block    0 ********************************************************
<Header> -----
 Block Offset: 0x00000000         Offsets: Lower      168 (0x00a8)
 Block: Size 8192  Version    9            Upper     8140 (0x1fcc)
 LSN:  logid      0 recoff 0x09dd7e08      Special  8192 (0x2000)
 Items:    2      td_count : 0x0004(4)     Free Space :  7972
 potrntial_freespace:0x1f24               pagesize_version:0x2009
 Checksum: 0x5328  XID_base : 0x57217      Flags: 0x0040
 Length (including item array): 56

        *****************************************************
UHeap Page TD information, nTDSlots = 4
        TD Slot #1, xid:57220, urp:0
        TD Slot #2, xid:35184372089108, urp:3750035457
        TD Slot #3, xid:0, urp:2305843009234534400
        TD Slot #4, xid:65536, urp:0
<Data> -----
 Item   1 -- Length: 26  Offset: 8166 (0x1fe6)  Flags: NORMAL
COPY: aa
 Item   2 -- Length: 26  Offset: 8140 (0x1fcc)  Flags: NORMAL
COPY: {"key1": "value1", "key2": "value2"}


*** End of File Encountered. Last Block Read: 0 ***
```

5.17 gs_rescue_tui

背景信息

gs_rescue_tui工具集成了gs_rescue工具的解析单表的功能，给用户提供TUI界面（Text-based User Interface，文本用户界面），在与用户的交互中，自动拼接解析系统表和用户表的gs_rescue命令，最终实现解析和恢复整个数据库的功能。

gs_rescue_tui工具提供恢复至本地文件和恢复至目标端两个功能。通过解析文件并落盘的方式，恢复至本地文件；通过解析并连接目标端数据库，通过集成COPY命令的方式，将数据恢复至目标端数据库。

 版权所有 © 华为云计算技术有限公司

⚠ 注意

- 使用场景：仅限当数据库备份损坏或无备份、备机异常无法主备切换、网络隔离或集群无法启动时，可以使用数据抢救工具对有损坏的数据文件进行解析恢复，恢复至本地csv文件或目标端数据库。
- 权限控制：为防止数据泄露，用户需要合理设置抢救数据导出文件的权限，保证权限最小化。
- 约束：
  - 源数据库和目标数据库的版本和数据库部署类型需要一致。且pg_database、pg_partition、pg_class、pg_attribute、pg_type、pg_namespace、pg_tablespace、gs_ilm、gs_ilm_object、gs_ilm_policy、pg_filenode.map文件存在且可正常访问，如系统表有损坏，解析出的数据可能有部分丢失。
  - gs_rescue_tui工具恢复ilm表，需要设置目标数据库为升级模式（upgrade_mode = 1）且开启ilm模式（enable_ilm =on）。建议执行完数据抢救后，恢复为原先的取值（upgrade_mode = 0，enable_ilm =off）。
- 故障端要求：数据库应处于离线状态。工具可以获取目标端数据库连接，正常连接目标端。
- 目标端要求：目标端的数据库磁盘空间足够。目标端用户需要有创建database，schema，table的权限，建议用低权限用户，操作数据抢救工具，guc参数shared_buffers最小为10G、session_timeout为0。
- 可恢复或解析的列说明：仅支持常用数据列类型。
- 可恢复或解析的表说明：仅支持astore heap表、ustore heap表、段页式、unlogged表、分区表、二级分区表、ilm表的数据解析，不支持临时表、透明加密表、全密态表、压缩表、内存表。不支持用户、索引和约束恢复(不在页面展示，解析的时候不会解析）。不支持用户自建的表空间恢复，不解析xlog/undo内数据。离线解析数据无法保证数据一致性。仅解析落盘数据。
- 恢复速度：数据抢救工具支持多张表的并发恢复。常用数据类型在16U 32G配置SSD磁盘的服务器上，对多张数据表进行并发恢复情况下，恢复自动入库速度不低于40MB/s。
- 后续版本升级对系统表更改，本产品需要适配。
- gs_rescue/gs_rescue_tui 解析的M-Compatibility模式数据文件仅支持在M-Compatibility模式数据库下进行数据使用，其他模式下数据文件不兼容。
- M-Compatibility模式下仅支持astore/ustore普通heap表、ASTORE段页式及分区表的数据解析。
- pdb以普通database形式恢复到目标库，只恢复数据，不恢复租户等其他信息。

操作说明

步骤1 Decode table to csv。

1. 输入gs_rescue_tui，选择1 Decode table to csv（解析数据文件并落盘的方式将csv文件恢复至本地）如图5-1所示。

 版权所有 © 华为云计算技术有限公司

图 5-1 Decode table to csv

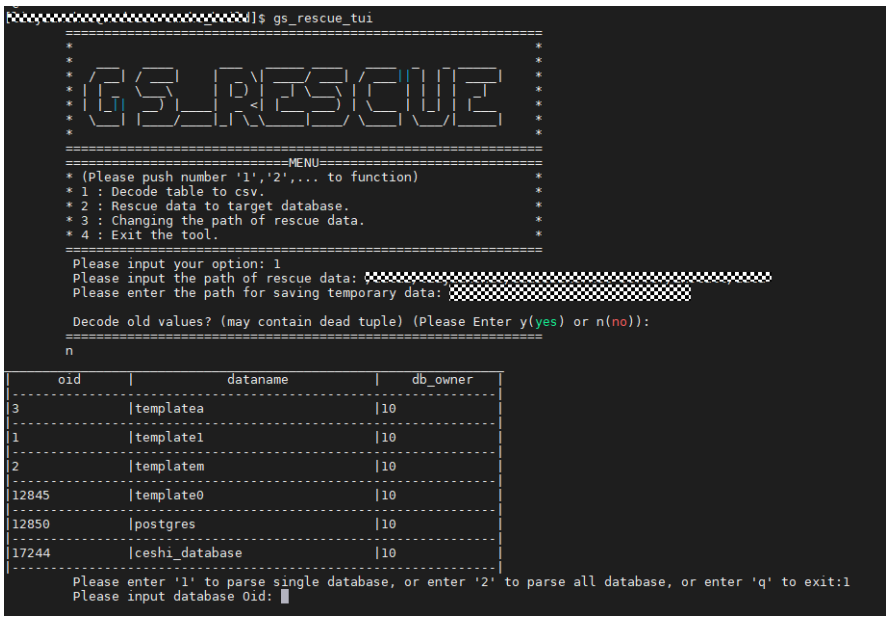


表 5-44 Decode table to csv 选项说明

选项	说明
1.Decode table to csv	将表解析为csv文件。
2.Rescue data to target database	将数据恢复到目标数据库。
3.Changing the path of rescue data	更改数据路径。
4.Exit the Tool	退出工具。

表 5-45 输入路径说明

选项	说明
Please input the path of rescue data	待恢复数据的存储路径。
Please enter the path for saving temporary data	数据的临时导出路径，此路径必须存在。

表 5-46 Decode old values 选项说明

选项	说明
Decode old values? (may contain dead tuple)	y：恢复数据文件中全部历史版本数据。 n：按照默认事务已提交的方式解析。

版权所有 © 华为云计算技术有限公司

表 5-47 解析单库或全库选项说明

选项	说明
enter '1' to parse single database	1：解析单个数据库。
enter '2' to parse all database	2：解析整个数据库。
enter 'q' to exit	q：返回上一层。

2. 启动程序后，输入数据库数据路径和csv解析文件保存路径。
3. 页面会展示所有的数据库，请根据实际情况，选择entire database或schema level或single table，解析对应范围的数据，或返回上一级，参考图5-2。

图 5-2 decode database

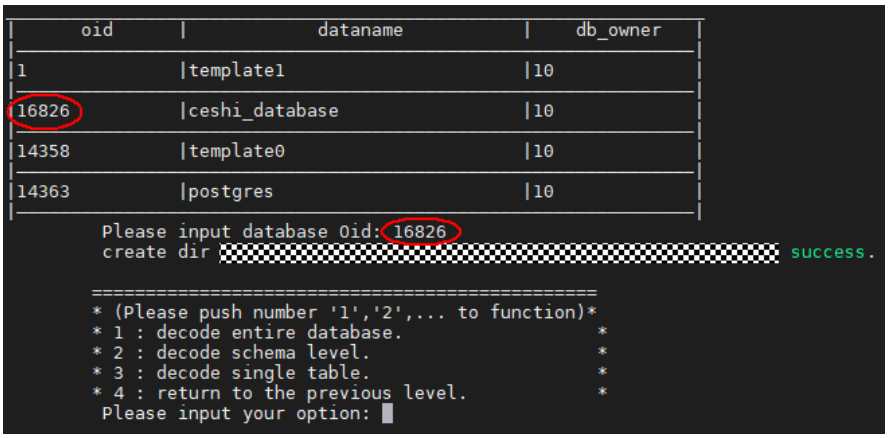


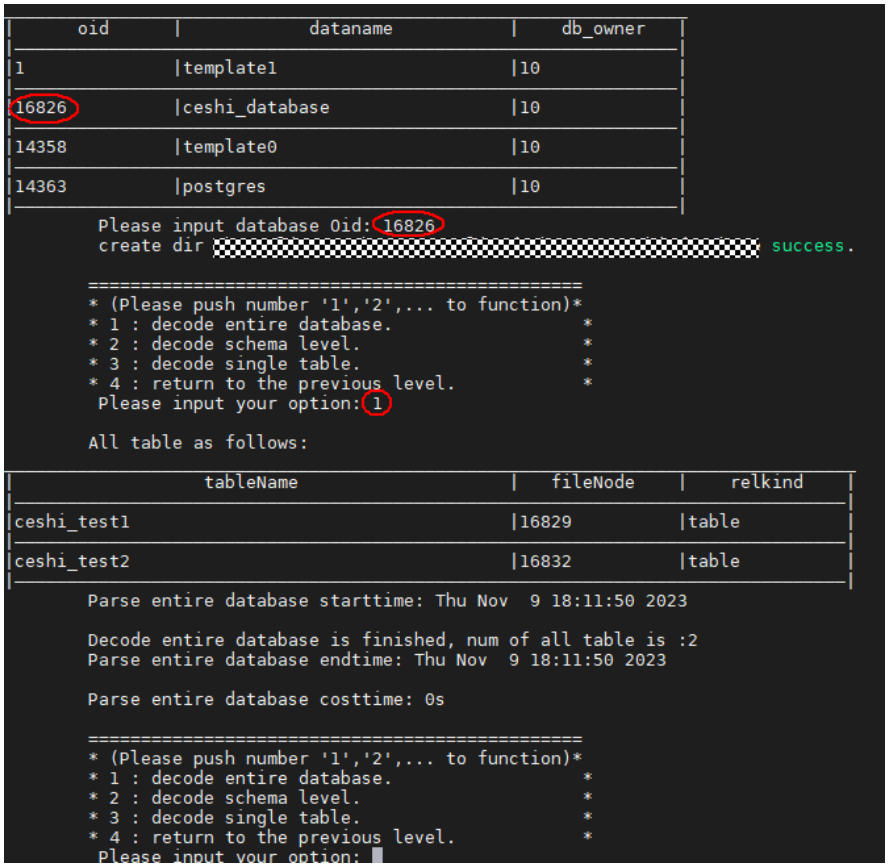
表 5-48 decode database 选项说明

选项	说明
1 : decode entire database	解析整个数据库。
2 : decode schema level	解析单个schema。
3 : decode single table	解析单表。
4 : return to the previous level	返回上一级。

- decode entire database。输入1，解析整个数据库，参考图5-3。

 版权所有 © 华为云计算技术有限公司

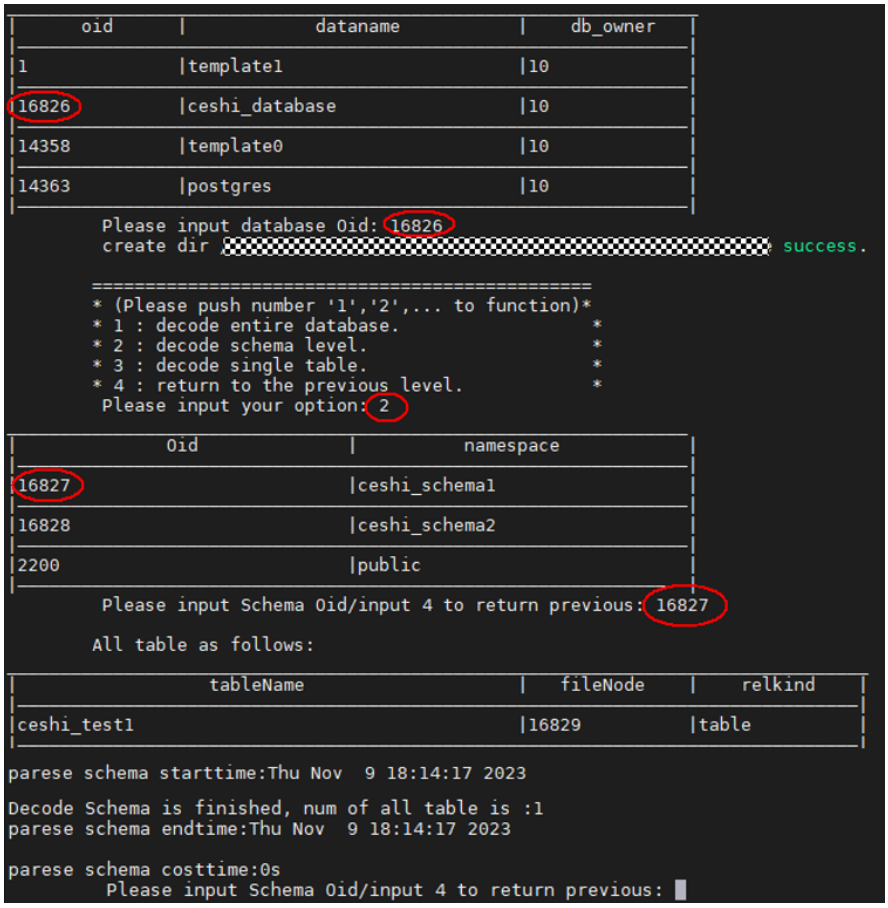
图 5-3 decode entire database



- decode schema level。输入2，展示所选数据库下的所有schema，根据提示选择需解析schema对应的oid，参考图5-4。

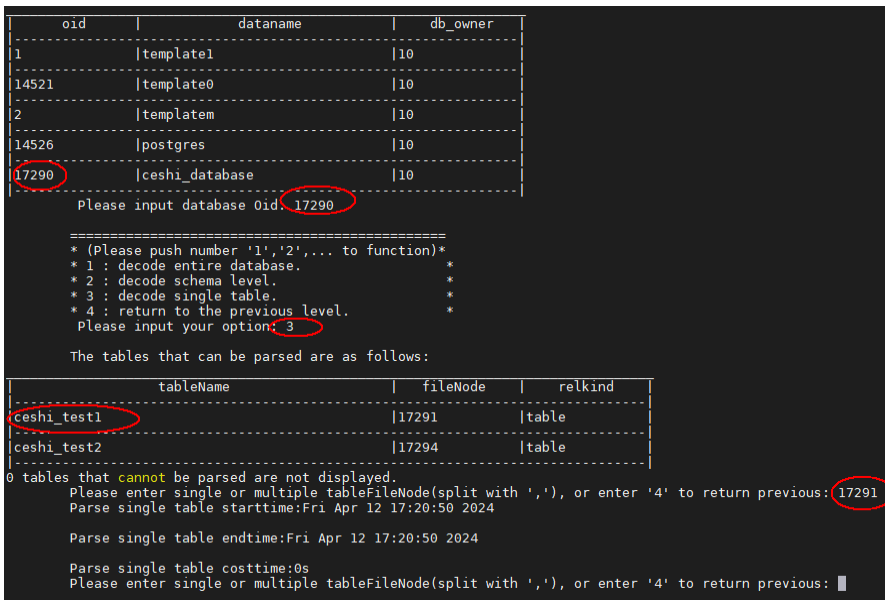
 版权所有 © 华为云计算技术有限公司

图 5-4 decode schema level



- decode single table。输入3，展示所选数据库下的所有table，根据提示选择操作，参考图5-5。

图 5-5 decode single table



- return to the previous level。输入4，返回上一级。

 版权所有 © 华为云计算技术有限公司

说明

导出的数据文件存放在数据临时导出路径下的各个对应数据库名下data目录下的schema_oid中，其中分区表在partition_table_filenode目录下。解析的表命名方式为：table_filenode.csv。

table_filenode、table_name、schema_oid、schema_name、partition_table_filenode对应关系可参考：

fieldOidAndNode.txt前六列为：table_oid、table_filenode、table_name、schema_oid、table_type、partition_table_filenode。

FinnalSchemaOid.txt前两列为：schema_oid、schema_name。

步骤2 Rescue data to target database。输入gs_rescue_tui，选择2 Rescue data to target database（将数据文件恢复至目标端）。

说明

如果程序退出后，执行**步骤2**（Rescue data to target database），需要先执行**步骤3**（Changing the path of rescue data），修改数据库数据路径和csv保存路径。

1. 输入目标端数据库信息，包括Database Name，Target Port，Target Host IP，User Name，Password，whether to enable SSLSOCKET，参考**图5-6**。

图 5-6 Rescue data to target database

须知

使用场景：当使用工具恢复M-Compatibility数据库时，目标端数据库信息Database Name需要填写非M-Compatibility数据库名称。

版权所有 © 华为云计算技术有限公司

表 5-49 Whether clear target table data 选项说明

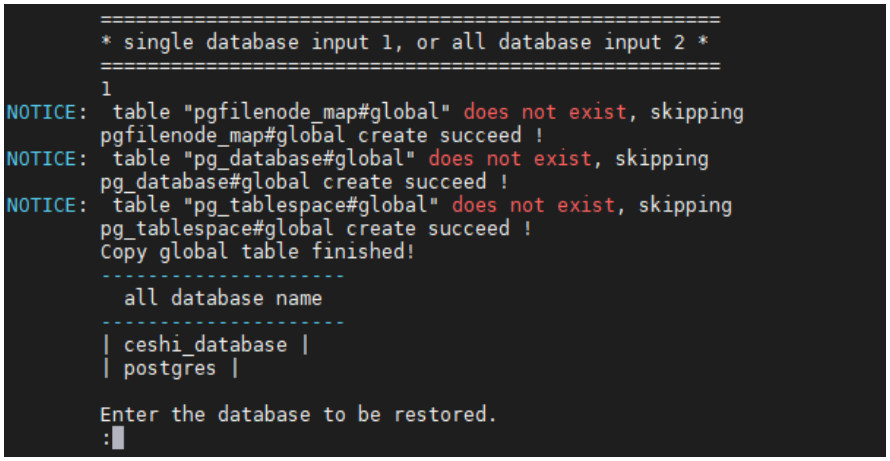
选项	说明
Whether clear target table data	y：建表之前，会清除目标端表数据。 n：建表之前，不会清除目标端表数据。

表 5-50 Rescue data to target database 选项说明

选项	说明
Please Enter Database Name	目标database名称。
Please Enter Target Port	目标数据库端口号。
Please Enter Target Host IP	目标数据库IP地址，用于登录数据库执行copy命令恢复数据。集中式主备场景下建议输入主机IP地址。
Please Enter User Name	目标数据库用户。
Please Enter Password	目标数据库密码。
Whether to enable SSLSOCKET(Please Enter y(yes) or n(no))	是否开启SSL SOCKET连接。

2. 解析所有database。输入2，等待数据库恢复。
3. 解析单个database。输入1，展示所有database，请选择需要解析的database name，参考图5-7。

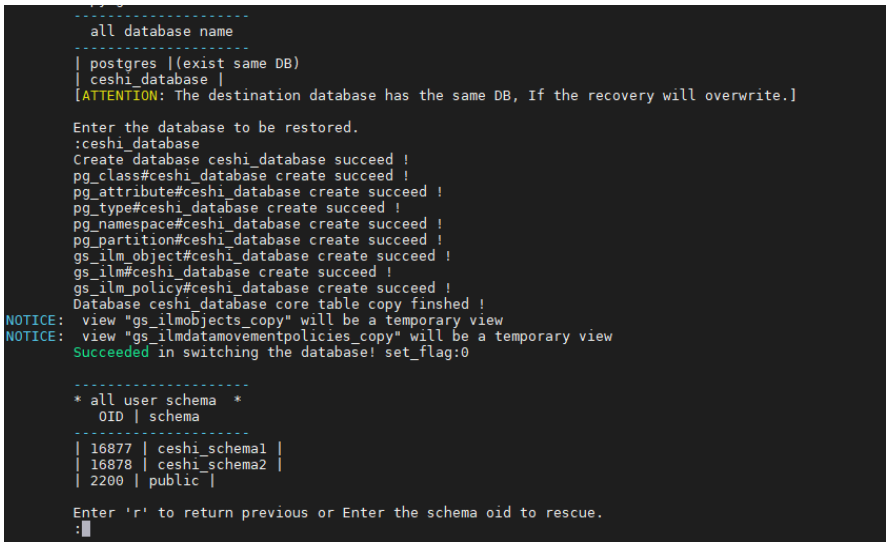
图 5-7 解析单个 database



4. 输入database name，展示所有schema，选择需要解析的schema_name，参考图5-8。

版权所有 © 华为云计算技术有限公司

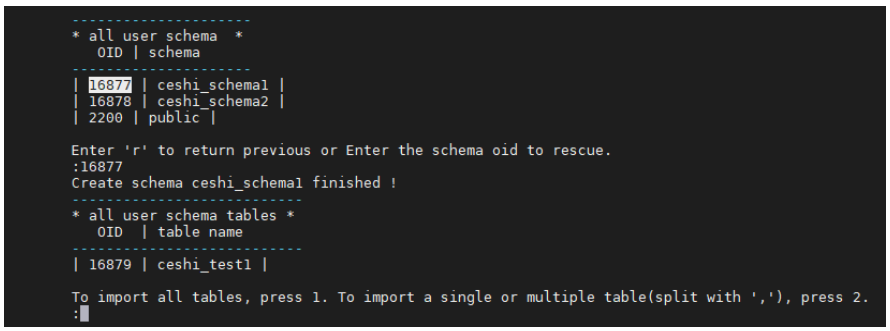
图 5-8 输入 database name



说明：‘r’返回上一层。

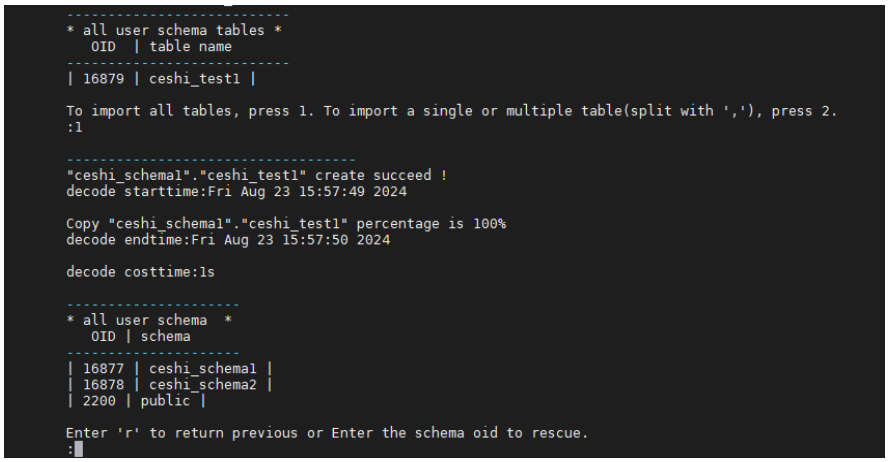
5. 输入schema的oid，展示所有的表，选择恢复所有表或者单表，参考图5-9。

图 5-9 输入 schema 的 oid



- 选择1，恢复上一步所选schema下所有表，参考图5-10。

  图 5-10 恢复上一步所选 schema 下所有表



 版权所有 © 华为云计算技术有限公司

- 选择2，恢复上一步所选schema下的单表或多表，参考图5-11和图5-12。

图 5-11 恢复上一步所选 schema 下的单表

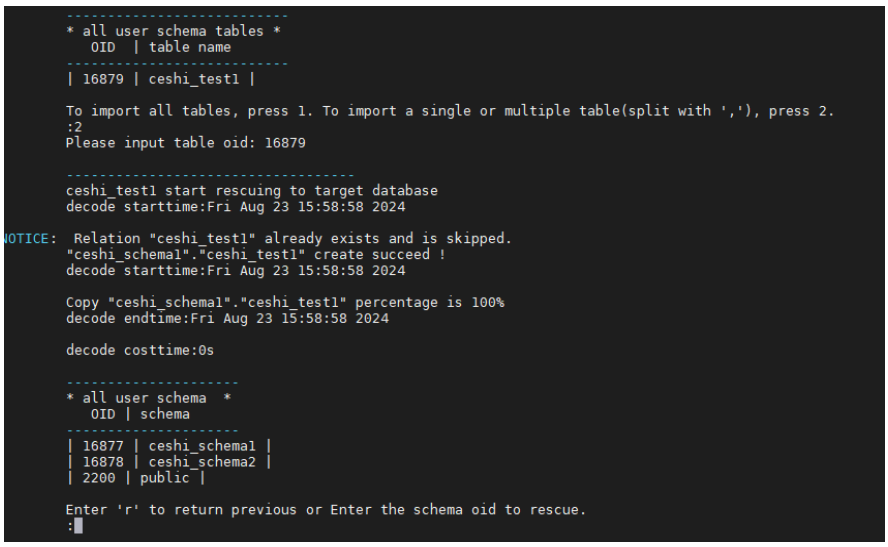
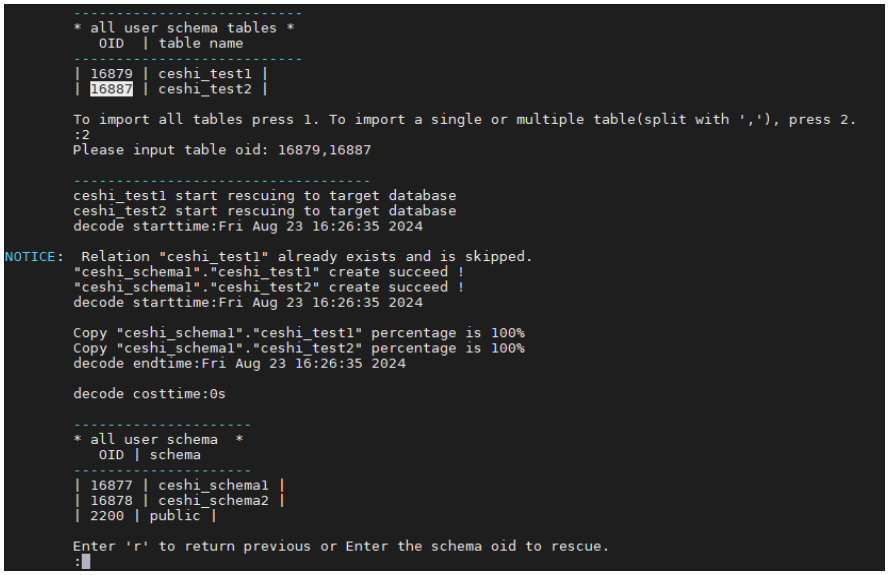


图 5-12 恢复上一步所选 schema 下的多表

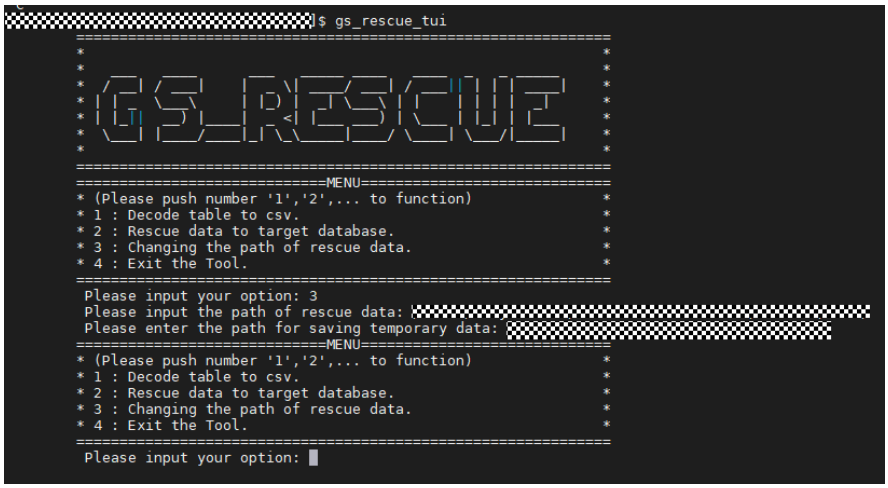


说明：如果恢复完成，可CRTL+C退出工具。如果恢复过程中，目标段数据库连接异常，需要根据提示重新输入IP地址，进行断点续传。

步骤3 Changing the path of rescue data。输入gs_rescue_tui，选择3 Changing the path of rescue data（修改数据库数据路径和csv保存路径），参考图5-13。

 版权所有 © 华为云计算技术有限公司

图 5-13 Changing the path of rescue data



步骤4 Exit the Tool。输入gs_rescue_tui，选择4 Exit the Tool（退出工具）。

----结束

5.18 gs_shrink

> **说明**
>
> 集中式仅支持降副本操作，且不支持单独调用。
>
> 集中式单节点不支持该工具。

背景信息

随着业务场景和数据量的变化，需要数据库实例节点具有可扩展性以及可收缩性。GaussDB提供了gs_shrink工具来帮助完成对数据库实例节点进行降副本。

降副本前提条件

- 数据库实例状态为Normal，重分布状态为No，balanced状态为Yes。如果缩容副本为级联备，实例状态支持Normal或Degraded，约束为除级联备外其他主备正常。
- 数据库集群业务压力较低且不能锁定集群。
- QUORUM 集群：3副本及大于3副本的QUORUM集群可执行降副本操作，降副本最大可将QUORUM集群降为2副本集群。
- DCF 集群：4副本及大于4副本的DCF集群可执行降副本操作，降副本最大可将DCF集群降为3副本集群。

降副本注意事项

- 指定IP进行降副本操作仅支持一个分片内减少副本数大于1的场景，等于1时默认降最后一个副本，指定IP为其他副本时不生效。

语法

- 物理数据库实例缩容

 版权所有 © 华为云计算技术有限公司

