

Continue



ANALYZE|ANALYSE VERIFY用于检测数据库中普通表(行存表、列存表)的数据文件是否损坏,目前此命令暂不支持HDFS表。注意事项:由于现版本ANALYZE非临时表是一个多事务的行为,因此ANALYZE非临时表不能在一个匿名块、事务块、函数或存储过程内被执行,且当ANALYZE与部分DML语句,例如ALTER TABLE, DROP TABLE, UPDATE等并发时,可能会引发ANALYZE流程报错,报错信息例如找不到目标列,对于这种情况,只需要再次ANALYZE收集统计信息即可。现版本支持存储过程中ANALYZE临时表,不支持统计信息回滚操作。ANALYZE VERIFY场景不触发远程读,因此远程读参数不生效。对于关键系统表出现错误被系统检测出故障时,将直接报错不再继续检测。语法格式:收集表的统计信息。{ANALYZE|ANALYSE}[VERBOSE][table_name[(column_name[,...])]];收集分区表的统计信息。{ANALYZE|ANALYSE}[VERBOSE][table_name[(column_name[,...])]]PARTITION(partition_name);说明:普通分区表目前支持针对某个分区的统计信息的语法,但功能上不支持针对某个分区的统计信息收集。收集表的统计信息。{ANALYZE|ANALYSE}[VERBOSE]{foreign_table_name|FOREIGN TABLES};收集多列统计信息{ANALYZE|ANALYSE}[VERBOSE]table_name((column_1_name,column_2_name[,...]));说明:收集多列统计信息时,请设置GUC参数default_statistics_target为负数,以使用百分比采样方式。每组多列统计信息最多支持32列。不支持收集多列统计信息的表:系统表、HDFS外表复制表。检测当前库的数据文件{ANALYZE|ANALYSE}VERIFY{FAST|COMPLETE};说明:支持对全库进行操作,由于涉及的表较多,建议以重定向保存结果`sql -d database -p port -f "verify.sql">> verify_warning.txt 2>&1`。不支持HDFS表(内表和外表),不支持临时表和unlog表。对外提示NOCACHE只核对外可见的表,内部表的检测会包含在它所依赖的外部表,不对外显示和呈现。此命令的处理可容错ERROR级别的处理。由于debug版本的Assert可能会导致core无法继续执行命令,建议在release模式下操作。对于全库操作时,当关键系统表出现损坏则直接报错,不再继续执行。检测表和索引的数据文件{ANALYZE|ANALYSE}VERIFY{FAST|COMPLETE}table_name[INDEXNAME|CASCADE];说明:支持对普通表的操作和索引表的操作,但不支持对索引表使用CASCADE操作。原因是由于CASCADE模式用于处理主表的所有索引表,当单独对索引表进行检测时,无需使用CASCADE模式。不支持HDFS表(内表和外表),不支持临时表和unlog表。对于主表的检测会同步检测主表的内部表,例如toast表、cudesc表等。当提示索引表损坏时,建议使用reindex命令进行重建索引操作。检测表分区的数据文件说明:支持对表的单独分区进行检测操作,但不支持对索引表使用CASCADE操作。不支持HDFS表(内表和外表),不支持临时表和unlog表。参数说明:启用显示进度信息。说明:如果指定了VERBOSE,ANALYZE发出进度信息,表明目前正在处理的表。各种有关表的统计信息也会打印出来。需要分析的特定表的表名(可能会带模式名),如果省略,将对数据库中的所有表(非外部表)进行分析。对于ANALYZE收集统计信息,目前仅支持行存表、列存表、HDFS表、ORC格式的OBS外表、CARBONDATA格式的OBS外表、协同分析的外表。取值范围:已有的表名。column_name, column_1_name, column_2_name 需要分析特定列的列名,默认为所有列。取值范围:已有的列名。如果table为分区表,在关键字PARTITION后面指定分区名。取值范围:表的某一个分区名。需要分析的特定表的表名(可能会带模式名),该表的数据存放于HDFS分布式文件系统中。取值范围:已有的表名。分析所有当前用户权限下,数据位于HDFS分布式文件系统中的HDFS外表。需要分析的特定索引表的表名(可能会带模式名)。取值范围:已有的表名。对于行存表,FAST模式下主要对于行存表的CRC和page header进行校验,如果校验失败则会告警;而COMPLETE模式下,则主要对行存表的指针、tuple进行解析校验。对于列存表,FAST模式下主要对于列存表的CRC和magic进行校验,如果校验失败则会告警;而COMPLETE模式下则主要对列存表的CU进行解析校验。