HBASE

What is HBase?

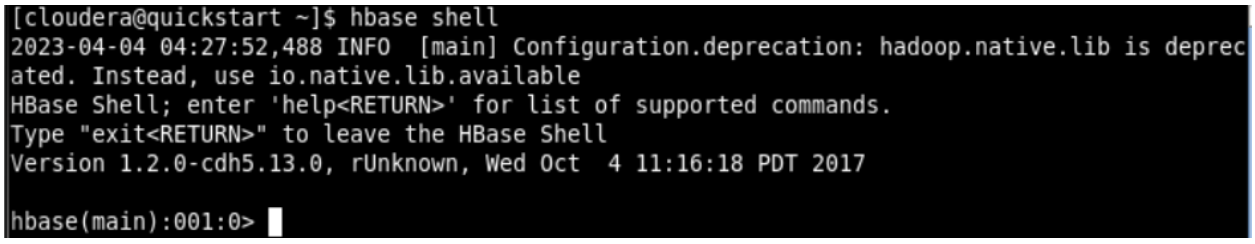
HBase is a column-oriented non-relational database management system that runs on top of Hadoop Distributed File System (HDFS). HBase provides a fault-tolerant way of storing sparse data sets, which are common in many big data use cases. It is well suited for real time data processing or random read/write access to large volumes of data.
Unlike relational database systems, HBase does not support a structured query language like SQL; in fact, HBase isn't a relational data store at all. HBase applications are written in JavaTM much like a typical Apache MapReduce application. HBase does support writing applications in Apache Avro, REST and Thrift.

HBase Shell

HBase contains a shell using which you can communicate with HBase. HBase uses the Hadoop File System to store its data. It will have a master server and region servers. The data storage will be in the form of regions (tables). These regions will be split up and stored in region servers.
The master server manages these region servers and all these tasks take place on HDFS. Given below are some of the commands supported by HBase Shell.
We can start the HBase interactive shell using "HBase shell" command as shown below.

Practicals

1. Start with HBase Shell.



2. Check for HBase Prompt. If not, restart the services.
   $ sudo su – This commands is to become super user
   $ service hbase-master restart – This commands is to restart hbase-master services
   $ service hbase-regionserver restart – This commands is to restart hbase-regionserver services.

3. Create a HBase table named ‘Student’ with column families StudID, StudName and StudAddress. Also cross check the creation and schema details of the table.

```
hbase(main):001:0> create 'student','studid','studname','studaddress'
0 row(s) in 1.5720 seconds

=> Hbase::Table - student
```

```
hbase(main):005:0> list
TABLE
student
1 row(s) in 0.0050 seconds

=> ["student"]
hbase(main):006:0> describe 'student'
Table student is ENABLED
student
COLUMN FAMILIES DESCRIPTION
{NAME => 'studaddress', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS
=> '1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE =
> '65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
{NAME => 'studid', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS => '1
', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '65
536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
{NAME => 'studname', DATA_BLOCK_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATION_SCOPE => '0', VERSIONS =>
'1', COMPRESSION => 'NONE', MIN_VERSIONS => '0', TTL => 'FOREVER', KEEP_DELETED_CELLS => 'FALSE', BLOCKSIZE => '
65536', IN_MEMORY => 'false', BLOCKCACHE => 'true'}
3 row(s) in 0.0910 seconds
```

