《大数据导论》实验报告

实验题目	华为特色实验-MRS

一、实验目的
 通过本次实验，了解华为云 MRS，并熟悉使用。

二、实验项目内容
1. 申请华为云免费资源 MRS
2. 通过弹性公网 IP 访问 Manager
3. 下载客户端到主节点
4. 主节点安装客户端
5. 主节点使用客户端
6. 案例实现

三、实验过程或算法（源程序）
1. 申请华为云免费资源 MRS
 保持默认+勾选全部组件(Hadoop,Spark,HBase,Hive,Flink)+立即购买，过程较为简单，在此仅展示结果。

设置密码

启动集群比较费时，大概需要 1 个多小时。

2. 通过弹性公网IP 访问 Manager
（1）集群信息，在集群列表中单击指定的集群名称，进入集群信息页面。

（2）前往 Manager，然后购买弹性公网 IP，注意选择按需付费

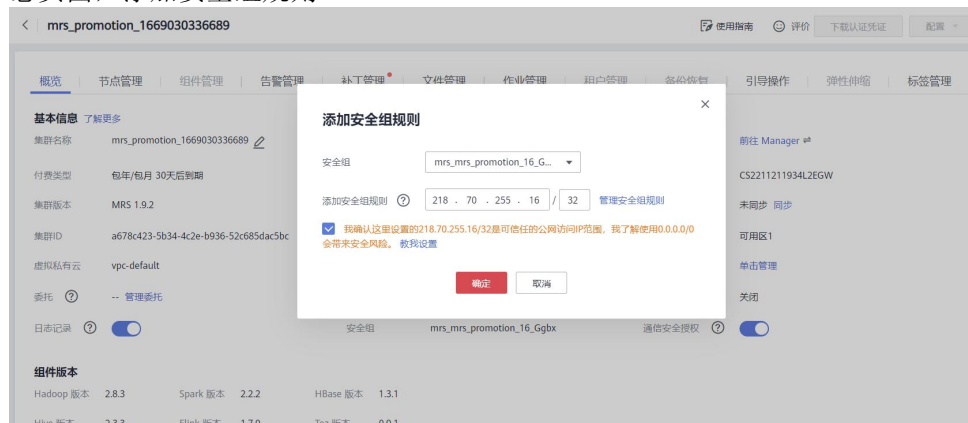此处一定要记得**手动**放通 9022 端口，否则无法访问后续页面：

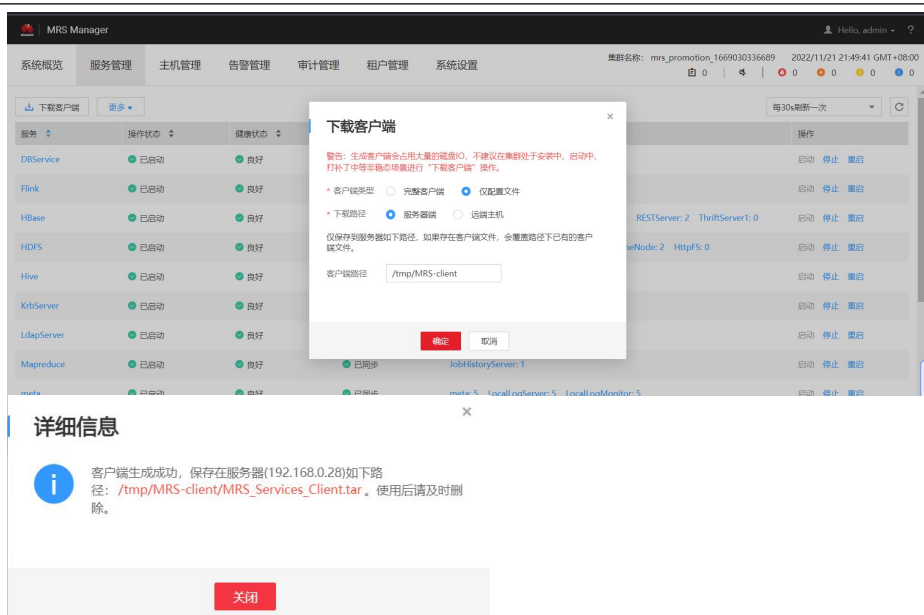（3）进入 MRS Manager 登录页面

输入默认用户名"admin"及创建集群时设置的密码，登录进入：
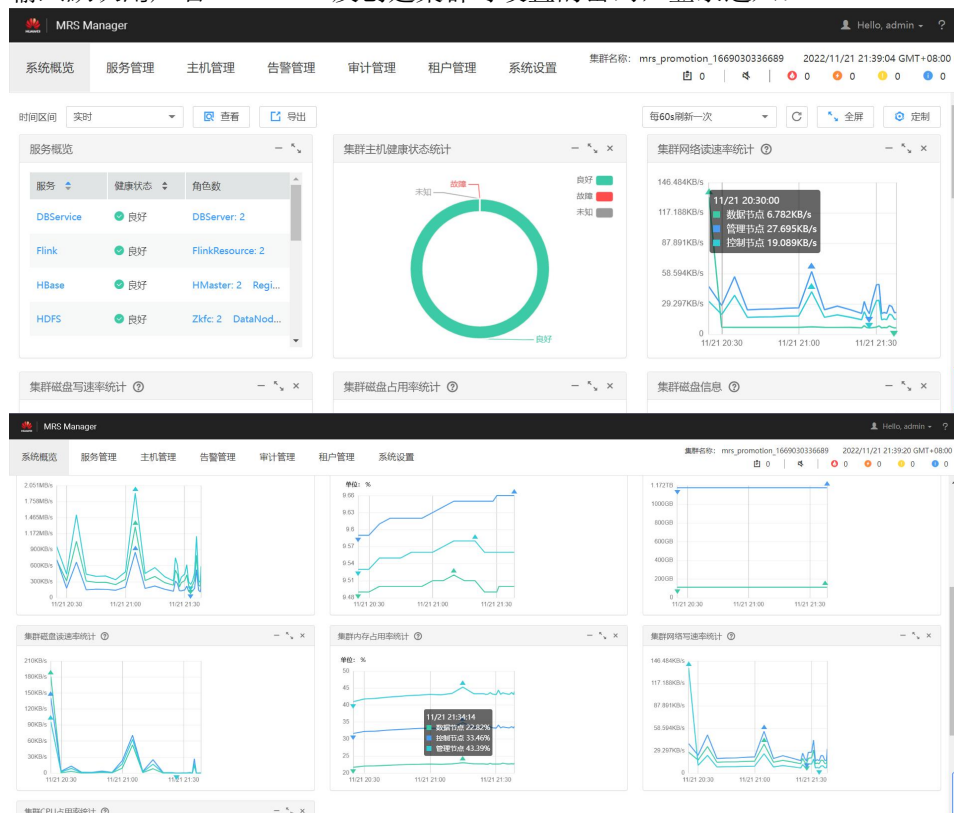
（4）在 MRS 管理控制台，在"现有集群"列表，单击指定的集群名称，进入集群信息页面，添加安全组规则

3．下载客户端到主节点

(1) 访问集群 Manager，点击"服务管理"

(2) 下载客户端，"客户端类型"选择"仅配置文件"，"下载路径"选择"服务器端"，单击"确定"开始生成客户端配置文件，文件生成后默认保存在主管理节点"/tmp/MRS-client"。

4. 主节点安装客户端

（1）MRS页面->现有集群->单击集群名称进入-点击节点管理，节点如图：

（2）此处选择...，单击并远程登录：

记得选择 VNC 登录而不是 CloudShell

输入账号和密码：

```
Hint: Num Lock on

node-master1aVxV login: root
Password:
Last login: Thu Jan  1 08:00:10 on

[root@node-master1aVxV ~]# _
```

此处需注意，直接输密码会导致一直 incorrect

（3）确认 MRS Manager 的主管理节点

```
Hint: Num Lock on

node-master2BZiG login: root
Password:
Last login: Thu Jan  1 08:00:10 on

[root@node-master2BZiG ~]# sudo su - root
Last login: Mon Nov 21 23:56:41 CST 2022 on tty1
[root@node-master2BZiG ~]# su - omm
Last login: Mon Nov 21 20:06:03 CST 2022
[omm@node-master2BZiG ~]$ sh ${BIGDATA_HOME}/om-0.0.1/sbin/status-oms.sh
```

