2. There is one database. Let’s say it is hosted locally and one of the team members migrates it to AWS or GCP. How can one confirm that the copied data is the same as the original data?

What would be the check points?

Imagine that data from table is of the form: List<Map<String,String>>

Answer:

Then for verifying that the copied data across a locally hosted database to AWS or GCP is the same as the original, check at the following checkpoints:

Data Validation Checkpoints:

- **Record Count Check:** For verifying that the number of records in each table of the original and migrated databases are the same.
- **Checksum/Hashing:** Calculate a hash or checksum (like MD5 or SHA-256) of each row or table and determine if the hash values between the original and copied data are identical.
- **Data Integrity Comparison:** Compare data line by line on the source and target databases.
- **Field-wise comparison:** Verify that data type as well as value for every field is same across the original and migrated tables.
- **Schema Comparison:** Simply put, check that the schema is identical for both sides, including the schemas of the tables, columns, data types, indexes, constraints, etc.

Java Code Example

- Below is a Java code snippet to compare data from two databases for a specific table in the form List<Map<String, String>>.

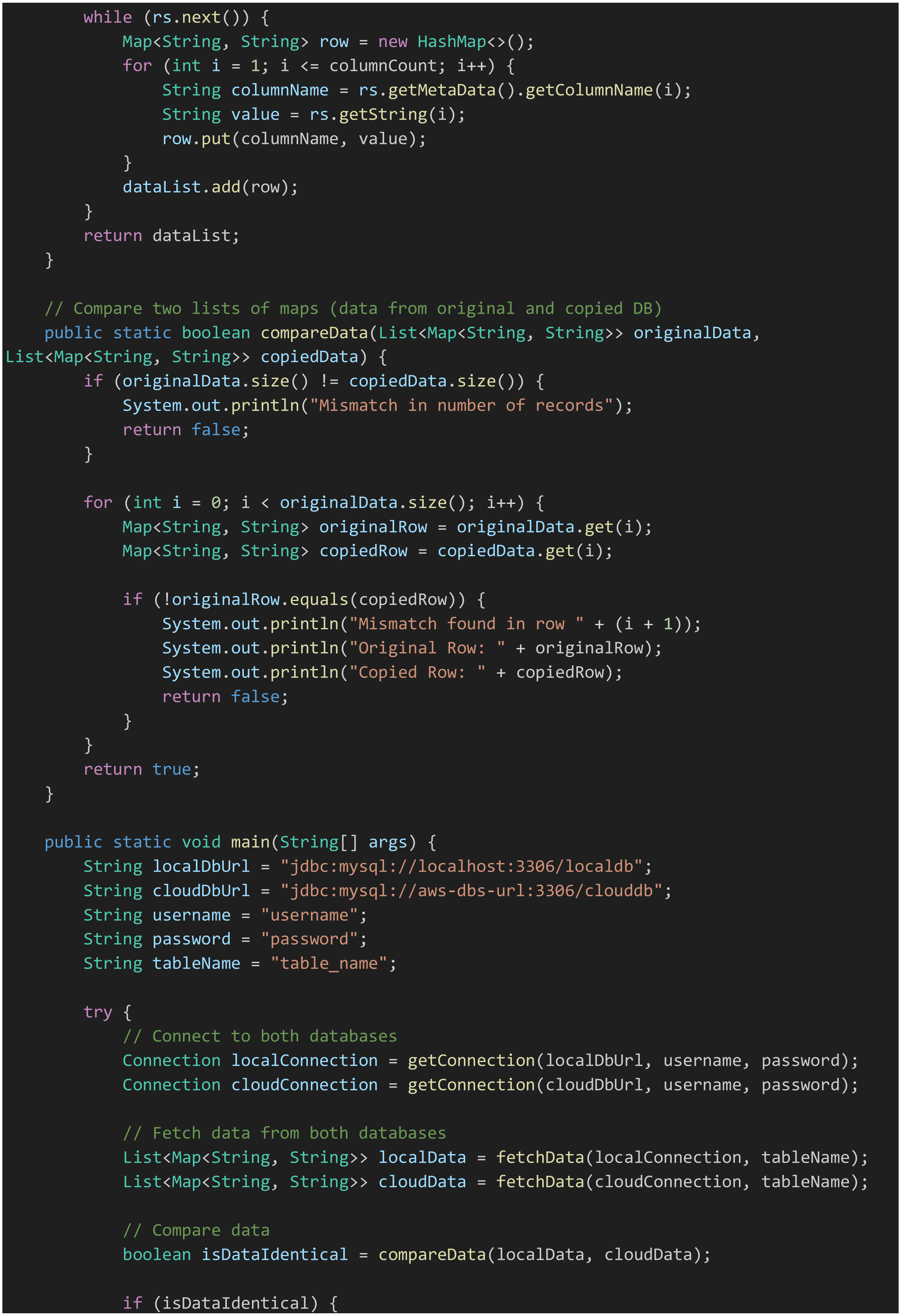
```
import java.sql.Connection;
import java.sql.DriverManager;
import java.sql.ResultSet;
import java.sql.Statement;
import java.util.*;

public class DatabaseComparer {

    // Helper method to get database connection
    public static Connection getConnection(String url, String user, String password)
throws Exception {
        return DriverManager.getConnection(url, user, password);
    }

    // Fetch data from a given table
    public static List<Map<String, String>> fetchData(Connection connection, String
tableName) throws Exception {
        List<Map<String, String>> dataList = new ArrayList<>();
        Statement stmt = connection.createStatement();
        ResultSet rs = stmt.executeQuery("SELECT * FROM " + tableName);

        int columnCount = rs.getMetaData().getColumnCount();
```



```
        while (rs.next()) {
            Map<String, String> row = new HashMap<>();
            for (int i = 1; i <= columnCount; i++) {
                String columnName = rs.getMetaData().getColumnName(i);
                String value = rs.getString(i);
                row.put(columnName, value);
            }
            dataList.add(row);
        }
        return dataList;
    }

