Установка и настройка Logstash

Это руководство объясняет как загрузить данные из реляционной БД в Elasticsearch с помощью Logstash и плагина JDBC.

Предварительные требования

- установленная и настроенная БД `ltoolslogs` в СУБД MSSQL или PostgreSQL;
- установленная и настроенная платформа Elasticsearch и Kibana;
- настроенная политика жизненного цикла индексов (index lifecycle policy) с именем `ltoolslogs` в Elasticsearch.

Установка Logstash под Windows Server

Архив Logstash идёт в комплекте поставки. Также может быть скачан с официального сайта Elasticsearch.

Распаковываем файл `logstash-8.6.2-windows-x86_64.zip` , например, в `C:\logstash-8.6.2` . В дальнейшем этот каталог будет называться `LS_ROOT` .

Настройка поддержки экранирования в конфигах

В файле `LS_ROOT\config\logstash.yml` необходимо добавить строку

```
config.support_escapes: true
```

Настройка шаблона индекса

Копируем файл `ltoolslogs-template.conf` из коплекта поставки в `LS_ROOT\conf` . В этом файле можно изменить шаблоны имён индексов и имя политики жизненного цикла индексов:

```
{
  "index_patterns": ["ltools-logs-*"],
  "template": {
    "settings": {
      "index": {
        "lifecycle": {
          "name": "ltoolslogs"
        }
      }
    }
  }
}
```

Настройка для MSSQL

Установка драйвера JDBC

Архив драйвера JDBC для MSSQL идёт в комплекте поставки. Также может быть скачан с официального сайта Microsoft.

Распаковывает файл `sqljdbc_12.2.0.0_enu.zip`, например, в `LS_ROOT\sqljdbc_12.2`. В дальнейшем этот каталог будет называться `JDBC_ROOT`.

Настройка конвейера для таблицы Logs

Копируем файл `ltools-logs-mssql.conf` из комплекта поставки в `LS_ROOT\config\ltools-logs.conf`. В этом файле необходимо заменить следующие параметры:

Параметр	Описание
`JDBC_ROOT`	Каталог, в котором находится драйвер JDBC для MSSQL
`<DB-SERVER-HOST>`	Имя хоста или IP-адрес сервера СУБД
`<DB-USER-NAME>`	Имя пользователя СУБД
`<DB-USER-PASSWORD>`	Пароль пользователя СУБД
`<ES-HOST>`	Имя хоста или IP-адрес сервера Elasticsearch
`<ES-PORT>`	Порт сервера Elasticsearch, обычно 9200
`<ES-USER>`	Имя пользователя Elasticsearch, обычно `elastic`
`<ES-PASSWORD>`	Пароль пользователя Elasticsearch

Параметр	Описание
<SCHEDULE>	Расписание запуска загрузки, например, `0 * * * * *` - каждую минуту

```
input {
  jdbc {
    jdbc_driver_library => "JDBC_ROOT/enu/mssql-jdbc-12.2.0.jre11.jar"
    jdbc_driver_class => "com.microsoft.sqlserver.jdbc.SQLServerDriver"
    jdbc_connection_string => "jdbc:sqlserver://<DB-SERVER-HOST>:1433;databaseName=ltoolslogs;encrypt=false;"
    jdbc_user => "<DB-USER-NAME>"
    jdbc_password => "<DB-USER-PASSWORD>"
    jdbc_paging_enabled => true
    jdbc_default_timezone => "UTC"
    tracking_column => "unix_ts_in_secs"
    use_column_value => true
    tracking_column_type => "numeric"
    schedule => "<SCHEDULE>"
    statement => "SELECT *
      ,DATEDIFF_BIG(ms, '1970-01-01 00:00:00', OrchTimestampUtc) AS unix_ts_in_secs
    FROM
      [ltoolslogs].[dbo].[Logs] WITH(NOLOCK)
    WHERE
      (DATEDIFF_BIG(ms, '1970-01-01 00:00:00', OrchTimestampUtc) > :sql_last_value
            AND OrchTimestampUtc < getutcdate())"
  }
}
filter {
  mutate {
    copy => { "id" => "[@metadata][_id]"}
    copy => { "orchtimestamputc" => "@timestamp" }
    remove_field => ["id", "@version", "unix_ts_in_secs", "orchtimestamputc", "signature"]
  }
}
output {
#  stdout { codec => "rubydebug"}
  elasticsearch {
    hosts => ["<ES-HOST>:<ES-PORT>"]
    ssl => true
    ssl_certificate_verification => false
    user => "<ES-USER>"
    password => "<ES-PASSWORD>"
    index => "ltools-logs-%{+YYYY-MM-dd}"
    manage_template => true
    template => "LS_ROOT/config/ltoolslogs-template.conf"
    template_name => "ltoolslogs"
    template_overwrite => true
  }
}
```