```
gs_shrink [-t contraction] {-h HOSTNAME | -f hostfile} [--parallel-jobs=NUMBER] [-l LOGFILE] [--redis-mode=MODE] [--resource-level=LEVL] [--join-tables=FILE]
```

- 重分布资源管控
  ```
  gs_shrink -t resourcectl --resource-level=f/h/m/l
  ```
- 重分布参数设置
  ```
  gs_shrink -t parametercfg [--lockwait-timeout=MSEC] [--lockwait-interval=<SECS>] [--trylock-threshold=NUMBER] [--enable-cancel=Boolean] [--last-catchup-threshold=MSEC] [--catchup-times=NUMBER] [--write-error-mode=Boolean] [--catchup-query-dop=NUMBER] [--parallel-catchup-threshold=NUMBER] [--parallel-reindex-jobs=NUMBER][--redis-retry-times=NUMBER][--redis-retry-interval=NUMBER]
  ```
- 逻辑数据库实例缩容
  ```
  gs_shrink {-h HOSTNAME | -f hostfile} [--parallel-jobs=NUMBER] [-l LOGFILE] [--nodegroup-name=LCGROUPNAME] [--checkdisk]
  ```
- 加速数据库实例缩容
  ```
  gs_shrink {-h HOSTNAME | -f hostfile} [--parallel-jobs=NUMBER] [-l LOGFILE] [--nodegroup-name=LCGROUPNAME]
  ```
- 加速数据库实例回滚
  ```
  gs_shrink {-h HOSTNAME | -f hostfile} [-l LOGFILE]
  ```
- 数据库实例降副本
  ```
  gs_shrink -t reduce --offline [--reduce-ips=IPS] [--reduce-times=1] [--reduce-other]
  ```
- 显示帮助信息
  ```
  gs_shrink -? | --help
  ```
- 显示版本号信息
  ```
  gs_shrink -V | --version
  ```
- 接收数据库管理员密码
  ```
  echo "password" | gs_shrink [-t [ reduce ]] {-h HOSTNAME | -f hostfile} [--parallel-jobs=NUMBER] [-l LOGFILE]
  ```

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

- 通用参数
  - -t

    gs_shrink命令参数类型。

    取值范围：reduce（降副本），contraction（缩容），resourcectl（资源管控），parametercfg（重分布参数设置）。

    说明

    不传入-t参数时，默认执行缩容操作。
  - -l

    指定日志文件及存放路径。

    默认值：$GAUSSLOG/om/gs_shrink-YYYY-MM-DD_hhmmss.log
  - -?, --help

    显示帮助信息。
  - -V, --version

    显示版本号信息。
  - 调用链接口参数

版权所有 © 华为云计算技术有限公司

--trace-id

每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。

取值范围：1-128位的字符串，允许的字符为数字，字母和'-'。

- 缩容参数
  - -h HOSTNAME

    指定需要缩容的DN环节点名称，中间用逗号分隔。
  - -f hostfile

    指定需要缩容的DN环节点名称文件，文件每行为一个DN环节点名称，中间用逗号分隔。

    说明

    文件内容如下：
    host1,host2,host3
    host4,host5,host6
    -h参数和-f参数只能同时指定一个。
  - --parallel-jobs=NUMBER

    指定数据重分布过程的并发执行任务个数。

    说明

    如果不指定--parallel-jobs参数，则数据重分布默认值为1，为非并发执行方式。

    取值范围：[1, 8]，推荐并发任务数为4。
  - --join-tables=*FILE*

    多表切换模式，输入为具有join关系的用户表配置文件。

    取值范围：指定的文件路径，指定文件内容不能超过1000行。

    说明

    指定文件格式。每一行格式为：database schema1 tablename1 schema2 tablename2 ...（每行的多张表构成一个分组（group））。例如：

    ```
    testdb public test1 public test2
    testdb public test3 public test4 public test5
    ```

    注意：

    - 不支持database、schema、tablename数据对象名称存在空格，schema和tablename名称中存在大写字母或者特殊字符时需要加双引号进行转义。
    - 每一行schema tablename必须成对出现，唯一确定一张表。
    - 一张表不能出现在多个分组中，每组的表数量必须大于等于2，小于等于5。
    - 如果存在一个schema里的tablename为其他schema的名字时，会正常处理这张表不会报错。
  - --redis-mode=MODE

    指定数据重分布的执行模式。

    取值范围：

    - read-only：只读模式。
    - insert：插入模式，正在进行重分布的表支持并发插入（insert）以及导入（Gds导入）。

    如果不指定，默认为read-only模式。

 版权所有 © 华为云计算技术有限公司

> **说明**
>
> 和扩容一样，在缩容重分布过程中，为了保障集群业务性能，也会为集群创建名为“redisuser”的数据库用户，此用户只在缩容重分布阶段存在，重分布完成后会被删除。此用户会拥有所有管理员权限来执行重分布。

- 资源管控参数：
  - --resource-level=RESLEVEL

    指定数据重分布的资源等级。

    取值范围：

    - f：100%资源
    - h：50%资源
    - m：30%资源
    - l：10%资源

    默认值：l

    > **说明**
    >
    > 当资源管控参数设置为f时，进入快速重分布场景，会充分使用资源，通过并发操作提升数据重分布的性能。此时，如果shared_buffers参数值设置过小（小于推荐值），会影响此场景下的性能优化效果。

- 重分布参数设置接口参数

  配置对应的全局系统表参数供重分布工具和缩容脚本查询使用，可以在重分布开始前和重分布过程中使用，缩容成功后清除所有参数配置。不设置对应参数则保留上次调用的结果，未设置过的参数重分布开始时写入默认值。

  - --lockwait-timeout

    锁超时时间。如果等锁线程在该时间内无法拿到锁，则报锁超时失败。触发CANCEL模式拿锁时，实际使用的锁超时时间为此参数设置的2倍。

    对应参数：lockwait_timeout

    取值范围：[1,1200000]

    单位：毫秒

    默认值：4000

  - --lockwait-interval

    等锁步长。连续拿锁超时的场景，如果此轮等锁超时，间隔该参数指定的时长进行下一轮尝试。

    对应参数：lockwait_interval

    取值范围：[0,300]

    单位：秒

    默认值：5

  - --trylock-threshold

    等锁尝试次数上限。需和enable_cancel参数同时使用，如果到达该上限还因为等锁超时无法成功获取到锁，则下次拿锁使用CANCEL模式。未开启enable_cancel不生效。

    对应参数：trylock_threshold

    取值范围：[1,INT_MAX], -1（无上限）

 版权所有 © 华为云计算技术有限公司

单位：次

默认值：3

- --enable-cancel

  是否允许用户业务快速失败。即是否允许达到等锁尝试次数上限或切表事务获取用户表8级锁时，快速失败用户业务从而保证重分布成功拿锁。

  对应参数：enable_cancel

  取值范围：[false,true]

  默认值：true

警告

- 此模式默认开启，如果用户业务没有连接级重试机制，建议手动关闭此参数。
- 触发CANCEL模式拿锁时，会发送FATAL级报错，用户业务不是在锁队列中等待而是中断连接。

- --last-catchup-threshold

  最后一轮追增时间上限。即最后一轮完成delete+insert追增耗时。

  对应参数：last_catchup_threshold

  取值范围：[1000,60000]

  单位：毫秒

  默认值：5000

- --catchup-times

  追增轮次上限。超出上限后强制进行最后一轮追增。

  对应参数：catchup_times

  取值范围：[1,INT_MAX], -1

  单位：次

  默认值：-1（表示无上限）

- --write-error-mode

  写报错模式。true表示写报错模式。false表示采用全阻塞模式进行最后一轮追增。

  对应参数：write_error_mode

  取值范围：[false,true]

  默认值：true

警告

- 此模式默认开启，如果用户业务没有会话级重试机制，建议手动关闭此参数。
- 开启写报错模式后，用户写业务会有ERROR级报错，而不是在锁队列中等待。
- 开启此模式后，为避免长时间处于写报错状态，可能触发重分布进程主动退出和自动拉起特性，此时缩容时间会变长，具体请联系管理员处理。

版权所有 © 华为云计算技术有限公司

- --catchup-query-dop

  追增并行扫描度。

  对应参数：catchup_query_dop

  取值范围：[1,32]

  单位：个

  默认值：4

  注：该参数设置过小可能导致追增性能下降，设置过大占用资源较多。

- --parallel-catchup-threshold

  多表重分布模式中开启并行追增所需的耗时阈值，一个分组中所有表追增耗时小于该值时，开启并行追增。

  对应参数：parallel_catchup_threshold

  取值范围：[1000,1800000]

  单位：毫秒

  默认值：300000

- --parallel-reindex-jobs

  同时临时表创建索引时的并发度，即多少个索引同时创建。

  对应参数：parallel_reindex_jobs

  取值范围：[1,64]

  单位：个

  默认值：2

- --redis-retry-times

  自动拉起重分布进程的次数上限，触发重分布进程主动退出时使用，拉起次数超出上限本轮重分布失败结束。

  对应参数：redis_retry_times

  取值范围：[0,100]

  单位：次

  默认值：5

- --redis-retry-interval

  自动拉起重分布进程的时间间隔，触发重分布进程主动退出时使用。

  对应参数：redis_retry_interval

  取值范围：[0,36000]

  单位：秒

  默认值：600

- 降副本参数
  - --offline

    降副本模式，表示离线降副本。

  - --reduce-ips=IPS

    被降副本的节点ip范围，多个使用逗号分隔。

  - --reduce-times=1

    降副本次数，默认降1个副本。

  - --reduce-other

 版权所有 © 华为云计算技术有限公司

支持删除节点上的ETCD/CMS。

示例

数据库实例降副本操作：

```
[Ruby@pekpomdev00063 script]$ gs_shrink -t reduce --offline
{'action': 'reduce', 'online': False}
Init global info successfully.
Create step directory successfully.
Cluster status is : Normal
Cluster balanced is : Yes
Cluster redistributing is : No
Check reduce replica condition successfully.
Successfully backup config file.
Start getting all standby datanode.
[<gspylib.common.DbClusterInfo.instanceInfo instance at 0x7fbae9872368>,
<gspylib.common.DbClusterInfo.instanceInfo instance at 0x7fbae986cdd0>,
<gspylib.common.DbClusterInfo.instanceInfo instance at 0x7fbae93f7320>]
[['10.244.45.69'[['10.244.44.156']][['10.244.44.156', ][, , [['10.244.44.252''10.244.45.69']]]
'10.244.44.252']
]]
Ready to perform sql: select group_members from pgxc_group;
Ready to perform sql: select group_members from pgxc_group;
 Ready to perform sql: select group_members from pgxc_group;
Execute sql command output result is : 16384 16387 16390
node_name in pgxc_node is: dn_6007_6008_6009
Ready to perform sql: select oid from pgxc_node where node_name='dn_6007_6008_6009' and
node_host='172.31.1.103' and node_type='S';
Execute sql command output result is : 16384 16387 16390
node_name in pgxc_node is: dn_6004_6005_6006
Ready to perform sql: select oid from pgxc_node where node_name='dn_6004_6005_6006' and
node_host='172.31.8.236' and node_type='S';
Execute sql command output result is : 16384 16387 16390
node_name in pgxc_node is: dn_6001_6002_6003
Ready to perform sql: select oid from pgxc_node where node_name='dn_6001_6002_6003' and
node_host='172.31.8.236' and node_type='S';
Get node_name nad listen_ip oid in pgxc_node, output result is [16389]
Get instance [10.244.44.252] oid successfully. [16389]
Clear async standby datanode.[10.244.44.252]
Get node_name nad listen_ip oid in pgxc_node, output result is [16386]
Get instance [10.244.44.252] oid successfully. [16386]
Clear async standby datanode.[10.244.44.252]
Get node_name nad listen_ip oid in pgxc_node, output result is [16392]
Get instance [10.244.45.69] oid successfully. [16392]
Clear async standby datanode.[10.244.45.69]
Get reduce datanode successfully.
 ['10.244.44.252', '10.244.45.69']
Stop reduce instance [/var/lib/engine/data1/data/dn_5] success.
Stop reduce instance [/var/lib/engine/data1/data/dn_3] success.
Stop reduce instance [/var/lib/engine/data1/data/dn_9] success.
Start generate new cluster_static_config.
Distribute cluster static config file successfully.
New cluster static config file distribute successfully.
Start update gaussdb.conf in datanode instance.
Update gaussdb.conf successfully.
node_name: dn_6007_6008_6009
listenips: ['172.31.1.103']
 hasips:['10.244.45.69']
Ready to perform sql: delete from pgxc_node where node_host='172.31.1.103' and
node_name='dn_6007_6008_6009' and node_type='S';
Clear node in pgxc_node table successfully.
node_name: dn_6001_6002_6003
listenips: ['172.31.8.236']
 hasips:['10.244.44.252']
Ready to perform sql: delete from pgxc_node where node_host='172.31.8.236' and
node_name='dn_6001_6002_6003' and node_type='S';
Clear node in pgxc_node table successfully.
node_name: dn_6004_6005_6006
```

 版权所有 © 华为云计算技术有限公司

```
listenips: ['172.31.8.236']
 hasips:['10.244.44.252']
Ready to perform sql: delete from pgxc_node where node_host='172.31.8.236' and
node_name='dn_6004_6005_6006' and node_type='S';
Clear node in pgxc_node table successfully.
Update pgxc_node table successfully.
Start update gs_hba.conf .
Update gs_hba successfully.
Starting new cluster.
Clear dynamic config file in cluster successfully.
Kill cluster manager process successfully.
...
Start cluster successfully. Cluster already normal.
Restart cluster finish.Reset dynamic config file successfully.
============================
Reduce replica successfully.
['10.244.44.252', '10.244.45.69']
============================
Clean backup file successfully.
```

相关命令

gs_expand

5.19 gs_sshexkey

说明

系统内部工具，不支持单独调用。

背景信息

数据库实例在安装过程中，需要在数据库实例中的节点间执行命令，传送文件等操作。因此，安装前需要确保互信是连通的。GaussDB提供了gs_sshexkey工具来帮助用户建立互信。

须知

root用户互信可能会存在安全隐患，因此建议用户在使用完成后，立即删除各主机上root用户的互信。

该接口出于安全考虑，前置完成后会自动删除。

使用gs_sshexkey工具建立互信成功之后，会在用户目录下面生成gaussdb_tmp目录，该目录用于存放ssh-agent进程文件gauss_socket_tmp和加密私钥的密码短语文件（随机生成32位密码短语，通过encrypt工具加密），同时用户下会拉起两个ssh-agent进程，此进程用于挂载私钥自动输入对应的密码短语。进程拉起完毕之后，需要source用户的环境变量，才能使互信功能生效。

前提条件

- 确保ssh服务打开。
- 确保ssh端口不会被防火墙关闭。
- 确保xml文件中各主机名称和IP配置正确。
- 确保所有机器节点间网络畅通。

版权所有 © 华为云计算技术有限公司

- 如果为普通用户建立互信，需要提前在各主机创建相同用户并设置密码。
- 执行命令前，需要检查是否所有数据库实例节点都支持expect命令，如果不支持请自行下载expect工具，建议使用expect version 5.45。
- 如果各主机安装并启动了SELinux服务，需要确保/root和/home目录安全上下文为默认值（root目录：system_u:object_r:home_root_t:s0，home目录：system_u:object_r:admin_home_t:s0）或者关闭掉SELinux服务。

  检查系统SELinux状态的方法：执行命令getenforce，如果返回结果是Enforcing ，说明SELinux安装并启用。

  检查目录安全上下文的命令：

  ```
  ls -ldZ  /root | awk '{print $4}'
  ls -ldZ  /home | awk '{print $4}'
  ```

  恢复目录安全上下文命令：

  ```
  restorecon -r -v /home/
  restorecon -r -v /root/
  ```

语法

- 建立互信

  ```
  gs_sshexkey -f HOSTFILE [...] [--skip-hostname-set] [-l LOGFILE]
  ```

- 显示帮助信息

  ```
  gs_sshexkey -? | --help
  ```

- 显示版本号信息

  ```
  gs_sshexkey -V | --version
  ```

参数说明

- -f

  主机列表，列出所有需要建立互信主机的IP。

  说明

  确保hostfile文件中只配置主机IP(每行只写一个IP)，不包含其它信息。

- -l

  指定日志文件的保存路径。

  取值范围：任意存在的可访问的绝对路径。

- --skip-hostname-set

  是否将“-f”参数文件中IP与其hostname的映射关系写入“/etc/hosts”文件中。默认写入，如果指定该参数则不写入。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

示例

如下是为root用户建立互信的示例。

- 用户密码相同情况下，非交互式模式使用以下命令建立互信。

  ```
  ./gs_sshexkey -f /opt/software/hostfile
  Please enter password for current user[root].
  ```

 版权所有 © 华为云计算技术有限公司

```
Password:
Checking network information.
All nodes in the network are Normal.
Successfully checked network information.
Creating SSH trust.
Creating the local key file.
Appending local ID to authorized_keys.
Successfully appended local ID to authorized_keys.
Updating the known_hosts file.
Successfully updated the known_hosts file.
Appending authorized_key on the remote node.
Successfully appended authorized_key on all remote node.
Checking common authentication file content.
Successfully checked common authentication content.
Distributing SSH trust file to all node.
Successfully distributed SSH trust file to all node.
Verifying SSH trust on all hosts.
Successfully verified SSH trust on all hosts.
Successfully created SSH trust.
```

- 用户密码相同情况下，交互式模式使用以下命令建立互信。

```
gs_sshexkey -f /opt/software/hostfile
Please enter password for current user[root].
Password:
Checking network information.
All nodes in the network are Normal.
Successfully checked network information.
Creating SSH trust.
Creating the local key file.
Appending local ID to authorized_keys.
Successfully appended local ID to authorized_keys.
Updating the known_hosts file.
Successfully updated the known_hosts file.
Appending authorized_key on the remote node.
Successfully appended authorized_key on all remote node.
Checking common authentication file content.
Successfully checked common authentication content.
Distributing SSH trust file to all node.
Successfully distributed SSH trust file to all node.
Verifying SSH trust on all hosts.
Successfully verified SSH trust on all hosts.
Successfully created SSH trust.
```

5.20 gs_switch_ddb

功能介绍

gs_switch_ddb工具用于将不带ETCD（DCC）数据库切换到ETCD数据库，或者将带ETCD数据库切换到不带ETCD（DCC）数据库。

前提条件

- 数据库各个实例状态正常，调用节点不能被删，仅支持集中式。
- 需提供目标拓扑信息的XML，该XML中需存在ETCD信息。切换DCC时，该XML中ETCD拓扑和原集群保持一致，仅更改目标CMS信息；切换ETCD时，ETCD拓扑需满足安装条件。
- 若DCC切换到ETCD，目标CMS拓扑需满足原有安装条件；若ETCD切换到DCC，目标CMS拓扑除满足原有安装条件之外，每个AZ至少有一个CMS。

版权所有 © 华为云计算技术有限公司

语法

- 切换ETCD
  ```
  gs_switch_ddb -t switch_etcd -X XMLFILE
  ```
- 切换DCC
  ```
  gs_switch_ddb -t switch_dcc -X XMLFILE
  ```
- 帮助信息
  ```
  gs_switch_ddb -? | --help
  ```
- 版本信息
  ```
  gs_switch_ddb -V | --version
  ```

参数说明

- 通用参数
  - -t
    gs_switch_ddb命令参数类型。
    取值范围：switch_etcd，switch_dcc。
  - -X
    数据库配置文件绝对路径。
    取值范围：clusterconfig.xml的文件路径。
  - -?, --help
    帮助信息，无传入参数。
  - -V, --version
    版本信息，无传入参数。
- 调用链参数
  - --trace-id
    标识id，用于追踪进程调用。

5.21 gs_uninstall

背景信息

GaussDB提供了gs_uninstall工具来帮助完成数据库实例的卸载。

语法

- 卸载数据库实例
  ```
  gs_uninstall [--delete-data] [-L] [-l LOGFILE]
  ```
- 显示帮助信息
  ```
  gs_uninstall -? | --help
  ```
- 显示版本号信息
  ```
  gs_uninstall -V | --version
  ```

参数说明

- --delete-data
  删除数据文件。

 版权所有 © 华为云计算技术有限公司

- -L

  只卸载本地主机。如果数据库实例内某主机做单点卸载后，数据库实例不能再做全量卸载。

- -l

  指定日志文件名及可访问的绝对路径。在内部会自动给日志名添加一个时间戳。

  - 当既不明确指定-l，又不在XML文件中配置gaussdbLogPath时，默认值为“$GAUSSLOG/om/gs_uninstall-YYYY-MM-DD_hhmmss.log”。
  - 当不明确指定-l，但在XML文件中配置了gaussdbLogPath时，默认值为“gaussdbLogPath/用户名/om/gs_uninstall-YYYY-MM-DD_hhmmss.log”。

  说明

  由于在执行gs_uninstall时，系统会自动删除数据库实例相关目录。因此建议用户通过该参数指定日志文件到非GaussDB数据库实例相关路径。

- -?, --help

  显示帮助信息。

- -V, --version

  显示版本号信息。

相关命令

gs_install，gs_postuninstall

5.22 gs_upgradectl

背景信息

用户会根据GaussDB提供的新特性，确定是否对现有系统进行升级。

升级分为就地升级、灰度升级和滚动升级。升级方式的策略又分为大版本升级和小版本升级。用户挑选升级方式后，系统会自动判断并选择合适的升级策略。

- 就地升级：就地升级需要停止业务进行，会一次性升级数据库实例中所有节点。
- 灰度升级：灰度升级支持全业务操作，可以选择部分节点先升级，进入灰度观察期观察新版本是否运行正常，然后再升级剩余节点，完成全部升级操作。
- 滚动升级：滚动升级支持全业务操作，采用滚动升级dn分片的方式对数据库实例进行升级。用户可以指定部分DN分片先升级，观察升级后的分片是否运行正常，然后再升级剩余分片，完成全部升级操作，也可以一次升级所有分片。

注意事项

- 就地升级需要停止业务，灰度升级支持全业务操作。
- 数据库实例运行正常且主DN的数据完全同步到备DN。
- 升级操作不能和主机替换同时执行。不支持虚拟IP。升级过程中不允许打开kerberos开关。
- 选择灰度升级方式进行升级时，灰度节点升级成功后不支持实例修复、节点修复、roach、温备、修改IP和hotpatch；全部节点升级成功后提交前不支持温备、修改IP、hotpatch。

版权所有 © 华为云计算技术有限公司

- 请不要修改安装包中解压出来的version.cfg文件。
- 升级成功后，原归档日志文件将失效。
- 升级成功后，原来的二进制目录将会被删除，请不要在二进制目录中存放个人数据文件。
- 如果就地升级过程中出现异常导致升级失败，并且自动回滚失败时，需要用户自动执行回滚命令进行手动回滚。如果灰度升级过程中出现异常导致升级失败，需用户手动回滚，并且必须回滚成功后才能进行下一次升级。两种升级方式回滚成功后，升级过程中设置的GUC参数将失效。
- 升级过程中，不允许对wal_level，max_connections，max_prepared_transactions和max_locks_per_transaction这四个GUC参数的值进行修改。
- DN磁盘使用率低于85%时才可以执行升级操作。
- 在业务压力大的情况下进行灰度升级，建议开启流控。
- 滚动升级注意事项：
  - 滚动升级要求数据库实例是独立部署的，每个实例都部署在单独的节点上。
  - 滚动升级中，每次升级一个分片都会产生不超过10s的业务中断。
- 升级支持以下故障场景：
  - 少数派VM故障。
  - 少数派进程故障。
  - 多数派VM故障或者多数派进程故障，支持回滚操作。回滚成功后，不保证数据库实例可用。
  - 因“安装目录\bin”下的gaussdb二进制文件损坏导致的多数派进程故障，支持小版本强制升级（内核版本号不变的升级方式为小版本升级，否则就是大版本升级。内核版本号在升级文件压缩包中的version.cfg的第二行查看）。
- 少数派VM故障或少数派进程故障升级约束：
  - 下发升级命令之前，需保证数据库实例状态至少为degrade。
  - 需要下发增加“--ignoreVMFaults”参数的升级命令。
  - 升级过程中，若正在进行升级主流程的VM节点故障，升级报错为正常报错，重新将升级命令下发至正常节点即可。
  - 通过增加“--ignoreVMFaults”参数，被隔离的节点不会被升级。升级完成后，需要通过节点替换完成数据库实例修复。被隔离的进程会忽略掉无法启动的错误。
- 多数派进程故障升级约束：
  - 只支持小版本升级，不支持大版本升级。
  - 升级不带有修复功能，升级后不保证数据库实例一定可用。
  - 强制升级未提交或回滚前，即使数据库实例恢复正常，也无法进行正常的升级操作。
  - 由于“安装目录\bin”下的gaussdb二进制文件损坏导致数据库实例unavailable时支持升级，由于VM级故障导致数据库实例unavailable时不支持升级。

语法

- 显示帮助信息

```
gs_upgradectl -? | --help
```

 版权所有 © 华为云计算技术有限公司

- 显示版本号信息
  gs_upgradectl -V | --version
- 选择升级策略
  gs_upgradectl -t chose-strategy [-l *LOGFILE*]

  说明

  当前支持就地升级、灰度升级和滚动升级。
- 升级数据库实例
  gs_upgradectl -t auto-upgrade -X *XMLFILE* {-g *NODENUMBER* | -h *HOSTNAME* | --continue | --inplace} [-l *LOGFILE*] [--rolling]
  [--grey] [--begin=*ShardStartNumber*] [--end=*ShardEndNumber*] [--force] [--ignoreVMFaults=*HOSTNAME/instance*][--intercluster-rollingupgrade]
- 回滚升级
  gs_upgradectl -t auto-rollback -X *XMLFILE* [-l *LOGFILE*] [--force] [--hadr_force] [--ignoreVMFaults=*HOSTNAME/instance*][--intercluster-rollingupgrade]
- 升级提交
  gs_upgradectl -t commit-upgrade -X *XMLFILE* [-l *LOGFILE*] [--force_commit] [--skip_rebalance_set][--intercluster-rollingupgrade]

  说明

  - 升级提交操作只适用于华为GaussDB数据库的就地升级、灰度升级和滚动升级。
  - 一旦提交操作完成，则不能再执行回滚操作。
- 组件升级
  gs_upgradectl -t component-upgrade -X XMLFILE --component-type=component_type --component-version=component_version [-l *LOGFILE*]
- 组件回滚
  gs_upgradectl component-rollback -X XMLFILE --component-type=component_type [-l *LOGFILE*]
- 接收数据库管理员密码
  echo "*password*" | gs_upgradectl -t [chose-strategy|auto-upgrade|auto-rollback|commit-upgrade|component-upgrade|component-rollback]

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

参数说明

- -t

  gs_upgradectl命令的类型。

  取值范围：chose-strategy、auto-upgrade、auto-rollback、component-upgrade、component-rollback和commit-upgrade。
- -l

  用于记录升级或回滚过程中的日志信息。

  取值范围：任意存在的可访问的绝对路径。

  默认值：$GAUSSLOG/om/gs_upgradectl-YYYY-MM-DD_hhmmss.log
- -?, --help

  显示帮助信息。
- -V, --version

  显示版本号信息。

 版权所有 © 华为云计算技术有限公司

- -X

  指定数据库实例配置文件。

  取值范围：XML文件的存储路径。

- -g

  指定灰度升级先升级的部分节点个数。

  取值范围：选择先灰度升级数据库实例中部分节点时，-g指定节点个数不能超过5；选择灰度升级数据库实例中全部节点时，-g后面指定数据库实例全部节点个数。例如：数据库实例中有7个节点，当选择先灰度升级数据库实例中部分节点时，-g [1-5]；当选择灰度升级数据库实例中全部7个节点时，-g 7。

- -h

  指定灰度升级先升级的部分节点的主机名称。

  取值范围：数据库实例中节点的主机名称，当选择多个节点升级时，用逗号分隔开来。例如：数据库实例中有4个节点，4个节点的主机名称分别为linux1、linux2、linux3和linux4，当选择先灰度升级数据库实例中部分节点linux1和linux2时，-h linux1,linux2。

- --inplace

  使用就地升级方式来进行升级操作。

- --continue

  使用灰度升级来进行升级操作，在升级完部分节点后，使用该参数升级全部剩余节点。

- --force

  当数据库实例状态不正常，无法支持正常回滚时，用此参数进行强制回滚操作。

  当存在因gaussdb二进制文件损坏导致的多数派进程故障时，用此参数进行小版本强制升级。

- --begin

  分片升级时，指定本次升级从第几个分片开始。

- --end

  分片升级时，指定本次升级到第几个分片结束。

- --rolling

  采用AZ升级的方式对实例进行滚动升级。

- --grey

  分片或者AZ升级时，使用灰度升级方式来进行升级操作。

- --force_commit

  同城双中心容灾，两地三中心跨Region容灾或同城双中心高可用支持基于流式复制的异地容灾，使用强制提交方式来提交主数据库实例操作。

- --hadr_force

  容灾数据库实例回滚的一种逃生方式，因容灾关系断连故障时使用。用于同城双中心容灾，两地三中心跨Region容灾和同城双中心高可用支持基于流式复制的异地容灾，当主备数据库实例容灾关系断连，使用容灾强制回滚方式来回滚主数据库实例操作，使其主数据库实例跳过从灾备数据库实例build升级upgrade_phase_info文件的操作。

- --skip_rebalance_set

 版权所有 © 华为云计算技术有限公司

升级提交时是否执行重平衡（cm_ctl switchover -a）。默认执行重平衡（不支持cm_ctl switchover -a命令的场景不执行重平衡）。如果指定该参数则不执行重平衡。

- --ignoreVMFaults=*nodeName1,nodeName2,.../instancename1,instancename2,...*

  少数派VM或少数派进程存在故障时进行升级，使用该参数提供数据库实例中需要忽略的故障的节点名列表或有进程故障的节点名列表，或者故障的实例名列表。

- --intercluster-rollingupgrade

  容灾集群间滚动升级标识。该参数仅支持在容灾场景下的逻辑复制关系中使用。

- --component-type

  组件升级/回滚时指定待升级的组件类型，当前仅有logical_decoding类型。

- --component-version

  组件升级的目标版本号。

示例

示例一：就地升级、灰度升级和滚动升级前使用新包进行前置操作。

