

CloudTable 操作指南

1 任务介绍

表格存储服务 (CloudTable Service , 简称 CloudTable) 是华为云基于 Apache HBase 提供的分布式、可伸缩、全托管的 KeyValue 数据存储服务，它提供了高性能的随机读写能力，适用于海量结构化数据、半结构化数据以及时序数据的存储和查询应用。

本章节提供从零开始使用 HBase 的操作指导：通过 HBase Shell 命令实现创建表，往表中插入数据，修改表，读取表数据，删除表中数据以及删除表的功能。

2 任务执行

2.1 创建集群

本章节为您介绍如何在 CloudTable 管理控制台创建一个新的集群。

操作步骤

步骤 1 登录表格存储服务管理控制台。

步骤 2 单击“免费试用”，进入“创建集群”页面。(21 天实战营的同学请勿点击“购买集群”，以免产生扣费)

创建集群相关配置样例如下：

- 计费模式：默认即可。

- 地域：默认即可。
- 可用分区：默认即可。
- 名称：输入集群名称，例如：“cloudtable-demo”。
- 存储配额：存储配额默认为 800GB。申请更多配额请单击“申请扩大配额”。
- 虚拟私有云：您可以选择使用已有的虚拟私有云网络。如果没有虚拟私有云，请单击“查看虚拟私有云”进入虚拟私有云管理控制台，创建一个虚拟私有云，创建完成后，重新进入表格存储服务管理控制台，在下拉框中选择新创建的虚拟私有云。创建虚拟私有云的详细操作，请参见《虚拟私有云用户指南》。
- 子网：创建虚拟私有云时会默认创建一个子网，您可以选择对应的子网名。
- 安全组：您可以选择默认安全组规则，并请确认所选安全组中存在如图 2-1 所示规则。
 - 方向：入方向
 - 类型：IPv4
 - 协议：Any
 - 端口范围/ICMP 类型：Any
 - 远端：本地安全组

图2-1 安全组规则

方向	类型	协议	端口范围/ICMP类型	远端
入方向	IPv4	Any	Any	Sys-default(9e0c54d7-26cf-4a42-9...

如果该规则不存在，则按图 2-2 所示创建如上所述安全组。

图2-2 添加规则

×

添加规则

*

方向:

入方向

出方向

*

协议:

Any

*

源地址:

IP地址

安全组

Sys-default(6e0c54...

确定

取消

如果您担心默认安全组存在网络安全风险，您可以在安全组中设置尽可能少的开放端口，但请确认所选安全组中存在如表 2-1 所示的安全组规则。

表2-1 自定义安全规则组

方向	协议	端口范围 /ICMP 类型	源地址/安全组	用途
出方向	ANY	ANY	ANY	出方向放行
入方向	TCP	22	0.0.0.0/0	ECS 远程登录端口 (Linux)
	TCP	3389	0.0.0.0/0	ECS 远程桌面服务

方向	协议	端口范围 /ICMP 类型	源地址/安全组	用途
				(Windows)
	TCP	16000	CloudTable 集群所在的安全组	HMaster RPC 端口
	TCP	16010		HMaster Web UI 端口
	TCP	16020		RegionServer RPC 端口
	TCP	16030		RegionServer web UI 端口
	TCP	2181		监听 ZooKeeper 客户端连接监听端口
	TCP	2888		Follower 连接监听端口
	TCP	3888		ZooKeeper 选举端口
	TCP	10101		RegionServer JMX 入口
	TCP	10102		HMaster JMX 入口
	TCP	2000		Hagent 访问端口
	TCP	8085		RestServer 访问端口
	TCP	9600		RestServer 访问端口
	TCP	4242		OpenTSDB 访问端口
	UDP	8649		Gmond 端口
	TCP	8649		Gmetad 端口

- CloudTable 版本：默认即可。
- HBase 版本：选择 1.3.1。

- IAM 统一身份认证：默认为关闭，默认即可。
- 高级特性：默认不勾选 “OpenTSDB 2.3.0”、GeoMesa，可自行选择是否勾选。本示例选择不勾选 “OpenTSDB 2.3.0”、GeoMesa。
- 计算单元数量：默认为 2，默认即可。

