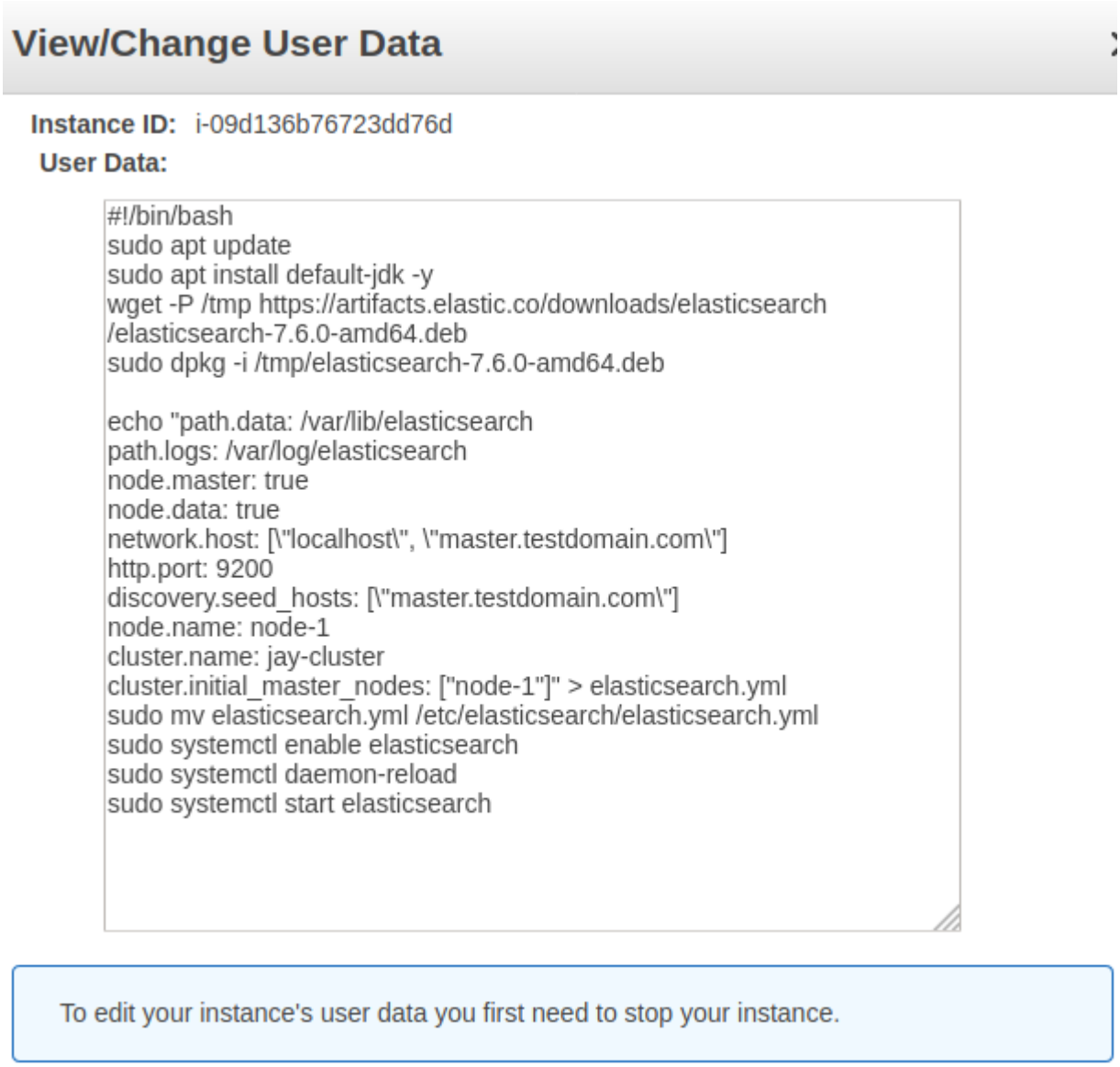
ELK Stack

1) Elasticsearch 2 node cluster(1st master+data, 2nd data+kibana node) setup through userdata