```
./gs_preinstall -U roach -G users -X /data/xml/3node_3c3d_1m2s_etcd.xml --alarm-type=5
Parsing the configuration file.
Successfully parsed the configuration file.
Installing the tools on the local node.
Successfully installed the tools on the local node.
Setting pssh path Successfully set core path.
Distributing package.
Begin to distribute package to tool path.
Successfully distribute package to tool path.
Begin to distribute package to package path.
Successfully distribute package to package path.
Successfully distributed package.
Installing the tools in the cluster.
Successfully installed the tools in the cluster.
Checking hostname mapping.
Successfully checked hostname mapping.
Checking OS version.
Successfully checked OS version.
Creating cluster's path.
Successfully created cluster's path.
Setting SCTP service.
Successfully set SCTP service.
Set and check OS parameter.
Setting OS parameters.
Successfully set OS parameters.
Set and check OS parameter completed.
Preparing CRON service.
Successfully prepared CRON service.
Preparing SSH service.
Successfully prepared SSH service.
Setting user environmental variables.
Successfully set user environmental variables.
Configuring alarms on the cluster nodes.
Successfully configured alarms on the cluster nodes.
Setting the dynamic link library.
Successfully set the dynamic link library.
Setting pssh path Successfully set pssh path.
Setting Cgroup. Successfully set Cgroup.
Set ARM Optimization.
Successfully set ARM Optimization.
Setting finish flag.
Successfully set finish flag.
Preinstallation succeeded.
```

 版权所有 © 华为云计算技术有限公司

示例二：使用gs_upgradectl脚本执行灰度升级，选择数据库实例中任意一个节点先升级。

```
gs_upgradectl -t auto-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml -g 1
Static configuration matched with old static configuration files.
Update the cluster with new product version 'GaussDB300'.
Old cluster app path is /data/gauss/app_e67b8bcd
Distributing xml configure file.
Successfully distributed xml configure file.
Checking upgrade environment.
Successfully checked upgrade environment.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Start to do health check.
Successfully checked cluster status.
Upgrade nodes ['hghproach00003'].
Performing grey rollback.
Successfully Cleaned new install path.
Rollback succeeded.
The directory /data/gauss/app_e67b8bcd will be deleted after commit-upgrade, please make sure there is no personal data.
Installing new binary.
Updating catalog.
Successfully updated catalog.
Sync cluster configuration.
Successfully synced cluster configuration.
Switch symbolic link to new binary directory.
Successfully switch symbolic link to new binary directory.
Create checkpoint before switching.
Time to create checkpoint: 3s
Switching all db processes.
Time to switch all process version: 1m33s
The nodes ['hghproach00003'] have been successfully upgraded to new version. Then do health check.
Start to do health check.
Successfully checked cluster status.
Successfully upgrade nodes.
```

示例三：使用gs_upgradectl脚本执行灰度升级，选择数据库实例中指定的一个节点先升级。

```
gs_upgradectl -t auto-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml -h hghproach00003
Static configuration matched with old static configuration files.
Update the cluster with new product version 'GaussDB300'.
Old cluster app path is /data/gauss/app_e67b8bcd
Distributing xml configure file.
Successfully distributed xml configure file.
Checking upgrade environment.
Successfully checked upgrade environment.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Start to do health check.
Successfully checked cluster status.
Upgrade nodes ['hghproach00003'].
Performing grey rollback.
Successfully Cleaned new install path.
Rollback succeeded.
The directory /data/gauss/app_e67b8bcd will be deleted after commit-upgrade, please make sure there is no personal data.
Installing new binary.
Updating catalog.
Successfully updated catalog.
Sync cluster configuration.
Successfully synced cluster configuration.
Switch symbolic link to new binary directory.
Successfully switch symbolic link to new binary directory.
Create checkpoint before switching.
Time to create checkpoint: 3s
```


版权所有 © 华为云计算技术有限公司

```
Switching all db processes.
Time to switch all process version: 1m33s
The nodes ['hghproach00003'] have been successfully upgraded to new version. Then do health check.
Start to do health check.
Successfully checked cluster status.
Successfully upgrade nodes.
```

示例四：使用gs_upgradectl脚本执行灰度升级，先选择数据库实例中一个节点先升级，升级成功后使用--continue参数升级数据库实例所有剩余节点。

```
gs_upgradectl -t auto-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml --continue
Static configuration matched with old static configuration files.
Update the cluster with new product version 'GaussDB300'.
Old cluster app path is /data/gauss/app_e67b8bcd
Distributing xml configure file.
Successfully distributed xml configure file.
Checking upgrade environment.
Successfully checked upgrade environment.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Start to do health check.
Successfully checked cluster status.
From this step, you can use -h to upgrade again if failed.
Sync cluster configuration.
Successfully synced cluster configuration.
Successfully backup hotpatch config file.
Switch symbolic link to new binary directory.
Successfully switch symbolic link to new binary directory.
Create checkpoint before switching.
Time to create checkpoint: 4s
Switching all db processes.
Time to switch all process version: 2m9s
The nodes ['hghproach00001', 'hghproach00002'] have been successfully upgraded to new version. Then do
health check.
Start to do health check.
Successfully checked cluster status.
Upgrade main process has been finished, user can do some check now.
Once the check done, please execute following command to commit upgrade:

    gs_upgradectl -t commit-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml

Successfully upgrade nodes.
```

示例五：使用gs_upgradectl脚本执行灰度升级自动回滚，撤销已经成功/失败的升级操作（灰度升级回滚）。

```
gs_upgradectl -t auto-rollback -X /data/xml/3node_3c3d_1m2s_etcd.xml
Static configuration matched with old static configuration files.
Old cluster app path is /data/gauss/app_e67b8bcd
Performing inplace rollback.
Distributing xml configure file.
Successfully distributed xml configure file.
Checking static configuration files.
Successfully checked static configuration files.
Switch symbolic link to old binary directory.
Successfully switch symbolic link to old binary directory.
Rollbacking catalog.
Stopping node.
Successfully stopped node.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Stopping node.
Successfully stopped node.
Successfully Rollbacked catalog.
Deleting om_monitor.
Successfully deleted om_monitor.
Restoring application and configurations.
```

 版权所有 © 华为云计算技术有限公司

```
Successfully restored application and configuration.
Restoring cluster configuration.
Successfully restored cluster configuration.
Starting the om_monitor process.
Successfully started om_monitor process.
Clean up backup catalog files.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Successfully Cleaned new install path.
Rollback succeeded.
```

示例六：使用gs_upgradectl脚本执行灰度升级提交（灰度升级提交）。

```
gs_upgradectl -t commit-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml
Old cluster app path is /data/gauss/app_e67b8bcd
Successfully Cleaned old install path.
Commit binary upgrade succeeded.
```

示例七：使用gs_upgradectl脚本执行就地升级。

```
gs_upgradectl -t auto-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml --inplace
Static configuration matched with old static configuration files.
Update the cluster with new product version 'GaussDB300'.
Old cluster app path is /data/gauss/app_e67b8bcd
Static configuration matched with old static configuration files.
Distributing xml configure file.
Successfully distributed xml configure file.
Performing inplace rollback.
Distributing xml configure file.
Successfully distributed xml configure file.
Rollback succeeded.
Checking upgrade environment.
Successfully checked upgrade environment.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Start to do health check.
Successfully checked cluster status.
Backing up current application and configurations.
Successfully backed up current application and configurations.
Stopping old clusters.
Successfully stopped old clusters.
Backing up cluster configuration.
Successfully backed up cluster configuration.
Installing new binary.
Restoring cluster configuration.
Successfully restored cluster configuration.
Modifying the socket path.
Successfully modified socket path.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Updating catalog.
Successfully updated catalog.
Stopping node.
Successfully stopped node.
Switch symbolic link to new binary directory.
Successfully switch symbolic link to new binary directory.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Start to do health check.
Successfully checked cluster status.
Upgrade main process has been finished, user can do some check now.
Once the check done, please execute following command to commit upgrade:

  gs_upgradectl -t commit-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml
```

 版权所有 © 华为云计算技术有限公司

示例八：使用gs_upgradectl脚本执行自动回滚，撤销已经成功/失败的升级操作（就地升级回滚）。

```
gs_upgradectl -t auto-rollback -X /data/xml/3node_3c3d_1m2s_etcd.xml
Static configuration matched with old static configuration files.
Old cluster app path is /data/gauss/app_e67b8bcd
Performing inplace rollback.
Distributing xml configure file.
Successfully distributed xml configure file.
Checking static configuration files.
Successfully checked static configuration files.
Switch symbolic link to old binary directory.
Successfully switch symbolic link to old binary directory.
Rollbacking catalog.
Stopping node.
Successfully stopped node.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Stopping node.
Successfully stopped node.
Successfully Rollbacked catalog.
Deleting om_monitor.
Successfully deleted om_monitor.
Restoring application and configurations.
Successfully restored application and configuration.
Restoring cluster configuration.
Successfully restored cluster configuration.
Starting the om_monitor process.
Successfully started om_monitor process.
Clean up backup catalog files.
Starting cluster.
Waiting for the cluster status to become normal.
The cluster status is normal.
Successfully Cleaned new install path.
Rollback succeeded.
```

示例九：使用gs_upgradectl脚本执行就地升级后提交（就地升级提交）。

```
gs_upgradectl -t commit-upgrade -X /data/xml/3node_3c3d_1m2s_etcd.xml
Old cluster app path is /data/gauss/app_e67b8bcd
Successfully Cleaned old install path.
Commit binary upgrade succeeded.
```

示例十：组件升级。

```
echo "pwd" | gs_upgradectl -t component-upgrade  -X /usr/local/tool/clusterConfig.xml --component-type
logical_decoding --component-version 1.95999
Init global infos in the command parsing phase
Successfully init global infos in the command parsing phase
NOTICE: the clusterType is dual-standby-dorado
Start to upgrade the component.
Successfully to do component-upgrade.
```

5.23 gns_ctl

背景信息

GaussDB数据库进行计划内维护时（switchover、restart DN、重启节点和重启数据库实例操作），OM相关接口将调用gns_ctl命令完成该次操作（而不是cm_ctl）；gns_server将通知JDBC客户端数据库实例即将进行变更，JDBC客户端将继续向服务端发送正在运行的事务、而暂存新事务，不再向数据库发送，待正在运行的事务完成后挂起和DN的连接。事务运行结束后，JDBC向GNS发送确认消息，并在维护完成后，JDBC开始恢复与DN的连接。

版权所有 © 华为云计算技术有限公司

前提条件

- 数据库实例已安装GNS组件。
- GNS组件配置正确，支持计划内无损透明功能。

语法

- 停止数据库实例、节点或者DN

  gns_ctl stop [-n *<NODE_ID>* [-D *<DATA_DIR>*]]
- 主备切换

  gns_ctl switchover -n *<NODE_ID>* -D *<DATA_DIR>*

  说明

  1. 执行gns_ctl 后不会自动重启DN、节点或者数据库实例，需要执行cm_ctl start命令手动启动。
  2. gns_ctl只能在安装有gns_server的节点上使用。
- 显示帮助信息

  ```
  gns_ctl -? | -h | --help
  ```
- 显示版本号信息

  ```
  gns_ctl -V | --version
  ```
- 显示数据库实例中gns_server和JDBC客户端的连接信息

  ```
  gns_ctl query
  nodeId          gnsInstId       port            listenNum       jdbcConnectedNum
  1               20001           25100           1024            0
  2               20002           25100           1024            0
  3               20003           25100           1024            0
  ```

参数说明

表 5-51 公共参数说明

参数	参数说明	取值范围
-n *<NODE_ID>*	DN节点ID。	无符号整型。
-D *<DATA_DIR>*	指定DN的路径。	绝对路径。

- stop 参数

  stop 后不跟参数时表示停止数据库实例，只带-n参数时表示停止节点，同时带-n和-D参数时表示停止DN。
- switchover 参数

  需要同时带-n和-D参数，将-n指定节点上的-D指定绝对路径下的备DN升为主DN。

版权所有 © 华为云计算技术有限公司

5.24 gstrace

功能介绍

gstrace是GaussDB提供的用来侦听内核代码执行路径，记录内核数据结构，分析代码性能的工具。Trace的有限点位和数据在版本中被固化，无法动态添加和删除。

警告

1. 对内核dump指定内存变量的数据用于诊断分析，不存在直接指定任意寄存器或者内存地址的行为。读取的内存地址均是在开发阶段固化，没有任意地读取或者修改寄存器或内存的操作。
2. Trace点可能涉及敏感数据，收集trace信息前需要同用户协商，授权和许可后方可收集。
3. GaussDB不会在敏感信息上打点，不会TRACE和用户相关的数据。
4. Trace仅用于诊断目的，开启trace将对性能产生一定影响，影响的大小视负载的高低，trace的模块而不同。
5. Trace工具的权限为0700，仅限于数据库用户读、写和执行。

说明

如果进程异常终止，/dev/shm/ 目录下将会有gstrace_trace_cfg_*残留，可以手动清除。

语法

```
gstrace [start|stop|config|dump|detailcodepath|analyze] [-p <port>][-s <BUFFER_SIZE>] [-f <DUMP_FILE>] [-o <OUTPUT_FILE>] [-t <STEP_SIZE>]
```

参数说明

表 5-52 gs_trace 参数说明

参数	说明
start	开始记录trace。
stop	停止trace，释放存储信息的共享内存trace buffer。注意：trace buffer中已捕获trace信息会丢失。
config	显示trace的配置信息。
dump	将共享内存中的trace信息写入指定文件。若没有启动trace将报错。
detail	将dump出来的二进制文件解析成文本文件，显示trace点的线程、时间信息。
codepath	提取dump文件中的函数调用信息，按照调用栈的方式显示。

版权所有 © 华为云计算技术有限公司

参数	说明
analyze	统计各个函数的执行次数、总耗时、平均耗时、最长耗时和最短耗时。
-p PORT	指定启动trace功能的实例侦听的端口号。
-f DUMP_FILE	指定dump导出的trace文件。
-o OUTPUT_FILE	指定写入的文件。
-t STEP_SIZE	指定分片分析的时间跨度（秒）,可选。将生成单独的{OUTPUT_FILE}.step文件。
-m MASK	指定哪些模块、函数需要追踪，如果参数为空，则追踪添加trace接口的所有模块的所有函数。 格式:：[comp…][ALL].[func…][ALL] 描述： 1. comp是模块列表，使用‘,’分隔，例如：executer,kernel。也可以使用序号，例如：executer,2 2. func是function列表，使用‘,’分隔。例如：sql_execute_ddl, sql_execute_dcl。也可以使用序号指代模块中第几个function，例如：2 3. 设置成ALL表示所有模块或所有function被trace。 示例： 1. executer,kernel.ALL：所有定义在executer和kernel模块下的function被trace。 2. executer,kernel.sql_execute_ddl：定义在executer 下的Function sql_execute_ddl被trace. kernel模块被忽略，因为下面没有该function。 3. executer,kernel.1：定义在模块executer和 kernel 下的第一个function被trace。 ALL.1,2：定义在所有模块下的第一，二个function被trace。
-s BUFFER_SIZE	指定用于trace功能的共享内存大小，默认为1G。如果指定的BUFFER_SIZE小于最小值2048，系统将自动调整为最小值。如果指定的BUFFER_SIZE不是2^N（二的N次方），则向下对齐2^N；例如：指定BUFFER_SIZE=3072，由于2^11<3072<2^12，系统将调整为2048。

示例

1. 启动trace
   ```
   gstrace start -p 8000
   ```
2. 停止trace
   ```
   gstrace stop -p 8000
   ```
3. 查看trace配置
   ```
   gstrace config -p 8000
   ```
4. dump trace
   ```
   gstrace dump -p 8000 -o /data/8000.dump
   ```

 版权所有 © 华为云计算技术有限公司

5. 解析dump detail信息

```
gstrace detail -f /data/8000.dump -o /data/8000.detail
1       ENTRY, Pid: 41853, Tid: 41911, Function: StartTransaction MicroSecond:902914 Fri Sep  2
11:50:19 2022
2       EXIT, Pid: 41853, Tid: 41911, Function: StartTransaction MicroSecond:903018 Fri Sep  2 11:50:19
2022
```

6. 解析dump codepath

```
gstrace codepath -f /data/8000.dump -o /data/8000.codepath
1        StartTransaction ENTRY
2        StartTransaction EXIT 104.000000
3        StartTransaction ENTRY
4        StartTransaction EXIT 114.000000
pid: 41853  tid: 41923
15       StartTransaction ENTRY
16       StartTransaction EXIT 100.000000
```

7. 分析全局性能

```
gstrace analyze -f /data/8000.dump -o /data/8000.perf
    MODULE                  FUNCTION      #CALL ELAPSETIME(total) ELAPSETIME(avg.)
ELAPSETIME(max.) ELAPSETIME(min.) SEQUENCE(max.) SEQUENCE(min.)
    access              StartTransaction        11          1235          112          176
66           20           8
```

8. 分析分片性能

```
gstrace analyze -f /data/8000.dump -o /data/8000.perf -t 1
    MODULE                  FUNCTION      #CALL ELAPSETIME(total) ELAPSETIME(avg.)
ELAPSETIME(max.) ELAPSETIME(min.) SEQUENCE(max.) SEQUENCE(min.)
    access              StartTransaction        11          1235          112          176
66           20           8
```

5.25 kadmin.local

功能介绍

为数据库实例认证服务Kerberos提供命令行接口，直接接入kdc服务的数据库，进行增加、删除、修改Kerberos用户等操作。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

5.26 kdb5_util

功能介绍

为数据库实例认证服务Kerberos提供数据库管理操作的工具，可以新建、销毁、导入和导出数据库。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

 版权所有 © 华为云计算技术有限公司

5.27 kdestroy

功能介绍

为数据库实例认证服务Kerberos提供销毁用户的授权票据操作。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

5.28 kinit

功能介绍

为数据库实例认证服务Kerberos提供特定用户获取和缓存TGT能力。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

5.29 klist

功能介绍

为数据库实例认证服务Kerberos提供列出凭证缓存中的用户和票据能力。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

5.30 krb5kdc

功能介绍

为数据库实例认证服务Kerberos提供列认证和密钥管理服务。

参数说明

Kerberos工具为开源第三方提供，具体参数说明请参考Kerberos官方文档：**https://web.mit.edu/kerberos/krb5-1.17/doc/admin/admin_commands/index.html**

版权所有 © 华为云计算技术有限公司

5.31 om_monitor

功能介绍

om_monitor的主要功能是注册在各个物理节点上的定时任务，负责数据库管理组件cm_agent的启动。

参数说明

无。

5.32 omRollback.py

说明

系统内部工具，不支持单独调用。

功能介绍

omRollback.py是在底层服务使用agent模式（区别于ssh模式的安装数据库的方式）下部署数据库时，升级过程中回滚代码的工具。

语法

```
python3 omRollback.py
```

- 接收数据库管理员密码
  ```
  echo "password" |  python3 omRollback.py
  ```

  说明

  云上环境中，对于未开启Peer认证方式的数据库系统，可以使用此方式传入数据库管理员用户的密码，用于连接数据库进行运维操作。

前提条件

- 数据库已经进入升级过程且升级未完成。
- 数据库中不存在故障节点。
- 数据库安装python3.7.4/3.8.5且依赖包齐全。

5.33 gs_controldata

功能介绍

gs_controldata工具用来显示数据库目录下control文件的信息。

前提条件

确保数据目录正确，以及相应数据目录下的pg_control文件存在。

版权所有 © 华为云计算技术有限公司

语法

- 查看control文件信息
  ```
  gs_controldata [OPTION]... [<DATA_DIR>]
  ```
- 显示版本号信息
  ```
  gs_controldata -V | --version
  ```
- 显示帮助信息
  ```
  gs_controldata
  -? | --help
  ```

参数说明

- <DATA_DIR>

  查看control文件信息时所用的参数，<DATA_DIR>为实例对应的数据库目录。

OPTION取值如下所示

- -I, –instance-id=INSTANCE_ID

  共享存储参数，打印指定实例的control文件信息，该参数不指定时打印全部实例的control文件信息。
- --enable-dss

  共享存储参数，开启共享存储功能。
- --socketpath=SOCKETPATH

  共享存储参数，dss实例进程使用的socket文件路径，仅支持绝对路径。
- -P, --pdb_control

  打印PDB控制文件信息。如果没有预先安装初始化多租特性，则不存在PDB控制文件，尝试打印会出现报错误。
- -V, --version

  显示版本信息。
- -?,--help

  打印帮助信息。

5.34 gs_resetxlog

功能介绍

gs_resetxlog是一个重新设置数据库事务文件的工具。使用该工具后，数据库有可能会含有不完整的数据，属于高危操作。只有在数据库无法启动、用户强烈期望数据库启动的场景下使用。除此之外不建议使用。

⚠ 注意

通过gs_resetxlog重新设置之前，需要先停止数据库。

前提条件

保证数据库目录存在且路径正确。

版权所有 © 华为云计算技术有限公司

语法

```
gs_resetxlog [OPTION]... <DATA_DIR>
```

参数说明

OPTION取值如下所示：

- <DATA_DIR>
  需要修改的数据库目录，确保路径正确。
- -e XIDEPOCH
  设置下一个事务id。
- -f
  强制更新。
- -l xlogfile
  为新的事务日志指定最小的WAL起始位置。
- -m XID
  设置下一个multitransaction ID。
- -n
  不更新，仅显示配置文件的值。
- -o OID
  设置下一个OID。
- -O OFFSET
  设置下一个multitransaction的偏移量。
- -V, --version
  显示版本信息。
- -x XID
  设置下一个事务ID。
- -?, --help
  打印帮助信息。

5.35 pscp

基本功能

用于执行数据库节点间的访问，发送文件给远程节点。

语法

```
python3 pscp [OPTIONS] localPath remote
```

参数说明

- --help
  显示帮助信息。

版权所有 © 华为云计算技术有限公司

- -H HOSTNAME
  主机列表。
- -h HOSTFILE
  写入主机列表的文件全路径。
- -t TIMEOUT
  连接超时时间，单位秒，默认300s。
- -p PARALLEL
  最大并行个数，默认256。
- -o OUTDIR
  执行结果输出全文件路径。
- -e ERRDIR
  错误信息输出全文件路径。
- -r
  递归复制目录。
- -v
  打开诊断消息
- -s
  显示执行结果。
- -x EXTRA
  补充的scp的参数
- -i
  显示所有节点的输出和错误。
- -O OPT
  补充的scp的参数。
- --trace-id
  日志id号。

5.36 pssh

基本功能

用于执行数据库节点间的访问，发送命令给远程节点。

说明

pssh是一个内部工具，且使用规格受限。当om agent侧到沙箱内执行命令时，为了兼容转义字符等情况，执行命令时带了-f参数，不做转义符展开，故此接口不支持展开转义符。

语法

```
python3 pssh [OPTIONS]
```

 版权所有 © 华为云计算技术有限公司

参数说明

- --help
  显示帮助信息。
- -H HOSTNAME
  主机列表。
- -h HOSTFILE
  写入主机列表的文件全路径。
- -t TIMEOUT
  连接超时时间，单位秒，默认300s。
- -r RETRY
  重试次数。
- -p PARALLEL
  最大并行个数，默认256。
- -o OUTDIR
  执行结果输出全文件路径。
- -e ERRDIR
  错误信息输出全文件路径。
- -P
  线上远端主机执行的命令。
- -s
  显示执行结果。
- -x EXTRA
  线下远端主机执行的命令
- -i
  显示所有节点的输出和错误。
- -O OPT
  补充的线下模式ssh的参数。
- -b --b64encode
  云上模式执行远端命令时是否需要对命令进行base64转码操作，线下模式此参数无效。
- --trace-id
  日志id号。
- -b, --b64encode
  是否进行Base64解码命令在代理模式下，此参数在SSH方式无效。
- --mock-cmd
  通过pssh工具执行的cmd中包含敏感信息时，可通过添加--mock-pwd参数掩盖进程中显示的cmd信息。

 版权所有 © 华为云计算技术有限公司

5.37 py_pstree.py

说明

系统内部工具，不支持单独调用。

功能介绍

py_pstree.py用来获取指定进程的所有父进程或者所有子进程。

语法

- 获取指定进程所有父进程列表
  ```
  py_pstree.py -p PID
  ```
- 获取指定进程所有子进程列表但不包括指定的进程
  ```
  py_ps_tree.py -c PID
  ```
- 获取指定进程所有子进程包括指定进程
  ```
  py_pstree.py -sc PID
  ```
- 获取帮助信息
  ```
  py_pstree.py -h | --help
  ```
- 显示版本号信息
  ```
  py_pstree.py --version
  ```

参数说明

- -p,--parents

  获取指定进程所有父进程列表。
- -c,--children

  获取指定进程所有子进程列表。
- -s,--self

  获取指定进程所有子进程列表中是否包括指定进程本身。
- -h,--help

  显示帮助信息。
- --version

  显示版本号信息。

5.38 LogicalStandbyDR.py

说明

容灾场景下，对备集群做复制关系切换使用。可以将备集群从物理复制转为逻辑复制，也可以将逻辑复制转为物理复制。

功能介绍

LogicalStandbyDR.py用来在容灾场景下对备集群做物理转逻辑或逻辑转物理操作。

版权所有 © 华为云计算技术有限公司

语法

- 物理转逻辑第一步，需要在备集群执行。
  ```
  echo password | LogicalStandbyDR.py -t switch_logical_standby --step prepare_standby
  ```
- 物理转逻辑第二步，需要在主集群执行并获取返回的uuid。
  ```
  echo password | LogicalStandbyDR.py -t switch_logical_standby --step prepare_primary
  ```
- 物理转逻辑第三步，需要在备集群执行同时要将第二步的uuid作为参数传入。
  ```
  echo password | LogicalStandbyDR.py -t switch_logical_standby --step do_switch --uuid uuid1
  ```
- 上述三步执行成功后完成容灾备集群的物理转逻辑操作，同时支持在上述失败后执行取消操作。
  ```
  echo password | LogicalStandbyDR.py -t cancel_logical_standby
  ```
- 支持逻辑转物理操作，需要在备集群执行。
  ```
  echo password | LogicalStandbyDR.py -t switch_physical_standby
  ```

参数说明

- -t [switch_logical_standby|cancel_logical_standby|switch_physical_standby]

对备集群复制关系切换的类型。

- --step [perpare_standby|prepare_primary|do_switch]

物理转逻辑时的步骤参数。

- --uuid

物理转逻辑时备集群需要校验的uuid，由主集群产生。

- --timeout

备集群复制关系转换的超时时间。

- --trace-id

每次om请求的唯一标识，可根据om操作日志中的trace-id识别调用链从而进行故障定界。可选参数。取值范围：1~128位的字符串，允许的字符为数字、字母和'-'。

 版权所有 © 华为云计算技术有限公司

6 统一数据库管理工具

CM（Cluster Manager）是一款数据库管理模块。支持自定义资源监控，提供了数据库主备的状态监控、网络通信故障监控、文件系统故障监控、故障自动主备切换等能力。提供了丰富的数据库管理能力，如节点、实例级的启停，数据库实例状态查询、主备切换和日志管理等。

说明

沙箱部署场景，所有工具需要进入到沙箱内部才能正确执行。

进入沙箱命令参考：chroot /var/chroot

6.1 特性介绍

集群管理，负责实例故障后的高可用切换能力。组件详见表6-1。

表 6-1 组件介绍

组件（二进制名）	描述
OMM（om_monitor）	管理服务组件，由crontab定时任务控制拉起（周期1分钟），OMM主要负责OMM、etcd、cm_agent进程的保活及启停。其故障会产生如下影响： etcd、cm_agent进程保活及启停能力丢失。
etcd（etcd）	管理服务组件，由OMM拉起（周期1秒），etcd是集群高可用的顶层组件，能力可以协助CMS选主、持久化集群仲裁信息。其故障会产生如下影响： CMS无主，进而导致集群无法查看状态、无法进行高可用切换（DCF自仲裁模式除外）。

版权所有 © 华为云计算技术有限公司

组件（二进制名）	描述
DDB（distribute database）	管理服务组件，用于CM Server自仲裁选主和持久化存储仲裁数据。其故障会产生如下影响： ● CM Server无主，进而导致集群无法查看状态、无法进行高可用切换（DCF自仲裁模式除外）。 CM Server可以选择使用ETCD或者DCC，由参数**ddb_type**控制， 单节点集中式不支持。 ● DCC（Distributed Configuration Center）：分布式配置中心，基于DCF（基于Paxos协议的高可用）实现的一个状态机，用于实现集群中配置信息的管理。 ● ETCD：高度一致的分布式键值存储组件。
CM Agent（cm_agent）	管理服务组件，由OMM拉起（周期1秒），CMA主要负责CMS、DN进程的保活及启停，仲裁指标采集、仲裁命令执行等。其故障会产生如下影响： ● CMS、DN进程保活及启停能力丢失。 ● 节点脱管，实例级故障检测能力丢失，被管理实例状态未知。 **说明** 当CMA故障或CMA与CMS网络断连后，此节点处于脱管状态，CMS主（集群仲裁大脑）无法接收到此节点上所有实例的心跳，此时即使被管理实例是正常的，也有可能发生高可用倒换。倒换原则： ● 参数instance_heartbeat_timeout：DN超过此参数无上报消息时，状态置为UNKNOWN，进而触发无主仲裁。 ● 参数cma_fault_kill_instance_timeout：DN主超过此参数无上报消息时，CMS认定此实例脱管时间过长，会通过pssh远程执行中止命令，尝试中止该实例。 以上参数默认值和关闭等功能的介绍可参考**cm_server参数**。
CM Server（cm_server）	管理服务组件，由CMA拉起（周期1秒），是集群的仲裁中心，依赖etcd（或启用DCC）实现自选主。CMS无主会产生以下影响： ● 集群状态无法查看。 ● 整个集群的高可用切换能力丢失（DCF自仲裁模式除外）。 ● 容灾barrier无法推进。 ● DCC模式下，无法持久化集群信息。
DN（gaussdb）	数据服务组件，由CMA拉起（周期1秒）。其故障会产生以下影响： ● 异步备故障：集群无影响。 ● 同步备故障：数据服务卡顿，卡顿时间取决于异步备到同步的日志量。 ● 主故障：数据服务不可用。 ● 多数派故障：数据服务卡顿直至恢复多数派或降副本。

版权所有 © 华为云计算技术有限公司

须知

CM特性如果涉及网络且未单独说明的场景，是指集群管理面IP（CM Agent、CMS所在网络平面IP）和数据面（数据库组件通信平面）的场景。集群管理面IP和数据面的场景支持检测和倒换，其他网络平面故障暂不支持检测和倒换。

cm_agent

cm_agent是部署在数据库每个主机上，用来启停和监控各个数据库实例进程的数据库管理组件。

主要功能有：

- 数据库实例启动和停止时负责拉起和停止本主机上部署的实例进程。
- 监控本主机上运行的实例状态并将状态上报发送给CM Server。
- 执行CM Server仲裁下发的命令。

故障影响：

- 无法上报本主机上部署的实例状态，cm_server会将这些实例视作故障处理。

命令说明：

- 显示帮助信息

```
cm_agent -? | -h | --help

cm_agent is a utility to monitor an instance.

