coreservlets.com – Hadoop Course

HBase Java API

In this exercise, you will have a chance to develop code based on HBase Java API. Your code will create a table, insert multiple records, retrieve data and finally drop the table.
Approx. Time: 45 minutes

Perform

Write Java code that will exercise HTable and HBaseAdmin classes to perform the following tasks:

1. Create table called 'Book' whose schema will able to house book's title, description, author's first and last names. Book's title and description should be grouped as they will be saved and retrieved together. Author's first and last name should also be grouped.
2. Add the following information to the 'Book' table:

ID	Title	Description	Author's first name	Author's last name
1	Faster than the speed love	Long book about love.	Brian	Dog
2	Long day	Story about Monday.	Emily	Blue
3	Flying Car	Novel about airplanes.	Phil	High

3. Retrieve and print to the screen an entire record with id '1'
4. Only retrieve title and description for record with id '2' and print to the screen
5. Change the last name of the author for the record with title 'Long Day' to 'Happy'. Display the record on the screen to verify the change.
6. Drop the table 'Book'

Expected Output

Your output should be similar to something like this:

```
------------------------------
1: Created table [Book] with 2 families: info and author
------------------------------
2: Saving rows:
Saved row with id [1]
Saved row with id [2]
Saved row with id [3]
------------------------------
3: Result with rowId [1], title=Faster than the speed love, description=Long book
about love., first name=Brian, last name=Dog
------------------------------
4: Result with rowId [2], title=Long day, description=Story about Monday.
------------------------------
5: Result with rowId [2], last name=Happy
```

```
------------------------------
12/01/16 16:37:10 INFO client.HBaseAdmin: Started disable of Book
12/01/16 16:37:12 INFO client.HBaseAdmin: Disabled Book
12/01/16 16:37:12 INFO client.HBaseAdmin: Deleted Book
```

Hints/Suggestions

1. Re-use Configuration and HTable instances. Why is that a good idea?
2. Write all 8 tasks in the same Java file
3. Create a method to print a single Result instance
4. Create a method to save a single record/row
5. Use *static* construct for all the schema names such as table name, families and even columns. Create these as byte arrays to avoid unnecessary conversions.

Solution

1. The code can be found in the Solutions project:

   `hbase.javaAPI.JavaAPISolution.java`

2. To execute

   ```
   $ yarn jar $PLAY_AREA/Solutions.jar hbase.javaAPI.JavaAPISolution
   ```