Master+Data node userdata

View/Change User Data

Instance ID: i-0010b837c4db9ad66

User Data:

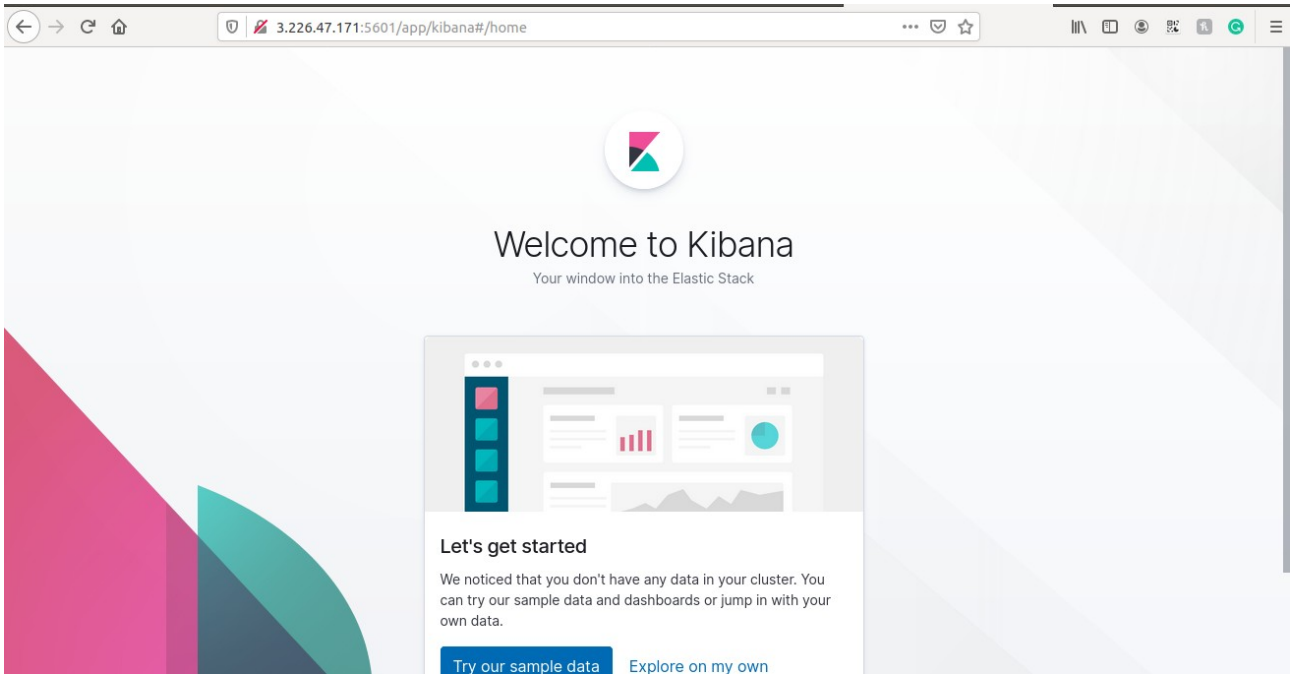
```
#!/bin/bash
sudo apt update
sudo apt install default-jdk -y
wget -P /tmp https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-
7.6.0-amd64.deb
sudo dpkg -i /tmp/elasticsearch-7.6.0-amd64.deb

echo "path.data: /var/lib/elasticsearch
path.logs: /var/log/elasticsearch
node.master: false
node.data: true
network.host: [\"localhost\", \"node.testdomain.com\"]
http.port: 9200
discovery.seed_hosts: [\"master.testdomain.com\"]
node.name: node-2
cluster.name: jay-cluster
cluster.initial_master_nodes: ["node-1"]" > elasticsearch.yml
sudo mv elasticsearch.yml /etc/elasticsearch/elasticsearch.yml
sudo systemctl enable elasticsearch
sudo systemctl daemon-reload
sudo systemctl start elasticsearch
wget https://artifacts.elastic.co/downloads/kibana/kibana-7.6.0-amd64.deb
dpkg -i kibana-7.6.0-amd64.deb
echo 'server.port: 5601
server.host: "0.0.0.0"
elasticsearch.hosts: ["http://localhost:9200"]' > kibana.yml
sudo mv kibana.yml /etc/kibana/kibana.yml
sudo systemctl enable kibana
sudo systemctl daemon-reload
sudo systemctl start kibana
```