Usage:
  cm_agent
  cm_agent 0
  cm_agent 1
  cm_agent 2
  cm_agent 3
  cm_agent normal
  cm_agent abnormal

Common options:
  -?, -h, --help        show this help, then exit
  -V, --version         output version information, then exit

location of the log information options:
  0                     LOG_DESTION_FILE
  1                     LOG_DESTION_SYSLOG
  2                     LOG_DESTION_FILE
  3                     LOG_DESTION_DEV_NULL

started mode options:
  normal                cm_agent is started normally
  abnormal               cm_agent is started by killed
```

- 显示版本号信息

```
cm_agent -V | --version
```

- 启动方式

```
cm_agent | normal | abnormal | 0 | 1 | 2 | 3
```

参数说明：

 版权所有 © 华为云计算技术有限公司

表 6-2 参数说明

参数	参数说明	参数描述
-V, --version	显示版本号信息。	-
-?, -h,--help	显示关于CM Agent命令行参数的帮助信息。	-
0	记录在设定的日志文件中。	日志信息记录的位置选项。
1	记录在syslog文件中。	日志信息记录的位置选项。
2	记录在设定的日志文件中。	日志信息记录的位置选项。
3	空文件，即不记录日志信息。	日志信息记录的位置选项。
normal	正常退出后，启动CM Agent。该启动模式用于集群启动，集群升级和回滚阶段。	启动模式选项（由om_monitor 指定的 CM Agent 启动模式）。
abnormal	异常退出后，启动CM Agent，默认启动模式。	启动模式选项（由om_monitor 指定的 CM Agent 启动模式）。

cm_server

cm_server是用来进行数据库实例管理和实例仲裁的组件。

主要功能有：

- 接收各个节点上cm_agent发送的数据库各实例状态。
- 提供数据库实例整体状态的查询功能。
- 监控实例的状态变化并进行仲裁命令的下发。

DCC模式：

DCC是自研分布式配置中心组件，底层基于DCF分布式共识框架实现，支持KV存储。具有高可用能力，cm_server可在dcc模式下实现自选主。

故障影响：

- cm_server主故障后，会自动进行主备切换。cm_server全部故障后，数据库实例管理和实例仲裁功能丧失，数据库实例内其他组件按原状态继续运行。

命令说明：

- 显示帮助信息

```
cm_server -? | -h | --help

cm_server is a utility to arbitrate an instance.
```

 版权所有 © 华为云计算技术有限公司

```
Usage:
  cm_server
  cm_server 0
  cm_server 1
  cm_server 2
  cm_server 3

Common options:
  -?, -h, --help          show this help, then exit
  -V, --version           output version information, then exit

location of the log information options:
  0                       LOG_DESTION_FILE
  1                       LOG_DESTION_SYSLOG
  2                       LOG_DESTION_FILE
  3                       LOG_DESTION_DEV_NULL
```

- 显示版本号信息

```
cm_server -V | --version
```

- 启动方式

```
cm_server | 0 | 1 | 2 | 3
```

参数说明：

表 6-3 参数说明

参数	参数说明	参数描述
-V, --version	显示版本号信息。	-
-?, -h,--help	显示关于CM Server命令行参数的帮助信息。	-
0	记录在设定的日志文件中。	日志信息记录的位置选项。
1	记录在syslog文件中。	日志信息记录的位置选项。
2	记录在设定的日志文件中。	日志信息记录的位置选项。
3	空文件，即不记录日志信息。	日志信息记录的位置选项。

CM 支持日志管理

- 简介：CM支持对$GAUSSLOG下的日志进行压缩和删除。
- 管理对象：$GAUSSLOG路径下**表1 日志关键字**中对应的日志。
- 压缩删除时机：enable_log_compress开启后，间隔log_threshold_check_interval。
- 日志压缩：
  - 对于单个目录下相同工具名的日志，除时间戳最新的日志外，剩余日志全部压缩。
  - 日志压缩后与压缩前路径保持一致，压缩文件命名格式为"工具名-日志创建时间.log.gz"。

版权所有 © 华为云计算技术有限公司

- 日志删除：
  - 当所有日志总大小超过（log_max_size*95/100）MB时，根据压缩日志文件名时间，依次删除历史压缩日志，直到日志总大小小于（log_max_size*95/100）MB。
  - 当所有日志总个数超过log_max_count，根据压缩日志文件名时间，依次删除超过保留天数log_saved_days的压缩日志，直到日志总个数小于log_max_count。
  - 日志压缩能力受内存限制，最多只能检测到30000个日志文件。总日志量超过30000个文件时，无法保证日志能被正常压缩及删除。可以通过调整log_saved_days和log_threshold_check_interval快速清理已压缩日志文件。

表 6-4 日志关键字

默认值
"cm_client-", "cm_ctl-", "gs_clean-", "gs_ctl-", "gs_guc-", "gs_dump-", "gs_dumpall-", "gs_restore-", "gs_upgrade-", "gs_initdb-", "gs_check-",
"cm_agent-", "system_call-", "cm_server-", "om_monitor-", "gs_local-", "gs_preinstall-", "gs_install-", "gs_replace-", "gs_uninstall-", "gs_om-", "pssh-", "pscp-",
"gs_upgradectl-", "gs_expand-", "gs_shrink-", "gs_postuninstall-", "gs_backup-", "gs_checkos-", "gs_collector-", "GaussReplace-", "GaussOM-", "gaussdb-",
"roach_agent-", "roach_controller-", "sync-", "gaussdb-", "sessionstat-", "system_alarm-", "pg_perf-", "slow_query_log-", "asp-", "etcd-", "etcd_", "gs_cgroup-",
"gs_hotpatch-", "cmd_sender-", "uploader-", "checkRunStatus-", "ffic_gaussdb-", "key_event-", "gs_streaming_disaster_recovery-", "mem_log-", "remote_client-", "system_stat-", "gs_change_mode-", "system_stat-", "plsqlffic_gaussdb-"

级联备仲裁

cm支持包含级联备组网下的DN仲裁。

主要功能有：

- CMA以pending方式拉起级联备，CMS下发notify消息，级联备pending DN升成级联备。
- 级联备支持switchover，只支持switchover –n –D（级联备 > 备机，备机 > 级联备），switchover -a只能恢复第一个节点为主DN，内部执行两次switchover操作（级联备 > 备机 > 主机）。
- 不支持failover操作。
- 支持级联备组网下的升降副本功能。
- 说明：
  - 级联备默认存在于DN的同步列表中。
  - 级联备不是正常副本。

 版权所有 © 华为云计算技术有限公司

- 约束条件：
  - 最小组网：一主两备一级联。
  - 副本数：最多8个副本。
  - 级联备不检查磁盘使用率。
  - 级联备组网不支持强启功能。

按分片自动升降副本

- 正常降副本功能
  - 简介：两AZ+仲裁AZ部署方式支持自动升降副本功能，修改DN的同步列表。
  - 描述：金融、银行业对业务需要极高的容灾能力。当一个分片故障了半数及以上节点，因同步备中有节点发生故障，无法执行写操作，DN执行写操作会超时。为了降低对业务的影响，需要对分片上可用节点进行降副本操作，当检测到故障恢复后，自动触发升副本操作。
  - 约束条件：
    - 基础保障（最多降至）：一主一备。
    - 部署要求：两AZ+仲裁AZ，正常副本数大于3。
    - 前提要求：DN主存在，或者DCF主存在，或者剩余副本数能选主成功。
    - 升级、扩容阶段，或者ETCD不可用时不会进行降副本操作。
    - 只有半数以上DN发生故障（DN状态是down，down状态一般为管理面网络不通或进程不在），且状态持续，才会进行降副本操作，升副本需要等待半数故障恢复后，且状态持续，才会自动升副本。只有当上一轮降副本操作执行成功后，才能进行下一轮降副本操作，不支持二次故障。
    - 不支持故障跳转，比如四个副本，第一次（3, 4）故障后，（1, 2）进行完降副本，第二次故障（1, 2），恢复（3, 4），此时数据库实例不可用，无法选出主，且不能对（3, 4）进行降副本。
    - 在升降副本结束后，才能执行switchover，且switchover只能切换到同步列表中的备DN上。
    - 故障节点恢复后，需要跟主机同步达到99%，才会被重新加入到主机的同步列表中。
    - 容备场景不开启降副本功能，dorado实例不开启降副本功能。
    - 升降副本过程中，不允许切换级联备机。
- 降副本至一主零备
  - 简介：多副本实例支持升降副本功能，修改DN同步列表。
  - 描述：当一个分片所有备机都故障后，无法执行业务。为了降低对业务的影响，需要对分片上可用节点进行降副本操作，当检测到故障恢复后，自动触发升副本操作。
  - 功能开启：设置参数请参见 **enable_synclist_single_inst**。
  - 约束条件：

 版权所有 © 华为云计算技术有限公司

- 前提要求：DN主存在，或者DCF主存在。
- 升级、扩容阶段，或者ETCD不可用时，不会进行降副本操作。
- 只有所有备机都发生故障（DN状态是unknown, 或者主DN上没有备机的连接）， 且状态稳定时，自动触发降副本操作，当检测到备机恢复后，自动触发升副本操作。只有当上一轮降副本操作执行成功后，才能进行下一轮降副本操作，不支持二次故障。
- 不支持故障跳转，比如四个副本，第一次（3, 4）故障后，（1, 2）进行完降副本，第二次故障（1, 2），恢复（3, 4），此时数据库实例不可用，无法选出主，且不能对（3, 4）进行降副本。
- 在升降副本结束后，才能执行switchover，且switchover只能切换到同步列表中的备DN上。
- 故障节点恢复后，需要跟主机同步达到99%，才会被重新加入到主机的同步列表中。

  容备场景不开启降副本功能，dorado实例不开启降副本功能。
- DCF自动模式不开启降副本功能。
- 为了保证数据不丢失选主只能在同步列表中选主，如果不在同步列表里的DN，无法选主。降副本到一主零备，如果原主故障，则无法选主，只能等待原主恢复再进行选主。
- 单副本实例不开启降副本功能。
- 升降副本过程中，不允许切换级联备机。
- 只支持DCF手动模式。

危险

降为一副本运行，可能只有主机上保存完整日志，当主机无法恢复，存在丢失数据风险。出现无主情况，需要恢复，请联系华为工程师。

- **备注说明**
  - 正常副本：能够升主的DN副本，包括主机，同步备DN，不包括级联备。
  - CM负责对DN进程状态监控，当监控到进程状态是故障或者正常，自动触发升降副本功能。
  - 同步列表：
    - 存在正常备机时：
      - synchronous_standby_names = 'ANY NODE NUM(*同步列表*)'，其中NUM = (同步列表个数 - 级联备个数) / 2。
      - synchronous_commit = 'on'。
    - 不存在正常备机时：
      - synchronous_commit = 'local'。

版权所有 © 华为云计算技术有限公司

        - CM监控synchronous_standby_names，synchronous_commit 参数后，此参数有CM维护：
            - synchronous_standby_names同步列表跟cm保持一致。
            - synchronous_commit只存在local和on两个类型。
    - 级联备DN是默认存放于DN的同步列表中的。
    - RTO时间：30s（不包含选主时间）。

CM 支持 DN 仲裁

CM支持的DN仲裁模式主要分为Quorum模式和DCF模式，默认模式为Quorum模式。

- **Quorum模式**：基于多数派模式仲裁，选出同步备。
    - 简介：CM基于Quorum模式进行仲裁，当DN分片处于无主场景时，CM在多数派DN redo或预解析完成后(因为只有当redo回放或者预解析结束，DN的term 和 lsn 值才是可信任的)，选择 term和lsn最大的节点（同步备）发送failover升主。
    - 选主流程：
        - **lock1**命令：
            - 功能：

              1. 等待备DN与主DN断连。
              2. 等待备DN redo回放 或者 预解析结束，待升主DN redo回放完成。

            - 约束：

              1. 需要主备DN断连。
              2. 需要 DN redo 回放结束。

        - **force lock1** 命令：
            - 功能：

              1. 强制断开备DN与主DN之间的连接，并把原主DN进程停止。
              2. 等待备DN redo回放结束。

            - 约束：

              当下发 lock1 时间达到阈值**cma_fault_kill_instance_timeout**，且存在多数派备机是normal，才会下发force lock1，强制断开DN主备断连，并停止原主DN进程。
    - 约束：最小满足一主两备实例。
- **DCF模式**：
    - **自动选主模式**：基于paxos 协议

      简介：DCF模式自动选主，在此场景下，CM不再进行对DN选主，只负责数据采集，和假死检测等。不支持强启、强切、自动降副本和网络检测等能力。

      CM配置：enable_dcf=ON、dn_arbitrate_mode=paxos。

      DN配置：enable_dcf=ON、dcf_run_mode=0。
    - **手动选主模式**：基于Quorum模式

      简介：DCF模式下的手动选主，在此场景下，恢复CM对DN的选主，不再由DCF自动选主，此时可以执行Quorum模式下的switchover命令。

 版权所有 © 华为云计算技术有限公司

CM配置：enable_dcf=ON、dn_arbitrate_mode=quorum。

DN配置：enable_dcf=ON、dcf_run_mode=1。

- 总体约束：最小满足一主两备实例。
- 默认安装：DCF手动选主模式。
- 模式切换：参考上述配置，支持在线切换。请参见“OM服务化 > OM Agent接口 > 数据库实例管理相关接口 > 数据库实例模式切换”。

  - paxos -> quorum 模式：

    ```
    ssh: gs_om -t change_mode -m quorum
    agent:
    curl  --header "Content-Type: application/json" -X "POST" --data
    '{"consistencyProtocol":"quorum"}' http://ip:port/v1/om/agent/mgr/cluster/change_mode
    ```

  - quorum -> paxos 模式：

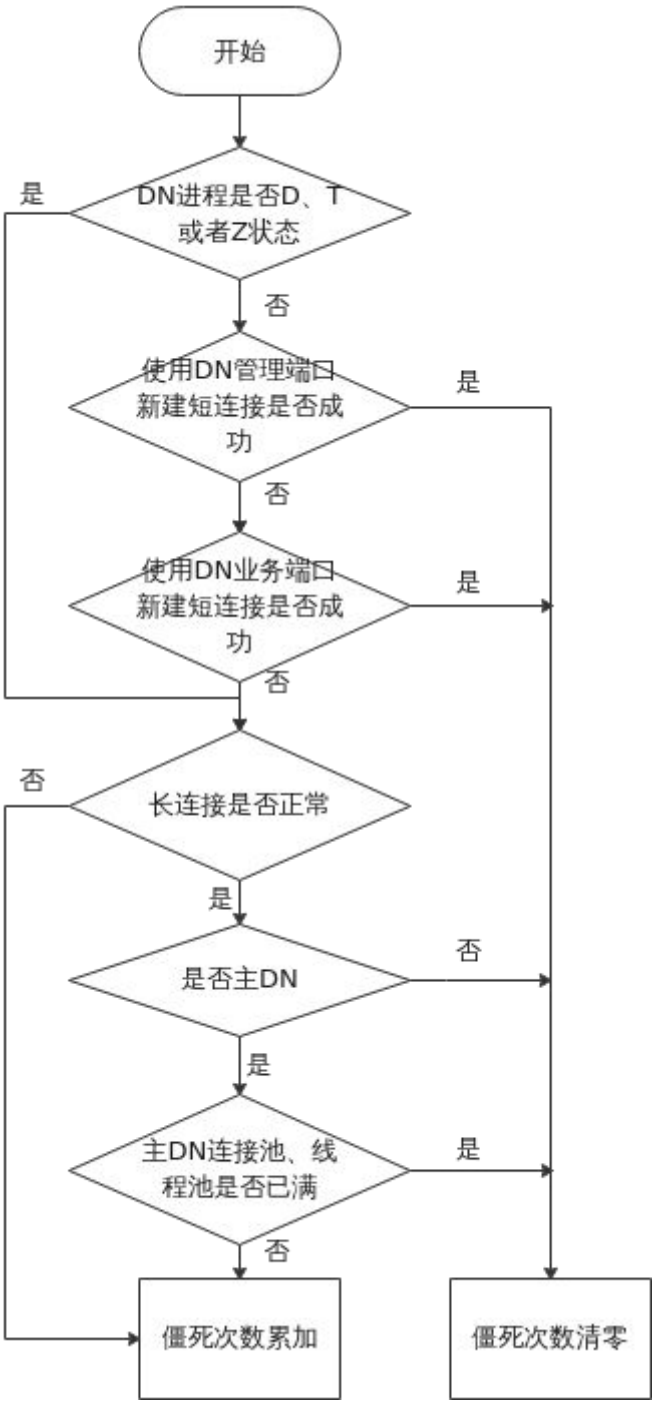
    ```
    ssh: gs_om -t change_mode -m dcf
    agent:
    curl  --header "Content-Type: application/json" -X "POST" --data
    '{"consistencyProtocol":"paxos"}' http://ip:port/v1/om/agent/mgr/cluster/change_mode
    ```

CM 支持 DN 僵死检测

- 简介：当DN进程处于D、T、Z或者连接超时等僵死状态时，DN无法及时处理业务请求，导致业务中断，影响业务正常功能。CM支持根据配置值定时检测DN是否处于僵死状态，对连续处于僵死状态的DN，达到配置次数后做重启恢复操作。
- 僵死判断主要分为D/T/Z状态判断和连接状态判断，检测流程：

 版权所有 © 华为云计算技术有限公司

图 6-1 DN 僵死检测流程



说明

- 长连接：cm_agent为了检测并上报DN状态到cm_server，使用管理端口与DN建立连接。在cm_agent进程中一直保持该连接。
- 短连接：cm_agent为了检测DN是否处于僵死状态，新建与DN的连接，僵死检测流程结束，关闭该连接。

- 僵死倒换规则：

 版权所有 © 华为云计算技术有限公司

a. 通过僵死检测算法，检测到僵死，然后进行僵死次数累加（发现正常则僵死次数清零），检测周期：

- 开启enable_e2e_rto开关：

  检测到1次僵死后会立即将实例状态置为UNKNOWN，进而触发选主仲裁流程。

- 不开启enable_e2e_rto开关：

  控制参数：agent_phony_dead_check_interval（默认值10s）、phony_dead_effective_time（默认值5次）

  > 连接类僵死默认(10 * 5)秒后标记为僵死，触发重启。

  > 由于实例正常运行也会出现D状态，执行gstack等定位动作会出现T状态，为避免这种场景导致的误判和core生成不完整，D/T状态僵死规格为持续3分钟状态未变化，则触发重启。

  > coredump时也会有D/T状态，检测到数据库core标记后，此时不会做重启动作，但是会将实例置为UNKNOWN，进而触发选主仲裁流程。

b. 当一个实例反复僵死时，会触发僵死抑制，抑制规则为：

控制参数：instance_phony_dead_restart_interval（默认值6h）

同一实例，默认6h内只能因僵死被重启1次。第二次僵死时会统计与上一次僵死的时间间隔，超过6h才会执行重启动作。

自定义资源

- 资源配置

  资源配置文件cm_resource.json，文件包含所有自定义资源的相关属性，可以通过cm_ctl res命令对配置文件修改。不支持动态生效，修改配置文件后需要重启CM才能生效。

- 客户端

  CM提供客户端动态库给资源进行集成，提供数据库实例状态查询、状态变更通知、数据库实例锁能力。

- 自动启停资源

  资源需要提供脚本，脚本包含启停、检测等能力，脚本路径需要配置在资源配置文件中。

- 手动启停资源

  可以通过cm_ctl start/stop -n -I命令实现资源实例的启停操作，详细请参见**cm_ctl工具介绍**。

- 自定义资源状态

  自定义资源有四种状态：online，offline，deleted，unknown。可以通过cm_ctl查询。

CM 支持 VIP 仲裁

简介：VIP是虚拟IP，随主DN状态进行动态绑定和切换，即如果DN主发生了切换，原主DN上的虚拟IP可能动态绑定到新主DN上，用户可以只通过该IP与数据库连接，不用感知数据库在哪一个节点上。

仲裁流程：

新主：

 版权所有 © 华为云计算技术有限公司

1. 在新主上用别名（网卡名：DN端口号）方式绑定VIP。
2. 把VIP添加到数据库配置项“listen_addresses”中。

原主：

1. 在原主上解绑VIP。
2. 把VIP从数据库配置项“listen_addresses”中删除。

功能开启：在数据库实例节点上每一个cm_agent数据目录下生成cm_resource.json配置文件，文件中包含VIP的json字段，可通过cm_ctl工具命令生成VIP配置文件，命令请参见**•执行res命令：**，可以通过cm_ctl show方法查看VIP信息。

说明

在DN选主成功后，5s内VIP绑定生效。

约束：

- 不支持容灾数据库实例。
- 不支持两节点三副本以上部署。
- VIP(网卡名：DN端口号)不允许被其他非虚拟IP占用。
- 每个节点上cm_resource.json配置文件必须一致。
- 只支持重启方式加载。
- 允许配置多个VIP，但是每个实例最多支持6个。
- 需要可用VIP，即需要可对外提供服务的VIP（跟现有网卡处于同一网段中）。
- 手动生成VIP配置文件步骤，在每个节点上都要生成该VIP配置文件，且每个节点要求一致，需要重启数据库实例才能生效（不建议手动操作）。

```
cm_ctl res --add --res_name="资源名" --
res_attr="resources_type=VIP,float_ip=IP1,gate_way=GATEWAY"
cm_ctl res --edit --res_name="资源名" --add_inst="node_id=NODE1,res_instance_id=6001" --
inst_attr="base_ip=IP2"
cm_ctl res --edit --res_name="资源名"  --add_inst="node_id=NODE2,res_instance_id=6002" --
inst_attr="base_ip=IP2"
cm_ctl res --edit --res_name="资源名"  --add_inst="node_id=NODE3,res_instance_id=6003" --
inst_attr="base_ip=IP2"
在数据库gs_hba.conf 以sha256 方式添加float_ip
```

警告

由于需要动态绑定和解绑VIP，需要对网卡进行操作，故而需要对ifconfig进行提权，允许数据库实例用户使用sudo ifconfig方式绑定和解绑VIP，此操作可能带来窃听、篡改、仿冒和拒绝服务等安全风险。提权操作，请参见**ifconfig提权步骤**。

IPv6场景下，还需要对IP命令进行提权，允许数据库实例用户使用sudo ip方式绑定和解绑VIP，此操作可能带来窃听、篡改、仿冒和拒绝服务等安全风险。提权操作，请参见**6.4.4-IP命令提权步骤**。

CM 支持急救室能力

急救室，是一个高度抽象的、更加纯粹的框架模型，由于故障存在假阳性和假阴性情况，因此需要框架更加的灵活，以应对不同场景和需求下的高可用。当前具备线程心跳自检、检测磁盘I/O hang（由于磁盘故障，导致无法进行数据写入和读取等）问题并自动隔离主DN的能力、系统资源监控能力、网络探测能力和支持异步日志功能。

版权所有 © 华为云计算技术有限公司

- 线程心跳自检能力
  - 介绍：新增自检模块，对CM实例关键线程自检。。
  - 主要流程：

    CMA、CMS新增线程自检能力。自检线程仅依赖CPU，无任何网络或磁盘操作。如果自检线程运行一周所需时间超过**thread_effective_time** 后，则此进程处于hang状态，进程会自动退出，并触发仲裁（CMS重新自选主，CMA进程重启）。
  - 规格：
    - 在默认参数情况下，RTO时间为3分钟，可通过修改cm_agent参数**thread_effective_time**和cm_server参数**thread_effective_time**、**cma_fault_kill_instance_timeout**优化RTO时间，参数设置太小，容易发生误切换，不建议改小。
  - 约束：
    - ETCD是仲裁的中心组件，本身具备一定的自仲裁能力，若ETCD故障、无法自动恢复时，则CM不再提供服务。
- 磁盘检测能力
  - 规格：
    - 只检测DN、CN、GTM、CM（CMA、CMS）的数据盘、日志盘以及DN的表空间所在盘。
    - 在默认参数下，读写磁盘的hang故障的RTO时间为60秒，其他操作（例如open）的hang故障的RTO时间为3分钟。可通过修改cm_agent参数**disk_check_hung_timeout**、**disk_check_block_timeout**优化RTO时间，参数设置太小，容易发生误切换，不建议改小。
    - 如果业务模型中涉及大量两阶段事务，那么故障后可能会有未执行完的两阶段事务存在集群中，该事务无法执行下去，并且会阻塞其他两阶段事务的执行。此场景下非两阶段事务可以在RTO规格内恢复，但受两阶段事务残留影响的业务完全恢复时间会增加，具体增加的时间取决于gs_clean清理残留事务的时间，关于残留事务清理相关介绍请参见《工具参考》中“服务端工具 > gs_clean”章节。
  - 检测规则
    - 检测通过读写临时文件实现，针对文件会有四个操作：open（打开文件），write（写文件），read（读文件），unlink（删除文件）。其中open、write、read任意一个操作失败或者执行超时，就会判定磁盘发生故障或处于hang状态，并触发仲裁。相反，unlink操作失败或执行超时，则不会作为触发仲裁的条件。
    - 非DIRECT读写：每隔1秒采用非DIRECT的方式向磁盘中写入1B的数据，如果读写失败，通过错误码判断失败的原因，其中如果是EIO（I/O错误）、EROFS（文件系统只读）、ENOSPC（磁盘空间不足）、EACCES（权限不足）这些原因，会判定故障，其他类型错误暂不处理。
    - DIRECT读写：默认每隔5秒（可配）采用DIRECT的方式向磁盘中写入512B（如果被检测磁盘为条带化，则写入256KB）的数据，如果写入操作阻塞超过20s（如果无法判断被检测磁盘是否为条带化，则超时时间为60s，写入数据的大小为256KB）没有返回，则判定磁盘处于I/O hang。

 版权所有 © 华为云计算技术有限公司

  - 说明：
    - 写入的数据大小可配，配置的写入大小需要根据磁盘的对齐规则进行对齐，如果配置时没对齐，cm会自动向上对齐。在磁盘LVM条带化场景，为保障写入检测能覆盖到所有盘，建议写入大小根据磁盘数量和数据分块大小动态调整。
    - I/O hang超时时间可配，需要根据磁盘性能进行配置，性能差的盘，超时时间可以适当配置长一些，进而减少业务I/O对磁盘检测的影响，相应的，hang场景RTO将会变长。配置成0则关闭hang检测。
    - 写入大小过大、检测周期过短时，可能对磁盘I/O有一定的争抢，进而可能影响数据库性能。
- 系统资源监控能力
  - 介绍：新增资源监控模块。监控CPU、内存、磁盘I/O这些指标，并将资源信息输出到日志中，输出频率根据使用率动态变化。
  - 日志输出频率和资源使用率对应关系：

资源使用率范围	输出频率
[95, 100]	1s/次
[90,95)	10s/次
[80,90)	1min/次
[0,80)	10min/次

  - 使用率记录：

    系统资源使用率会输出到cm_agent的日志中，格式为：
    - CPU：
      ```
      [CPU STAT] cpu usage: %.2f%%.
      ```
    - 内存：
      ```
      [MEMORY STAT] memory system usage: %.2f%%, application usage: %.2f%%.
      [MEMORY STAT] details info: memTotal(%lu), memFree(%lu), memAvailable(%lu).
      ```
    - 磁盘I/O：
      ```
      [DISK STAT] device(%s) IO util: %lu%%.
      ```
  - 规格：
    - 默认开启检测，可以通过参数**sys_check_interval**关闭。
      - 磁盘检测，只检测集群内所有实例用到的磁盘。
      - 磁盘使用率超过了高磁盘I/O的阈值**high_disk_threshold**，会连续打印十次（每次间隔1s）iostat信息输出到日志文件system_stat-current.log中。
- 支持异步日志能力
  - 介绍：新增异步日志。减少日志盘故障对CM进程的影响。
  - 异步日志标识：
    ```
    ASYN LOG: asyn log init success
    ```
  - 规格：

 版权所有 © 华为云计算技术有限公司

    - 默认开启异步日志，可以通过参数**asyn_log_enable**关闭。
    - 只有cm_agent和cm_server进程支持异步日志，其他CM进程，例如cm_ctl、om_monitor进程不支持异步日志。
    - 磁盘故障时，会根据参数**asyn_log_fault_solution**对日志进行过滤，默认是将所有日志保存到日志缓冲队列中，当缓冲队列达到设定上限时，后续的日志将无法记录，日志缓冲队列大小可以通过参数**asyn_log_max_buf_size**控制。

- 急救室动脉检测模型
  - 基本原理：通过构建动脉检测模型，突破识别动脉问题，给出对应的措施，提升数据库高可用能力。
    - 急诊室框架是基于规则的自动化处理引擎。
    - 提供可配置的急诊室策略，包括动脉定义、指标定义、诊断策略以及处理动作。
    - 执行引擎通过配置的策略进行指标的采集，仲裁诊断，以及处理动作的下发。
  - 功能使用：
    - 使用cm_ctl命令开启和修改检测规则。
      - cm_ctl pull --cm_meta_config, 从cms主获取最新的元数据文件。
      - cm_ctl push --cm_meta_config，上传本地的元数据文件给cms主。
    - 关闭急诊室检测：

      cm_ctl push --cm_meta_config，只包含版本号（"version"）和 元数据版本号（"meta_version"）的元数据文件。

      说明

      - 此功能依赖磁盘检测能力，需要同步打开才可使用。
      - 关闭急诊室功能前，请先调用cm_ctl pull命令，获取最新的急诊室元数据文件，备份后再关闭此功能。
      - 升级和扩容阶段不做处理。强起状态不支持急诊室功能。
      - 急诊室功能当前只支持磁盘IO检测，且仅支持数据目录和日志目录所在磁盘。
      - 仲裁组件（包括CMA、CMS、OM_MONITOR和ETCD等）急诊室不作处理。
      - 降副本至一主零备后，不支持急诊室功能。单副本不支持急诊室功能。
      - 急诊室功能暂不支持自定义脚本。
      - 急诊室功能仲裁需要一定周期，如果采集周期内出现无效数据，则此阶段不再仲裁。
      - 由于磁盘IO检测是一种复杂多变的场景，此急诊室场景下的磁盘检测仅是一种辅助手段。精确判断则需要从硬件角度，由磁盘厂商提供专业手段进行检测。
  - 元数据文件介绍：
    - 存放位置：cm_agent数据目录下cm_meta_cma_config.json，cm_meta_cm_ctl_config.json

版权所有 © 华为云计算技术有限公司

- 文件格式：