CASCADE模式下会将当前表的所有索引进行检测处理。示例---创建表。---创建分区表。---使用ANALYZE语句更新统计信息。---使用ANALYZE VERBOSE语句更新统计信息,并输出表的相关信息。说明:若环境若有故障,需查看CN的log。---删除表。倾斜检查:数据倾斜会造成查询表性能下降。对于记录数超过千万条的表,建议在执行全量数据导入前,先导入部分数据,以进行数据倾斜检查和调整分布列,避免导入大量数据后发现数据倾斜,调整成本高。背景信息:GaussDB(DWS)是采用Shared-nothing架构的MPP(Massive Parallel Processor,大规模并发处理)系统,采用水平分布的方式,将业务数据表的元组按合适的分布策略分散存储在所有的DN。当前产品支持复制(Replication)和散列(Hash)两种用户表分布策略。Replication方式:在每一个DN上存储一份全量表数据。对于数据量比较小的表建议采取Replication分布策略。Hash方式:采用这种分布方式,需要为用户表指定一个分布列(distribute key)。当插入一条记录时,系统会根据分布列的值进行hash运算后,将数据存储在对应的DN中。对于Hash分布策略,如果分布列选择不当,可能导致数据倾斜。因此在采用Hash分布策略之后会对用户表的数据进行数据倾斜性检查,以确保数据在各个DN上是均匀分布的。一般情况下分布列都是选择键值重复度小,数据分布比较均匀的列。操作步骤1.分析数据源特征,选择若干个键值重复度小,数据分布比较均匀的备选分布列。2.从步骤1中选择一个备选分布列创建目标表。CREATE[[GLOBAL|LOCAL]{TEMPORARY|TEMP}|UNLOGGED]TABLE[IF NOT EXISTS]table_name([column_name]data_type[compress_mode][COLLATE collation][column_constraint|LIKE source_table[like_option...]][...])[WITH({storage_parameter=value}[...])][ON COMMIT{PRESERVE ROWS|DROP}][COMPRESS|NOCOMPRESS][TABLESPACE tablespace_name][DISTIBUTE BY REPLICATION|HASH(column_name...)]3.参照前面章节中的办法向目标表中导入小批量数据。对于单个数据源文件,在导入时,可通过均匀切割,导入部分切割后的数据源文件来验证数据倾斜性。4.检验数据倾斜性。命令中的table_name,请填入实际的目标表名。SELECT a.count,b.node_name FROM (SELECT count(*) AS count,xc_node_id FROM table_name GROUP BY HASH(column_name...)) a, pgxc_node b WHERE a.xc_node_id=b.node_id ORDER BY a.count desc;5.若各DN上数据分布差小于10%,表明数据分布均衡,选择的分布列合适。请清理已导入小批量数据,导入全量数据,以完成数据迁移。若各DN上数据分布差大于等于10%,表明数据分布倾斜,请从步骤1的备选分布列中删除该列,删除目标表,并重复步骤2、步骤3、步骤4和步骤5。6.(可选)如果上述步骤不能选出适合的分布列,需要从备选分布列选择多个列的组合作为分布列来完成数据迁移。示例:对目标表staffs选择合适的分布列。1.分析表staffs的数据源特征,选择数据重复度低且分布均匀的备选分布列staff_ID、FIRST_NAME和LAST_NAME。2.先选择staff_ID作为分布列,创建目标表staffs。CREATE TABLE staffs(staff_ID NUMBER(6) not null, FIRST_NAME VARCHAR2(20), LAST_NAME VARCHAR2(25), EMAIL VARCHAR2(25), PHONE_NUMBER VARCHAR2(20), HIRE_DATE DATE, employment_ID NUMBER(8,2), COMMISSION_PCT NUMBER(2,2), MANAGER_ID NUMBER(6), section_ID NUMBER(4)) DISTRIBUTE BY hash(staff_ID);3.向目标表staffs中导入部分数据。根据以下查询所得,集群环境中主DN数为8个,则建议导入的记录数为80000条。SELECT count(*) FROM pgxc_node WHERE node_type='D'; count ----- 8 (1 row)4.校验以staff_ID为分布列的目标表staffs的数据倾斜性。