Kudu+Impala 新手手册

通过该手册，你将学习到如何在 **CDH5.4.7+**版本(集群操作系统 **CentOS6.5**)上部署 Kudu；如何在 Kudu 上集成 Impala；如何通过 Impala 创建/更新 Kudu 表。

安装 Kudu

安装 CSD 文件 (使得用户可以通过 Cloudera Manager **添加服务向导**部署 Kudu)
登录部署 Cloudera Manager 的机器：

```
wgethttp://archive.cloudera.com/beta/kudu/csd/KUDU-0.5.0.jar
mv KUDU-0.5.0.jar /opt/cloudera/csd/
chowncloudera-scm:cloudera-scm /opt/cloudera/csd/KUDU-0.5.0.jar
chmod 644 /opt/cloudera/csd/KUDU-0.5.0.jar
servicecloudera-scm-server restart
```

安装 Kudu Parcel 包
Parcel 包下载地址：

```
http://archive.cloudera.com/beta/kudu/parcels/latest/
```

通过 Cloudera Manager 管理界面(Parcel 管理界面)下载、分发、激活 Kudu Parcel 包：

KUDU	0.5.0-1.kudu0.5.0.p0.223 ⓘ	Distributed, Activated	Deactivate

在 Cloudera Manager 管理界面添加 Kudu 服务
1. 配置 Kudu 实例

2. 配置 Kudu 参数(主要是存储 WAL、数据的文件目录)

Add Service Wizard

Review Changes

Kudu Master WAL Directory
fs_wal_dir

Master Default Group C

/mnt/sdb/kudu/wal

Missing required value: Kudu Master WAL Directory

Kudu Master Data Directories
fs_data_dirs

Master Default Group C

/mnt/sdb/kudu/master **+ −**

/mnt/sdc/kudu/master **+ −**

Missing required value: Kudu Master Data Directories

Kudu Tablet Server WAL Directory
fs_wal_dir

Tablet Server Default Group C

/mnt/sdb/kudu/wal

Missing required value: Kudu Tablet Server WAL Directory

Kudu Tablet Server Data Directories
fs_data_dirs

Tablet Server Default Group C

/mnt/sdb/kudu/tablet **+ −**

/mnt/sdc/kudu/tablet **+ −**

Missing required value: Kudu Tablet Server Data Directories

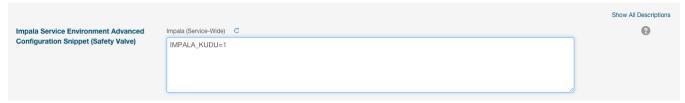启动 Kudu 服务

安装 Impala on Kudu

卸载原先安装的 Impala 服务(stop> Delete)

重新下载 Impala on Kudu 的 Parcel 包，下载地址：

http://archive.cloudera.com/beta/impala-kudu/parcels/latest/

通过 Cloudera Manager 管理界面(Parcel 管理界面)下载、分发、激活 Impala on Kudu Parcel 包：

| IMPALA_KUDU | 2.3.0-1.cdh5.5.0.p0.15 | Distributed, Activated | Deactivate |

通过添加服务向导添加 Impala 服务，添加完毕后，修改 Impala 配置

Show All Descriptions

Impala Service Environment Advanced Configuration Snippet (Safety Valve)