回显信息如下：

```
192-168-0-28                 pms              Normal          Normal
   Single_active
192-168-0-28                 tomcat           Normal          Normal
   Single_active
192-168-0-51                 acs              Stopped         Normal
   Single_active
192-168-0-51                 aos              Stopped         Normal
   Single_active
192-168-0-51                 controller       Stopped         Normal
   Single_active
192-168-0-51                 executor         Stopped         Normal
   Single_active
192-168-0-51                 floatip          Stopped         Normal
   Single_active
192-168-0-51                 fms              Stopped         Normal
   Single_active
192-168-0-51                 gaussDB          Standby_normal  Normal
   Active_standby
192-168-0-51                 heartBeatCheck   Stopped         Normal
   Single_active
192-168-0-51                 httpd            Stopped         Normal
   Single_active
192-168-0-51                 iam              Stopped         Normal
   Single_active
192-168-0-51                 knox             Normal          Normal
   Double_active
192-168-0-51                 ntp              Standby_normal  Normal
   Active_standby
192-168-0-51                 okerberos        Normal          Normal
   Double_active
192-168-0-51                 oldap            Standby_normal  Normal
   Active_standby
192-168-0-51                 pms              Stopped         Normal
   Single_active
192-168-0-51                 tomcat           Stopped         Normal
   Single_active
```

没有显示全部，无法确定主节点和备节点，因此将其输出到文本文件

```
[omm@node-master2BZiG ~]$ nohup sh ${BIGDATA_HOME}/om-0.0.1/sbin/status-oms.sh > myout.txt 2>&1 &
[1] 18037
[omm@node-master2BZiG ~]$ vim myout.txt
```

按 i 进入编辑模式，上滑至顶部可以看到主备点信息：

```
nohup: ignoring input
HAMode
double

NodeName                        HostName                           HAVersion     StartTime
   HAActive     HAAllResOK        HARunPhase
192-168-0-51                    node-master2BZiG.mrs-qgnh.com      V100R001C01   2022-11-21 20:07:17
   standby      normal            Deactived
192-168-0-28                    node-master1aVxV.mrs-qgnh.com      V100R001C01   2022-11-21 20:07:17
   active       normal            Actived

NodeName                        ResName                           ResStatus     ResHAStatus
   ResType
192-168-0-51                    acs                               Stopped       Normal
   Single_active
192-168-0-51                    aos                               Stopped       Normal
```

192-168-0-28 为主管理节点，192-168-0-51 为备管理节点

（4）已经在 omm 用户中了，该指令不用执行

（5）切换到客户端安装目录

```
[omm@node-master2BZiG ~]$ cd /opt/client
[omm@node-master2BZiG client]$
```

（6）更新主管理节点的客户端配置

如图，更新成功。

5. 主节点使用客户端

(1) 在已更新客户端的主管理节点，执行 cd/opt/client 命令切换到客户端目录。

(2) 执行 source bigdata_env 命令配置环境变量

(3) 当前集群未启用Kerberos 认证，跳过

(4) 执行 HBase 组件的客户端，hbase shell

```
[root@node-master1aVxV client]# cd /opt/client
[root@node-master1aVxV client]# source bigdata_env
[root@node-master1aVxV client]# hbase shell
```