```
{
"version": "1",
   "metrics": [
      {
         "name": "disk_check",
         "check_type": "INNER",
         "inner_check_type": "DISK_CHECK",
         "gather_scope": "INST",
         "check_interval": 30,
         "reserve_count": 10,
         "check_timeout": 60,
         "calu_count": 10,
         "calu_method": "AVG"
      },
      {
         "name": "io_check",
         "check_type": "INNER",
          "inner_check_type": "IO_CHECK",
         "gather_scope": "INST",
         "check_interval": 30,
         "reserve_count": 10,
         "check_timeout": 60,
         "calu_count": 10,
         "calu_method": "AVG"
      }
   ],
   "artery": [
      {
         "src": "DN",
         "dest": "DISK",
         "decision": {
            "metric_name": "io_check",
            "metric_place": "DEST",
            "conditions": [
               {
                  "compare": "GREAT",
                  "value": 1000,
                  "action_type": "decision",
                  "action_place": "SRC",
                  "decision": {
                     "metric_name": "disk_check",
                     "metric_place": "DEST",
                     "conditions": [
                        {
                           "compare": "GREAT",
                           "value": 15,
                           "action_type": "RESTART",
                           "action_place": "SRC"
                        }
                     ]
                  }
               },
               {
                  "compare": "GREAT",
                  "value": 600,
                  "action_type": "ALARM",
                  "action_place": "SRC"
               },
               {
                  "compare": "LESS",
                  "value": 100,
                  "action_type": "DISALARM",
                  "action_place": "SRC"
               }
            ]
         }
      }
```


版权所有 © 华为云计算技术有限公司

```
  ]
}
```

- 文件参数介绍：

表 6-5 元数据文件参数介绍

参数名	参数含义	参数类型	取值范围
version	版本号	数值类型以字符串形式存放	[1, 2^64-1]
metrics	指标	数组	NA
name	指标名字	字符串，不可重复	1~32个字节
check_type	检查类型	字符串	"INNER"
inner_check_type	内部检查类型，跟 check_type "INNER" 对应	字符串	"DISK_CHECK"、"IO_CHECK"
gather_scope	数据采集范围	字符串	"INST"
check_interval	检查周期	数值类型	[1, 86400]，单位为秒
reserve_count	采集数据缓存个数	数值类型	[1, 30]
check_timeout	检查超时时间（保留字段）	数值类型	[1, 300]，单位为秒
calu_count	计算数值缓存个数	数值类型	[1, 30]
calu_method	计算方法	字符串	"AVG"、"ORG"、"MEDIAN"
artery	动脉	数组	NA
src	动脉源端	字符串	"DN"
dest	动脉目的端	字符串	"DISK"
decision	决策	数组	NA
metric_name	指标名字，用于关联指标	字符串	只能是指标数组内存在的
metric_place	指标采集点	字符串	"SRC"、"DEST"
conditions	判断条件	数组	NA

版权所有 © 华为云计算技术有限公司

参数名	参数含义	参数类型	取值范围
compare	判断比较方法	字符串	"LESS"、"LESS_EQ"、"EQUAL"、"GREAT"，"GREAT_EQ"
value	比较阈值	数值类型	NA
action_place	需要被下发操作命令的实例点	字符串	"SRC"、"DEST"
action_type	执行命令方式	字符串	"DISALARM"、"ALARM"、"SWITCHOVER"、"RESTART"、"DECISION"

- 磁盘IO检测：
  - 通过检测磁盘IO状态，判断数据库业务是否受影响。
  - 仲裁原理：
    - 连续检测20次，都检测到await数值超过800，则上报告警“ALM_ATI_EmergencyThresholdAlarmNotice”。
    - 连续检测36次，都检测到await数值超过1000，且读写时间超过15s，触发主备切换仲裁。
    - 以上数值规则可通过cm_ctl命令动态修改。
  - 检测频率：默认30s检测一次。
  - 底噪情况：CPU无明显增加。内存增加说明：一主两备集中式场景，CMA和CMS内存共增加100KB。每增加一个实例，内存增加10KB。
- AZ网络拓扑检测
  - 说明：当发生AZ级数据面网络平面断网，**expect_stop_az_name**指定的AZ上的实例会被停止，默认是与CMS主不同AZ上的实例会被停止。
  - 相关参数：
    - **enable_check_az_network**：是否开启AZ级数据面网络检测。
    - **expect_stop_az_name**：发生AZ级数据面网络断网后，期望停止所在AZ上的实例进程。默认为空时，停止与CMS主不同AZ上的实例。
    - **check_az_network_concurrency**：AZ级数据面网络检测最大并发数。
  - 约束：
    - 只支持az级数据面网络平面断网，即需要CMS主AZ上的所有实例都与另一AZ上的所有实例发生断网。
    - 数据面网络跟管理面网络不在同一个平面上。

版权所有 © 华为云计算技术有限公司

- 网络探测 数据面网络拓扑检测
  - 检测方式：本功能是通过TCP建连方式监控DN之间的数据面连通性（每2秒进行一次TCP建连探测，若TCP建连超过5秒，则认为连接失败）。本特性只对DN->DN之间的端口检测，DN->DN 之间的监听端口为DN的（base+1）端口。除以上场景外，不支持其他端口检测。
  - 仲裁方式：
    - CMA根据网络连接情况，当DN之间存在连接失败超过topology_network_collecting_failed_times次数时，则上报网络拓扑检测异常（ALM_AI_AbnormalTopologyConnect）告警。
    - 对分片内不满足连接多数派的DN主进行切主或降副本，保障集群可用。
    - 同分片DN之间如果需要切主或者降副本，以倒换优先，如无更优的DN主选择，则尝试降副本。
  - 约束条件：
    - 暂不支持容灾灾备集群数据面网络仲裁，可以检测网络异常，并上报网络拓扑检测异常告警ALM_AI_AbnormalTopologyConnect。
    - 单节点集群不做检测。
    - 强启后不支持网络仲裁。
    - 升级扩容阶段不做仲裁。
    - 启动过程中不做仲裁。
    - 默认配置下，可在1分钟内告警并触发DN切换。

CM 支持强切功能

强切：顾名思义为强制主备切换。在DN无主仲裁过程中，CM通过下发finishredo命令，强制DN停止回放，可实现快速主备切换的能力。

- 触发方式
  - 手动触发

    命令：cm_ctl finishredo
  - 自动触发

    通过设置cm_server参数**force_promote**和**switch_rto**开启功能
- 功能开启
  - 设置DN参数：force_promote，wal_keep_segments，recovery_time_target
  - 设置CN参数：audit_enabled，audit_system_object
- 约束：

  带logger组网不支持强切功能。
- DCF模式操作步骤：

步骤1 确保集群状态是Normal。

步骤2 选举模式设置为手动模式。

```
ssh 线下模式：
cm_ctl set --param --server -k dn_arbitrate_mode=quorum
cm_ctl reload --param --server
gs_guc reload -Z datanode -N all -I all -c "dcf_run_mode=1"
```

 版权所有 © 华为云计算技术有限公司

步骤3 开启强切功能，设置CM、 DN、CN相关参数。

----结束

⚠ 警告

该功能属于高风险操作，请谨慎执行。

CM&DCF 联合仲裁

DCF仲裁选主是根据Paxos协议各个节点分别发起选主，节点间相互独立，没有全局视角，在一些复杂的场景（例如半数节点故障自动降副本）DCF无法实现仲裁处理。联合仲裁的目的是借助CM的全局视角实现DCF仲裁能力增强，解决复杂场景的问题。主要优势包括：

- 一致性选主：任何故障不会发生“脑裂”双主现象

  在一个高可用系统中，当联系着的节点断开联系时，本来为一个整体的系统，分裂成两个独立节点，两个节点开始争抢共享资源造成系统混乱、数据损坏的现象，称为“脑裂”。传统模式下存在“脑裂”双主问题或新主升主后因无备机而不可用等问题。CM&DCF联合仲裁模式下，DCF在选主时已经与多数派节点达成一致，选主成功即意味着多数备机已认可新主，因此不会存在“脑裂”双主现象。

- 一致性日志复制：任意节点故障、主备切换不触发数据重建

  传统主备切换（包括同步备节点）很容易触发日志分叉而重建数据，恢复耗时久，甚至会出现选主错误，丢失数据。CM&DCF联合仲裁模式下主备切换不触发数据重建。

- 一致性增删节点：支持升级、扩容、临时运维等场景

  为保证正确性，DN的增删对集群限制较多，无法达到自动化DN移动的基础要求；若对集群不做限制则会产生复杂的异常场景，很难保证正确性。CM&DCF联合仲裁模式支持动态增删DN节点，协议保证过程的一致性和正确性，支持升级、扩容和临时运维等场景。

CM 支持硬件故障处理

硬件故障处理属于端到端全栈高可用能力。数据库运维平台负责检测硬件故障，如果发生故障，数据库运维平台会通知CM，CM收到消息后会根据故障的级别，进行相应的故障处理。CM支持硬件故障处理RTO<=65s，时间从数据库运维平台下发故障消息开始算。RTO（Recovery Time Objective）指故障发生到业务恢复的时间。

说明

故障json串大小限制为1KB。

数据库运维平台管控下发频率不高于1次/s，频率太高会加大CMS的消息处理负载，影响CMS处理消息的效率。

DDB不可用时，不做仲裁。

版权所有 © 华为云计算技术有限公司

表 6-6 硬件故障级别和处理策略

故障级别	故障场景	处理策略
100	• 系统掉电 • 未知原因导致系统重启 • 系统控制命令导致系统重启 • DIMM MCE错误 • CPU MCE错误 • CPU AER错误	主动停止节点上的所有进程。

- 当CMS没有主的场景下，硬件故障级别和处理策略：

  故障级别：100，处理策略：尝试调用stop命令停止故障节点，如果stop命令执行不成功，则故障无法处理。
- 当数据库已经存在故障，且降副本成了1主1备，硬件故障级别和处理策略：

  故障级别：100，处理策略：cm依旧会停止故障节点，这种场景可能会导致集群不可用。

CM 适配小规格典型故障场景

- CM在4U16G环境可正常运行。
- 约束：仅支持在一主两备集中式带etcd组网下底噪优化：
  - 单核CPU总使用率由7.25%（ETCD+CMS+CMA）降低至4.5%。
  - 内存总使用率由440M（ETCD+CMS+CMA）降低至240M以下。

6.2 cm_ctl 工具介绍

功能介绍

cm_ctl是GaussDB提供的用来控制数据库实例服务的工具。该工具主要供OM调用，及数据库实例服务自恢复时使用。

cm_ctl的主要功能有：

- 启动数据库实例服务、AZ的所有实例、单个主机上的所有实例或单独启动某个实例进程。
- 停止数据库实例服务、AZ的所有实例、单个主机上的所有实例或单独停止某个DN实例进程。
- 重启逻辑数据库实例服务。
- 查询数据库实例状态或者单个主机的状态。
- 切换主备实例或重置实例状态。
- 重建备机。
- 查看数据库实例配置文件。

 版权所有 © 华为云计算技术有限公司

- 设置日志级别、一主多备数据库实例部署下cm_server的仲裁模式、AZ之间的切换模式。
- 获取日志级别，一主多备数据库实例部署下cm_server的仲裁模式、AZ之间的切换模式。
- 检测实例进程状态。

与cm_ctl工具相关的文件：

- cluster_manual_start

  该文件是数据库实例启停标志文件。文件位于*$GAUSSHOME/bin*下。其中，GAUSSHOME为环境变量。启动数据库实例时，cm_ctl会删除该文件；停止数据库实例时，cm_ctl会生成该文件，并向文件写入停止模式。
- instance_manual_start_X（X是实例编号）

  该文件是单个实例启停标志文件。文件位于*$GAUSSHOME/bin*下。其中，GAUSSHOME为环境变量。启动实例时，cm_ctl会删除该文件；停止实例时，cm_ctl会生成该文件，并向文件写入停止模式。

cm_ctl的相关约束：

- 灾备数据库实例下使用build及switchover命令无法生效。
- 一主一备一logger实例finishredo命令无法生效。
- cm_ctl日志文件无法写入时，会将日志信息在终端屏幕输出。

命令说明

命令格式中常用的几个符号含义如下：

- 尖括号 < >：必选参数，实际使用时应将其替换为所需要的参数。
- 大括号 { }：必选参数，使用时只能选择括号内的参数。
- 方括号 [ ]：可选参数，在命令中根据需要加以取舍。
- 竖线 |：用于分隔多个互斥参数，含义为“或”，使用时只能选择一个。

cm_ctl命令的格式如下：

- cm_ctl <模式参数> [option参数] [公共参数]

cm_ctl参数可分为如下几类：

- 模式参数，详细请参见 **表 模式参数**。
- 公共参数，详细请参见 **表 公共参数**。
- 各模式的option参数：
  - switchover模式的参数，详细请参见 **表 switchover参数**。
  - build模式的参数，详细请参见 **表 build参数**。
  - check模式的参数，详细请参见 **表 check参数**。
  - start模式的参数，详细请参见 **表6-27**。
  - stop模式的参数，详细请参见 **表 stop参数**。
  - query模式的参数，详细请参见 **表 query参数**。
  - view模式的参数，详细请参见 **表 view参数**。
  - set模式的参数，详细请参见 **表 set参数** **表 set cm参数**。

 版权所有 © 华为云计算技术有限公司

- get模式的参数，详情请参见 **表 get参数**。
- hotpatch模式的参数，详情请参见 **表 hotpatch参数**
- setrunmode模式的参数，详细请参见 **表setrunmode参数**。
- changerole模式的参数，详细请参见 **表changerole参数**。
- changemember功能的参数，详细请参见 **表 changemember参数**。
- reload模式的参数，详细请参见 **表 reload 参数**。
- list模式的参数，详细请参见 **表 list参数**。
- encrypt模式的参数，详细请参见 **表 encrypt参数**。
- ddb模式的参数，详细请参见 **表6-23**。
- switch模式的参数，详细请参见 **表 switch参数**。
- res功能的参数，详细请参见 **表6-28**。
- notify功能的参数，详细请参见 **表 notify参数**。
- disablecm功能的参数，详细请参见 **表24 disablecm\enablecm参数**。
- pull功能的参数，详细请参见 **表25 pull参数**。
- push功能的参数，详细请参见 **表26 push参数**。

表 6-7 模式参数

参数	参数说明
start	一主多备、小型化数据库部署模式下启动数据库实例服务、单个主机上的所有实例或单独启动某个DN实例进程，或者直接启动整个AZ。
switchover	一主多备数据库部署模式下切换数据库主备实例、实现AZ之间的相互切换。
finishredo	所有备机停止回放，每个分片中选择一个强制升主。 **注意** 该参数属于高风险操作，请谨慎执行。
build	重建备实例。
check	检测实例进程运行状态，用户无需关注，不建议使用。
stop	一主多备、小型化数据库部署模式下停止数据库实例服务、单个主机上的所有实例或单独停止某个DN实例进程。或者直接停止整个AZ。
query	一主多备、小型化数据库部署模式下查询数据库实例状态或者单个主机的状态。
view	查看数据库实例配置文件。
set	设置日志级别，一主多备数据库部署模式下cm_server的仲裁模式、AZ之间的切换模式。
set --param	设置cm参数，具体参数可以参见 **CM配置参数介绍**。
get	获取日志级别，一主多备数据库部署模式下cm_server的仲裁模式、AZ之间的切换模式。

版权所有 © 华为云计算技术有限公司

参数	参数说明
hotpatch	对cm_server执行热补丁操作。
setrunmode	DCF部署方式下，设置DCF投票数，主要用于DCF强启。
changerole	DCF模式下，将角色为primary的修改为passive或者follower。
changemember	DCF模式下，改变指定DCF节点属性，包括节点角色、节点所在的逻辑组、节点的选举优先级等。
reload	在线加载数据库实例静态配置文件，用户无需关注，不建议使用。
reload --param	加载可以动态生效的cm参数，部分参数不支持reload，只能重启cm才能生效。
list --param	列出cm_agent或cm_server所有的参数。
encrypt	对输入的密码进行加密操作，密码支持8~15位，且必须包含三种字符（数字，字母，符号）。
ddb	执行对DCC组件的配置数据存取命令。CM依赖DCC组件对配置数据分布式存取，实现数据库实例配置管理高可用能力。
switch	执行ddb模式的切换。 **注意** • ETCD模式切换到DCC模式之前需要参见 **安全设计** 手动生成证书，切换完成后，需要通过cm_ctl打开enable_ssl开关（修改enable_ssl参数）。 • 如果有会涉及到ETCD的操作，例如节点替换，扩容，备份恢复等。在操作之前， 必须先确保切回到ETCD模式再执行相关操作。
res	在节点上进行资源配置的新增、修改、查询、删除。
show	展示数据库实例浮动ip。
res	执行自定义资源相关命令。
notify	执行通知cm消息的命令。
pull	从cms主获取最新的急诊室元数据文件
push	把本地急诊室元数据文件上传给cms主

表 6-8 公共参数

参数	参数说明
-D <DATA_DIR>	指定实例数据目录。
-n <NODE_ID>	指定节点。
-t <SECS>	指定超时时间。超时后，会退出并报错。
-l <FILE_NAME>	结果输出到指定文件。
-z <AZ_NAME>	指定AZ名称。

版权所有 © 华为云计算技术有限公司

参数	参数说明
-V, --version	打印cm_ctl版本信息，然后退出。
-?, -h,--help	显示关于cm_ctl命令行参数的帮助信息，然后退出。

表 6-9 query 参数

参数	参数说明
-v	显示详细数据库实例状态。 **说明** 数据库实例有如下几种状态： ● Normal：表示数据库实例可用，且数据有冗余备份。所有进程都在运行，主备关系正常。 ● Degraded：表示数据库实例可用，但数据没有冗余备份。 ● Unavailable：表示数据库实例不可用。
-C	按主备关系成对显示数据库实例状态。 **说明** -C参数需要和-v参数一起使用才能按主备关系成对显示数据库实例详细状态信息，使用-C时，必须指定-v参数。
-r	显示备机redo状态。 **说明** 使用-r时，必须指定-v参数。
-s	显示导致各个节点主备实例数量不均衡的实例。 **说明** -s参数需要和-v以及-C参数一起使用才能按主备关系成对显示导致各个节点主备实例数量不均衡的实例，使用-s参数时，必须指定-C、-v参数。
-S	显示数据库实例启动时的状态检查结果。 **说明** -S参数需要和-v以及-C参数一起使用才能按显示数据库实例的状态检查结果。有以下三种状态： ● Normal：表示数据库实例可用，且数据有冗余备份。所有进程都在运行，主备关系正常。 ● Degraded：表示数据库实例可用，但数据没有冗余备份。 ● Unavailable：表示数据库实例不可用。
-F	显示各个节点Fenced UDF状态。 **说明** -F参数需要和-v以及-C参数一起使用才能显示各个节点Fenced UDF状态，使用-F参数时，必须指定-C、-v参数。
-i	显示物理节点ip。 **说明** -i参数需要和-v以及-C参数一起使用。
-d	显示实例数据目录。 **说明** -d参数需要和-v以及-C参数一起使用。

版权所有 © 华为云计算技术有限公司

参数	参数说明
-z ALL	显示数据库实例所有实例AZ名称。 **说明** -z参数需要和-v以及-C参数一起使用，并且-z后面需要跟参数ALL。
-x	显示数据库实例所有异常实例。 **说明** -x参数需要和-v以及-C参数一起使用。
-p	显示数据库实例所有DN端口。 **说明** -p参数需要和-v以及-C参数一起使用。
--minorityAz	只查询指定AZ的cms和etcd。 **说明** 此参数会忽略非指定AZ的cms、etcd节点，可以在少数派场景提高查询速度。
-g	显示备份和恢复群集信息。

表 6-10 stop 参数

参数	参数说明
-m <SHUTDOWN_MODE>	指定停止模式，停止模式有以下几种： ● smart（s）：等待用户业务结束后，停止所有数据库实例。 ● fast（f）：不等待用户业务结束，指定数据库实例退出。 ● immediate（i）：不等待用户业务结束，指定数据库实例强制退出。
-I	指定需要停止的资源实例号，可以通过cm_ctl query -Cv查看资源实例和实例号的映射关系。
--server	DCC模式下，停止独立启动的cm_server。

表 6-11 build 参数

参数	参数说明
-f	强制重建备机。
-b full	指定进行全量build。不指定情况下，对于一主多备数据库实例部署模式进行auto build。auto build指：先调用增量build，失败之后调用全量build。

版权所有 © 华为云计算技术有限公司

表 6-12 check 参数

参数	参数说明
-B <BIN_NAME>	指定进程名，其进程名包括“cm_agent”、“gaussdb”和“cm_server”。
-T <DATA_PATH>	指定实例数据目录。

表 6-13 set 参数

参数	参数说明
--log_level=<LOG_LEVEL>	设置主cm_server日志级别。共分为DEBUG5、DEBUG1、WARNING、LOG、ERROR和FATAL六个级别，日志打印信息级别越来越高。设置日志级别越高，输出日志信息就越少。
--cm_arbitration_mode=<ARBITRATION_MODE>	一主多备功能，设置cm_server的仲裁模式。共有MAJORITY、MINORITY两种模式，MAJORITY为多数派模式，MINORITY为少数派模式。少数派模式适用于一主多备数据库部署并且只有AZ3存活时，此时cm_server可以进行正常的仲裁业务；多数派模式适用于一主多备数据库部署并且各个组件（cm_server、etcd和dn）存活数量大于一半的场景。数据库实例正常情况下默认为多数派模式。
--cm_switchover_az_mode=<SWITCHOVER_AZ_MODE>	一主多备功能，设置AZ之间的自动切换开关。共有NON_AUTO、AUTO两种模式，NON_AUTO为非自动切换模式，AUTO为自动切换模式。AUTO模式由主cm_server自动控制AZ1和AZ2之间的DN实例切换。
--cmsPromoteMode=<CMS_PROMOTE_MODE>	设置cms的升主方式，共有AUTO、PRIMARY_F两种模式，AUTO为默认自选主方式，PRIMARY_F会强制使-I指定的节点升主，无论当前是否有主，因此，有可能会导致cms存在多主情况。

表 6-14 set cm 参数

参数	参数说明
--param	表明需要设置cm参数，不带此参数则不能执行设置cm参数。
--agent | --server	此参数为必带参数，表明当前需要设置cm_server还是cm_agent的参数。
-n	指定需要修改哪一个节点的参数，如果不指定，即不加-n参数，则默认会修改所有节点的参数。
-k "<parameter>=<value>"	指定设置的参数和参数的值，只能设置已经存在的参数，不支持增加或删减参数。

版权所有 © 华为云计算技术有限公司

表 6-15 get 参数

参数	参数说明
--log_level=<LOG_LEVEL>	获取主cm_server日志级别。共分为DEBUG5、DEBUG1、WARNING、LOG、ERROR和FATAL六个级别，日志打印信息级别越来越高。设置日志级别越高，输出日志信息就越少。
--cm_arbitration_mode=<ARBITRATION_MODE>	一主多备功能，获取cm_server的仲裁模式。共有MAJORITY、MINORITY两种模式，MAJORITY为多数派模式，MINORITY为少数派模式。少数派模式适用于一主多备数据库部署并且只有AZ3存活时，此时cm_server可以进行正常的仲裁业务；多数派模式适用于一主多备数据库部署并且各个组件（cm_server、etcd和dn）存活数量大于一半的场景。数据库实例正常情况下默认为多数派模式。
--cm_switchover_az_mode=<SWITCHOVER_AZ_MODE>	一主多备功能，获取AZ之间的自动切换开关。共有NON_AUTO、AUTO两种模式，NON_AUTO为非自动切换模式，AUTO为自动切换模式。AUTO模式由主cm_server自动控制AZ1和AZ2之间的DN实例切换。

表 6-16 switchover 参数

参数	参数说明
-A	将所有DN实例统一从主切换到备。
-a	重置节点状态为初始配置状态。 **说明** switchover为维护操作：确保数据库实例状态正常，所有业务结束，并使用pg_get_senders_catchup_time视图查询无主备追赶后，再进行switchover操作。
-f	指定进行-f类型switchover。 **说明** ● switchover为维护操作：确保数据库实例状态正常，所有业务结束，并使用pg_get_senders_catchup_time视图查询无主备追赶后，再进行switchover操作。 ● 使用方式：cm_ctl switchover -n <NODE_ID> -D <DATA_DIR> -f。

版权所有 © 华为云计算技术有限公司

表 6-17 view 参数

参数	参数说明
-v	显示数据库实例所有节点的静态配置详细信息。 **说明** 相对于cm_ctl view命令，-v 显示内容增加了cmserver、dn组件的编号显示，如下：cmseverInstanceID，datanodeInstanceID。
-N	只显示本节点的静态配置信息，即执行cm_ctl view命令的节点的信息，N表示Native。

表 6-18 hotpatch 参数

参数	参数说明	取值范围
-E <HOTPATCH_COMMAND>	指定需要执行的热补丁操作。	• load： 加载热补丁 • unload：卸载热补丁 • active：激活热补丁 • deactive：去激活热补丁 • info：显示热补丁信息 • list：显示热补丁列表
-P <HOTPATCH_NAME>	指定需要执行操作的热补丁名称，需要完整的路径名。	字符串

表 6-19 reload 参数

参数	参数说明
--param	表明需要加载cm参数，不带此参数则不能执行加载cm参数。
--agent | --server	表明当前需要动态加载cm_server还是cm_agent的参数。

表 6-20 list 参数

参数	参数说明
--param	此参数为必带参数，表明需要列出cm参数信息
--agent | --server	此参数为必带参数，表明当前需要查看cm_server还是cm_agent的参数。

版权所有 © 华为云计算技术有限公司

表 6-21 switch 参数

参数	参数说明
--ddb_type=<DDB_TYPE>	选择需要切换到哪一个模式（etcd或dcc）。
--commit	执行切换会使数据库实例无法选主，需要执行commit恢复数据库实例。
--rollback	回滚操作，执行切换失败需要执行回滚操作。

表 6-22 encrypt 参数

参数	参数说明
-M	指定加密类型，支持server、client。默认类型为server。
-D	指定生成的加密密码文件路径。

表 6-23 ddb 参数

参数	参数说明
--put <key> <value>	往dcc中插入键值对，如果键值对已存在则会修改键key所对应的值value。
--get <key>	查询dcc中key对应的value。
--delete <key>	删除dcc中指定的键值对。
--prefix	get或者delete后添加prefix参数，可以实现模糊匹配查询和删除。
--cluster_info	获取数据库实例信息。
--leader_info	获取主节点信息。
--help，-h	显示dcc命令帮助信息。
--version，-v	显示dcc版本信息。
--member_list	查询dcc节点信息。
--member_add -n <NODE_ID> -i IP -p <PORT> -r <ROLE>	向dcc集群中增加节点。
--member_remove -n <NODE_ID>	从dcc集群中删除节点。

版权所有 © 华为云计算技术有限公司

参数	参数说明
--change_role -n <NODE_ID> -r <ROLE>	改变节点的角色。
--build_info	查询dcc构建信息。
--fixleader -fixval {0 | 1}	修改dcc定主状态。

表 6-24 setrunmode 参数

参数	参数说明	取值范围
--xmode	指定DCF的运行模式。	● normal：正常模式。 ● minority：少数派模式，需要--votenum指定投票数。
--votenum	指定DCF少数派运行方式的投票数。	正整数，不高于DCF总副本数。

表 6-25 changerole 参数

参数	参数说明	取值范围
--role	集中式DN数据库实例，DCF模式下节点角色。	● passive：passive角色。 ● follower：follow角色。

表 6-26 changemember 参数

参数	参数说明	取值范围
--role	集中式DN数据库实例，DCF模式下节点角色。	● passive：passive角色。 ● follower：follow角色。
--group	DCF模式下，修改group的值。	0~2147483647
--priority	DCF模式下，修改priority的值。	0~2147483647

版权所有 © 华为云计算技术有限公司

表 6-27 start 参数

参数	参数说明
-I	指定需要启动的资源实例号，可以通过cm_ctl query -Cv查看资源实例和实例号的映射关系。
--server	DCC模式下，单独启动cm_server。

表 6-28 res 参数

参数	参数说明
--add --res_name=<name> --res_attr=<res_info>	新增资源配置，name为新增资源的名称，res_info为新增资源的信息。
--del --res_name=<name>	删除资源配置，name为删除资源的名称。
--edit --res_name=<name> --res_attr=<res_info>	修改资源配置，name为被修改资源的名称，res_info为被修改的资源信息。
--edit --res_name=<name> --add_inst=<inst_info> [--inst_attr=<inst_addr>]	新增资源实例，name为新增实例的资源名称，inst_info为新增实例的信息，inst_addr为新增实例的属性。
--edit --res_name=<name> --edit_inst=<inst_info> [--inst_attr=<inst_attr>]	修改资源实例，name为新增实例的资源名称，inst_info为新增实例的信息，inst_attr为新增实例属性。
--edit --res_name=<name> --del_inst=<inst_info>	删除资源实例，name为删除实例的资源名称，inst_info为删除实例的信息。
--check	检查资源配置。不合理处会显示Warning，配置不可用才会报错。
--list	展示资源配置。
--list --res_name=<name>	展示资源配置，name为该资源名称。
--list --res_name=<name> --list_inst	展示资源实例配置，name为该资源名称。

表 6-29 notify 参数

参数	参数说明
--fault_occur	通知cm，发生了一个或多个故障。
--fault_recover	通知cm，一个节点的故障已经恢复。

版权所有 © 华为云计算技术有限公司

参数	参数说明

--detail	发生故障的详细信息。故障以json串的形式表示，json串格式如下：
	{
	"faultList": [{
	// 必要参数
	"alarmId": "", // 故障ID
	"faultNodeId": "", // 故障节点号
	"faultGrade": "" // 故障级别
	}, { ...
	}]
	}

表 6-30 disablecm\enablecm 参数

参数	参数说明
--all	所有实例（DN、CMS）的自动启停和集群仲裁。
--start	所有实例（DN、CMS）的自动启动。
--stop	所有实例（DN、CMS）的自动停止。
--arbitrate	集群仲裁能力。参数可选“DN_ARBITRATE”。不附加参数默认选择可选参数内的全部仲裁能力。

表 6-31 pull 参数

参数	参数说明
--cm_meta_config	从cms主，获取最新的急诊室元数据文件。

表 6-32 push 参数

参数	参数说明
--cm_meta_config	把本地急诊室元数据文件上传给cms主。

命令参考

- 启动实例。
  cm_ctl start [-z <AZ_NAME> [--cm_arbitration_mode=<ARBITRATION_MODE>]] | [-n <NODE_ID> [-D <DATA_DIR> | -I <RESOURCE_INSTANCE_ID>]] [-t <SECS>]
- 数据库主备倒换。
  cm_ctl switchover [-z <AZ_NAME>] | [-n <NODE_ID> -D <DATA_DIR> [-f]] | [-a] | [-A] [-t <SECS>]

 版权所有 © 华为云计算技术有限公司

- 所有备机停止回放，每个分片中选择一个强制升主。
  cm_ctl finishredo
- 重建备DN。
  cm_ctl build -n <NODE_ID> -D <DATA_DIR> [-t <SECS>] [-f] [-b full]
- 检测实例进程运行状态。
  cm_ctl check -B <BIN_NAME> -T <DATA_PATH>
- 停止实例。
  cm_ctl stop [[-z <AZ_NAME>] | [-n <NODE_ID> [-D <DATA_DIR> | -I <RESOURCE_INSTANCE_ID>]]] [-t <SECS>] [-m <SHUTDOWN_MODE>]
- 查询数据库实例状态。
  cm_ctl query [-z ALL] [-l <FILE_NAME>] [-v [-C [-s] [-S] [-d] [-i] [-F] [-x] [-p]] | [-r]] [-t <SECS>] [--minorityAz=<AZ_NAME>]
- 查看数据库实例配置文件。
  cm_ctl view [-v | -N | -n <NODE_ID> | -c] [-l <FILE_NAME>]
- 设置参数。
  cm_ctl set [--log_level=<LOG_LEVEL>] [--cm_arbitration_mode=<ARBITRATION_MODE>] [--cm_switchover_az_mode=<SWITCHOVER_AZ_MODE>]
- 设置CM参数：
  cm_ctl set --param {--agent | --server} [-n <NODE_ID>] -k "<PARAMETER>=<VALUE>"
- 获取参数。
  cm_ctl get [--log_level] [--cm_arbitration_mode] [--cm_switchover_az_mode]
- 对cm_server执行热补丁操作。
  cm_ctl hotpatch -E <PATCH_COMMAND> -P <PATCH_NAME>
- 设置DCF投票数

  cm_ctl setrunmode -n <NODE_ID> -D <DATA_DIR> {--xmode=normal | --xmode=minority --votenum=<NUM>}
- 改变dcf角色信息
  cm_ctl changerole --role={PASSIVE | FOLLOWER} -n <NODE_ID> -D <DATA_DIR> [-t <SECS>]
- 改变dcf节点属性
  cm_ctl changemember --role={PASSIVE | FOLLOWER} [--group=xx] [--priority=xx] -n <NODE_ID> -D <DATA_DIR> [-t <SECS>]
- 动态加载静态配置文件
  cm_ctl reload
- 动态加载CM参数：
  cm_ctl reload --param {--agent | --server}
- 列出所有CM参数：
  cm_ctl list --param {--agent | --server}
- 加密：
  cm_ctl encrypt [-M <MODE>] -D <DATA_DIR>
- 执行ddb命令行
  cm_ctl ddb "<DCC_CMD>"
  设置：cm_ctl ddb "--put <key> <value>"
  删除：cm_ctl ddb "--delete <key>"
  查看dcc命令帮助信息：cm_ctl ddb "--help"
- 切换ddb模式：
  cm_ctl switch {--ddb_type=<DDB_TYPE> | --commit | --rollback}
- 执行res命令：
  - 新增资源：
    cm_ctl res --add --res_name=<name> --res_attr=<res_info>
  - 删除资源：
    cm_ctl res --del --res_name=<name>

 版权所有 © 华为云计算技术有限公司

  - 修改资源：
    ```
    cm_ctl res --edit --res_name=<name>  --res_attr=<res_info> [--inst_attr=<res_attr>]
    ```
  - 新增资源实例：
    ```
    cm_ctl res --edit --res_name=<name>  --add_inst=<inst_info> [--inst_attr=<res_attr>]
    ```
  - 修改资源实例：
    ```
    cm_ctl res --edit --res_name=<name> --edit_inst=<inst_info> [--inst_attr=<res_attr>]
    ```
  - 删除资源实例：
    ```
    cm_ctl res --edit --res_name=<name>  --del_inst=<inst_info>
    ```
  - 检查资源：
    ```
    cm_ctl res --check
    ```
  - 展示所有资源基本配置：
    ```
    cm_ctl res --list
    ```
  - 展示单个资源配置:
    ```
    cm_ctl res --list --res_name=<name>
    ```
  - 展示单个资源实例配置：
    ```
    cm_ctl res --list --res_name=<name> --list_inst
    ```
- 执行show命令：
  ```
  cm_ctl show
  ```
- 执行notify命令：
  - 通知cm发生故障
    ```
    cm_ctl notify --fault_occur --detail="<fault_info_json>"
    ```
  - 通知cm故障恢复
    ```
    cm_ctl notify --fault_recover -n <node_id>
    ```
- 开启和修改急救室功能命令：
  - 获取最新的元数据信息：
    ```
    cm_ctl pull --cm_meta_config
    ```
  - 上传本地最新的元数据信息：
    ```
    cm_ctl push --cm_meta_config
    ```

6.3 cm_ctl 工具使用指南

6.3.1 查询集群状态

命令格式