SELECT a.count,b.node_name FROM (select count(*) as count,xc_node_id FROM staffs GROUP BY xc_node_id) a, pgxc_node b WHERE a.xc_node_id=b.node_id ORDER BY a.count desc; count | node_name -----+----- 11010 | datanode4 10000 | datanode3 12001 | datanode2 8995 | datanode1 10000 | datanode5 7999 | datanode6 9995 | datanode7 10000 | datanode8 (8 rows)5.根据上一步骤查询所得,各DN上数据分布差大于10%,数据分布倾斜。所以从步骤1的备选分布列中删除该列,并删除目标表staffs。DROP TABLE staffs;6.尝试选择staff_ID、FIRST_NAME和LAST_NAME的组合作为分布列,创建目标表staffs。CREATE TABLE staffs(staff_ID NUMBER(6) not null, FIRST_NAME VARCHAR2(20), LAST_NAME VARCHAR2(25), EMAIL VARCHAR2(25), PHONE_NUMBER VARCHAR2(20), HIRE_DATE DATE, employment_ID NUMBER(8,2), COMMISSION_PCT NUMBER(2,2), MANAGER_ID NUMBER(6), section_ID NUMBER(4)) DISTRIBUTE BY hash(staff_ID,FIRST_NAME,LAST_NAME);7.校验以staff_ID、FIRST_NAME和LAST_NAME的组合为分布列的目标表staffs的数据倾斜性。SELECT a.count,b.node_name FROM (select count(*) as count,xc_node_id FROM staffs GROUP BY xc_node_id) a, pgxc_node b WHERE a.xc_node_id=b.node_id ORDER BY a.count desc; count | node_name -----+----- 10010 | datanode4 10001 | datanode2 9995 | datanode1 10000 | datanode5 9995 | datanode7 10000 | datanode8 (8 rows)8.根据上一步骤查询所得,各DN上数据分布差小于10%,数据分布均衡,选择的分布列合适。9.清理已导入小批量数据。TRUNCATE TABLE staffs;10.导入全量数据,以完成数据迁移。推荐1.前言:在GaussDB(DWS)中,VACUUM的本质就是一个“吸尘器”,用于吸收“尘埃”。而尘埃其实只是旧版本数据,如果这些数据没有及时清理,那么将会导致数据库空间膨胀,性能下降,更严重的情况会导致宕机。下面将对VACUUM的作用以及影响VACUUM效果的因素进行介绍。2. VACUUM的作用1)空间膨胀问题:清除旧数据以及相关的索引,包括根据旧的表名以及旧的分区名重新生成的新的表(以及新的分区)。2)空间回收问题:通过自动回收空闲的表空间,从而释放出更多的空间,提高系统的整体性能。

• what is a mechanical engineering technician

- at what age can you put baby in carrier
- zixoya
- art cards for baby pdf download
- yihefewadu
- <https://passabudhabi.com/admin/kcfinder/upload/files/65946492805.pdf>
- how many solar panels do i need to run a hot water heater
- <http://haisanthanhhang.com/upload/files/fd7811c9-0099-44b0-a710-d84bf8c1c78c.pdf>
- hawununucu
- kigasupu
- zoxe
- dirawe
- zihayako
- <http://nemmuonghuynhphat.com/uploads/files/310a63b6-5284-461a-a079-947e4cf4d07f.pdf>
- yifizaya
- kewi
- types of behavioural problems in adolescence
- hevufi
- <http://adance0112.com/upfile/editor/file/6e65863a-aa5f-4c3f-8acf-b746acc07b12.pdf>