Data node +kibana userdata

☐	testdomain.com.	NS	ns-1536.awsdns-00.co.uk. ns-0.awsdns-00.com. ns-1024.awsdns-00.org. ns-512.awsdns-00.net.	-
☐	testdomain.com.	SOA	ns-1536.awsdns-00.co.uk. awsdns-hostmaster.amaz	-
☐	master.testdomain.com.	A	172.31.24.67	-
☐	node.testdomain.com.	A	172.31.24.189	-

```
ubuntu@ip-172-31-24-189:/etc/kibana$ curl -GET http://localhost:9200/_cat/indices?v
health status index                    uuid                   pri rep docs.count docs.deleted store.size pri.store.size
green  open   .kibana_task_manager_1   Y-2dep5zTJSKbLYa6MwrPA   1   1          2            6     79.9kb         42.8kb
green  open   .apm-agent-configuration XZSoWAmnS-WQFBDNhPc1jQ   1   1          0            0       460b          230b
green  open   .kibana_1                LZtF2mgLTSWlvNdvCBZ9Vw   1   1          4            0       37kb         18.5kb
ubuntu@ip-172-31-24-189:/etc/kibana$
```

```
Last login: Fri Apr 24 00:18:13 2020 from 157.37.98.103
ubuntu@ip-172-31-24-189:~$ curl -GET http://localhost:9200/_cat/nodes?v
ip            heap.percent ram.percent cpu load_1m load_5m load_15m node.role master name
172.31.24.67            11          86   0    0.00    0.00     0.00 dim       *      node-1
172.31.24.189           13          86   0    0.00    0.03     0.09 dil       -      node-2
```



2) Write regex for Apache and Nginx logs
Install Apache and Nginx and enable error and access logs for both
Make separate conf for apache and nginx logs and include them in td-agent.conf
Parse the logs with proper access and error logs format (check on web for different fields on logs)
Make a separate index for nginx and apache logs on kibana also.

