北风网 IBEIFENG.COM

大数据Hadoop高薪直通车课程

HBase高级使用

讲师：轩宇（北风网版权所有)

课程大纲

1. HBase 表的设计
2. HBase 表属性
3. HBase 表管理
4. 集成 Hive使用
5. HBase 实战案例

课程大纲

1. HBase 表的设计
2. HBase 表属性
3. HBase 表管理
4. 集成 Hive使用
5. HBase 实战案例

Create Table

```
hbase(main):007:0> help 'create'
Creates a table. Pass a table name, and a set of column family
specifications (at least one), and, optionally, table configuration.
Column specification can be a simple string (name), or a dictionary
(dictionaries are described below in main help output), necessarily
including NAME attribute.
Examples:

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}
```

Create Table

◆ 创建表时，每个列簇指定多个属性

```
hbase(main):014:0> create 't2',{NAME => 'f1', VERSIONS => 1, TTL => 2592000, BLOCKCACHE => true}
0 row(s) in 0.1600 seconds

=> Hbase::Table - t2
hbase(main):015:0> describe 't2'
DESCRIPTION                                                          ENABLED
 't2', {NAME => 'f1', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW true
 ', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MI
 N_VERSIONS => '0', TTL => '2592000 SECONDS (30 DAYS)', KEEP_DELETED_CEL
 LS => 'false', BLOCKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE =
 > 'true'}
1 row(s) in 0.0480 seconds
```

◆ 创建表时，多个列簇

```
hbase(main):020:0> create 't2',{NAME => 'f1'}, {NAME => 'f2'}
0 row(s) in 0.1650 seconds

=> Hbase::Table - t2
hbase(main):021:0> describe 't2'
DESCRIPTION                                                          ENABLED
 't2', {NAME => 'f1', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW true
 ', REPLICATION_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MI
 N_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'false', BLO
 CKSIZE => '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}, {NAME =
 > 'f2', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATIO
 N_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN_VERSIONS =>
  '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'false', BLOCKSIZE => '65
 536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
```

Table Split

```
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',
CONFIGURATION => {'hbase.hregion.scan.loadColumnFamiliesOnDemand' => 'true'}}
```

create 't1', 'f1', SPLITS => ['10', '20', '30', '40']

create 't1', 'f1', SPLITS_FILE => 'splits.txt'

create 't1', 'f1', {NUMREGIONS => 15, SPLITALGO => 'HexStringSplit'}

表的设计

- 表的预分区
- 设计表（依据【话单数据分析】）
  - 表rowkey设计
  - 索引表

本课程版权归北风网所有

欢迎访问我们的官方网站

www.ibeifeng.com