```
cm_ctl query [-v [-C [-s] [-S] [-d] [-i] [-F] [-x] [-p] [-z ALL]] | [-r]] [--minorityAz=<AZ_NAME>] [-g] [-l <FILE_NAME>] [-t <SECS>]
```

参数说明

cm_ctl query可单独使用，表示只显示集群状态，也可以根据实际情况和中括号[]内参数结合使用。

- -v：显示详细集群状态。
- -C：按主备关系成对显示集群状态。-C参数需要和-v参数一起使用。
- -s：显示导致各个节点主备实例数量不均衡的实例。-s参数需要和-v以及-C参数一起使用。
- -S：显示集群启动时的状态检查结果。-S参数需要和-v以及-C参数一起使用。
- -d：显示实例数据目录。-d参数需要和-v以及-C参数一起使用。

 版权所有 © 华为云计算技术有限公司

- -i：显示物理节点ip。-i参数需要和-v以及-C参数一起使用。
- -F：显示各个节点Fenced UDF状态。-F参数需要和-v以及-C参数一起使用。
- -x：显示集群所有异常实例。-x参数需要和-v以及-C参数一起使用。
- -p：显示数据库实例所有DN端口。-p参数需要和-v以及-C参数一起使用。
- -z ALL：显示集群所有实例AZ名称。-z参数需要和-v以及-C参数一起使用，并且-z后面需要跟参数ALL。
- -r：显示备机redo状态。使用-r时，必须指定-v参数。-r参数需要和-v参数一起使用。
- --minorityAz=<AZ_NAME>：只查询指定AZ的cms和etcd，AZ_NAME为指定AZ名称，可通过cm_ctl query -Cvz ALL命令查询到。此参数会忽略非指定AZ的cms、etcd节点，可以在少数派场景提高查询速度。
- -g：显示备份和恢复群集信息。
- -l <FILE_NAME>：结果输出到指定文件，<FILE_NAME>为用户设置的文件名称。
- -t <SECS>：指定超时时间。超时后，会退出并报错。不设置该参数时，cm_ctl query命令默认超时时间为60s。

常用命令

- 按主备关系成对显示集群详细状态信息：

```
cm_ctl query -Cv
```

执行命令结果如下：

```
[  CMServer State   ]

node            instance state
---------------------------------
1  pekphis351066 1       Primary
2  pekphis351067 2       Standby
3  pekphis351068 3       Standby

[    ETCD State     ]

node            instance state
-----------------------------------------
1  pekphis351066 7001     StateFollower
2  pekphis351067 7002     StateLeader
3  pekphis351068 7003     StateFollower

[   Cluster State   ]

cluster_state   : Normal
redistributing  : No
balanced        : Yes
current_az      : AZ_ALL

[  Datanode State   ]

node            instance state          | node            instance state          | node            instance state
-----------------------------------------------------------------------------------------------------------------------------------
----------------
1  pekphis351066 6001     P Primary Normal | 2  pekphis351067 6002     S Standby Normal | 3
pekphis351068 6003     S Standby Normal
```


版权所有 © 华为云计算技术有限公司

说明

- 节点名称为node列，本示例中node分别为1、2、3。
- 实例名称为instance列，本示例中CMServer组件的instance分别为1、2、3。同理，ETCD组件的instance分别为7001、7002、7003，DN组件的instance分别6001、6002、6003。
- 各个组件的状态为state列所展示。
- 集群状态为Cluster State所展示。数据库实例状态分别为：Normal，表示数据库实例可用，且数据有冗余备份，所有进程都在运行，主备关系正常；Degraded，表示数据库实例可用，但数据没有冗余备份；Unavailable，表示数据库实例不可用。

- 按主备关系成对显示集群详细状态信息，并查询实例目录路径和实例AZ名称：

```
cm_ctl query -Cvdz ALL
```

执行命令结果如下：

```
[  CMServer State   ]

node                instance                              state
------------------------------------------------------------------------
AZ1 1  pekphis351066 1    /data1/omm/cluster/data/cm/cm_server Primary
AZ2 2  pekphis351067 2    /data1/omm/cluster/data/cm/cm_server Standby
AZ3 3  pekphis351068 3    /data1/omm/cluster/data/cm/cm_server Standby

[    ETCD State     ]

node                instance                       state
------------------------------------------------------------------------
AZ1 1  pekphis351066 7001 /data1/omm/cluster/data/etcd StateFollower
AZ2 2  pekphis351067 7002 /data1/omm/cluster/data/etcd StateLeader
AZ3 3  pekphis351068 7003 /data1/omm/cluster/data/etcd StateFollower

[   Cluster State   ]

cluster_state   : Normal
redistributing  : No
balanced        : Yes
current_az      : AZ_ALL

[  Datanode State   ]

node                instance                           state          | node
instance                           state          | node              instance                           state
------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------
--
AZ1 1  pekphis351066 6001 /data1/omm/cluster/data/datanode1p P Primary Normal | AZ2 2
pekphis351067 6002 /data1/omm/cluster/data/datanode1s S Standby Normal | AZ3 3  pekphis351068
6003 /data1/omm/cluster/data/datanode1s S Standby Normal
```

说明

比起上条cm_ctl query -Cv命令，该命令增加-d和-z ALL参数，查询结果即增加实例目录路径、实例AZ名称信息。

- 实例目录路径位于instance名称后，本示例中CMServer组件的实例1、2、3的目录路径均为/data1/omm/cluster/data/cm/cm_server。ETCD、DN组件的目录路径位置展示同理。
- 实例AZ名称位于node名称前，本示例中AZ分别为AZ1、AZ2、AZ3。

- 显示备机redo状态：

```
cm_ctl query -rv
```

执行命令结果如下：

```
------------------------------------------------------------------------

cluster_state           : Normal
```


版权所有 © 华为云计算技术有限公司

```
redistributing            : No
balanced                  : Yes

-----------------------------------------------------------------------

-----------------------------------------------------------------------

-----------------------------------------------------------------------

node                      : 2
instance_id               : 6002
node_ip                   : 172.31.20.71
data_path                 : /data1/omm/cluster/data/datanode1s
type                      : Datanode
instance_state            : Standby
is_by_query               : 1
primary_flush_location    : 00000000/100E8AD0
standby_received_location : 00000000/100E8AD0
standby_replay_location   : 00000000/100E8AD0
static_connections        : 2
HA_state                  : Normal
reason                    : Normal
redo_start_location       : 00000000/05000028
min_recovery_location     : 00000000/100E8A48
read_location             : 00000000/100E8A88
last_replayed_end_location : 00000000/100E8AD0
recovery_done_location    : 00000000/06000058
local_max_lsn             : 00000000/100E8AD0
read_xlog_io_counter      : 220102
read_xlog_io_total_dur    : 14767657
read_data_io_counter      : 135
read_data_io_total_dur    : 2242
write_data_io_counter     : 0
write_data_io_total_dur   : 0
process_pending_counter   : 779630
process_pending_total_dur : 888622
apply_counter             : 791945
apply_total_dur           : 67530080
speed(est.)               : 0 KB/s
delay(est.)               : 0 s
senderPercent             : 0%
receiverPercent           : 100%
worker_info               :
id  q_use   q_max_use  rec_cnt
0   0       4094       351176
1   0       255        270636
2   0       141        154917
3   0       3          2901

node                      : 3
instance_id               : 6003
node_ip                   : 172.31.16.100
data_path                 : /data1/omm/cluster/data/datanode1s
type                      : Datanode
instance_state            : Standby
is_by_query               : 1
primary_flush_location    : 00000000/100E8AD0
standby_received_location : 00000000/100E8AD0
standby_replay_location   : 00000000/100E8AD0
static_connections        : 2
HA_state                  : Normal
reason                    : Normal
redo_start_location       : 00000000/03876F70
min_recovery_location     : 00000000/100E8AD0
read_location             : 00000000/100E8A88
last_replayed_end_location : 00000000/100E8AD0
recovery_done_location    : 00000000/04000058
local_max_lsn             : 00000000/100E8AD0
read_xlog_io_counter      : 210365
```


版权所有 © 华为云计算技术有限公司

```
read_xlog_io_total_dur          : 8463548
read_data_io_counter            : 135
read_data_io_total_dur          : 98189
write_data_io_counter           : 0
write_data_io_total_dur         : 0
process_pending_counter         : 779630
process_pending_total_dur       : 818020
apply_counter                   : 791950
apply_total_dur                 : 85874445
speed(est.)                     : 0 KB/s
delay(est.)                     : 0 s
senderPercent                   : 0%
receiverPercent                 : 100%
worker_info                     :
id  q_use   q_max_use  rec_cnt
0   0       4094       351176
1   0       248        270636
2   0       422        154917
3   0       3          2901

-----------------------------------------------------------------------
```

执行命令参数说明如下表：

参数	参数说明
is_by_query	是否通过SQL查询。
primary_flush_location	主机下刷日志的位置。
standby_received_location	备机收到日志的位置。
standby_replay_location	备机回放日志的位置。
static_connections	连接统计。
redo_start_location	回放开始的位置。
min_recovery_location	最小恢复点位置。
read_location	当前回放读取日志的位置。
last_replayed_end_location	上一次回放到的位置。
recovery_done_location	回放结束的位置。
local_max_lsn	本地最大日志位置。
read_xlog_io_counter	读日志的次数。
read_xlog_io_total_dur	读日志的总耗时。
read_data_io_counter	读数据的次数。
read_data_io_total_dur	读数据的总耗时。
write_data_io_counter	写数据的次数。
write_data_io_total_dur	写数据的总耗时。
process_pending_counter	回放主线程日志分发的次数。
process_pending_total_dur	回放主线程日志分发的总耗时。
apply_counter	日志回放同步的次数。

版权所有 © 华为云计算技术有限公司

参数	参数说明
apply_total_dur	日志回放同步的总耗时。
worker_info	回放线程的信息。
id	回放线程序号。
q_use	回放线程队列的使用量。
q_max_use	回放线程队列的最大容量。
rec_cnt	回放线程处理日志总个数。

6.3.2 启停集群

启动集群

命令格式

```
cm_ctl start [-z <AZ_NAME> [--cm_arbitration_mode=<ARBITRATION_MODE>]] | [--server] | [-n <NODE_ID> [-D <DATA_DIR> | -I <RESOURCE_INSTANCE_ID>]] [-t <SECS>]
```

参数说明

cm_ctl start可单独使用，表示启动整个集群，也可以根据实际情况和中括号[]内参数结合使用。

- -z <AZ_NAME>：启动指定AZ，<AZ_NAME>为AZ名称，可通过cm_ctl query -Cvz ALL命令查询。
- --cm_arbitration_mode=<ARBITRATION_MODE>：以多数派或者少数派模式启动指定AZ。<ARBITRATION_MODE>共有MAJORITY、MINORITY两种模式，MAJORITY为多数派模式，MINORITY为少数派模式。--cm_arbitration_mode参数需要和-z参数一起使用。
- -n <NODE_ID>：启动指定节点，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。
- -D <DATA_DIR>：启动指定实例，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- -I <RESOURCE_INSTANCE_ID>：指定需要停止的资源实例号，可以通过cm_ctl query -Cv查看资源实例和实例号的映射关系。
- -t <SECS>：指定超时时间。超时后，会退出并报错。不设置该参数时，cm_ctl start命令默认超时时间为600s。
- --server：独立启动cm_server。只能在集群停止后执行。

常用命令

- 启动集群：
  ```
  cm_ctl start
  ```
- 启动指定AZ：
  ```
  cm_ctl start -z AZ1
  ```
- 启动指定节点：
  ```
  cm_ctl start -n 1
  ```
- 启动指定实例：
  ```
  cm_ctl start -n 1 -D /data1/omm/cluster/data/datanode1p
  ```


版权所有 © 华为云计算技术有限公司

- DCC模式单独启动cm_server
  ```
  cm_ctl start --server
  ```

停止集群

命令格式

```
cm_ctl stop [[-z <AZ_NAME>] | [--server] | [-n <NODE_ID> [-D <DATA_DIR> | -I <RESOURCE_INSTANCE_ID>]]] [-m <SHUTDOWN_MODE>] [-t <SECS>]
```

参数说明

cm_ctl stop可单独使用，表示停止整个集群 ，也可以根据实际情况和中括号[]内参数结合使用。

- -z <AZ_NAME>：停止指定AZ，<AZ_NAME>为AZ名称，可通过cm_ctl query -Cvz ALL命令查询。
- -n <NODE_ID>：停止指定节点，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。
- -D <DATA_DIR>：停止指定实例，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- -I：指定需要停止的资源实例号，可以通过cm_ctl query -Cv查看资源实例和实例号的映射关系。
- -m <SHUTDOWN_MODE>：指定停止模式；停止模式有以下几种：
  - smart（s）：等待用户业务结束后，停止所有数据库实例。
  - fast（f）：不等待用户业务结束，指定数据库实例退出。
  - immediate（i）：不等待用户业务结束，指定数据库实例强制退出。
- -t <SECS>：指定超时时间。超时后，会退出并报错。不设置该参数时，cm_ctl stop命令默认超时时间为1200s。
- --server：停止独立启动的cm_server。只能在独立启动cm_server后执行。

常用命令

- 停止集群：
  ```
  cm_ctl stop
  ```
- 停止指定AZ：
  ```
  cm_ctl stop -z AZ1
  ```
- 停止指定节点：
  ```
  cm_ctl stop -n 1
  ```
- 停止指定实例：
  ```
  cm_ctl stop -n 1 -D /data1/omm/cluster/data/datanode1p
  ```
- 强制停止集群：
  ```
  cm_ctl stop -m i
  ```
- 停止独立启动的cm_server
  ```
  cm_ctl stop --server
  ```

6.3.3 数据库主备倒换

注意事项

switchover为维护操作，确保集群状态正常，所有业务结束，并使用pg_get_senders_catchup_time视图查询无主备追赶后，再进行switchover操作。

版权所有 © 华为云计算技术有限公司

命令格式

```
cm_ctl switchover [-z <AZ_NAME>] | [-n <NODE_ID> -D <DATA_DIR> [-f]] | [-a] | [-A] [-t <SECS>]
```

参数说明

- -z <AZ_NAME>：将所有GTM、DN主实例切换到指定AZ，<AZ_NAME>为AZ名称，可通过cm_ctl query -Cvz ALL命令查询。
- -n <NODE_ID>：将指定的实例切换为主实例，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。-n参数需要和-D参数一起使用。
- -D <DATA_DIR>：将指定的实例切换为主实例，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- -f：指定进行-f类型switchover。-f参数需要和-n以及-D参数一起使用。
- -a：重置集群主备关系为初始配置状态。
- -A：将所有DN实例统一从主切换到备，并选择相应分片的其它实例升主。
- -t <SECS>：指定超时时间。超时后，会退出并报错。不设置该参数时，cm_ctl switchover命令默认超时时间为120s。

常用命令

- 指定AZ的数据库主备倒换：
  ```
  cm_ctl switchover -z AZ1
  ```
- 指定实例的数据库主备倒换：
  ```
  cm_ctl switchover -n 1 -D /data1/omm/cluster/data/datanode1p
  ```
- 将所有DN实例统一从主切换到备，并选择相应分片的其它实例升主：
  ```
  cm_ctl switchover -A
  ```
- 重置集群主备关系为初始配置状态：
  ```
  cm_ctl switchover -a
  ```
- DCC模式下指定cm_server主备切换：
  ```
  cm_ctl switchover -n 1 -D /data1/omm/cluster/cm/cm_server
  ```

6.3.4 重建备 DN

命令格式

```
cm_ctl build [-c] | [-n <NODE_ID> -D <DATA_DIR> [-f] [-b full]] [-t <SECS>]
```

参数说明

- -c：重建 cm_server（将主节点的dcc数据目录复制到指定节点，只适用于一主一备模式）。
- -n <NODE_ID>：指定重建备DN，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。-n参数需要和-D参数一起使用。
- -D <DATA_DIR>：指定重建备DN，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- -f：强制重建备机。-f参数需要和-n以及-D参数一起使用。
- -b full：指定进行全量build。不指定情况下，对于一主多备集群部署模式进行auto build。auto build指：先调用增量build，失败之后调用全量build。-b参数需要和-n以及-D参数一起使用。
- -t <SECS>：指定超时时间。超时后，会退出并报错。

版权所有 © 华为云计算技术有限公司

命令约束

非强制build目标需要处于Down状态，或是处于NeedRepair状态的备机、级联备，且reason不为Disconnected或者Connecting。

常用命令

- 重建备DN：
  ```
  cm_ctl build -n 2 -D /data1/omm/cluster/data/datanode1s
  ```
- 重建 cm_server：
  ```
  cm_ctl build -c
  ```

6.3.5 检测实例进程运行状态

命令格式

```
cm_ctl check -B <BIN_NAME> -T <DATA_PATH>
```

参数说明

- -B <BIN_NAME>：指定进程名，其进程名包括“cm_agent”，“gaussdb”和“cm_server”。
- -T <DATA_PATH>：指定实例数据目录。

常用命令

- 检测cm_agent进程运行状态：
  ```
  cm_ctl check -B cm_agent -T /data1/omm/cluster/data/cm/cm_agent
  ```
- 检测cm_server进程运行状态：
  ```
  cm_ctl check -B cm_server -T /data1/omm/cluster/data/cm/cm_server
  ```

6.3.6 查看数据库实例配置文件

命令格式

```
cm_ctl view [-v] [-N] | [-n <NODE_ID>] [-c] [-l <FILE_NAME>]
```

参数说明

cm_ctl view可单独使用，表示显示集群静态配置信息，也可以根据实际情况和中括号[]内参数结合使用。

- -v：显示集群所有节点的静态配置详细信息。相对于cm_ctl view命令，-v 显示内容增加了cmserver、dn组件的编号显示。
- -N：只显示本节点的静态配置信息，即执行cm_ctl view命令的节点的信息，N表示Native。
- -n <NODE_ID>：显示指定节点的静态配置信息，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。
- -c：显示CN动态配置信息，CN是否被剔除。指定-c参数需要在部署有cm_server的节点执行。
- -l <FILE_NAME>：结果输出到指定文件，<FILE_NAME>为用户设置的文件名称。

版权所有 © 华为云计算技术有限公司

常用命令

- 查看集群配置文件：
  ```
  cm_ctl view
  ```
- 查看指定节点配置文件：
  ```
  cm_ctl view -n 1
  ```

6.3.7 设置、获取 CM_Server 相关参数

设置参数

命令格式

```
cm_ctl set [--log_level=<LOG_LEVEL>] [--cm_arbitration_mode=<ARBITRATION_MODE>] [--cm_switchover_az_mode=<SWITCHOVER_AZ_MODE>] [--cmsPromoteMode=<CMS_PROMOTE_MODE> -I <INSTANCE_ID>]
```

参数说明

- --log_level=<LOG_LEVEL>：设置主cm_server日志级别。共分为DEBUG5、DEBUG1、WARNING、LOG、ERROR和FATAL六个级别，日志打印信息级别越来越高。设置日志级别越高，输出日志信息就越少。
- --cm_arbitration_mode=<ARBITRATION_MODE>：一主多备功能，设置cm_server的仲裁模式。共有MAJORITY、MINORITY两种模式，MAJORITY为多数派模式，MINORITY为少数派模式。少数派模式适用于一主多备集群部署并且只有AZ3存活时，此时cm_server可以进行正常的仲裁业务；多数派模式适用于一主多备集群部署并且各个组件（cm_server，etcd，dn）存活数量大于一半的场景。集群正常情况下默认为多数派模式。
- --cm_switchover_az_mode=<SWITCHOVER_AZ_MODE>：一主多备功能，设置AZ之间的自动切换开关。共有NON_AUTO、AUTO两种模式，NON_AUTO为非自动切换模式，AUTO为自动切换模式。AUTO模式由主cm_server自动控制AZ1和AZ2之间的DN实例切换。
- --cmsPromoteMode=<CMS_PROMOTE_MODE> -I <INSTANCE_ID>：设置cms的升主方式，共有AUTO、PRIMARY_F两种模式，AUTO为默认自选主方式，PRIMARY_F会强制使-I指定的节点升主，无论当前是否有主，因此，有可能会导致cms存在多主情况。<INSTANCE_ID>为实例名称，可通过cm_ctl query -Cv命令查询。

常用命令

- 设置主cm_server日志级别为DEBUG5：
  ```
  cm_ctl set --log_level=DEBUG5
  ```
- 设置cm_server的仲裁模式为多数派模式：
  ```
  cm_ctl set --cm_arbitration_mode=MAJORITY
  ```
- 设置AZ之间的自动切换开关为自动切换模式：
  ```
  cm_ctl set --cm_switchover_az_mode=AUTO
  ```
- 设置cms的升主方式为自选主方式：
  ```
  cm_ctl set --cmsPromoteMode=AUTO -I 1
  ```

获取参数

命令格式

```
cm_ctl get {--log_level | --cm_arbitration_mode | --cm_switchover_az_mode}
```

参数说明

 版权所有 © 华为云计算技术有限公司

- --log_level：获取主cm_server日志级别。共分为DEBUG5、DEBUG1、WARNING、LOG、ERROR和FATAL六个级别，日志打印信息级别越来越高。设置日志级别越高，输出日志信息就越少。
- --cm_arbitration_mode：一主多备功能，获取cm_server的仲裁模式。共有MAJORITY、MINORITY两种模式，MAJORITY为多数派模式，MINORITY为少数派模式。少数派模式适用于一主多备集群部署并且只有AZ3存活时，此时cm_server可以进行正常的仲裁业务；多数派模式适用于一主多备集群部署并且各个组件（cm_server、etcd和dn）存活数量大于一半的场景。集群正常情况下默认为多数派模式。
- --cm_switchover_az_mode：一主多备功能，获取AZ之间的自动切换开关。共有NON_AUTO、AUTO两种模式，NON_AUTO为非自动切换模式，AUTO为自动切换模式。AUTO模式由主cm_server自动控制AZ1和AZ2之间的DN实例切换。

常用命令

- 获取主cm_server日志级别：
  ```
  cm_ctl get --log_level
  ```
- 获取cm_server的仲裁模式：
  ```
  cm_ctl get --cm_arbitration_mode
  ```
- 获取AZ之间的自动切换开关：
  ```
  cm_ctl get --cm_switchover_az_mode
  ```

6.3.8 CM 参数相关操作

cm_agent相关参数可通过cm_agent数据目录下的cm_agent.conf文件查看，cm_server相关参数可通过cm_server数据目录下的cm_server.conf文件查看，详细请参见**CM配置参数介绍**。

设置 CM 参数

命令格式

```
cm_ctl set --param {--agent | --server} [-n <NODE_ID>] -k "[PARAMETER]=<value>"
```

参数说明

- --agent：设置cm_agent参数。
- --server：设置cm_server参数。
- -n <NODE_ID>：设置指定节点的CM参数，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。不设置该参数时，默认设置所有节点。
- -k "parameter='value'"：指定设置的参数和参数的值，只能设置已经存在的参数，不支持增加或删减参数。

常用命令

- 设置cm_agent参数incremental_build值为on：
  ```
  cm_ctl set --param --agent -k "incremental_build=on"
  ```
- 设置cm_server参数switch_rto值为600：
  ```
  cm_ctl set --param --server -k "switch_rto=600"
  ```

动态加载 CM 参数

命令格式

```
cm_ctl reload --param {--agent | --server}
```

参数说明

 版权所有 © 华为云计算技术有限公司

- --agent：动态加载cm_agent参数。
- --server：动态加载cm_server参数。

常用命令

- 动态加载cm_agent参数：
  ```
  cm_ctl reload --param --agent
  ```
- 动态加载cm_server参数：
  ```
  cm_ctl reload --param --server
  ```

列出所有 CM 参数

命令格式