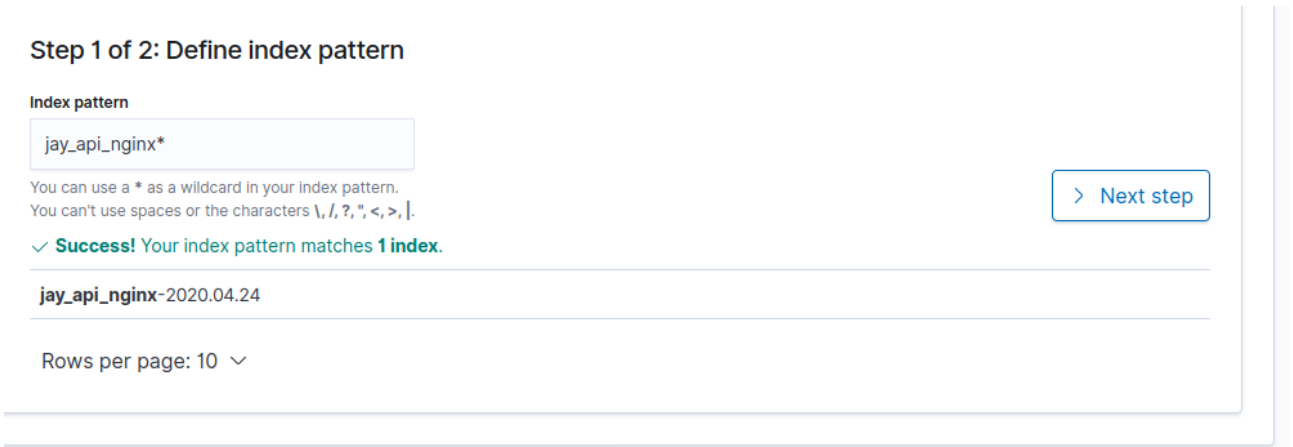
For nginx

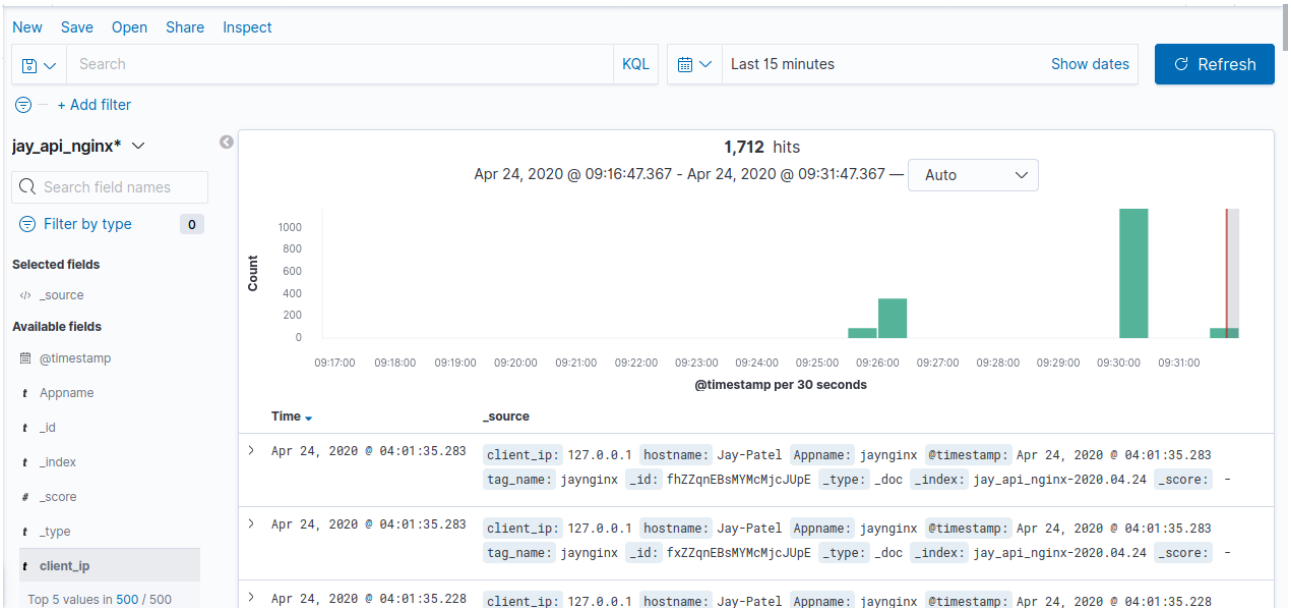
```
    ssl_prefer_server_ciphers on;

    ##
    # Logging Settings
    ##
    log_format jay '$remote_addr [$time_local]'' "$request" $status $body_bytes_sent';

    access_log /var/log/nginx/access.log;
    error_log /var/log/nginx/error.log;

    ##
    # Gzip Settings
    ##

    gzip on;

    # gzip_vary on;
    # gzip_proxied any;
```

```
jay@Jay-Patel:td-agent $ sudo cat /etc/nginx/sites-available/drupaljay
server {
        listen 127.0.0.1:80;

        root /var/www/drupal;

        index  index.php index.html index.htm;
        server_name drupaljay.com www.drupaljay.com;
         access_log /var/log/nginx/drupalaccess.log jay;


        location / {
                try_files $uri /index.php?$query_string;
    }

        location @rewrite {
               rewrite ^/(.*)$ /index.php?q=$1;
        }

         location ~ [^/]\.php(/|$) {
                 include snippets/fastcgi-php.conf;
                 fastcgi_pass unix:/var/run/php/php7.2-fpm.sock;
    }

    location ~ ^/sites/.*/files/styles/ {
               try_files $uri @rewrite;
        }

    location ~ ^(/[a-z\-]+)?/system/files/ {
        try_files $uri /index.php?$query_string;
        }



}
```