步骤 3 单击 “立即购买” 。

步骤 4 确认集群规格的订单详情信息，单击 “提交” ，完成集群的创建。

集群创建需要时间，所创集群的初始状态为 “创建中”，创建成功后状态更新为 “服务中”，请您耐心等待。

-----结束

2.2 准备弹性云服务器

在使用 HBase shell 访问集群或者访问集群的 Web UI 之前，需要在相同地域里准备一台弹性云服务器。

用户可以根据不同使用功能，选择对应镜像的弹性云服务器。

- 如果使用弹性云服务器通过 HBase shell 访问集群，建议使用 Linux 弹性云服务器。
- 如果使用弹性云服务器通过 Web UI 访问集群，建议使用 Windows 弹性云服务器。

准备弹性云服务器

购买 Linux 或 Windows 弹性云服务器的操作步骤请参见《弹性云服务器快速入门》中的 “购买并登录 Linux 弹性云服务器”。请参考：

https://support.huaweicloud.com/qs-ecs/zh-cn_topic_0021831611.html

购买的弹性云服务器需要满足如下要求。

- 弹性云服务器需要与 CloudTable 集群具有相同的地域、可用分区、虚拟私有云和子网。

虚拟私有云相关操作请参见《虚拟私有云用户指南》中“虚拟私有云和子网”。请参考：

https://support.huaweicloud.com/usermanual-vpc/zh-cn_topic_0013935842.html?utm_source=vpc_Growth_map&utm_medium=display&utm_campaign=help_center&utm_content=Growth_map

- 弹性云服务器需要与 CloudTable 集群具有相同的安全组。

安全组相关操作请参见《虚拟私有云用户指南》中“安全性 > 安全组”。请参考：

https://support.huaweicloud.com/usermanual-vpc/zh-cn_topic_0030969470.html

确认安全组是否存在如下规则：

- 方向：出方向
- 类型：IPv4
- 协议：Any
- 端口范围/ICMP 类型：Any
- 远端：0.0.0.0/0

如上述规则不存在，则按图 2-3 所示添加如下规则。

- 方向：出方向
- 协议：TCP
- 端口范围：80
- 目的地址：IP 地址，并设置为 0.0.0.0/0

图2-3 添加规则

添加规则

* 方向:

* 协议:

* 端口范围:

* 目的地址:

/ ?

- 购买弹性云服务器时需要选择现在购买“弹性 IP”，并绑定“弹性 IP”。

绑定“弹性 IP”的操作方法请参见《虚拟私有云用户指南》中“网络组件 > 弹性 IP”。

- 访问 Linux 弹性云服务器，需要允许用户使用密码方式登录 Linux 弹性云服务器（SSH 方式）。

具体登录操作步骤请参见《弹性云服务器用户指南》中的“实例 > 登录 Linux 弹性云服务器 > SSH 密码方式登录”。

- 访问 Windows 弹性云服务器。

具体登录操作步骤请参见《弹性云服务器用户指南》中的“实例 > 登录 Windows 弹性云服务器”。

2.3 使用 HBase Shell 访问集群

用户在弹性云服务器里可以通过[手动安装客户端](#)或[一键部署客户端](#)的方法，使用 HBase Shell 访问集群。推荐使用[一键部署客户端](#)的方法。

一键部署客户端

步骤 1 准备 Linux 弹性云服务器。

使用一键部署客户端工具，建议 Linux 弹性云服务器的操作系统类型为 EulerOS，CentOS，Ubuntu 和 SUSE。具体操作请参见 2.2 准备弹性云服务器章节。

步骤 2 下载客户端一键部署工具。

使用 SSH 登录工具（如 PuTTY）通过“弹性 IP”远程登录到 Linux 弹性云服务器，然后执行如下命令，获取客户端一键部署工具：

```
curl -O -k "http://cloudtable-  
publish.obs.myhwclouds.com/quick_start_hbase_shell.sh"
```

步骤 3 准备集群访问地址。

在集群管理列表中，查找对应集群“ZK 链接地址”，作为集群访问地址。

“ZK 链接地址”需要登录表格存储服务管理控制台，单击“集群管理”，在集群列表中搜索对应的集群名，并在“ZK 链接地址”列获取该集群的 ZK 链接地址，如图 2-4 所示。

图2-4 ZK 链接地址

集群名称	集群状...	任务状态	CloudTable...	创建时间	ZK链接地址	操作
------	--------	------	---------------	------	--------	----