结果如下：

```
2022-11-22 00:44:47,072 INFO  [main] zookeeper.ZooKeeper: Client environment:os.arch=aarch64
2022-11-22 00:44:47,074 INFO  [main] zookeeper.ZooKeeper: Client environment:os.version=4.19.36-vhulk1906.1.0.h288.eulerosv2r8
arch64
2022-11-22 00:44:47,076 INFO  [main] zookeeper.ZooKeeper: Client environment:user.name=root
2022-11-22 00:44:47,078 INFO  [main] zookeeper.ZooKeeper: Client environment:user.home=/root
2022-11-22 00:44:47,079 INFO  [main] zookeeper.ZooKeeper: Client environment:user.dir=/opt/client
2022-11-22 00:44:47,081 INFO  [main] zookeeper.ZooKeeper: Client environment:os.memory.free=208MB
2022-11-22 00:44:47,082 INFO  [main] zookeeper.ZooKeeper: Client environment:os.memory.max=3044MB
2022-11-22 00:44:47,083 INFO  [main] zookeeper.ZooKeeper: Client environment:os.memory.total=235MB
2022-11-22 00:44:47,088 INFO  [main] zookeeper.ZooKeeper: Initiating client connection, connectString=node-ana-coreimub.mrs-qg
.com:2181,node-master2bzig.mrs-qgnh.com:2181,node-master1avxv.mrs-qgnh.com:2181 sessionTimeout=90000 watcher=org.apache.hadoop
base.zookeeper.PendingWatcher02b960a7
2022-11-22 00:44:47,113 INFO  [main] zookeeper.ClientCnxn: zookeeper.request.timeout is not configured. Using default value 12
00.
2022-11-22 00:44:47,115 INFO  [main] zookeeper.ClientCnxn: zookeeper.client.bind.port.range is not configured.
2022-11-22 00:44:47,117 INFO  [main] zookeeper.ClientCnxn: zookeeper.client.bind.address is not configured.
2022-11-22 00:44:47,127 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] client.FourLetterWordMain: connecting to
de-ana-coreimub.mrs-qgnh.com 2181
2022-11-22 00:44:47,141 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] zookeeper.ClientCnxn: Got server principa
from the server and it is null
2022-11-22 00:44:47,144 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] zookeeper.ClientCnxn: Using server princi
l zookeeper/node-ana-coreimub.mrs-qgnh.com
2022-11-22 00:44:47,148 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] zookeeper.ClientCnxn: Opening socket conn
tion to server node-ana-coreimub.mrs-qgnh.com/192.168.0.32:2181. Will not attempt to authenticate using SASL (unknown error)
2022-11-22 00:44:47,157 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] zookeeper.ClientCnxn: Socket connection e
ablished, initiating session, client: /192.168.0.28:63738, server: node-ana-coreimub.mrs-qgnh.com/192.168.0.32:2181
2022-11-22 00:44:47,166 INFO  [main-SendThread(node-ana-coreimub.mrs-qgnh.com:2181)] zookeeper.ClientCnxn: Session establishme
 complete on server node-ana-coreimub.mrs-qgnh.com/192.168.0.32:2181, sessionid = 0x240000098fa710d6, negotiated timeout = 900
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.3.1-mrs-1.9.0, rUnknown, Sat Dec  7 14:43:10 CST 2019

hbase(main):001:0> _
```

6. 案例实现

本案例选取华为官方商城中部分数据构建电商数据库，如下所示：

| row key | base_info | | | sale_info | | |
	name	publish	OS	RAM	ROM	price
001	mate50	2022	HarmonyOS3	8GB	128GB	4999
002	P50	2021	HarmonyOS2	8GB	128GB	4488 3758
003	P50	2021	HarmonyOS2	8GB	256GB	4988 4258

（0）查看 create 命令的帮助文件

```
Create a table with namespace=ns1 and table qualifier=t1
  hbase> create 'ns1:t1', {NAME => 'f1', VERSIONS => 5}

Create a table with namespace=default and table qualifier=t1
  hbase> create 't1', {NAME => 'f1'}, {NAME => 'f2'}, {NAME => 'f3'}
  hbase> # The above in shorthand would be the following:
  hbase> create 't1', 'f1', 'f2', 'f3'
  hbase> create 't1', {NAME => 'f1', VERSIONS => 1, TTL => 2592000, BLOCKCACHE => true}
  hbase> create 't1', {NAME => 'f1', CONFIGURATION => {'hbase.hstore.blockingStoreFiles' => '10'}}
  hbase> create 't1', {NAME => 'f1', IS_MOB => true, MOB_THRESHOLD => 1000000, MOB_COMPACT_PARTITION_POLICY => 'weekly'}

Table configuration options can be put at the end.
Examples:

  hbase> create 'ns1:t1', 'f1', SPLITS => ['10', '20', '30', '40']
  hbase> create 't1', 'f1', SPLITS => ['10', '20', '30', '40']
  hbase> create 't1', 'f1', SPLITS_FILE => 'splits.txt', OWNER => 'johndoe'
  hbase> create 't1', {NAME => 'f1', VERSIONS => 5}, METADATA => { 'mykey' => 'myvalue' }
  hbase> # Optionally pre-split the table into NUMREGIONS, using
  hbase> # SPLITALGO ("HexStringSplit", "UniformSplit" or classname)
  hbase> create 't1', 'f1', {NUMREGIONS => 15, SPLITALGO => 'HexStringSplit'}
  hbase> create 't1', 'f1', {NUMREGIONS => 15, SPLITALGO => 'HexStringSplit', REGION_REPLICATION => 2, CONFIGURATION => {'hbas
hregion.scan.loadColumnFamiliesOnDemand' => 'true'}}
  hbase> create 't1', {NAME => 'f1', DFS_REPLICATION => 1}

You can also keep around a reference to the created table:

  hbase> t1 = create 't1', 'f1'

Which gives you a reference to the table named 't1', on which you can then
call methods.
```