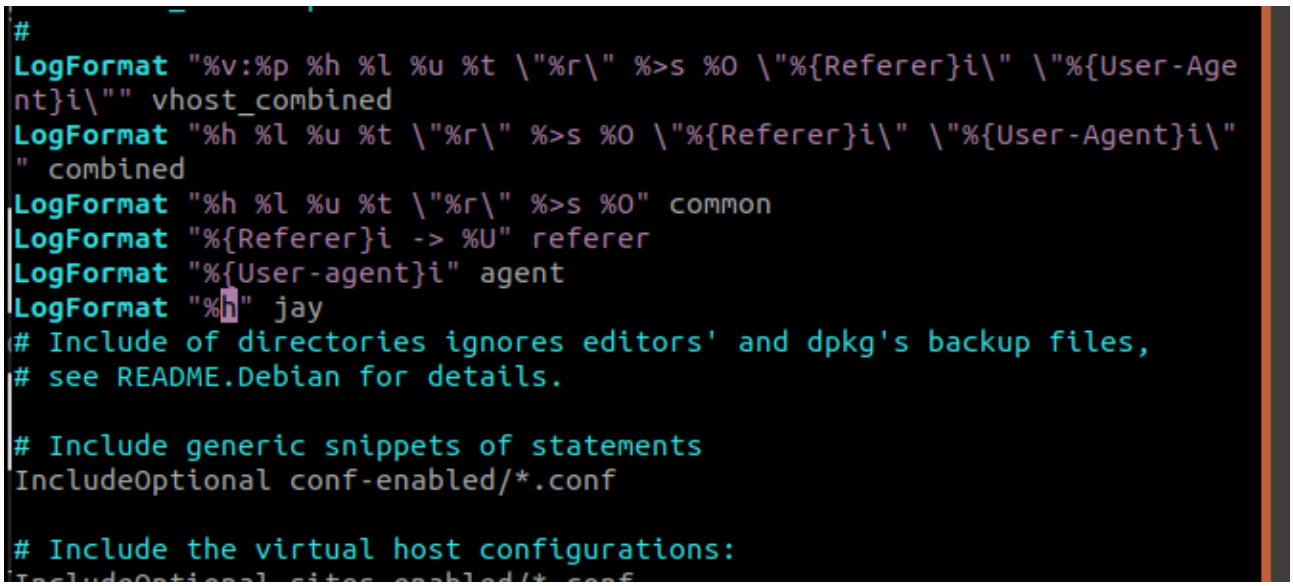
```
<source>
  @type tail
  read_from_head false
  pos_file /tmp/jaynginx.log
  path /var/log/nginx/drupalaccess.log
  <parse>
  @type regexp
  expression /^(?<client_ip>[0-9.]*)?/
  </parse>
 tag jayngin
</source>
<filter jaynginx>
 @type record_transformer
 remove_keys key_id
<record>
   hostname "#{Socket.gethostname}"
   Appname jaynginx
</record>
</filter>

<match jaynginx>
    @type forest
    subtype elasticsearch
      <template>
        host 192.168.225.53
        port 9200
        logstash_format true
        logstash_prefix jay_api_nginx
        include_tag_key true
           tag_key tag_name
          flush_interval 5s
```

```
</record>
</filter>

<match jaynginx>
    @type forest
    subtype elasticsearch
      <template>
        host 192.168.225.53
        port 9200
        logstash_format true
        logstash_prefix jay_api_nginx
        include_tag_key true
           tag_key tag_name
          flush_interval 5s
        buffer_type file
       # buffer_path /var/log/td-agent/buffer/${tag_parts[-2]}-buffer
        buffer_path /tmp/test-buffer1.log
        buffer_chunk_limit 100m
        buffer_queue_limit 256
        buffer_queue_full_action drop_oldest_chunk
        retry_wait 15.0
      </template>
</match>
```