Impala (Service-Wide) C

IMPALA_KUDU=1

重启 Impala 服务

ImpalaSQL 操作

修改 Impala-shell 二进制程序

```
alternatives --display impala-shell
impala-shell - status is auto.
link currently points to /opt/cloudera/parcels/CDH-5.4.7-1.cdh5.4.7.p0.3/bin/impala-shell
/opt/cloudera/parcels/CDH-5.4.7-1.cdh5.4.7.p0.3/bin/impala-shell - priority 10
/opt/cloudera/parcels/IMPALA_KUDU-2.3.0-1.cdh5.5.0.p0.15/bin/impala-shell - priority 5
Current `best' version is /opt/cloudera/parcels/CDH-5.4.7-1.cdh5.4.7.p0.3/bin/impala-shell.
alternatives --set impala-shell /opt/cloudera/parcels/IMPALA_KUDU-2.3.0-1.cdh5.5.0.p0.15/bin/impala-shell
alternatives --display impala-shell
impala-shell - status is manual.
link currently points to /opt/cloudera/parcels/IMPALA_KUDU-2.3.0-1.cdh5.5.0.p0.15/bin/impala-shell
/opt/cloudera/parcels/CDH-5.4.7-1.cdh5.4.7.p0.3/bin/impala-shell - priority 10
/opt/cloudera/parcels/IMPALA_KUDU-2.3.0-1.cdh5.5.0.p0.15/bin/impala-shell - priority 5
Current `best' version is /opt/cloudera/parcels/CDH-5.4.7-1.cdh5.4.7.p0.3/bin/impala-shell.
```
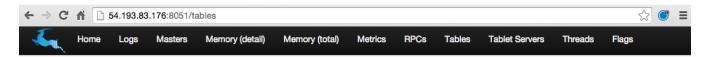
通过 Impala-shell 验证

```
select if (version() like '%KUDU%', "all set to go!", "check your configs");
Query: select if (version() like '%KUDU%', "all set to go!", "check your configs")
+------------------------------------------------------------------+
| if(version() like '%kudu%', 'all set to go!', 'check your configs') |
+------------------------------------------------------------------+
| all set to go!                                                   |
+------------------------------------------------------------------+
Fetched 1 row(s) in 0.01s
```

通过 Impala 创建 Kudu 表，要求构成 Primary Key 的列必须排在前面

```
create TABLE my_first_table (
id BIGINT,
name STRING
)
TBLPROPERTIES(
'storage_handler' = 'com.cloudera.kudu.hive.KuduStorageHandler',
'kudu.table_name' = 'my_first_table',
'kudu.master_addresses' = 'ip-172-31-28-144:7051',
'kudu.key_columns' = 'id'
);
```

创建表后，可以通过 Kudu 管理界面进行查看

Home Logs Masters Memory (detail) Memory (total) Metrics RPCs Tables Tablet Servers Threads Flags

Tables

Table Name	Table Id	State
my_first_table	15a057d933824bd88b5766a2ec87fa48	Running

通过 insert/update/delete 语句修改 Kudu 表

```
insert into my_first_table values (1, 'alex');
Query: insert into my_first_table values (1, 'alex')
Inserted 1 row(s) in 0.46s
insert into my_first_table values (2, 'dol');
Query: insert into my_first_table values (2, 'dol')
Inserted 1 row(s) in 0.11s
insert into my_first_table values (3, 'tonny');
Query: insert into my_first_table values (3, 'cheng')
Inserted 1 row(s) in 0.11s
select * from my_first_table;
Query: select * from my_first_table
+----+-------+
| id | name  |
+----+-------+
| 2  | dol   |
| 3  | cheng |
| 1  | alex  |
+----+-------+
Fetched 3 row(s) in 0.67s

delete from my_first_table where id = 3;
Query: delete from my_first_table where id = 3
Fetched 0 row(s) in 0.15s
select * from my_first_table;
Query: select * from my_first_table
+----+------+
| id | name |
+----+------+
| 2  | dol  |
| 1  | alex |
+----+------+
Fetched 2 row(s) in 0.13s
```

```
updatemy_first_table set name = 'ding' where id = 1;
Query: update my_first_table set name = 'ding' where id = 1
Fetched 0 row(s) in 0.13s
select * from my_first_table;
Query: select * from my_first_table
+----+------+
| id | name |
+----+------+
| 1  | ding |
| 2  | dol  |
+----+------+
Fetched 2 row(s) in 0.13s
```

Enjoy Kudu+Impala！

更多内容：http://www.cloudera.com/content/www/en-us/documentation/betas/kudu/0-5-0/topics/kudu.html