# 概述

# 安装

## 编译安装

1、设置Makefile

cd dbgen

cp makefile.suite Make file

vim Makefile

修改CC、DATABASE、MACHINE、WORKLOAD参数的值：

################

## CHANGE NAME OF ANSI COMPILER HERE

################

CC= gcc

# Current values for DATABASE are: INFORMIX, DB2, ORACLE,

# SQLSERVER, SYBASE, TDAT (Teradata)

# Current values for MACHINE are: ATT, DOS, HP, IBM, ICL, MVS,

# SGI, SUN, U2200, VMS, LINUX, WIN32

# Current values for WORKLOAD are: TPCH

DATABASE= MYSQL

MACHINE = LINUX

WORKLOAD = TPCH

2、修改tpcd.h

vim tpcd.h

添加如下宏定义。

#ifdef MYSQL

#define GEN\_QUERY\_PLAN ""

#define START\_TRAN "START TRANSACTION"

#define END\_TRAN "COMMIT"

#define SET\_OUTPUT ""

#define SET\_ROWCOUNT "limit %d;\n"

#define SET\_DBASE "use %s;\n"

#endif

3、make

4、拷贝文件

mkdir tpch-100g

cd tpch-100g

cp ../dists.dss ./

cp ../dbgen ./

5、生成测试数据

生成100g的测试数据，用30个并发生成30个chunk

date

for i in {1..30}

do

./dbgen -vf -s 100 -S $i -C 30 &

done

wait

./dbgen -vf -s 100 -T r

./dbgen -vf -s 100 -T n

date

6、生成测试sql

生成测试的sql

cp -r ../queries ./

cp ../qgen ./queries

cp ../dists.dss ./queries

cd ./queries

运行下面的脚本

#!/usr/bin/bash

for i in {1..22}

do

./qgen -N -d $i -s 1 > Q"$i".sql

done

7、导入数据（load\_data.sh）

#!/bin/bash

dbname="tpch100g"

dir=`pwd`

opts=" -h -P -utest -ptest123"

echo "Creating tables"

mysql $opts -e "DROP DATABASE IF EXISTS $dbname"

mysql $opts -e "CREATE DATABASE $dbname"

opts="$opts $dbname"

mysql $opts < tpch-create-mysql.sql

echo "Creating indexes"

# mysql $opts < tpch-alter-mysql.sql

# mysql $opts < tpch-create-index-polar.sql

mysql $opts < tpch-create-index-tdsql3.sql

date

for tbl in nation region

do

echo "Importing table: $tbl"

mysql $opts -e "set tdsql\_bulk\_load\_allow\_unsorted=1;set tdsql\_bulk\_load = 1;LOAD DATA INFILE '${dir}/${tbl}.tbl' INTO TABLE $tbl FIELDS TERMINATED BY '|';"

done

for tbl in part customer supplier partsupp orders lineitem

do

for i in {1..30}

do

echo "Importing table: $tbl"

mysql $opts -e "set tdsql\_bulk\_load\_allow\_unsorted=1;set tdsql\_bulk\_load = 1;LOAD DATA INFILE '${dir}/${tbl}.tbl.$i' INTO TABLE $tbl FIELDS TERMINATED BY '|';" &

done

done

wait

date

8、创建视图：

create view revenue (supplier\_no, total\_revenue) as

select

l\_suppkey,

sum(l\_extendedprice \* (1 - l\_discount))

from

lineitem

where

l\_shipdate >= date '1994-08-01'

and l\_shipdate < date '1994-08-01' + interval '3' month

group by

l\_suppkey;

9、执行analyze

ANALYZE table customer;

ANALYZE table nation;

ANALYZE table orders;

ANALYZE table part;

ANALYZE table partsupp;

ANALYZE table region;

ANALYZE table supplier;

ANALYZE table lineitem;

ANALYZE table customer UPDATE HISTOGRAM on C\_CUSTKEY, C\_NATIONKEY, C\_ACCTBAL;

ANALYZE table lineitem UPDATE HISTOGRAM on L\_ORDERKEY, L\_PARTKEY, L\_SUPPKEY, L\_LINENUMBER, L\_QUANTITY, L\_EXTENDEDPRICE, L\_DISCOUNT, L\_TAX, L\_SHIPDATE, L\_COMMITDATE, L\_RECEIPTDATE;

ANALYZE table nation UPDATE HISTOGRAM on N\_NATIONKEY, N\_REGIONKEY, N\_NAME;

ANALYZE table orders UPDATE HISTOGRAM on O\_ORDERKEY, O\_CUSTKEY, O\_TOTALPRICE, O\_ORDERDATE, O\_ORDERPRIORITY, O\_SHIPPRIORITY, o\_orderstatus;

ANALYZE table part UPDATE HISTOGRAM on P\_PARTKEY, P\_TYPE, P\_SIZE, P\_RETAILPRICE;

ANALYZE table partsupp update histogram on PS\_PARTKEY, PS\_SUPPKEY, PS\_AVAILQTY;

ANALYZE table region update histogram on R\_REGIONKEY, R\_NAME;

ANALYZE table supplier update histogram on S\_SUPPKEY, S\_NATIONKEY, S\_ACCTBAL;

## 修改配置

secure-file-priv = ""

optimizer\_switch='engine\_condition\_pushdown=off,batched\_key\_access=off'

innodb\_adaptive\_hash\_index=0

innodb\_buffer\_pool\_size = 102400M

max\_heap\_table\_size = 10737418240

tmp\_table\_size = 10737418240

上述的securti-file-priv需要修改为对应的SQLEngine的目录，否则会报错：The MySQL server is running with the --secure-file priv option so it cannot execute this statement

# 测试模型

参考：

TPCH 22条SQL语句分析：<https://developer.aliyun.com/article/149715>

从TPC-H分析论文学习优化器的挑战与应对思路：

<https://zhuanlan.zhihu.com/p/608897702>

# 使用