（1）创建华为手机存储表，包含了基础信息和销售信息两个列族

```
hbase(main):003:0> create 'huawei_phones',{NAME => 'base_info', VERSIONS =>2},{NAME => 'sale_info', VERSIONS =>10}
2022-11-22 01:22:54,900 INFO  [main] client.HBaseAdmin: Created huawei_phones
0 row(s) in 1.6480 seconds

=> Hbase::Table - huawei_phones
hbase(main):004:0> _
```

可以查看表的信息：

```
hbase(main):004:0> describe 'huawei_phones'
Table huawei_phones is ENABLED
huawei_phones
COLUMN FAMILIES DESCRIPTION
{NAME => 'base_info', BLOOMFILTER => 'ROW', VERSIONS => '2', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_E
ODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', REP
CATION_SCOPE => '0'}
{NAME => 'sale_info', BLOOMFILTER => 'ROW', VERSIONS => '10', IN_MEMORY => 'false', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_
CODING => 'NONE', TTL => 'FOREVER', COMPRESSION => 'NONE', MIN_VERSIONS => '0', BLOCKCACHE => 'true', BLOCKSIZE => '65536', RE
ICATION_SCOPE => '0'}
2 row(s) in 0.1640 seconds
```

（2）存入 mate50

```
hbase(main):005:0> put 'huawei_phones','001','base_info:name','mate50'
put 'hu0 row(s) in 0.1230 seconds

hbase(main):006:0> put 'huawei_phones','001','base_info:publish','2022'
0 row(s) in 0.0170 seconds

hbase(main):007:0> put 'huawei_phones','001','base_info:os','harmonyos3'
p0 row(s) in 0.0200 seconds

uhbase(main):008:0> put 'huawei_phones','001','sale_info:ram','8'
0 row(s) in 0.0060 seconds

hbase(main):009:0> put 'huawei_phones','001','sale_info:rom','128'
p0 row(s) in 0.0160 seconds

uhbase(main):010:0> put 'huawei_phones','001','sale_info:price','4999'
0 row(s) in 0.0170 seconds
```

（3）存入 p50 128G

```
hbase(main):011:0> put 'huawei_phones','002','base_info:name','p50'
0 row(s) in 0.0110 seconds

phbase(main):012:0> put 'huawei_phones','002','base_info:publish','2021'
p0 row(s) in 0.0200 seconds

uhbase(main):013:0> put 'huawei_phones','002','base_info:os','harmonyos2'
pu0 row(s) in 0.0300 seconds

hbase(main):014:0> put 'huawei_phones','002','sale_info:ram','8'
0 row(s) in 0.0060 seconds

phbase(main):015:0> put 'huawei_phones','002','sale_info:rom','128'
put 0 row(s) in 0.0360 seconds

'hbase(main):016:0> put 'huawei_phones','002','sale_info:price','4488'
0 row(s) in 0.0560 seconds
```

（4）存入 p50 256G

```
hbase(main):017:0> put 'huawei_phones','003','base_info:name','p50'
p0 row(s) in 0.0120 seconds

hbase(main):018:0> put 'huawei_phones','003','base_info:publish','2021'
0 row(s) in 0.0140 seconds
p
hbase(main):019:0> put 'huawei_phones','003','base_info:os','harmonyos2'
p0 row(s) in 0.0150 seconds

uhbase(main):020:0> put 'huawei_phones','003','sale_info:ram','8'
0 row(s) in 0.0110 seconds

puhbase(main):021:0> put 'huawei_phones','003','sale_info:rom','256'
p0 row(s) in 0.0170 seconds

hbase(main):022:0> put 'huawei_phones','003','sale_info:price','4988'
0 row(s) in 0.0340 seconds
```

（5）更新 p50 的双十一折扣价格

```
hbase(main):023:0> put 'huawei_phones','002','sale_info:price','3758'
pu0 row(s) in 0.0240 seconds

hbase(main):024:0> put 'huawei_phones','003','sale_info:price','4988'
0 row(s) in 0.0120 seconds
```

（6）查询 p50 的所有数据

```
hbase(main):025:0> scan 'huawei_phones', FILTER=>"ColumnPrefixFilter('name') AND ValueFilter(=,'binary:p50')"
ROW                        COLUMN+CELL
 002                       column=base_info:name, timestamp=1669051600628, value=p50
 003                       column=base_info:name, timestamp=1669051834387, value=p50
2 row(s) in 0.1330 seconds
```