4. Insert following data into the student table.
(Syntax: put 'Student','S01','StudName:FName','Amit')
S01,Amit Prakash Vaze, Mulund
S02,Sumit Neeraj Kulkarni, Devi Dayal Road, Mulund W
S03,Alka Deepak Yadav, PK Road, Thane, Maharashtra
S04,Amit Neeraj Yadav, Goshala Road, Mulund W, Maharashtra
S05,Veena Pushkar Mhatre, JN Road, Thane W, Maharashtra

```
hbase(main):011:0> put 'student','01','studname:fname','amit'
0 row(s) in 0.0090 seconds

hbase(main):012:0> put 'student','01','studname:mname','prakash'
0 row(s) in 0.0060 seconds

hbase(main):013:0> put 'student','01','studname:lname','vaze'
0 row(s) in 0.0030 seconds

hbase(main):014:0> put 'student','01','studaddress:place','mulund'
0 row(s) in 0.0100 seconds
```

```
hbase(main):015:0> put 'student','02','studname:fname','siddhesh'
0 row(s) in 0.0040 seconds

hbase(main):016:0> put 'student','02','studname:mname','ravi'
0 row(s) in 0.0040 seconds

hbase(main):017:0> put 'student','02','studname:lname','mane'
0 row(s) in 0.0100 seconds

hbase(main):018:0> put 'student','02','studaddress:place','ghatkopar'
0 row(s) in 0.0040 seconds

hbase(main):019:0> put 'student','03','studname:fname','jaya'
0 row(s) in 0.0040 seconds

hbase(main):020:0> put 'student','03','studname:mname','shivaji'
0 row(s) in 0.0030 seconds

hbase(main):021:0> put 'student','03','studname:lname','phadale'
0 row(s) in 0.0050 seconds

hbase(main):022:0> put 'student','03','studaddress:place','vikhroli'
0 row(s) in 0.0050 seconds

hbase(main):023:0> put 'student','04','studname:fname','shiv'
0 row(s) in 0.0030 seconds

hbase(main):024:0> put 'student','04','studname:mname','shankar'
0 row(s) in 0.0050 seconds

hbase(main):025:0> put 'student','04','studname:lname','prasad'
0 row(s) in 0.0050 seconds

hbase(main):026:0> put 'student','04','studaddress:place','thane'
0 row(s) in 0.0060 seconds
```

5. Display all rows from the Student table. (scan command)

```
hbase(main):027:0> scan 'student'
ROW                         COLUMN+CELL
 01                         column=studaddress:place, timestamp=1681379382487, value=mulund
 01                         column=studname:fname, timestamp=1681379180288, value=amit
 01                         column=studname:lname, timestamp=1681379244088, value=vaze
 01                         column=studname:mname, timestamp=1681379230520, value=prakash
 02                         column=studaddress:place, timestamp=1681379465083, value=ghatkopar
 02                         column=studname:fname, timestamp=1681379425587, value=siddhesh
 02                         column=studname:lname, timestamp=1681379453530, value=mane
 02                         column=studname:mname, timestamp=1681379442548, value=ravi
 03                         column=studaddress:place, timestamp=1681379534174, value=vikhroli
 03                         column=studname:fname, timestamp=1681379486772, value=jaya
 03                         column=studname:lname, timestamp=1681379509188, value=phadale
 03                         column=studname:mname, timestamp=1681379498534, value=shivaji
 04                         column=studaddress:place, timestamp=1681379615685, value=thane
 04                         column=studname:fname, timestamp=1681379585883, value=shiv
 04                         column=studname:lname, timestamp=1681379607028, value=prasad
 04                         column=studname:mname, timestamp=1681379595142, value=shankar
4 row(s) in 0.0430 seconds
```

6. Display address list of all students. (get command)

```
hbase(main):032:0> get 'student','01', {COLUMN => 'studaddress:place'}
COLUMN                      CELL
 studaddress:place          timestamp=1681379382487, value=mulund
1 row(s) in 0.0100 seconds

hbase(main):033:0>
hbase(main):034:0* get 'student','02', {COLUMN => 'studaddress:place'}
COLUMN                      CELL
 studaddress:place          timestamp=1681379465083, value=ghatkopar
1 row(s) in 0.0030 seconds

hbase(main):035:0> get 'student','03', {COLUMN => 'studaddress:place'}
COLUMN                      CELL
 studaddress:place          timestamp=1681379534174, value=vikhroli
1 row(s) in 0.0030 seconds

hbase(main):036:0> get 'student','04', {COLUMN => 'studaddress:place'}
COLUMN                      CELL
 studaddress:place          timestamp=1681379615685, value=thane
1 row(s) in 0.0030 seconds
```