Настройка для PostgreSQL

Установка драйвера JDBC для PostgreSQL

Архив драйвера JDBC для PostgreSQL идёт в комплекте поставки. Также может быть скачан с официального сайта PostgreSQL.

Копируем файл `postgresql-42.5.4.jar`, например, в `LS_ROOT\postgresqljdbc`. В дальнейшем этот каталог будет называться `JDBC_ROOT`.

Настройка конвейера для таблицы Logs

Копируем файл `ltools-logs-pgsql.conf` из комплекта поставки в `LS_ROOT\config\ltools-logs.conf`. В этом файле необходимо заменить следующие параметры:

Параметр	Описание
`JDBC_ROOT`	Каталог, в котором находится драйвер JDBC для MSSQL
`<DB-SERVER-HOST>`	Имя хоста или IP-адрес сервера СУБД
`<DB-USER-NAME>`	Имя пользователя СУБД
`<DB-USER-PASSWORD>`	Пароль пользователя СУБД
`<ES-HOST>`	Имя хоста или IP-адрес сервера Elasticsearch
`<ES-PORT>`	Порт сервера Elasticsearch, обычно 9200
`<ES-USER>`	Имя пользователя Elasticsearch, обычно `elastic`
`<ES-PASSWORD>`	Пароль пользователя Elasticsearch
`<SCHEDULE>`	Расписание запуска загрузки, например, `0 * * * * *` - каждую минуту

```
input {
  jdbc {
    jdbc_driver_library => "JDBC_ROOT/postgresql-42.5.4.jar"
    jdbc_driver_class => "org.postgresql.Driver"
    jdbc_connection_string => "jdbc:postgresql://<DB-SERVER-HOST>:5432/ltoolslogs"
    jdbc_user => "<DB-USER-NAME>"
    jdbc_password => "<DB-USER-PASSWORD>"
    jdbc_paging_enabled => true
    jdbc_default_timezone => "UTC"
    tracking_column => "unix_ts_in_secs"
    use_column_value => true
    tracking_column_type => "numeric"
    schedule => "<SCHEDULE>"
    statement => "SELECT *
      ,EXTRACT(EPOCH FROM \"OrchTimestampUtc\") AS unix_ts_in_secs
    FROM
      public.\"Logs\"
    WHERE
      EXTRACT(EPOCH FROM \"OrchTimestampUtc\") > :sql_last_value
        AND \"OrchTimestampUtc\" < NOW()"
  }
}
filter {
  mutate {
    copy => { "id" => "[@metadata][_id]"}
    copy => { "orchtimestamputc" => "@timestamp" }
    remove_field => ["id", "@version", "unix_ts_in_secs", "orchtimestamputc", "signature"]
  }
}
output {
#  stdout { codec => "rubydebug"}
  elasticsearch {
    hosts => ["<ES-HOST>:<ES-PORT>"]
    ssl => true
    ssl_certificate_verification => false
    user => "<ES-USER>"
    password => "<ES-PASSWORD>"
    index => "ltools-logs-%{+YYYY-MM-dd}"
    manage_template => true
    template => "LS_ROOT/config/ltoolslogs-template.conf"
    template_name => "ltoolslogs"
    template_overwrite => true
  }
}
```

Настройка Logstash как сервиса Windows

Настройка файла pipelines.yml

Файл `LS_ROOT\config\pipelines.yml` должен содержать описание активных конвейеров:

```
- pipeline.id: ltools-logs
  queue.type: persisted
  path.config: "config/ltools-logs.conf"
```

Установка и настройка NSSM

Архив NSSM идёт в комплекте поставки. Также может быть скачан с официального сайта NSSM.