步骤 4 使用工具部署客户端。

请将以下命令中的“\$zookeeper_address”参数替换为[步骤 3](#)获取的 ZK 链接地址，然后在弹性云服务器命令行窗口，执行该命令一键部署客户端。

```
source quick_start_hbase_shell.sh $zookeeper_address
```


步骤 5 启动 Shell 访问集群。

执行 “bin/hbase shell” , 启动 Shell 访问集群。

-----结束

2.4 HBase Shell 命令简介

本节介绍常用的 HBase Shell 命令。更多的 HBase Shell 命令，请参见 <https://learnhbase.wordpress.com/2013/03/02/hbase-shell-commands/>。

1. 启动 HBase shell。

进入 HBase 目录下执行如下命令，进入 HBase shell。

```
./bin/hbase shell
```

2. 获取帮助。

在 HBase shell 执行 “help” 命令，将会返回所有命令信息，以及基本命令的说明和使用方法。

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

3. 创建表。

执行 “create” 命令创建一张表，创建表时必须指定表的名称和 ColumnFamily 名称。

```
hbase(main):007:0> create 'cloudtable','cf'
```

```
0 row(s) in 1.5530 seconds
```

```
=> Hbase::Table - cloudtable
```

4. 查询表。

执行 “list” 命令查询指定表，必须指定查询表的名称。

```
hbase(main):009:0> list 'cloudtable'
```

```
TABLE
```

```
cloudtable
```

```
1 row(s) in 0.0060 seconds
```

```
=> ["cloudtable"]
```

5. 插入记录。

执行“put”命令往指定表插入一条记录，需要指定表的名称，主键，自定义列，以及插入的具体值。

```
hbase(main):004:0> put 'cloudtable','row1','cf:a','value1'  
0 row(s) in 0.2720 seconds
```

命令中各参数分别代表如下含义：

- cloudtable：表的名称
- row1：主键
- cf:a：自定义的列
- value1：插入的值

6. 扫描记录。

执行“scan”命令扫描表，扫描时需要指定表的名称，即可以全表扫描，也可以指定范围扫描。

```
hbase(main):001:0> scan 'cloudtable'  
ROW  
COLUMN+CELL  
row1  
column=cf:a, timestamp=1504866237162, value=value1  
1 row(s) in 0.2420 seconds
```

7. 查询单条记录。

执行“get”命令查询单条记录，必须指定查询表的名称和主键。

```
hbase(main):001:0> get 'cloudtable','row1'  
COLUMN  
CELL  
cf:a  
timestamp=1504866237162, value=value1  
1 row(s) in 0.2280 seconds
```

8. 禁用表。

如果需要修改表，比如删除表，修改表的设置，需要先对表进行禁用。执行 “disable” 命令禁用表。在禁用表时执行操作会报 ERROR，并提示表被禁用。

```
hbase(main):002:0> disable 'cloudtable'  
0 row(s) in 2.3550 seconds
```

9. 使能表。

与禁用表相对应的，在禁用表后，如果需要使用表，就需要执行 “enable” 命令使能表。

```
hbase(main):004:0> enable 'cloudtable'  
0 row(s) in 1.2500 seconds
```

10. 删除表。

当不再需要表时，可以执行 “drop” 命令删除表。删除表之前需要禁用表，否则会报 ERROR，并提示表被使能。删除表将会导致数据丢失，请谨慎操作。

```
hbase(main):007:0> disable 'cloudtable'  
0 row(s) in 2.2380 seconds  
  
hbase(main):008:0> drop 'cloudtable'  
0 row(s) in 1.2600 seconds
```

11. 退出 HBase shell。

执行 “quit” 命令退出 HBase Shell。

```
hbase(main):009:0> quit
```

2.5 GeoMesa 命令行简介

本节介绍常用的 GeoMesa 命令。更多的 GeoMesa 命令，请参见 http://www.geomesa.org/documentation/user/accumulo/commandline_tools.html。

12. 查看 classpath

执行 “classpath” 命令，将会返回当前命令行工具的所有 classpath 信息。

```
bin/geomesa-hbase classpath
```

13. 创建表

执行 “create-schema” 命令创建表，创建表时至少要指定目录名称与表名称，以及表规格。

```
bin/geomesa-hbase create-schema -c geomesa -f test -s
Who:String,What:java.lang.Long,When:Date,*Where:Point:srid=43
26,Why:String
```

14. 描述表

执行 “describe-schema” 命令获取表描述信息，描述表信息时必须指定目录名称与表名称。

```
bin/geomesa-hbase describe-schema -c geomesa -f test
```