```
cm_ctl list --param {--agent | --server}
```

参数说明

- --agent：列出cm_agent参数。
- --server：列出cm_server参数。

常用命令

- 列出cm_agent参数：
  ```
  cm_ctl list --param --agent
  ```
- 列出cm_server参数：
  ```
  cm_ctl list --param --server
  ```

6.3.9 CM_Server 热补丁升级操作

注意事项

数据库热补丁升级操作会调用cm_ctl hotpatch命令，以实现CM_Server的热补丁升级，不建议用户使用该命令进行热补丁升级操作。

命令格式

```
cm_ctl hotpatch -E <PATCH_COMMAND> -P <PATCH_NAME>
```

参数说明

- -E <HOTPATCH_COMMAND>：指定需要执行的热补丁操作。
  <HOTPATCH_COMMAND>有以下取值范围：
  - load： 加载热补丁
  - unload：卸载热补丁
  - active：激活热补丁
  - deactive：去激活热补丁
  - info：显示热补丁信息
  - list：显示热补丁列表
- -P <HOTPATCH_NAME>：指定需要执行操作的热补丁名称，需要完整的路径名。

常用命令

- 显示热补丁列表：

 版权所有 © 华为云计算技术有限公司

```
cm_ctl hotpatch -E list
```

- 加载热补丁：
```
cm_ctl hotpatch -E load -P /data1/omm/cluster/app/bin/hotpatch/GaussDB-Kernel_xxx.pat
```
- 激活热补丁：
```
cm_ctl hotpatch -E active -P /data1/omm/cluster/app/bin/hotpatch/GaussDB-Kernel_xxx.pat
```
- 显示热补丁信息：
```
cm_ctl hotpatch -E info -P /data1/omm/cluster/app/bin/hotpatch/GaussDB-Kernel_xxx.pat
```
- 去激活热补丁：
```
cm_ctl hotpatch -E deactive -P /data1/omm/cluster/app/bin/hotpatch/GaussDB-Kernel_xxx.pat
```
- 卸载热补丁：
```
cm_ctl hotpatch -E unload -P /data1/omm/cluster/app/bin/hotpatch/GaussDB-Kernel_xxx.pat
```

6.3.10 加密

cm_ctl encrypt用于加密自认证ssl证书的私钥密码，对输入的密码进行加密操作，密码支持8~15位，且必须包含三种字符（数字，字母，符号）。

命令格式

```
cm_ctl encrypt [-M <MODE>] -D <DATA_DIR>
```

参数说明

- -M：指定加密类型，支持server、client。不设置此参数时，默认类型为server。
- -D：指定生成的加密密码文件路径。

常用命令

指定client加密：
```
cm_ctl encrypt -M client -D /data1/omm/cluster/data/datanode1p
```

6.3.11 ddb 相关操作

ddb是指分布式数据存储，例如dcc、etcd。ddb相关操作主要用于存储CM的仲裁信息，通过cm_ctl ddb命令查看ddb运行状态、设置和获取键值对，通过cm_ctl switch命令切换ddb模式。

执行 ddb 命令行

命令格式
```
cm_ctl ddb "[--put <key> <value>] | [--get <key> | --delete <key> [--prefix]] | [--cluster_info] | [--leader_info] | [--help] | [--version]"
```

参数说明

- --put <key> <value>：往ddb中插入键值对，如果键值对已存在则会修改键key所对应的值value。
- --get <key>：查询ddb中key对应的value。
- --delete <key>：删除ddb中指定的键值对。
- --prefix：get或者delete后添加prefix参数，可以实现模糊匹配查询和删除。
- --cluster_info：获取数据库实例信息。
- --leader_info：获取主节点信息。

 版权所有 © 华为云计算技术有限公司

- --help：显示ddb命令帮助信息。
- --version：显示ddb版本信息。
- --member_list : 查询数据库实例中所有节点的信息。
- --member_add ： 向数据库实例中增加某节点。
- --member_remove：删除数据库实例中某节点。
- --change_role：改变某节点的角色。
- --change_weight：改变某节点投票的权重。
- --setrunmode：设置运行模式，可以启动少数派模式或恢复正常模式。
- --fixleader -fixval {0 | 1}：修改定主, 1:设置定主,0:取消定主。
- --build_info：查询dcc的build信息。

常用命令

- 插入键值对：
  ```
  cm_ctl ddb "--put /test/key1 value1"
  ```
- 查询键值对：
  ```
  cm_ctl ddb "--get /test/key1"
  ```
- 删除键值对：
  ```
  cm_ctl ddb "--delete /test/key1"
  ```
- 获取数据库实例信息：
  ```
  cm_ctl ddb "--cluster_info"
  ```
- 获取主节点信息：
  ```
  cm_ctl ddb "--leader_info"
  ```
- 查看ddb命令帮助信息：
  ```
  cm_ctl ddb "--help"
  ```
- 显示ddb版本信息：
  ```
  cm_ctl ddb "--version"
  ```
- 查询数据库实例中所有节点的信息：
  ```
  cm_ctl ddb "--member_list"
  ```
- 向数据库实例中增加3号节点：
  ```
  cm_ctl ddb "--member_add -n 3 -i 127.0.0.1 -p 12766 -r FOLLOWER"
  ```
- 删除数据库实例中的3号节点：
  ```
  cm_ctl ddb "--member_remove -n 3"
  ```
- 改变1号节点的角色：
  ```
  cm_ctl ddb "--change_role -n 1 -r passive"
  ```
- 改变1号节点的投票权重为2：
  ```
  cm_ctl ddb "--change_weight -n 1 -w 2"
  ```
- 设置运行模式为少数派模式，投票数为2：
  ```
  cm_ctl ddb "--setrunmode -xmode minority -votenum 2"
  ```
- 设置运行模式为正常模式：
  ```
  cm_ctl ddb "--setrunmode -xmode normal"
  ```
- 设置定主：
  ```
  cm_ctl ddb "--fixleader -fixval 1"
  ```
- 取消定主：
  ```
  cm_ctl ddb "--fixleader -fixval 0"
  ```
- 查询dcc的build信息：
  ```
  cm_ctl ddb "--build_info"
  ```

 版权所有 © 华为云计算技术有限公司

切换 ddb 模式

命令格式

```
cm_ctl switch {--ddb_type=<DDB_TYPE> | --commit | --rollback}
```

参数说明

- --ddb_type=<DDB_TYPE>：选择需要切换到哪一个模式（ETCD或DCC），切换成功后需要执行--commit命令，切换失败后需要执行--rollback命令。
- --commit：执行切换会使集群无法选主，需要执行commit恢复集群。
- --rollback：回滚操作，执行切换失败需要执行回滚操作。

常用命令

- 切换ddb模式到ETCD：
  ```
  cm_ctl switch --ddb_type=ETCD
  ```
- 切换ddb模式到DCC：
  ```
  cm_ctl switch --ddb_type=DCC
  ```
- 提交ddb模式切换：
  ```
  cm_ctl switch --commit
  ```
- 回滚ddb模式切换：
  ```
  cm_ctl switch --rollback
  ```

6.3.12 设置 DCF 相关内容

设置 DCF 投票数

命令格式

```
cm_ctl setrunmode -n <NODE_ID> -D <DATA_DIR>  {--xmode=normal | --xmode=minority --votenum=<NUM>}
```

参数说明

该命令只用于DCF模式。

- -n <NODE_ID>：指定实例节点，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。-n参数需要和-D参数一起使用。
- -D <DATA_DIR>：指定实例路径，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- --xmode：指定DCF的运行模式。运行模式有以下取值范围：
  - normal：正常模式。
  - minority：少数派模式，需要--votenum指定投票数
- --votenum：指定DCF少数派运行方式的投票数。NUM取值范围为正整数，不高于DCF总副本数。

常用命令

- 设置DCF运行模式为正常模式：
  ```
  cm_ctl setrunmode -n 1 -D /data1/omm/cluster/data/datanode1p --xmode=normal
  ```
- 设置DCF运行模式为少数派模式，且投票数为2：
  ```
  cm_ctl setrunmode -n 1 -D /data1/omm/cluster/data/datanode1p --xmode=minority --votenum=2
  ```

改变 DCF 角色信息

命令格式

```
cm_ctl changerole --role={PASSIVE | FOLLOWER}  -n <NODE_ID> -D <DATA_DIR> [-t <SECS>]
```


版权所有 © 华为云计算技术有限公司

参数说明

该命令只用于DCF模式。

- -n <NODE_ID>：指定实例节点，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。-n参数需要和-D参数一起使用。
- -D <DATA_DIR>：指定实例路径，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- --role： 集中式DN数据库实例，DCF模式下节点角色。角色有以下取值范围：
  - passive：passive角色。
  - follower：follow角色。
- -t <SECS>：指定超时时间。超时后，会退出并报错。

常用命令

- 改变DCF角色为passive角色：
  ```
  cm_ctl changerole --role=PASSIVE -n 2 -D /data1/omm/cluster/data/datanode1s
  ```
- 改变DCF角色为follow角色：
  ```
  cm_ctl changerole --role=FOLLOWER -n 2 -D /data1/omm/cluster/data/datanode1s
  ```

改变 DCF 节点属性

命令格式

```
cm_ctl changemember --role={PASSIVE | FOLLOWER} [--group=<GROUP>] [--priority=<PRIORITY>] -n <NODE_ID> -D <DATA_DIR> [-t <SECS>]
```

参数说明

该命令只用于DCF模式。

- -n <NODE_ID>：指定实例节点，<NODE_ID>为节点名称，可通过cm_ctl query -Cv命令查询。-n参数需要和-D参数一起使用。
- -D <DATA_DIR>：指定实例路径，<DATA_DIR>为指定实例数据目录，可通过cm_ctl query -Cvd命令查询。-D参数需要和-n参数一起使用。
- --role：集中式DN数据库实例，DCF模式下节点角色。角色有以下取值范围：
  - passive：passive角色。
  - follower：follow角色。
- --group=<GROUP>：DCF模式下，修改group的值，其取值范围为0~2147483647。
- --priority=<PRIORITY>：DCF模式下，修改priority的值，其取值范围为0~2147483647。
- -t <SECS>：指定超时时间。超时后，会退出并报错。

常用命令

- 改变DCF角色为passive角色，同时修改group的值为3，修改priority的值为2：
  ```
  cm_ctl changemember --role=PASSIVE --group=3 --priority=2 -n 2 -D /data1/omm/cluster/data/datanode1s
  ```
- 改变DCF角色为follower角色，同时修改group的值为3，修改priority的值为2：
  ```
  cm_ctl changemember --role=FOLLOWER --group=3 --priority=2 -n 2 -D /data1/omm/cluster/data/datanode1s
  ```

6.3.13 一键停止仲裁

命令格式

 版权所有 © 华为云计算技术有限公司

```
cm_ctl disablecm {--all | --start | --stop | --arbitrate="<MODE>"} [-n <NODE_ID> -D <DATA_DIR>]
cm_ctl enablecm {--all | --start | --stop | --arbitrate} [-n <NODE_ID> -D <DATA_DIR>]
```

参数说明

- disablecm：禁用集群能力。
- enablecm：启用集群能力。
- --all：所有实例（DN、CMS）的自动启停和集群仲裁。
- --start：所有实例（DN、CMS）的自动启动。
- --stop：所有实例（DN、CMS）的自动停止。
- --arbitrate="<MODE>"：集群仲裁能力。参数可选“DN_ARBITRATE”。不附加参数默认选择可选参数内的全部仲裁能力。

常用命令

- 一键禁用所有启停仲裁
  ```
  cm_ctl disablecm --all
  ```
- 一键启用所有启停仲裁
  ```
  cm_ctl enablecm --all
  ```
- 禁用全部实例启动
  ```
  cm_ctl disablecm --start
  ```
- 启用用单个实例启动
  ```
  cm_ctl enablecm --start -n 1 -D /data1/omm/cluster/data/datanode1p
  ```
- 禁用dn仲裁
  ```
  cm_ctl disablecm --arbitrate="DN_ARBITRATE"
  ```
- 启用全部仲裁
  ```
  cm_ctl enablecm --arbitrate
  ```

6.3.14 其他

所有备机停止回放，每个分片中选择一个强制升主

命令格式

```
cm_ctl finishredo
```

6.4 安全设计

6.4.1 手动替换证书步骤

创建自认证证书

自认证证书生成请联系管理员处理。

对私钥密码的保护，使用cm encrypt工具，请参考**cm_ctl工具介绍**。

```
cm_ctl encrypt [-M <MODE>] -D <DATA_DIR>
```

步骤1 生成服务器端密钥因子server.key.cipher、server.key.rand。

```
cm_ctl encrypt -M server -D <DATA_DIR>
please enter the password:（使用密码需要与服务器私钥的保护密码相同）
```

步骤2 生成客户端密钥因子client.key.cipher、client.key.rand。

版权所有 © 华为云计算技术有限公司

```
cm_ctl encrypt -M client -D <DATA_DIR>
please enter the password:（使用密码需要与客户端私钥的保护密码相同）
```

----结束

6.4.2 证书使用指南

步骤1 使用证书时需配置cm_server参数为on（默认关闭）。

```
cm_ctl set --param --server -k enable_ssl="on"
```

步骤2 证书文件需要存在于所有节点的$$GAUSSHOME/share/sslcert/cm中，证书替换后重启数据库实例生效。

所需证书文件：server.crt、server.key、client.crt、client.key、cacert.pem、server.key.cipher、server.key.rand、client.key.cipher、client.key.rand

根证书、密钥、证书以及密钥密码加密文件的权限，需保证权限为400。如果权限不满足要求，则无法使用ssl。

```
chmod 400 cacert.pem
chmod 400 server.crt
chmod 400 server.key
chmod 400 server.key.cipher
chmod 400 server.key.rand
chmod 400 client.crt
chmod 400 client.key
chmod 400 client.key.cipher
chmod 400 client.key.rand
```

步骤3 证书有效期的检测周期为1天，可通过ssl_cert_expire_check_interval设置。证书有效期剩余90天时会开始产生告警，可通过ssl_cert_expire_alert_threshold设置（配置参数详见**ssl_cert_expire_alert_threshold**）。

步骤4 客户端和服务端的证书吊销列表分别为client.crl、server.crl。

----结束

6.4.3 ifconfig 提权步骤

步骤1 执行以下命令获取ifconfig绝对路径。

```
which ifconfig
```

步骤2 在权限文件/etc/sudoers中输入以下命令。

```
Cmnd_Alias COMMAND_FLAG = ifconfig 绝对路径
数据库实例用户名 ALL=(root) NOPASSWD: COMMAND_FLAG
```

可以缩小权限范围，例如：

```
Cmnd_Alias COMMAND_FLAG = /usr/sbin/ifconfig * netmask * up, /usr/sbin/ifconfig * down
数据库实例用户名 ALL=(root) NOPASSWD: COMMAND_FLAG
```

在实际使用中，需将/usr/sbin/ifconfig和数据库实例用户名修改为实际环境对应的ifconfig的绝对路径和数据库实例用户名。

----结束

6.4.4 IP 命令提权步骤

步骤1 执行以下命令获取ip绝对路径。

```
which ip
```

 版权所有 © 华为云计算技术有限公司

步骤2 执行以下命令修改权限文件/etc/sudoers。

```
Cmnd_Alias COMMAND_FLAG = ip 绝对路径
数据库实例用户名 ALL=(root) NOPASSWD: COMMAND_FLAG
```

可以缩小权限范围，例如：