查询 p50,128GB 的当前价格

```
hbase(main):026:0> get 'huawei_phones','002','sale_info:price'
COLUMN                          CELL
 sale_info:price                timestamp=1669051898072, value=3758
1 row(s) in 0.0370 seconds
```

（7）查看 p50,128GB 历史价格

```
hbase(main):027:0> get 'huawei_phones','002', {COLUMN=>'sale_info:price',VERSIONS=>2}
COLUMN                     CELL
 sale_info:price           timestamp=1669051898072, value=3758
 sale_info:price           timestamp=1669051771377, value=4488
1 row(s) in 0.0350 seconds
```

（8）查看全表

```
hbase(main):028:0> scan 'huawei_phones'
ROW                        COLUMN+CELL
 001                       column=base_info:name, timestamp=1669051473725, value=mate50
 001                       column=base_info:os, timestamp=1669051519489, value=harmonyos3
 001                       column=base_info:publish, timestamp=1669051478002, value=2022
 001                       column=sale_info:price, timestamp=1669051523856, value=4999
 001                       column=sale_info:ram, timestamp=1669051520483, value=8
 001                       column=sale_info:rom, timestamp=1669051521512, value=128
 002                       column=base_info:name, timestamp=1669051600628, value=p50
 002                       column=base_info:os, timestamp=1669051602955, value=harmonyos2
 002                       column=base_info:publish, timestamp=1669051601808, value=2021
 002                       column=sale_info:price, timestamp=1669051898072, value=3758
 002                       column=sale_info:ram, timestamp=1669051603924, value=8
 002                       column=sale_info:rom, timestamp=1669051604980, value=128
 003                       column=base_info:name, timestamp=1669051834387, value=p50
 003                       column=base_info:os, timestamp=1669051837517, value=harmonyos2
 003                       column=base_info:publish, timestamp=1669051836360, value=2021
 003                       column=sale_info:price, timestamp=1669051900313, value=4988
 003                       column=sale_info:ram, timestamp=1669051838485, value=8
 003                       column=sale_info:rom, timestamp=1669051839510, value=256
3 row(s) in 0.0640 seconds
```

（9）数据库删除表

```
hbase(main):029:0> disable 'huawei_phones'
drop2022-11-22 01:35:30,458 INFO  [main] client.HBaseAdmin: Started disable of huawei_phones
 'huawei_phones'2022-11-22 01:35:32,775 INFO  [main] client.HBaseAdmin: Disabled huawei_phones
0 row(s) in 2.4050 seconds

hbase(main):030:0> drop 'huawei_phones'
2022-11-22 01:35:36,095 INFO  [main] client.HBaseAdmin: Deleted huawei_phones
0 row(s) in 1.2640 seconds
```

到此，华为云 MRS 特色实验完成。

四、实验结果及分析和（或）源程序调试过程

实验结果

　　添加完数据后查看全表如下：

```
hbase(main):028:0> scan 'huawei_phones'
ROW                          COLUMN+CELL
 001                         column=base_info:name, timestamp=1669051473725, value=mate50
 001                         column=base_info:os, timestamp=1669051519489, value=harmonyos3
 001                         column=base_info:publish, timestamp=1669051478002, value=2022
 001                         column=sale_info:price, timestamp=1669051523856, value=4999
 001                         column=sale_info:ram, timestamp=1669051520483, value=8
 001                         column=sale_info:rom, timestamp=1669051521512, value=128
 002                         column=base_info:name, timestamp=1669051600628, value=p50
 002                         column=base_info:os, timestamp=1669051602955, value=harmonyos2
 002                         column=base_info:publish, timestamp=1669051601808, value=2021
 002                         column=sale_info:price, timestamp=1669051898072, value=3758
 002                         column=sale_info:ram, timestamp=1669051603924, value=8
 002                         column=sale_info:rom, timestamp=1669051604980, value=128
 003                         column=base_info:name, timestamp=1669051834387, value=p50
 003                         column=base_info:os, timestamp=1669051837517, value=harmonyos2
 003                         column=base_info:publish, timestamp=1669051836360, value=2021
 003                         column=sale_info:price, timestamp=1669051900313, value=4988
 003                         column=sale_info:ram, timestamp=1669051838485, value=8
 003                         column=sale_info:rom, timestamp=1669051839510, value=256
3 row(s) in 0.0640 seconds
```