Извлекаем из архива nssm-2.24-101-g897c7ad.zip файл `nssm-2.24-101-g897c7ad\win64\nssm.exe` в каталог `LS_ROOT\bin` . В консоли администратора выполняем следующее:

```
> LS_ROOT\bin\nssm.exe install logstash
```

В окне настройки сервиса указываем:

Закладка	Параметр	Значение
Application	Path	`LS_ROOT\bin\logstash.bat`
Application	Startup Directory	`LS_ROOT\bin`
Environment	Environment Variables	`TZ=UTC`

Нажимаем `Install Service` и `OK` в появившемся окне `Service 'logstash' installed successfully!` . В дальнейшем, службой Logstash можно управлять с помощью оснастки Службы (Services).

Проверка корректности настройки

Файлы журналов Logstash находятся в каталоге `LS_ROOT\logs` , основной файл для мониторинга - `logstash-plain.log` .

При успешной работе загрузки данных в Kibana можно увидеть наличие индексов `ltools-logs-YYYY-MM-dd` . В настройках представления (view) Kibana необходимо указать шаблон индексов (index pattern) `ltools-*` .

Установка Logstash под Centos Linux 8

Архив Logstash идёт в комплекте поставки. Также может быть скачан с официального сайта Elasticsearch.

Устанавливаем Logstash с помощью следующей команды:

```
# rpm -Uhv logstash-8.6.2-x86_64.rpm
```

Настройка поддержки экранирования в конфигах

В файле `/etc/logstash/logstash.yml` необходимо добавить строку

```
config.support_escapes: true
```

Настройка шаблона индекса

Копируем файл `ltoolslogs-template.conf` из коплекта поставки в `/etc/logstash` . В этом файле можно изменить шаблоны имён индексов и имя политики жизненного цикла индексов:

```
{
  "index_patterns": ["ltools-logs-*"],
  "template": {
    "settings": {
      "index": {
        "lifecycle": {
          "name": "ltoolslogs"
        }
      }
    }
  }
}
```

Настройка для MSSQL

Установка драйвера JDBC

Архив драйвера JDBC для MSSQL идёт в комплекте поставки. Также может быть скачан с официального сайта Microsoft.

Извлекаем из архива `sqljdbc_12.2.0.0_enu.tar.gz` файл `sqljdbc_12.2/enu/mssql-jdbc-12.2.0.jre11.jar` в `/usr/share/logstash/logstash-core/lib/jars` .

Настройка конвейера для таблицы Logs

Копируем файл `ltools-logs-mssql.conf` из комплекта поставки в `/etc/logstash/ltools-logs.conf`. В этом файле необходимо заменить следующие параметры:

Параметр	Описание
`<DB-SERVER-HOST>`	Имя хоста или IP-адрес сервера СУБД
`<DB-USER-NAME>`	Имя пользователя СУБД
`<DB-USER-PASSWORD>`	Пароль пользователя СУБД
`<ES-HOST>`	Имя хоста или IP-адрес сервера Elasticsearch
`<ES-PORT>`	Порт сервера Elasticsearch, обычно 9200
`<ES-USER>`	Имя пользователя Elasticsearch, обычно `elastic`
`<ES-PASSWORD>`	Пароль пользователя Elasticsearch
`<SCHEDULE>`	Расписание запуска загрузки, например, `0 * * * * *` - каждую минуту

```
input {
  jdbc {
    jdbc_driver_class => "com.microsoft.sqlserver.jdbc.SQLServerDriver"
    jdbc_connection_string => "jdbc:sqlserver://<DB-SERVER-HOST>:1433;databaseName=ltoolslogs;encrypt=false;"
    jdbc_user => "<DB-USER-NAME>"
    jdbc_password => "<DB-USER-PASSWORD>"
    jdbc_paging_enabled => true
    jdbc_default_timezone => "UTC"
    tracking_column => "unix_ts_in_secs"
    use_column_value => true
    tracking_column_type => "numeric"
    schedule => "<SCHEDULE>"
    statement => "SELECT *
      ,DATEDIFF_BIG(ms, '1970-01-01 00:00:00', OrchTimestampUtc) AS unix_ts_in_secs
    FROM
      [ltoolslogs].[dbo].[Logs] WITH(NOLOCK)
    WHERE
      (DATEDIFF_BIG(ms, '1970-01-01 00:00:00', OrchTimestampUtc) > :sql_last_value
            AND OrchTimestampUtc < getutcdate())"
  }
}
filter {
  mutate {
    copy => { "id" => "[@metadata][_id]"}
    copy => { "orchtimestamputc" => "@timestamp" }
    remove_field => ["id", "@version", "unix_ts_in_secs", "orchtimestamputc", "signature"]
  }
}
output {
#  stdout { codec =>  "rubydebug"}
  elasticsearch {
    hosts => ["<ES-HOST>:<ES-PORT>"]
    ssl => true
    ssl_certificate_verification => false
    user => "<ES-USER>"
    password => "<ES-PASSWORD>"
    index => "ltools-logs-%{+YYYY-MM-dd}"
    manage_template => true
    template => "/etc/logstash/ltoolslogs-template.conf"
    template_name => "ltoolslogs"
    template_overwrite => true
  }
}
```