15. 批量导入数据

执行 “ingest” 命令批量导入数据，导入时需要指定目录名称，表名称，表规格，以及相应的数据转换器等。

数据(车牌号，车辆颜色，经度，纬度，时间)：data.csv，并将数据表放在 data 文件夹中。

```
AAA,red,113.918417,22.505892,2017-04-09 18:03:46
BBB,white,113.960719,22.556511,2017-04-24 07:38:47
CCC,blue,114.088333,22.637222,2017-04-23 15:07:54
DDD,yellow,114.195456,22.596103,2017-04-21 21:27:06
EEE,black,113.897614,22.551331,2017-04-09 09:34:48
```

表结构定义：myschema.sft，并将 myschema.sft 放在 geomesa 命令行工具的 conf 文件夹中。

```
geomesa.sfts.cars = {
  attributes = [
    { name = "carid", type = "String", index = true }
    { name = "color", type = "String", index = false }
    { name = "time", type = "Date", index = false }
```

```

        { name = "geom", type = "Point", index = true, srid =
4326, default = true }
    ]
}

```

转换器定义：myconvertor.convert，并将 myconvertor.convert 放在 geomesa 命令行工具的 conf 文件夹中。

```

geomesa.converters.cars= {
    type    = "delimited-text",
    format  = "CSV",
    id-field = "$fid",
    fields = [
        { name = "fid",      transform = "concat($1,$5)" }
        { name = "carid",    transform = "$1::string" }
        { name = "color",    transform = "$2::string" }
        { name = "lon",      transform = "$3::double" }
        { name = "lat",      transform = "$4::double" }
        { name = "geom",     transform = "point($lon,$lat)" }
        { name = "time",     transform = "date('YYYY-MM-dd
HH:mm:ss',$5)" }
    ]
}

```

执行命令导入数据：

```

bin/geomesa-hbase ingest -c geomesa -C
conf/myconvertor.convert -s conf/myschema.sft data/data.csv

```

数据导入其他参数具体说明请参见：

<http://www.geomesa.org/documentation/user/accumulo/examples.html#ingesting-data>

16. 解释查询

执行 “explain” 命令获取指定查询语句执行计划的解释说明，解释语句时必须指定目录名称和表名称，以及给定查询语句。

```

bin/geomesa-hbase explain -c geomesa -f cars -q "carid = 'BBB'"

```

17. 统计分析

执行 “stats-analyze” 命令对数据表进行统计分析，同时还可以进一步执行 “stats-bounds”，“stats-count”，“stats-histogram”，“stats-top-k” 命令对数据表做更详细的统计。

```
bin/geomesa-hbase stats-analyze -c geomesa -f cars
bin/geomesa-hbase stats-bounds -c geomesa -f cars
bin/geomesa-hbase stats-count -c geomesa -f cars
bin/geomesa-hbase stats-histogram -c geomesa -f cars
bin/geomesa-hbase stats-top-k -c geomesa -f cars
```

18. 导出 feature

执行 “export” 命令导出 feature，导出时必须指定目录名称和表名称，同时还可以根据指定的查询语句进行导出。

```
bin/geomesa-hbase export -c geomesa -f cars -q "carid = 'BBB'"
```

19. 删除 feature

执行 “delete-features” 命令删除 feature，删除时必须指定目录名称和表名称，同时还可以根据指定的查询语句进行删除。

```
bin/geomesa-hbase delete-features -c geomesa -f cars -q "carid = 'BBB'"
```

20. 获取目录中的全部表的名称

执行 “get-type-names” 命令获取指定目录中的表名称。

```
bin/geomesa-hbase get-type-names -c geomesa
```

21. 删除表

执行 “remove-schema” 命令删除表，删除表示至少要指定表所在的目录与表名称。

```
bin/geomesa-hbase remove-schema -c geomesa -f test
bin/geomesa-hbase remove-schema -c geomesa -f cars
```

22. 删除目录

执行 “delete-catalog” 命令删除指定的目录。

```
bin/geomesa-hbase delete-catalog -c geomes
```

3 打卡任务

3.1 任务一

截取创建出来的 CloudTable 集群页面，请把华为云账号包含在截图中，如下图所示：



3.2 任务二

请在 HBase shell 上创建一个以华为云账号命名的表，如下图：