7. Display list of students.

```
hbase(main):038:0> get 'student','01'
COLUMN                  CELL
 studaddress:place      timestamp=1681379382487, value=mulund
 studname:fname         timestamp=1681379180288, value=amit
 studname:lname         timestamp=1681379244088, value=vaze
 studname:mname         timestamp=1681379230520, value=prakash
4 row(s) in 0.0150 seconds

hbase(main):039:0> get 'student','02'
COLUMN                  CELL
 studaddress:place      timestamp=1681379465083, value=ghatkopar
 studname:fname         timestamp=1681379425587, value=siddhesh
 studname:lname         timestamp=1681379453530, value=mane
 studname:mname         timestamp=1681379442548, value=ravi
4 row(s) in 0.0060 seconds

hbase(main):040:0> get 'student','03'
COLUMN                  CELL
 studaddress:place      timestamp=1681379534174, value=vikhroli
 studname:fname         timestamp=1681379486772, value=jaya
 studname:lname         timestamp=1681379509188, value=phadale
 studname:mname         timestamp=1681379498534, value=shivaji
4 row(s) in 0.0070 seconds

hbase(main):041:0> get 'student','04'
COLUMN                  CELL
 studaddress:place      timestamp=1681379615685, value=thane
 studname:fname         timestamp=1681379585883, value=shiv
 studname:lname         timestamp=1681379607028, value=prasad
 studname:mname         timestamp=1681379595142, value=shankar
4 row(s) in 0.0070 seconds
```

To get a particular list

```
hbase(main):043:0> get 'student','01','studaddress'
COLUMN                  CELL
 studaddress:place      timestamp=1681379382487, value=mulund
1 row(s) in 0.0020 seconds
```

8. Display total number of students.

```
hbase(main):044:0> count 'student'
4 row(s) in 0.0650 seconds

=> 4
```

9. Disable Student table and display the table schema using describe command.
   >disable ‘Student’

```
hbase(main):045:0> disable 'student'
0 row(s) in 2.2520 seconds
```

10. Check disabled status of the table using is-disabled command.
    >is_disabled ‘Student’

```
hbase(main):046:0> is_disabled 'student'
true
0 row(s) in 0.0240 seconds
```

11. Enable Student table and display the table schema.

```
hbase(main):047:0> enable 'student'
0 row(s) in 1.3040 seconds

hbase(main):048:0> scan 'student'
ROW                   COLUMN+CELL
 01                   column=studaddress:place, timestamp=1681379382487, value=mulund
 01                   column=studname:fname, timestamp=1681379180288, value=amit
 01                   column=studname:lname, timestamp=1681379244088, value=vaze
 01                   column=studname:mname, timestamp=1681379230520, value=prakash
 02                   column=studaddress:place, timestamp=1681379465083, value=ghatkop
                      ar
 02                   column=studname:fname, timestamp=1681379425587, value=siddhesh
 02                   column=studname:lname, timestamp=1681379453530, value=mane
 02                   column=studname:mname, timestamp=1681379442548, value=ravi
 03                   column=studaddress:place, timestamp=1681379534174, value=vikhrol
                      i
 03                   column=studname:fname, timestamp=1681379486772, value=jaya
 03                   column=studname:lname, timestamp=1681379509188, value=phadale
 03                   column=studname:mname, timestamp=1681379498534, value=shivaji
 04                   column=studaddress:place, timestamp=1681379615685, value=thane
 04                   column=studname:fname, timestamp=1681379585883, value=shiv
 04                   column=studname:lname, timestamp=1681379607028, value=prasad
 04                   column=studname:mname, timestamp=1681379595142, value=shankar
4 row(s) in 0.0340 seconds
```

