Near Real-time Analytics Dashboard with Tableau

Data Science Lab, The Department of Computer Science,
KSKV Kachchh University.

Web: http://cs.kutchuni.edu.in

Introduction:

Tableau is one of the great products when it comes to Data Visualization and Business Intelligence tool. Various companies have rule of thumb about tableau for visualization. When it comes to Big Data – larger volumes of data at that time it is not good practice to connect tableau directly with your Data lake, it would be great practice to connect tableau on clean / processed data, doing transformation in tableau is bad practice when volume is high because tableau always do transformation in memory.

I would like to propose Unique Big Data High Level Architecture:

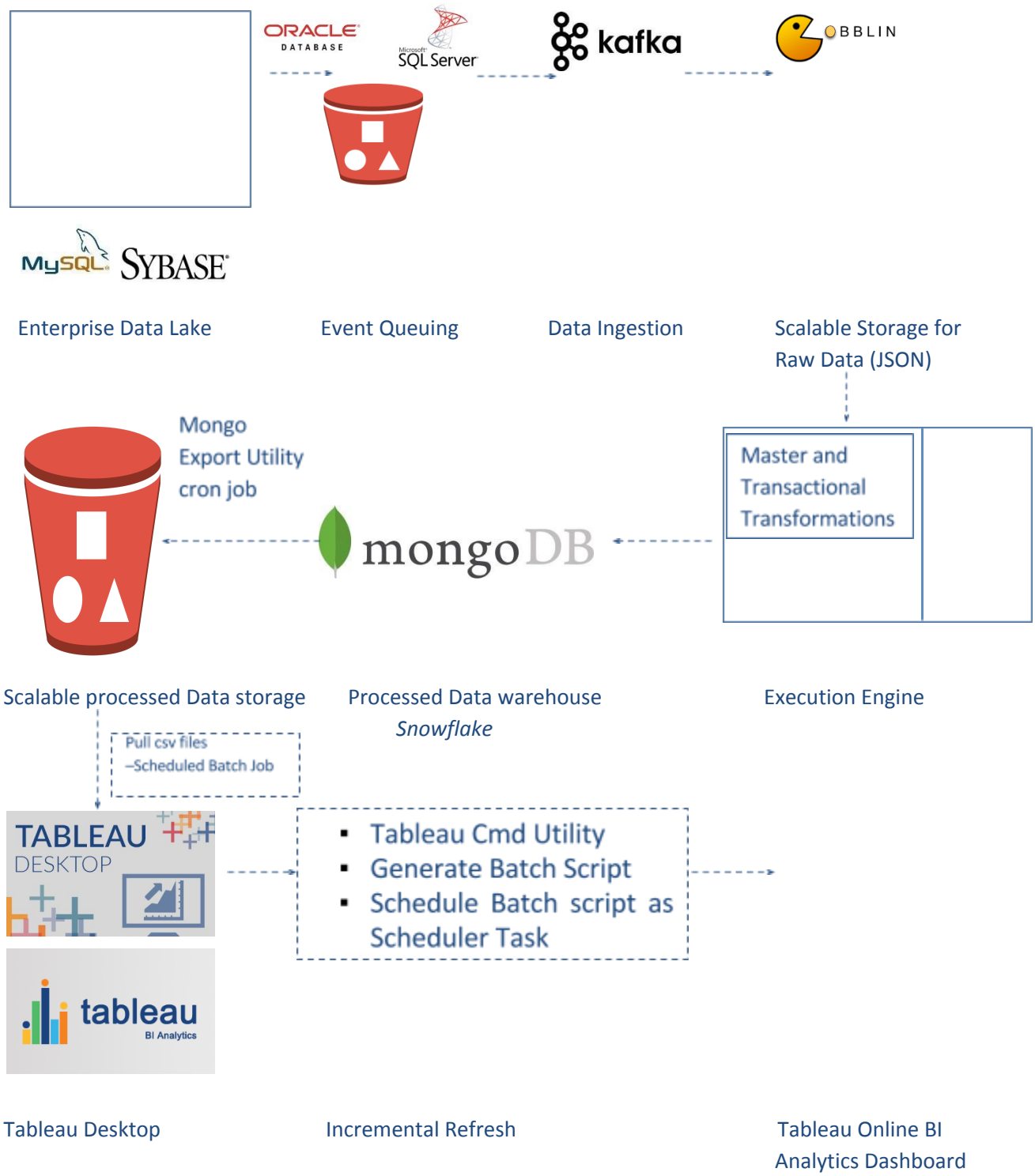

Figure 1: High Level Architecture – Near Real time Data Analytics with Tableau Online BI Tool

Tableau Data Extract Command-Line Utility

You can automate extract refresh tasks using the Tableau Data Extract Utility. This is a command-line utility that comes with Tableau Desktop, through which you can refresh published data sources or append data to them from a file.

Run the utility

1. Open the Command Prompt as an administrator and change to the Tableau Desktop bin directory. For example:
   ```
   cd C:\Program Files\Tableau\Tableau 10.1\bin
   ```
2. Use either of the following commands, adding parameters described in the tables below.
   ```
   tableau refreshextract
   tableau addfiletoextract
   ```

Example,

1.1 Generate Windows Batch script with below syntax example,

```
cd "C:\Program Files\Tableau\Tableau 10.1\bin"

tableau refreshextract --server https://online.tableau.com --username
your-email-address --password "your-password" --site
your-tableau-bi-tool-site-name  --project
"your-project-name-tableau-online" --datasource "sample-source"
--original-file "E:\folder\customer_level.csv"
```

Note: Every line of command should be on separate line and single command should be in single line.

1.2 Batch script with configuration property file.

1.2.1 Batch script

```
cd "C:\Program Files\Tableau\Tableau 10.1\bin"
tableau refreshextract --config-file "C:\tableau-utility\new-tableau.txt"
```

Note: Both commands must be on separate lines.

1.2.2 Configuration file

```
server=https://online.tableau.com
```

```
username=user-email-address
password=user-password
project=project-name-at-tableau-online-tool
datasource=data-source-name
source-username=user-email-address
source-password=user-password
```

Hence, If you would like to achieve 5 minutes of planned SLA. That means you want your Analytics Dashboard at Tableau Cloud Online must be refreshed by every 5 minutes. That's what you need to do.

References:

1) Tableau Data Extract Command-Line Utility
   [Online] https://onlinehelp.tableau.com/current/pro/desktop/en-us/extracting_TDE.html
2) Tableau Online: Keep Data fresh
   [Online] https://onlinehelp.tableau.com/current/online/en-us/to_keep_data_fresh.htm