最后数据库成功删除表：

```
hbase(main):029:0> disable 'huawei_phones'
drop2022-11-22 01:35:30,458 INFO  [main] client.HBaseAdmin: Started disable of huawei_phones
 'huawei_phones'2022-11-22 01:35:32,775 INFO  [main] client.HBaseAdmin: Disabled huawei_phones
0 row(s) in 2.4050 seconds

hbase(main):030:0> drop 'huawei_phones'
2022-11-22 01:35:36,095 INFO  [main] client.HBaseAdmin: Deleted huawei_phones
0 row(s) in 1.2640 seconds
```

调试过程

1. 前往 manager 失败，网页无法打开。

解决办法是手动放通 9022 端口；

2. 登录时一直 incorrect：

```
sudo su - root
Password:
Login incorrect

node-master1aVxV login:
```

解决办法是先输入用户名即"root"再输密码；
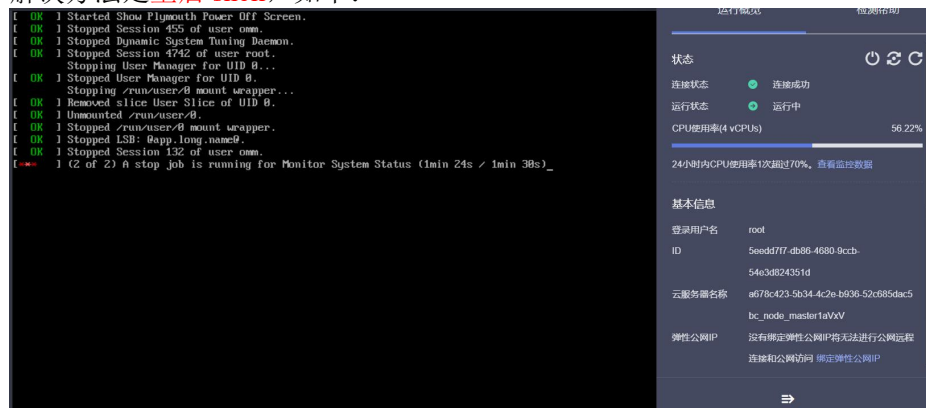
3. 更新主管理节点的客户端配置失败：

```
[omm@node-master2BZiG ~]$ cd /opt/client
[omm@node-master2BZiG client]$ sh refreshConfig.sh /opt/client /tmp/MRS-client/MRS_Services_Client.tar
[22-11-22 00:17:03]: The current user is not user root.
Failed to decompress outer client config.
[22-11-22 00:17:03]: Error: Failed to refresh components client config.
[omm@node-master2BZiG client]$
```

解决办法是重新登录 VNC 即可；

4. 查询指令不展示结果：


```
hbase(main):026:1> scan 'huawei_phones', FILTER=>"ColumnPrefixFilter('name') AND ValueFilter(=,'binary:p50'
hbase(main):027:1> get 'huawei_phones','002','sale_info:price'
hbase(main):028:1> get 'huawei_phones','002', {COLUMN=>'sale_info:price',VERSIONS=>2}
hbase(main):029:1>  scan 'huawei_phones'
hbase(main):030:1> disable 'huawei_phones'
hbase(main):031:1> drop 'huawei_phones'
hbase(main):032:1> _
```

解决办法是重启 shell，如下：

实验感悟：

　　本次实验中，我们认识了华为云的 MRS，并通过弹性公网 IP 访问 Manager，实现了主节点客户端的下载、安装和使用。并且我们在最后成功完成了案例的实现——利用华为官方商城中部分数据构建电商数据库，收获颇丰。

　　当然，我们在配置过程和实现过程中也遇到了很多困难，例如最开始前往 manager 失败，网页无法打开，研究了很久才发现需要考虑端口的问题，在本次实验中解决办法是手动放通 9022 端口。还有一些看似比较容易的地方也会出错，比如登录时直接输入密码，导致一直无法成功登录，这里要注意先输入默认的用户名 root；此外，还有更新主管理节点时的客户端配置失败和查询指令不展示结果这样的问题，它们都可以通过重启进行解决......

　　最终，在我们的尝试探索下，终于成功完成该次实验，对于华为云有了更加深刻的理解，也可以更好的使用它！