12. Demonstrate the deletion of column values and a row.
    - Deleting a specific cell

```
hbase(main):050:0> delete 'student','04','studname:mname',
hbase(main):051:0* 1681379595142
0 row(s) in 0.0480 seconds

hbase(main):052:0> scan 'student'
ROW                     COLUMN+CELL
 01                     column=studaddress:place, timestamp=1681379382487, value=mulund
 01                     column=studname:fname, timestamp=1681379180288, value=amit
 01                     column=studname:lname, timestamp=1681379244088, value=vaze
 01                     column=studname:mname, timestamp=1681379230520, value=prakash
 02                     column=studaddress:place, timestamp=1681379465083, value=ghatkop
                        ar
 02                     column=studname:fname, timestamp=1681379425587, value=siddhesh
 02                     column=studname:lname, timestamp=1681379453530, value=mane
 02                     column=studname:mname, timestamp=1681379442548, value=ravi
 03                     column=studaddress:place, timestamp=1681379534174, value=vikhrol
                        i
 03                     column=studname:fname, timestamp=1681379486772, value=jaya
 03                     column=studname:lname, timestamp=1681379509188, value=phadale
 03                     column=studname:mname, timestamp=1681379498534, value=shivaji
 04                     column=studaddress:place, timestamp=1681379615685, value=thane
 04                     column=studname:fname, timestamp=1681379585883, value=shiv
 04                     column=studname:lname, timestamp=1681379607028, value=prasad
4 row(s) in 0.0410 seconds
```

    - Deleting all cells

```
hbase(main):054:0> deleteall 'student','04'
0 row(s) in 0.0030 seconds

hbase(main):055:0> scan 'student'
ROW                     COLUMN+CELL
 01                     column=studaddress:place, timestamp=1681379382487, value=mulund
 01                     column=studname:fname, timestamp=1681379180288, value=amit
 01                     column=studname:lname, timestamp=1681379244088, value=vaze
 01                     column=studname:mname, timestamp=1681379230520, value=prakash
 02                     column=studaddress:place, timestamp=1681379465083, value=ghatkop
                        ar
 02                     column=studname:fname, timestamp=1681379425587, value=siddhesh
 02                     column=studname:lname, timestamp=1681379453530, value=mane
 02                     column=studname:mname, timestamp=1681379442548, value=ravi
 03                     column=studaddress:place, timestamp=1681379534174, value=vikhrol
                        i
 03                     column=studname:fname, timestamp=1681379486772, value=jaya
 03                     column=studname:lname, timestamp=1681379509188, value=phadale
 03                     column=studname:mname, timestamp=1681379498534, value=shivaji
3 row(s) in 0.0150 seconds
```

13. Set the maximum number of cell changes to 5.
    >alter ‘Student’, Name=’NewColFamily’, VERSIONS=>5

```
hbase(main):056:0> alter 'student',NAME => 'studaddress', VERSIONS => 5
Updating all regions with the new schema...
0/1 regions updated.
1/1 regions updated.
Done.
0 row(s) in 3.0450 seconds
```

14. Delete column family ‘Address’ from Student table.
    >alter ‘Student’,‘delete’ => ‘Address’

```
hbase(main):059:0> alter 'student','delete' => 'studaddress'
Updating all regions with the new schema...
0/1 regions updated.
1/1 regions updated.
Done.
0 row(s) in 3.4250 seconds
```

15. Check the existence of the Student table.
    >exists ‘Student’

```
hbase(main):061:0> exists 'student'
Table student does exist
0 row(s) in 0.0980 seconds
```

16. Drop Table ‘Student’

```
hbase(main):068:0> list
TABLE
student
1 row(s) in 0.0130 seconds

=> ["student"]
hbase(main):069:0> disable 'student'
0 row(s) in 2.2750 seconds

hbase(main):070:0> drop 'student'
0 row(s) in 1.2830 seconds

hbase(main):071:0> list
TABLE
0 row(s) in 0.0120 seconds

=> []
```

```
hbase(main):073:0> exists 'student'
Table student does not exist
0 row(s) in 0.0190 seconds
```