Use Logstash to import data from Mysql to ES:

1. Download Logstash and mysql packages (installer is not required), set up the system environment variables properly. Then connect to mysql.
2. Download latest version of JDBC plugin, mysql-connector-java-5.1.47.
3. Create configuration file in your logstash directory /bin, e.g. default-logstash.conf.
4. Create drive file in the same directory, e.g. run\_default.bat

First, in the configuration file, set up the date pipeline and use mysql connector, there is three plugins in the file, **input{}: get the data source, filter{} (optional): modify the data and output{}: to the data destination**.   
The basic setup looks like:

**input{**

**jdbc{**

**jdbc\_driver\_library => “the path of the jdbc jar file”**

**jdbc\_driver\_class => “com.mysql.jdbc.Driver”**

**jdbc\_connection\_string => “jdbc:mysql://hostname of mysql: port/specific database”**

**jdbc\_user => “username of mysql”**

**jdbc\_password => “your password”**

**schedule => “\*\*\*\*\*\*”**

**}**

**}**

**output{**

**elasticsearch{**

**hosts => “elasticsearch host address(localhost:9200)”**

**index => “index name”**

**document\_type => “type name”**

**document\_id = “%{id}(if id field is existed in table)”**

**}**

**}**

Since in the latest version of ES, one index can only have one mapping, if the tables in one database have different mapping, you have to create different conf files. But if the mapping is the same, it is possible to use only one .conf, for example, add type field in input{} and use if loop to select different field if [type] == “xxx”.

Then add

**logstash –f default-logstash.conf**

in run\_default.bat

your\_directory/run\_default.bat to run logstash

Test the result in Kibana.

It is also possible to do that by Java…

1. What are the options for triggering it?

After setting up the pipeline, elasticsearch can update at the same time with mysql in theory, but it take time for logstash to process it.

1. How do we track records need to be updated?

Use sql\_last\_value in the statement field:

**statement => “SELECT id, mycolumn1, mycolumn2 FROM my\_table WHERE id >: sql\_last\_value”**

Then tell logsatsh which field is set to update sql\_last\_value:

**tracking\_column => id**

**use\_column\_value => true**

Use date filter plugin to filter input data’s timestamp:

https://www.elastic.co/guide/en/logstash-versioned-plugins/current/v3.1.9-plugins-filters-date.html

1. Can we do bidirectional updates?

No? But it is possible to export es data to json file using elasticdump tool and import them into Mysql…