Настройка для PostgreSQL

Установка драйвера JDBC для PostgreSQL

Архив драйвера JDBC для PostgreSQL идёт в комплекте поставки. Также может быть скачан с официального сайта PostgreSQL.

Копируем файл `postgresql-42.5.4.jar` в `/usr/share/logstash/logstash-core/lib/jars` .

Настройка конвейера для таблицы OrgEvents

Копируем файл `ltools-logs-pgsql.conf` из комплекта поставки в `/etc/logstash/ltools-logs.conf` . В этом файле необходимо заменить следующие параметры:

Параметр	Описание
`<DB-SERVER-HOST>`	Имя хоста или IP-адрес сервера СУБД
`<DB-USER-NAME>`	Имя пользователя СУБД
`<DB-USER-PASSWORD>`	Пароль пользователя СУБД
`<ES-HOST>`	Имя хоста или IP-адрес сервера Elasticsearch
`<ES-PORT>`	Порт сервера Elasticsearch, обычно 9200
`<ES-USER>`	Имя пользователя Elasticsearch, обычно `elastic`
`<ES-PASSWORD>`	Пароль пользователя Elasticsearch
`<SCHEDULE>`	Расписание запуска загрузки, например, `0 * * * * *` - каждую минуту

```
input {
  jdbc {
    jdbc_driver_class => "org.postgresql.Driver"
    jdbc_connection_string => "jdbc:postgresql://<DB-SERVER-HOST>:5432/ltoolslogs"
    jdbc_user => "<DB-USER-NAME>"
    jdbc_password => "<DB-USER-PASSWORD>"
    jdbc_paging_enabled => true
    jdbc_default_timezone => "UTC"
    tracking_column => "unix_ts_in_secs"
    use_column_value => true
    tracking_column_type => "numeric"
    schedule => "<SCHEDULE>"
    statement => "SELECT *
      ,EXTRACT(EPOCH FROM \"OrchTimestampUtc\") AS unix_ts_in_secs
    FROM
      public.\"Logs\"
    WHERE
      EXTRACT(EPOCH FROM \"OrchTimestampUtc\") > :sql_last_value
        AND \"OrchTimestampUtc\" < NOW()"
  }
}
filter {
  mutate {
    copy => { "id" => "[@metadata][_id]"}
    copy => { "orchtimestamputc" => "@timestamp" }
    remove_field => ["id", "@version", "unix_ts_in_secs", "orchtimestamputc", "signature"]
  }
}
output {
#  stdout { codec =>  "rubydebug"}
  elasticsearch {
    hosts => ["<ES-HOST>:<ES-PORT>"]
    ssl => true
    ssl_certificate_verification => false
    user => "<ES-USER>"
    password => "<ES-PASSWORD>"
    index => "ltools-logs-%{+YYYY-MM-dd}"
    manage_template => true
    template => "/etc/logstash/ltoolslogs-template.conf"
    template_name => "ltoolslogs"
    template_overwrite => true
  }
}
```

Настройка Logstash как сервиса

Настройка нескольких конвейеров

Файл `/etc/logstash/pipelines.yml` должен содержать описание активных конвейеров:

```
- pipeline.id: ltools-logs
  queue.type: persisted
  path.config: "/etc/logstash/ltools-logs.conf"
```

Управление службой Logstash

Запуск:

```
# systemctl start logstash.service
```

Перезапуск:

```
# systemctl restart logstash.service
```

Останов:

```
# systemctl stop logstash.service
```

Проверка корректности настройки

Файлы журналов Logstash находятся в каталоге `/var/log/logstash` , основной файл для мониторинга - `logstash-plain.log` .

При успешной работе загрузки данных в Kibana можно увидеть наличие индексов `ltools-logs-YYYY-MM-dd` . В настройках представления (view) Kibana необходимо указать шаблон индексов (index pattern) `ltools-*` .