    // Compare two lists of maps (data from original and copied DB)
    public static boolean compareData(List<Map<String, String>> originalData,
List<Map<String, String>> copiedData) {
        if (originalData.size() != copiedData.size()) {
            System.out.println("Mismatch in number of records");
            return false;
        }

        for (int i = 0; i < originalData.size(); i++) {
            Map<String, String> originalRow = originalData.get(i);
            Map<String, String> copiedRow = copiedData.get(i);

            if (!originalRow.equals(copiedRow)) {
                System.out.println("Mismatch found in row " + (i + 1));
                System.out.println("Original Row: " + originalRow);
                System.out.println("Copied Row: " + copiedRow);
                return false;
            }
        }
        return true;
    }

    public static void main(String[] args) {
        String localDbUrl = "jdbc:mysql://localhost:3306/localdb";
        String cloudDbUrl = "jdbc:mysql://aws-dbs-url:3306/clouddb";
        String username = "username";
        String password = "password";
        String tableName = "table_name";

        try {
            // Connect to both databases
            Connection localConnection = getConnection(localDbUrl, username, password);
            Connection cloudConnection = getConnection(cloudDbUrl, username, password);

            // Fetch data from both databases
            List<Map<String, String>> localData = fetchData(localConnection, tableName);
            List<Map<String, String>> cloudData = fetchData(cloudConnection, tableName);

            // Compare data
            boolean isDataIdentical = compareData(localData, cloudData);

            if (isDataIdentical) {
                System.out.println("Data is identical in both databases.");
            } else {
                System.out.println("Data mismatch found!");
            }

            // Close connections
            localConnection.close();
            cloudConnection.close();

        } catch (Exception e) {
            e.printStackTrace();
        }
    }
}
```

Steps in the Code:

1. **getConnection**: Establishes a connection to the database (local or cloud).
2. **fetchData**: Fetches all rows from a specified table in the form of a List<Map<String, String>>.
3. **compareData**: Compares the original data with the copied data row by row.
4. **Main Logic**:
   - Connect to both databases (local and cloud).
   - Fetch data from both databases.
   - Compare data row by row.
   - Display whether the data is identical or if there are any mismatches.