```
Cmnd_Alias COMMAND_FLAG = /usr/sbin/ip * add *, /usr/sbin/ip * del *
数据库实例用户名 ALL=(root) NOPASSWD: COMMAND_FLAG
```

在实际使用中，需将/usr/sbin/ip和数据库实例用户名修改为实际环境对应的ip的绝对路径和数据库实例用户名。

----结束

6.5 CM 配置参数介绍

cm_agent相关参数可通过cm_agent数据目录下的cm_agent.conf文件查看，cm_server相关参数可通过cm_server数据目录下的cm_server.conf文件查看。

6.5.1 cm_agent 参数

log_dir

参数说明：log_dir为存放cm_agent日志文件的目录。 必须是绝对路径。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --agent -k "log_dir='/log/dir'"。

取值范围：字符串，最大长度为1024。修改后需要重启cm_agent才能生效。参数修改请参见 **set cm参数** 进行设置。

默认值：'$GAUSSLOG/cm/cm_agent'，表示在$GAUSSLOG下对应的cm目录下生成cm_agent日志。

log_file_size

参数说明：控制日志文件的大小。当'cm_agent-xx-current.log'日志文件达到指定大小时，则重新创建一个日志文件记录日志信息。

取值范围：字符串，[0MB, 2047MB]，实际生效范围[1MB, 2047MB]，单位：MB。可不带单位设置，默认视为MB。如果需要带单位，仅能是"MB"。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：16

log_min_messages

参数说明：控制写到cm_agent日志文件中的消息级别。每个级别都包含排在它后面的所有级别中的信息。级别越低，服务器运行日志中记录的消息就越少。

取值范围：枚举类型，有效值有debug5、debug1、warning、error、log、fatal（不区分大小写）。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：warning

版权所有 © 华为云计算技术有限公司

incremental_build

参数说明：控制重建备DN模式是否为增量。打开这个开关，则增量重建备DN；否则，全量重建备DN。

取值范围：布尔型。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

- on、yes、true、1：表示增量重建备DN。
- off、no、false、0：表示全量重建备DN。

默认值：on

security_mode

参数说明：控制是否以安全模式启动DN。打开这个开关，则以安全模式启动DN；否则，以非安全模式启动DN。

取值范围：布尔型。修改后可以reload生效。参数修改请参见 **set cm参数** 进行设置。

- on、yes、true、1：表示以安全模式启动DN。
- off、no、false、0：表示以非安全模式启动DN。

默认值：off

upgrade_from

参数说明：就地升级过程中使用，用于标示升级前数据库的内部版本号，此参数禁止手动修改。

取值范围：非负整型，[0, 4294967295]。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：0

alarm_component

参数说明：在使用第一种告警方式时，设置用于处理告警内容的告警组件的位置。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --agent -k "alarm_component='/alarm/dir'"。

取值范围：字符串，最大长度为1024。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：/opt/huawei/snas/bin/snas_cm_cmd

alarm_report_interval

参数说明：指定告警上报的时间间隔。

取值范围：非负整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：1

版权所有 © 华为云计算技术有限公司

agent_heartbeat_timeout

参数说明：cm_server心跳超时时间。超时后，cm_agemt会断开已有的cm_server连接，重新尝试与cm_server建连。

取值范围：整型，[2, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：5

agent_connect_timeout

参数说明：cm_agent连接cm_server超时时间。超时后，表示建连失败。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：1

agent_connect_retries

参数说明：cm_agent连接cm_server尝试次数。

取值范围：整型，[0, 2147483647]。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：15

agent_kill_instance_timeout

参数说明：当cm_agent在无法连接cm_server主节点后，发起一次终止本节点上所有实例的操作之前，所需等待的时间间隔。

取值范围：整型，[0, 2147483647]。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：0

agent_report_interval

参数说明：cm_agent上报实例状态的时间间隔。

取值范围：整型，[1, 2147483647]。单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：1

alarm_report_max_count

参数说明：指定告警上报的最大次数。

取值范围：非负整型，[1, 2592000]。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：1

版权所有 © 华为云计算技术有限公司

agent_check_interval

参数说明：cm_agent查询实例状态的时间间隔。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：2

enable_log_compress

参数说明：控制压缩日志功能。

取值范围：布尔型。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

- on、yes、true、1：表示允许压缩日志。
- off、no、false、0：表示不允许压缩日志。

默认值：on

process_cpu_affinity

参数说明：控制是否以绑核优化模式启动主DN进程。配置该参数为0，则不进行绑核优化；否则，进行绑核优化，且物理CPU片数为2^n个。仅支持ARM。

取值范围：整型，[0, 2]。修改后需要重启数据库cm_agent才能生效，参数修改请参见 **set cm参数** 进行设置。

默认值：0

enable_xc_maintenance_mode

参数说明：在数据库为只读模式下，控制是否可以修改pgxc_node系统表。

取值范围：布尔型。修改后需要重启cm_agent才能生效。参数修改请参见 **set cm参数** 进行设置。

- on、yes、true、1：表示开启可以修改pgxc_node系统表功能。
- off、no、false、0：表示关闭可以修改pgxc_node系统表功能。

默认值：on

log_threshold_check_interval

参数说明：日志压缩和清除的时间间隔，每log_threshold_check_interval秒压缩和清理一次。

生效范围：请参见**CM支持日志管理**。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：1800

版权所有 © 华为云计算技术有限公司

log_max_size

参数说明：控制日志最大存储值，当日志总大小超过（log_max_size*95/100）MB时，根据压缩日志生成时间，依次删除历史压缩日志，直到日志总大小小于（log_max_size*95/100）MB。

生效范围：请参见**CM支持日志管理**。

取值范围：整型，[0, 2147483647]，单位：MB。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：10240

log_max_count

参数说明：可存储的最多日志数量，当日志总个数超过该值，根据压缩日志文件名时间，删除超过保留天数log_saved_days的压缩日志。

生效范围：请参见**CM支持日志管理**。

取值范围：整型，[0, 10000]，单位：个。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：10000

log_saved_days

参数说明：压缩日志保存的天数，压缩日志超过该值并且日志总个数超过log_max_count，删除压缩日志。

生效范围：请参见**CM支持日志管理**。

取值范围：整型，[0, 1000]，单位：天。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：90

须知

日志压缩能力受内存限制，最多只能检测到30000个日志文件。总日志量超过30000个文件时，无法保证日志能被正常压缩及删除。可以通过调整log_saved_days和log_threshold_check_interval快速清理已压缩日志文件。

agent_phony_dead_check_interval

参数说明：cm_agent检测进程是否僵死的时间间隔。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：10

版权所有 © 华为云计算技术有限公司

unix_socket_directory

参数说明：unix套接字的目录位置。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --agent -k "unix_socket_directory='/unix/dir'"。

取值范围：字符串，最大长度为1024。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：''

enable_dcf

参数说明：DCF模式开关。

取值范围：布尔型。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。请根据数据库实例实际的DCF运行模式进行更改，否则会导致数据库实例不可用。

- on、yes、true、1：表示启用dcf。
- off、no、false、0：表示不启用dcf。

默认值：off

disaster_recovery_type

参数说明：主备数据库灾备关系的类型。

取值范围：整型，[0, 2]。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

- 0：表示未搭建灾备关系。
- 1：表示搭建了dorado灾备关系。
- 2：表示搭建了流式灾备关系

默认值：0

agent_backup_open

参数说明：灾备模式设置，开启后CM按照灾备模式运行。

取值范围：整型，[0, 3]。修改后需要重启cm_agent才能生效。参数修改请参见 **set cm参数** 进行设置。

- 0：表示未搭建灾备关系。
- 1：表示搭建了dorado灾备关系。
- 2：表示搭建了流式灾备关系。
- 3：表示搭建了逻辑灾备关系。

默认值：0

environment_threshold

参数说明：环境资源监控阈值。

取值范围：string型。包含五个参数，分别为内存使用率阈值，cpu使用率阈值，磁盘IO阈值，实例的内存使用率阈值，实例的线程池使用率阈值，后两个参数只在分布式

版权所有 © 华为云计算技术有限公司

生效，前三个参数不区分集中式和分布式，超过阈值则会打印日志。修改后可以reload生效，参数修改请参见 **set cm参数** 进行设置。

默认值：(0,0,0,0,0)

max_connections_threshold

参数说明：使用DN管理端口和业务端口连接主DN失败后检测DN当前已使用的连接数，如果超过最大配置连接数的阈值，不做僵死处理，否则按照僵死流程处理。

取值范围：整型，[0, 100]。修改后可以reload生效。参数修改请参见 **set cm参数** 进行设置。

- 0表示关闭使用业务端口连接DN判定僵死流程，不打印DN堆栈，不判断DN已使用的连接数，只使用管理端口连接DN判断是否僵死。

默认值：90

thread_effective_time

参数说明：线程运行一轮的有效超时时间。超时说明此线程卡住，进程自动退出。

取值范围：整型，[0, 1200]，单位：秒。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

- 0：表示关闭线程监控能力。

注意：不建议调整参数值大小，如果调整请联系华为技术支持。

默认值：60

ping_fault_kill_inst_timeout

参数说明：主DN实例与所有备机ping不通后，重启该实例的超时时间。

取值范围：整型，[0, 1000]，单位：秒。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

备注说明：cm_server也会读取该参数，修改参数后，通知cm_server请参见**reload 参数**进行设置。

- 0：表示关闭ping fault处理能力。

默认值：30

high_cpu_threshold

参数说明：CPU高使用率的阈值，当CPU超过阈值，cm_agent会根据检测周期将CPU使用信息打印到日志中，默认：1s/次。

取值范围：整型，[0, 100]，单位：%。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：95

版权所有 © 华为云计算技术有限公司

high_mem_threshold

参数说明：内存高使用率的阈值，当内存超过阈值，cm_agent会根据检测周期将内存使用信息打印到日志中，默认：1s/次。

取值范围：整型，[0, 100]，单位：%。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：95

high_disk_threshold

参数说明：磁盘高IO的阈值，当IO超过阈值，cm_agent会根据检测周期将IO信息打印到日志中，默认：1s/次。并连续打印10次，频率为1s/次，将iostat -xm信息输出到system_stat日志中。

取值范围：整型，[0, 100]，单位：%。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：95

sys_check_interval

参数说明：系统资源检测周期，即CPU，内存，磁盘IO检测的间隔时间。配置成0则关闭检测。

取值范围：整型，[0, 10]，单位：秒。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：1

asyn_log_enable

参数说明：异步日志开关，打开后cm_agent启动异步日志。

取值范围：布尔型。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

- on、yes、true、1：表示cm_agent开启异步日志。
- off、no、false、0：表示cm_agent不开启异步日志，使用串行日志。

默认值：on

asyn_log_max_buf_size

参数说明：异步日志缓存大小，当异步日志发生堆积，且积压的日志超过了缓存大小，后续的日志无法记录。

取值范围：字符串，[1KB, 1048576KB]，单位：KB。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：1024

asyn_log_fault_solution

参数说明：识别到日志磁盘故障后，异步日志处理方案。当CM识别到日志盘故障或异步日志缓存溢出，CM会根据配置来处理异步日志。

版权所有 © 华为云计算技术有限公司

取值范围：整型，[0, 2]。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

- 0：表示不对日志做多余的处理，缓存区保留所有日志，直到缓存大小达到上限，达到上限后不会记录更多日志。
- 1：表示对日志进行筛选，识别到日志磁盘故障后，只将关键日志、告警和ERROR级别以上的日志保存到缓存区，直到缓存大小达到上限，达到上限后不会记录更多日志。
- 2：表示丢弃日志，识别到日志磁盘故障后，只将关键日志、告警保存到缓存区，直到缓存大小达到上限，达到上限后不会记录更多日志。

默认值：0

disk_check_hung_timeout

参数说明：磁盘I/O hang的超时时间，磁盘检测的单次I/O，如果超过这个时间，则认为磁盘已经I/O hang。。配置成0则关闭磁盘hang检测。

取值范围：整型，[0, 1000]，单位：秒。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：20，但如果安装时无法判断被检测磁盘是否为条带化，则会将默认值改为60。

disk_check_block_timeout

参数说明：磁盘因文件系统某些原因，导致open这类文件操作阻塞。这类操作如果阻塞超过此参数设定的时间，才会判定磁盘故障。配置为0，表示关闭此判断。

取值范围：整型，[0, 1000]，单位：秒。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：60

disk_check_interval

参数说明：控制磁盘I/O hang检测的流量，即最少需要在多长时间内写入指定的数据。默认检测周期是5秒，配置成0则关闭磁盘检测。

取值范围：整型，[0, 1000]，单位：秒。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：5

disk_check_io_size

参数说明：控制磁盘检测的流量，即在检测周期内向磁盘中写入的数据大小，这个是直接对磁盘的I/O操作。默认是写入512B的数据，如果检测的磁盘为条带化，则安装的时候会把默认参数改为262144B(256KB)。最大写入1G的数据。配置成0则关闭磁盘I/O hang检测。磁盘检测白名单包括：CM数据目录、DN数据目录、日志目录所在磁盘以及自定义表空间目录所在磁盘。

取值范围：整型，[0, 1073741824]，单位：B。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：512

版权所有 © 华为云计算技术有限公司

enable_e2e_rto

参数说明：端到端RTO开关，开启后僵死检测周期及网络检测超时时间将缩短，CM可以达到端到端RTO指标（单实例故障RTO<=10s，叠加故障RTO<=30s）。

取值范围：整型，[0, 1]。1表示开启，0表示关闭。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

默认值：0

process_unhealthy_alarm_time

参数说明：进程状态检测超时告警参数，当进程长时间处于某种异常状态(D/T/Z)时，超过此参数时间则触发进程异常告警。

取值范围：整型，[0, 2147483647]，单位：秒。0表示关闭此功能。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：50

enable_tblspc_disk_check

参数说明：是否开启表空间磁盘检测开关。

取值范围：布尔型。修改后需重启集群生效，参数修改请参见**set cm参数**进行设置。

- on、yes、true、1：表示开启表空间磁盘检测，即表空间所在磁盘hang，中止对应DN。
- off、no、false、0：表示不开启表空间磁盘检测。

默认值：on

max_network_package_loss

参数说明：检测DN主备之间丢包率，DN主丢包率超过阈值，触发重启操作

取值范围：整型，[0, 99]，单位：%。0表示关闭此功能。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

默认值：30

enable_auto_start

参数说明：控制cm_agent是否自动拉起DN。

取值范围：整型，[0, 1]。1表示开启自动拉起DN，0表示关闭自动拉起DN。参数修改请参见**set cm参数**进行设置。

默认值：1

enable_check_az_network

参数说明：是否开启az级数据面网络检测。

取值范围：布尔型。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

版权所有 © 华为云计算技术有限公司

- on、yes、true、1：表示开启az级数据面网络检测。
- off、no、false、0：表示关闭az级数据面网络检测。

默认值：on

expect_stop_az_name

参数说明：az级网络检测功能，期望停止的az名字。

取值范围：string型。取值为空时，与cmserver主相反的az上的实例停止。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：空

check_az_network_concurrency

参数说明：az级网络检测支持并发数。

取值范围：整型，[1,10]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：2

cm_hadr_process_type

参数说明：容灾流程标识。

取值范围：枚举。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

- none：表示当前无流程。
- failover：表示当前处于灾备集群升主流程。
- switchover_promote：表示主备集群倒换流程中灾备集群升主流程。
- switchover_demote：表示主备集群倒换流程中主集群降为灾备集群流程。

默认值：none

topology_network_collecting_failed_times

参数说明：数据面网络拓扑连续断连次数，断连失败次数超过此参数，则判断为数据面网络拓扑异常。需要与cms的topology_network_collecting_failed_times值保持一致。

取值范围：整型，[3，50]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：5

topology_network_collecting_interval

参数说明：数据面网络拓扑检测周期。配置成0则关闭网络拓扑检测。

取值范围：整型，[0，600]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：2

版权所有 © 华为云计算技术有限公司

6.5.2 cm_server 参数

log_dir

参数说明：log_dir为存放cm_server日志文件的目录。 它必须是绝对路径。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --server -k "log_dir='/log/dir'"。

取值范围：字符串，最大长度为1024。修改后需要重启cm_server才能生效。参数修改请参考 **set cm参数** 进行设置。

默认值：'$GAUSSLOG/cm/cm_server'，表示在$GAUSSLOG下对应的cm目录下生成cm_server日志。

log_file_size

参数说明：控制日志文件的大小。当'cm_server-xx-current.log'日志文件达到指定大小时，则重新创建一个日志文件记录日志信息。

取值范围：字符串，[0MB, 2047MB]，实际生效范围[1MB, 2047MB]，单位：MB。可不带单位设置，默认视为MB。如果需要带单位，仅能是"MB"。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：16

log_min_messages

参数说明：控制写到cm_server日志文件中的消息级别。每个级别都包含排在它后面的所有级别中的信息。级别越低，服务器运行日志中记录的消息就越少。

取值范围：枚举类型，有效值有debug5、debug1、log、warning、error、fatal（不区分大小写）。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：warning

thread_count

参数说明：agent线程池的线程数。

取值范围：字符串，长度大于0。修改后需要重启cm_server才能生效。参数修改请参考**set cm参数**进行设置。

该参数分为2个部分，'worker_thread_count, io_ratio'，这2个部分的具体含义如下：

- worker_thread_count：AgentWorker的线程数量，取值范围2~1000。
- io_ratio：io和worker线程的比例，值为n代表一个AgentIO线程对应n个AgentWorker线程，取值范围1~100。AgentIO线程最少一个。

默认值：(1000,1)

instance_heartbeat_timeout

参数说明：实例心跳超时时间。超时后，则表示该实例处于托管状态，实例状态显示unknown。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

版权所有 © 华为云计算技术有限公司

默认值：6

instance_failover_delay_timeout

参数说明：cm_server检测到主机宕机，延迟仲裁出新主机。超时后，cm_server会重新仲裁出实例主机。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：0

cmserver_ha_connect_timeout

参数说明：cm_server主备连接超时时间。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：2

cmserver_ha_heartbeat_timeout

参数说明：cm_server主备心跳超时时间。超时后，会触发cm_server重新自选主。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：6

cmserver_ha_status_interval

参数说明：cm_server主备同步状态信息间隔时间。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：1

cmserver_self_vote_timeout

参数说明：cm_server之间相互投票的超时时间。旧版本遗留参数，实际不生效。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：6

phony_dead_effective_time

参数说明：用于DN进程的僵死检测，当检测到的僵死次数大于该参数值，认为进程僵死，将进程重启。

取值范围：整型，[1, 2147483647]，单位：次。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：5

版权所有 © 华为云计算技术有限公司

cm_server_arbitrate_delay_base_time_out

参数说明：cm_server仲裁延迟基础时长。cm_server主断连后，仲裁启动计时开始，经过仲裁延迟时长后，将选出新的cm_server主。其中仲裁延迟时长由仲裁延迟基础时长、节点index（server ID序号）和增量时长共同决定。公式为：仲裁延迟时长=仲裁延迟基础时长+节点index*仲裁延迟增量时长参数。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：10

cm_server_arbitrate_delay_incrememtal_time_out

参数说明：cm_server仲裁延迟增量时长。cm_server主断连后，仲裁启动计时开始，经过仲裁延迟时长后，将选出新的cm_server主。其中仲裁延迟时长由仲裁延迟基础时长、节点index（server ID序号）和增量时长共同决定。公式为：仲裁延迟时长=仲裁延迟基础时长+节点index*仲裁延迟增量时长参数。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：3

alarm_component

参数说明：在使用第一种告警方式时，设置用于处理告警内容的告警组件的位置。参数修改请参考 **set cm参数** 进行设置。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --server -k "alarm_component='/alarm/dir'"。

取值范围：字符串，最大长度为1024。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：/opt/huawei/snas/bin/snas_cm_cmd

alarm_report_interval

参数说明：指定告警上报的时间间隔。

取值范围：非负整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：3

alarm_report_max_count

参数说明：指定告警上报的最大次数。

取值范围：非负整型，[1, 2592000]。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：1

版权所有 © 华为云计算技术有限公司

instance_keep_heartbeat_timeout

参数说明：cm_agent会定期检测实例状态并上报给cm_server，若实例状态长时间无法成功检测，累积次数超出该数值，则cm_server将下发命令给agent重启该实例。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：40

az_switchover_threshold（废弃）

参数说明：若一个AZ内DN分片的故障率（故障的DN分片数 / 总DN分片数 * 100%）超过该数值，则会触发AZ自动切换。在该版本中已废弃。

取值范围：整型，[1, 100]。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：100

az_check_and_arbitrate_interval（废弃）

参数说明：当某个AZ状态不正常时，会触发AZ自动切换，该参数是检测AZ状态的时间间隔。在该版本中已废弃。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：2

az_connect_check_interval（废弃）

参数说明：定时检测AZ间的网络连接，该参数表示连续两次检测之间的间隔时间。在该版本中已废弃。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：60

az_connect_check_delay_time（废弃）

参数说明：每次检测AZ间的网络连接时有多次重试，该参数表示两次重试之间的延迟时间。在该版本中已废弃。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：150

cmserver_demote_delay_on_etcd_fault

参数说明：因为etcd不健康而导致cm_server从主降为备的时间间隔。

取值范围：整型，[1, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：8

版权所有 © 华为云计算技术有限公司

instance_phony_dead_restart_interval

参数说明：当DN实例僵死时，会被cm_agent重启，相同的实例连续因僵死被杀时，其间隔时间不能小于该参数数值，否则cm_agent不会下发命令。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：21600

enable_transaction_read_only

参数说明：控制数据库是否为只读模式开关。

取值范围：布尔型，有效值有on、off、true、false、yes、no、1、0。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：on

datastorage_threshold_check_interval

参数说明：检测磁盘占用的时间间隔。间隔时间由用户指定，表示检测一次磁盘的间隔时间。

取值范围：整型，[1, 2592000]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：10

datastorage_threshold_value_check

参数说明：设置数据库只读模式的磁盘占用阈值，当数据目录所在磁盘占用超过这个阈值，自动将数据库设置为只读模式。如果磁盘占用继续增长并超过disk_full_threshold， 数据库会停止，详见**disk_full_threshold**参数。数据库只读后无法限制日志回放，应当关注集群磁盘容量，及时处理数据库只读相关告警，防止出现磁盘空间写满、出现无法快速恢复的场景。

取值范围：整型，[1, 99]，单位： %。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：85

datastorage_threshold_recover

参数说明：数据库取消只读模式的缓冲阈值。数据目录所在磁盘占用超过datastorage_threshold_value_check后，被设置为只读模式。当磁盘占用恢复到小于(datastorage_threshold_value_check - datastorage_threshold_recover后)，才会取消只读模式。datastorage_threshold_recover需要小于datastorage_threshold_value_check，否则会被视为0。

取值范围：整型，[0, 99]，单位： %。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：5

 版权所有 © 华为云计算技术有限公司

max_datastorage_threshold_check

参数说明：设置磁盘使用率的最大检测间隔时间。当用户手动修改只读模式参数后，会自动在指定间隔时间后开启磁盘检测操作。

取值范围：整型，[1, 2592000]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：43200

enable_az_auto_switchover

参数说明：AZ自动切换开关，若打开，则表示允许cm_server自动切换AZ。否则当发生dn故障等情况时，即使当前AZ已经不再可用，也不会自动切换到其他AZ上，除非手动执行切换命令。

取值范围：非负整型，0或1，0：开关关闭，1：开关打开。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：1

cm_auth_method

参数说明：CM模块端口认证方式，trust表示未配置端口认证，gss表示采用kerberos端口认证。必须注意的是：只有当kerberos服务端和客户端成功安装后才能修改为gss，否则CM模块无法正常通信，将影响数据库实例状态。

取值范围：枚举类型，有效值有trust、gss。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：trust

cm_krb_server_keyfile

参数说明：kerberos服务端key文件所在位置，需要配置为绝对路径。该文件通常为${GAUSSHOME}/kerberos路径下，以keytab格式结尾，文件名与数据库实例运行所在用户名相同。与上述cm_auth_method参数是配对的，当cm_auth_method参数修改为gss时，该参数也必须配置为正确路径，否则将影响数据库实例状态。通过cm_ctl设置绝对路径时需要将路径用''把路径包含起来，例如：cm_ctl set --param --server -k "cm_krb_server_keyfile='/krb/dir'"。

取值范围：字符串类型，修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：${GAUSSHOME}/kerberos/{UserName}.keytab，默认值无法生效，仅作为提示。

switch_rto

参数说明：cm_server强起逻辑等待时延。在force_promote被置为1时，当数据库实例的某一分片处于无主状态开始计时，等待该延迟时间后开始执行强起逻辑。

取值范围：整型，[60, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：600

版权所有 © 华为云计算技术有限公司

force_promote

参数说明：cm_server是否打开强起逻辑（指数据库实例状态为Unknown的时候以丢失部分数据为代价保证数据库实例基本功能可用）的开关。0代表功能关闭，1代表功能开启。

取值范围：整型，[0, 1]。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：0

enable_finishredo_retrieve

参数说明：cm_server强切后是否对redo切除的xLog进行数据找回的功能开关。置为on时，发生强切后进行数据自动找回。

取值范围：布尔型。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

- on、yes、true、1：表示开启。
- off、no、false、0：表示关闭。

默认值：off

backup_open

参数说明：灾备数据库实例设置，开启后CM按照灾备数据库实例模式运行。此参数为内部调用参数，不建议用户修改。

取值范围：整型，[0, 3]。修改后需要重启cm_server才能生效。非灾备数据库实例不能开启该参数。参数修改请参考 **set cm参数** 进行设置。

- 0：表示未搭建灾备关系。
- 1：表示搭建了dorado灾备关系。
- 2：表示搭建了流式灾备关系。
- 3：表示搭建了逻辑灾备关系。

默认值：0

enable_dcf

参数说明：DCF模式开关。

取值范围：布尔型。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。请根据数据库实例实际的DCF运行模式进行更改，更改时确保数据库实例状态Normal，否则可能导致数据库实例不可用。

- on、yes、true、1：表示启用dcf。
- off、no、false、0：表示不启用dcf。

默认值：off

ddb_type

参数说明：etcd，dcc模式切换开关。DCC（Distributed Configuration Center）：分布式配置中心，基于DCF（基于Paxos协议的高可用）实现的一个状态机，用于实现集群中配置信息的管理。ETCD：高度一致的分布式键值存储组件。

版权所有 © 华为云计算技术有限公司

取值范围：整型。修改后需要重启cm_server才能生效。参数修改请参考**set cm参数**进行设置。

- 0：表示模式为etcd。
- 1：表示模式为dcc。

默认值：0

enable_ssl

参数说明：ssl证书开关。

取值范围：布尔型。打开后使用ssl证书加密通信。修改后需要重启cm_server才能生效。参数修改请参考 **set cm参数** 进行设置。

默认值：

- on、yes、true、1：表示启用ssl。
- off、no、false、0：表示不启用ssl。
- **默认值**：off

> **须知**
>
> 出于安全性考虑，建议不要关闭该配置。关闭后cm将不使用加密通信，所有信息明文传播，可能带来窃听、篡改、冒充等安全风险。

ssl_cert_expire_alert_threshold

参数说明：ssl证书过期告警时间。

取值范围：整型，[7, 180]，单位：天。证书过期时间少于该时间时，上报证书即将过期告警。修改后需要重启cm_server才能生效，参数修改请参考 **set cm参数** 进行设置。

默认值：90

ssl_cert_expire_check_interval

参数说明：ssl证书过期检测周期。

取值范围：整型，[0, 2147483647]，单位：秒。修改后需要重启cm_server才能生效，参数修改请参考 **set cm参数** 进行设置。

默认值：86400

ddb_log_level

参数说明：设置ddb日志级别。

关闭日志：“NONE”，NONE表示关闭日志打印，不能与以下日志级别混合使用。

开启日志：“RUN_ERR|RUN_WAR|RUN_INF|DEBUG_ERR|DEBUG_WAR|DEBUG_INF|TRACE|PROFILE|OPER”日志级别可以从上述字符串中选取字符串并使用竖线组合使用，不能配置空串。

版权所有 © 华为云计算技术有限公司

取值范围：字符串，RUN_ERR|RUN_WAR|RUN_INF|DEBUG_ERR|DEBUG_WAR|DEBUG_INF|TRACE|PROFILE|OPER。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：RUN_ERR|RUN_WAR|DEBUG_ERR|OPER|RUN_INF|PROFILE

ddb_log_backup_file_count

参数说明：最大保存日志文件个数。

取值范围：整型，[1, 100]。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：10

ddb_max_log_file_size

参数说明：单条日志最大字节数。

取值范围：字符串，长度最大为1024，[1M, 1000M]。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：10

ddb_log_suppress_enable

参数说明：是否开启日志抑制功能。

取值范围：整型，0：关闭； 1：开启。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：0

ddb_election_timeout

参数说明：dcc 选举超时时间。

取值范围：整型，[1, 600]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：3

ddb_dcf_sync_nodes

参数说明：设置dcc同步复制的备机列表。

取值范围：字符串。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。支持按如下格式配置：

ANY *num_sync:node_id*[, ...]，其中*num_sync*:是ddb需要同步复制的备机数量，node_id是备机的序号。格式不满足要求，参数不会生效。

示例：ddb_dcf_sync_nodes='ANY 1:2,3'

默认值：''

 版权所有 © 华为云计算技术有限公司

ddb_dcf_extreme_mode

参数说明：设置dcc的日志提交策略。

取值范围：整型，[0, 2]。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

- 0：表示最大保护，多数派提交。
- 1：表示最大可用，备机故障时主机可直接提交，备机正常时多数派提交。
- 2：表示最大性能，本地提交。

默认值：0

delay_arbitrate_timeout

参数说明：设置等待跟主DN同AZ节点redo回放后升主的时间。

取值范围：整型，[0, 2147483647]，单位：秒。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

默认值：0

dn_arbitrate_mode

参数说明：DN仲裁模式。

取值范围：字符串。修改后可以reload生效。参数修改请参考 **set cm参数**进行设置。

- quorum
- paxos
- share_disk

说明

- dn_arbitrate_mode设置为share_disk后，不建议用户再对DN仲裁模式进行修改。
- DCF模式下，cm_ctl设置的DCF工作模式需要和GUC参数dcf_run_mode设置保持一致，即两者需要同步设置为DCF手动或自动模式。

默认值：quorum

cluster_starting_aribt_delay

参数说明：cm_server在数据库实例启动阶段，等待DN静态主升主的时间。

取值范围：整型，[1,2592000]，单位：秒。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

默认值：180

cma_fault_kill_instance_timeout

参数说明：DN实例无法上报，cmserver主内部状态置为unknown 后， 等待超时，cmserver 通过 pssh 命令远程发送kill操作，重启不上报DN实例进程, 且之后下发的lock1命令换成force lock1命令。

取值范围：整型，[0, 2147483647]，单位：秒。0 表示关闭。修改后可以reload生效，参数修改请参考 **set cm参数** 进行设置。

版权所有 © 华为云计算技术有限公司

说明：此参数默认不添加到cmserver配置文件中，需要手动添加。

注意：不配置此参数，默认开启。

默认值为：（cm_server_num * **agent_connect_timeout** * **agent_connect_retries** + **agent_heartbeat_timeout** + 10）

thread_effective_time

参数说明：线程运行一轮的有效超时时间。超时说明此线程卡住，进程自动退出。

取值范围：整型，[0, 1200]，单位：秒。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

- 0：表示关闭线程监控能力。

注意：不建议调整参数值大小，如果调整请联系华为技术支持。

默认值：20

enable_synclist_single_inst

参数说明：开启降副本降至一主零备功能。

取值范围：布尔型，开启后降副本会降至一主零备。参数设置错误，按默认值处理。修改后可以reload生效。参数修改请参考**set cm参数**进行设置。

- off：表示关闭降副本降至一主零备功能。
- on：表示开启降副本降至一主零备功能。

默认值：off

危险

打开此参数(on)， 在所有备机故障情况下，主机会自动转为本地提交方式，此时如果主机再故障，则可能产生不可恢复的影响。打开前请联系华为技术支持。

ctl_thread_count

参数说明：ctl线程池的线程数。

取值范围：字符串，长度大于0。修改后重启cm_server生效，参数修改请参考**set cm参数**进行设置。

该参数分为2个部分，'worker_thread_count, io_ratio'，这2个部分的具体含义如下：

- worker_thread_count：CtlWorker的线程数量，取值范围0~1000。取值为0时，CtlWorker的线程数量等于集群内节点数。
- io_ratio：io和worker线程的比例，值为n代表一个CtlIO线程对应n个CtlWorker线程，取值范围1~100。CtlIO线程最少一个。

默认值：（0,1）

asyn_log_enable

参数说明：异步日志开关，打开后cm_server启动异步日志。

版权所有 © 华为云计算技术有限公司

取值范围：布尔型。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

- on、yes、true、1：表示cm_server开启异步日志。
- off、no、false、0：表示cm_server不开启异步日志，使用串行日志。

默认值：on

asyn_log_max_buf_size

参数说明：异步日志缓存大小，当异步日志发生堆积，且积压的日志超过了缓存大小，后续的日志无法记录。

取值范围：整型，[1kB, 1048576kB]，单位：kB。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

默认值：1024

asyn_log_fault_solution

参数说明：识别到日志磁盘故障后，异步日志处理方案。当CM识别到日志盘故障或异步日志缓存溢出，CM会根据配置来处理异步日志。

取值范围：整型，[0, 2]。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

- 0：表示不对日志做多余的处理，缓存区保留所有日志，直到缓存大小达到上限，达到上限后不会记录更多日志。
- 1：表示对日志进行筛选，识别到日志磁盘故障后，只将关键日志、告警和ERROR级别以上的日志保存到缓存区，直到缓存大小达到上限，达到上限后不会记录更多日志。
- 2：表示丢弃日志，识别到日志磁盘故障后，只将关键日志、告警保存到缓存区，直到缓存大小达到上限，达到上限后不会记录更多日志。

默认值：0

enable_isolate

参数说明：隔离实例开关。

取值范围：整型，[0, 1]。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

- 1：表示开启。开启后，cm_server会自动隔离故障实例。
- 0：表示关闭。

默认值：1

enable_e2e_rto

参数说明：端到端RTO开关，开启后僵死检测周期及网络检测超时时间将缩短，CM可以达到端到端RTO指标（单实例故障RTO<=10s，叠加故障RTO<=30s）。

取值范围：整型，[0, 1]。1表示开启，0表示关闭。修改后可以reload生效，参数修改请参考**set cm参数**进行设置。

默认值：0

版权所有 © 华为云计算技术有限公司

alarm_ddb_unhealth_threshold

参数说明：ddb不健康告警上报阈值。

取值范围：整型，[0, 2592000]。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

默认值：5

disk_full_threshold

参数说明：在参数**enable_disk_full_check**开启下，设置数据库磁盘满的磁盘占用阈值，当数据目录所在磁盘占用超过这个阈值，会自动将数据库停止。当磁盘占用恢复到**datastorage_threshold_value_check**以下时，自动将数据库启动。
disk_full_threshold需要大于datastorage_threshold_value_check才能生效。

取值范围：整型，[1, 99]，单位：%。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

默认值：98

cm_hadr_process_type

参数说明：容灾流程标识。

取值范围：枚举。修改后可以reload生效，参数修改请参见**set cm参数**进行设置。

- none：表示当前无流程。
- failover：表示当前处于灾备集群升主流程。
- switchover_promote：表示主备集群倒换流程中灾备集群升主流程。
- switchover_demote：表示主备集群倒换流程中主集群降为灾备集群流程。

默认值：none

topology_network_arbitrate_wait_time

参数说明：网络拓扑仲裁等待网络拓扑数据稳定的时间。如果在该参数时间内，网络拓扑数据发生变化，则不会进程网络拓扑仲裁。

取值范围：整型，[5，600]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：10

topology_network_collecting_failed_times

参数说明：数据面网络拓扑连续断连次数，断连失败次数超过此参数，则判断为数据面网络拓扑异常。需要和cma的topology_network_collecting_failed_times值保持一致。

取值范围：整型，[3，50]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

默认值：5

版权所有 © 华为云计算技术有限公司

enable_topology_network_arbitrate

参数说明：网络拓扑仲裁开关。

取值范围：整型，[0，1]。修改后可以reload生效，参数修改请参见**设置CM参数**和**动态加载CM参数**进行设置。

- 1：表示开启。开启后，cm_server会自动进程网络拓扑仲裁。
- 0：表示关闭。

默认值：1

enable_disk_full_check

参数说明：磁盘 disk full（磁盘使用率超过**disk_full_threshold**，会停止故障节点的实例进程）检测开关，打开后，允许磁盘disk full check。

取值范围：布尔型。修改后可以reload生效。参数修改请参见**set cm参数**进行设置。

- on、yes、true、1：表示开启磁盘disk full check。
- off、no、false、0：表示关闭磁盘disk full check。

默认值：on

6.6 CM 错误日志信息参考

ERRMSG: "Fail to access the cluster static config file."

CMSTATE: c3000

CAUSE: "The cluster static config file is not generated or is manually deleted."

ACTION: "Please check the cluster static config file."

ERRMSG: "Fail to open the cluster static file."

CMSTATE: c3000

CAUSE: "The cluster static config file is not generated or is manually deleted."

ACTION: "Please check the cluster static config file."

ERRMSG: "Fail to read the cluster static file."

CMSTATE: c3001

CAUSE: "The cluster static file permission is insufficient."

ACTION: "Please check the cluster static config file."

ERRMSG: "Failed to read the static config file."

CMSTATE: c1000

CAUSE: "out of memeory."

版权所有 © 华为云计算技术有限公司

ACTION: "Please check the system memory and try again."

ERRMSG: "Could not find the current node in the cluster by the node id %u."
CMSTATE: c3002
CAUSE: "The static config file probably contained content error."
ACTION: "Please check static config file."

ERRMSG: "Failed to open the logic config file."
CMSTATE: c3000
CAUSE: "The logic config file is not generated or is manually deleted."
ACTION: "Please check the cluster static config file."

ERRMSG: "Fail to read the logic static config file."
CMSTATE: c3001
CAUSE: "The logic static config file permission is insufficient."
ACTION: "Please check the logic static config file."

ERRMSG: "Failed to open or read the static config file."
CMSTATE: c1000
CAUSE: "out of memeory."
ACTION: "Please check the system memory and try again."

ERRMSG: "Failed to open the log file '%s'."
CMSTATE: c3000
CAUSE: "Log file not found."
ACTION: "Please check the log file."

ERRMSG: "Failed to open the log file '%s'."
CMSTATE: c3000
CAUSE: "The log file permission is insufficient."
ACTION: "please check the log file."

ERRMSG: "Failed to open the dynamic config file '%s'."
CMSTATE: c3000
CAUSE: "The dynamic config file permission is insufficient."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the dynamic config file."

ERRMSG: "Failed to malloc memory, size = %lu."
CMSTATE: c1000
CAUSE: "out of memeory."
ACTION: "Please check the system memory and try again."

ERRMSG: "unrecognized AZ name '%s'."
CMSTATE: c3000
CAUSE: "The parameter(%s) entered by the user is incorrect."
ACTION: "Please check the parameter entered by the user and try again."

ERRMSG: "unrecognized minorityAz name '%s'."
CMSTATE: c3000
CAUSE: "The parameter(%s) entered by the user is incorrect."
ACTION: "Please check the parameter entered by the user and try again."

ERRMSG: "Get GAUSSHOME failed."
CMSTATE: c3000
CAUSE: "The environment variable('GAUSSHOME') is incorrectly configured."
ACTION: "Please check the environment variable('GAUSSHOME')."

ERRMSG: "Get current user name failed."
CMSTATE: c3000
CAUSE: "N/A"
ACTION: "Please check the environment."

ERRMSG: "-B option must be specified."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-T option must be specified.\n"
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "can't stop one node or instance with -m normal."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "can't stop one node or instance with -m resume."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "can't stop one availability zone with -m resume."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "log level or cm server arbitration mode must be specified."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "log level or cm server arbitration mode need not be specified."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-R is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-D is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n and -R are needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n and -D are needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "no operation specified."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "no cm directory specified."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "Please check the usage of switchover."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n and -z cannot be specified at the same time."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-m cannot be specified at the same time with -n or -z."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n node(%d) is invalid."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n node is needed."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "%s: -C is needed."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-z value must be 'ALL' when query mppdb cluster."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-v is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-C is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-Cv is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-L value must be 'ALL' when query logic cluster."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "unrecognized LC name '%s'."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n is needed."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "There is no '%s' information in cluster."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-D path is too long.\n"
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-D path is invalid."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-n node(%s) is invalid."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-R only support when the cluster is single-inst."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-t time is invalid."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "-votenum is invalid."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "unrecognized build mode."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "unrecognized build mode '%s'."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "too many command-line arguments (first is '%s')."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "unrecognized operation mode '%s'."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "no cm directory specified."
CMSTATE: c3000
CAUSE: "%s: The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "Failed to malloc memory."
CMSTATE: c1000
CAUSE: "out of memeory."
ACTION: "Please check the system memory and try again."

ERRMSG: "Failed to open etcd: %s."
CMSTATE: c4000
CAUSE: "Etcd is abnoraml."
ACTION: "Please check the Cluster Status and try again."

ERRMSG: "[PATCH-ERROR] hotpatch command or path set error."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "no standby datanode in single node cluster."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "restart logic cluster failed."
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."
ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "restart logic cluster failed"
CMSTATE: c3000
CAUSE: "The cmdline entered by the user is incorrect."

 版权所有 © 华为云计算技术有限公司

ACTION: "Please check the cmdline entered by the user(%s)."

ERRMSG: "The option parameter is not specified."

CMSTATE: c3000

CAUSE: "The cmdline entered by the user is incorrect."

ACTION: "Please check the cmdline entered by the user(%s)."

 版权所有 © 华为云计算技术有限公司

7 FAQ

7.1 执行命令报错“Failed to obtain the GPHOME”

问题现象

如果在root用户下执行工具命令时，出现如下信息：

```
Failed to obtain the GPHOME.
```

原因分析

可能是环境变量中GPHOME变量没有配置正确。需要检查环境变量GPHOME是否包含数据库XML中gaussdbToolPath路径。

操作步骤

使用如下命令检查$GPHOME路径。

```
echo $GPHOME
```

如果不是安装时默认的路径，请在配置文件中修改GPHOME的路径。

```
vim /etc/profile
```

 版权所有 © 华为云计算技术有限公司

8 gs_collector 工具支持收集的系统表和视图列表

表 8-1 gs_collector 工具支持收集的系统表和视图列表

Class	View Name
OS	pv_os_run_info
	pg_os_threads
Session /Thread	pv_session_stat
	pv_session_time
	dv_sessions
	dv_session_longops
	gs_session_cpu_statistics
	gs_session_memory_statistics
	pv_session_memory_context
	pv_session_memory
	pv_session_memory_detail
	pg_stat_activity
	pv_thread_memory_context
WLM	gs_all_control_group_info
	gs_get_control_group_info
	gs_cluster_resource_info
	pg_total_user_resource_info
	pg_total_user_resource_info_oid
	gs_wlm_cgroup_info

版权所有 © 华为云计算技术有限公司

Class	View Name
	gs_wlm_rebuild_user_resource_pool
	pg_resource_pool
	gs_wlm_resource_pool
	gs_wlm_user_info
	gs_wlm_workload_records
	pg_wlm_statistics
	pg_app_workloadgroup_mapping
Query-AP Operator	gs_wlm_operator_info
	gs_wlm_operator_history
	gs_wlm_operator_statistics
Query-AP	gs_wlm_session_info_all
	gs_wlm_session_history
	gs_wlm_session_info
	gs_wlm_user_session_info
	gs_wlm_session_statistics
	pg_session_iostat
	pg_session_wlmstat
Cache I/O	pg_statio_all_indexes
	pg_statio_all_sequences
	pg_statio_all_tables
	pg_statio_sys_indexes
	pg_statio_sys_sequences
	pg_statio_sys_tables
	pg_statio_user_indexes
	pg_statio_user_sequences
	pg_statio_user_tables
	gs_stat_db_cu
	gs_stat_session_cu
Memory	pg_total_memory_detail
	pg_shared_memory_detail

版权所有 © 华为云计算技术有限公司

Class	View Name
	pv_total_memory_detail
File	pv_file_stat
	pv_redo_stat
	pg_stat_bad_block
Object	pg_stat_user_tables
	pg_stat_user_indexes
	pg_stat_sys_tables
	pg_stat_sys_indexes
	pg_stat_all_tables
	pg_stat_all_indexes
	pg_stat_database
	pg_stat_database_conflicts
	pg_stat_xact_all_tables
	pg_stat_xact_sys_tables
	pg_stat_xact_user_tables
	pg_stat_xact_user_functions
Lock	pg_locks
Utility	pg_stat_replication
	pg_replication_slots
	pg_stat_bgwriter
	pg_pooler_status
Transaction	pg_running_xacts
	pg_prepared_xacts
Comm	pg_comm_delay
	pg_comm_recv_stream
	pg_comm_send_stream
	pg_comm_status
Waits	pg_thread_wait_status

版权所有 © 华为云计算技术有限公司