New Save Open Share Inspect
Search
KQL
Last 15 minutes
Show dates
Refresh
+ Add filter
jay_api_nginx*
Search field names
Filter by type 0
Selected fields
_source
Available fields
@timestamp
Appname
_id
_index
_score
_type
client_ip
Top 5 values in 500 / 500
1,712 hits
Apr 24, 2020 @ 09:16:47.367 - Apr 24, 2020 @ 09:31:47.367 —
Auto
Count
1000
800
600
400
200
0
09:17:00 09:18:00 09:19:00 09:20:00 09:21:00 09:22:00 09:23:00 09:24:00 09:25:00 09:26:00 09:27:00 09:28:00 09:29:00 09:30:00 09:31:00
@timestamp per 30 seconds
Time
_source
Apr 24, 2020 @ 04:01:35.283 client_ip: 127.0.0.1 hostname: Jay-Patel Appname: jaynginx @timestamp: Apr 24, 2020 @ 04:01:35.283 tag_name: jaynginx _id: fhZZqnEBsMYMcMjcJUpE _type: _doc _index: jay_api_nginx-2020.04.24 _score: -
Apr 24, 2020 @ 04:01:35.283 client_ip: 127.0.0.1 hostname: Jay-Patel Appname: jaynginx @timestamp: Apr 24, 2020 @ 04:01:35.283 tag_name: jaynginx _id: fxZZqnEBsMYMcMjcJUpE _type: _doc _index: jay_api_nginx-2020.04.24 _score: -
Apr 24, 2020 @ 04:01:35.228 client_ip: 127.0.0.1 hostname: Jay-Patel Appname: jaynginx @timestamp: Apr 24, 2020 @ 04:01:35.228

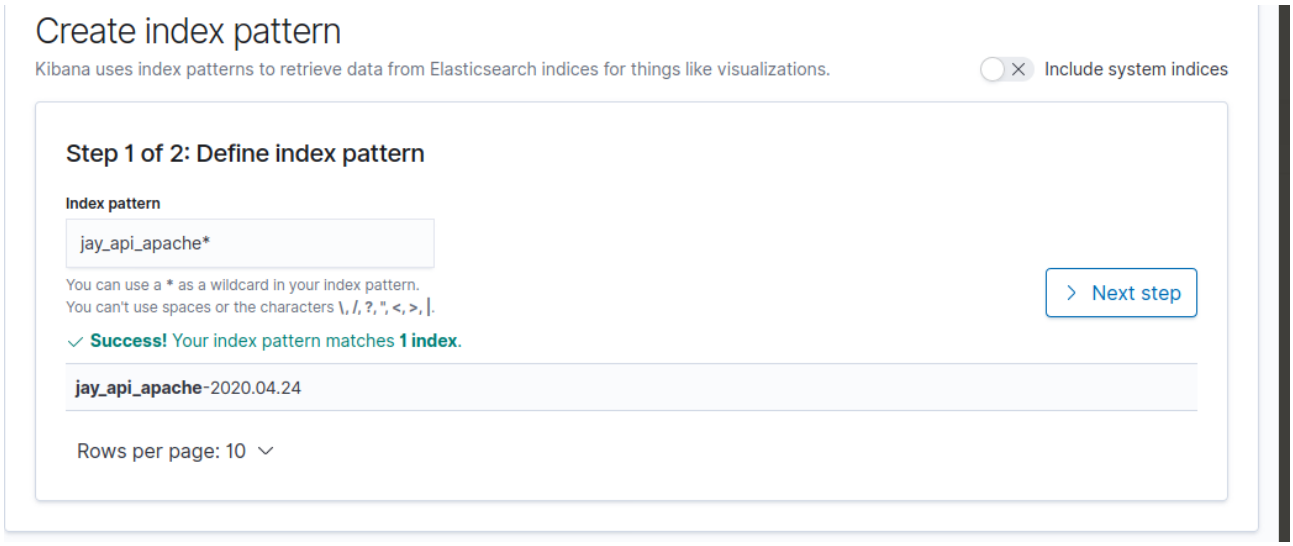
For Apache

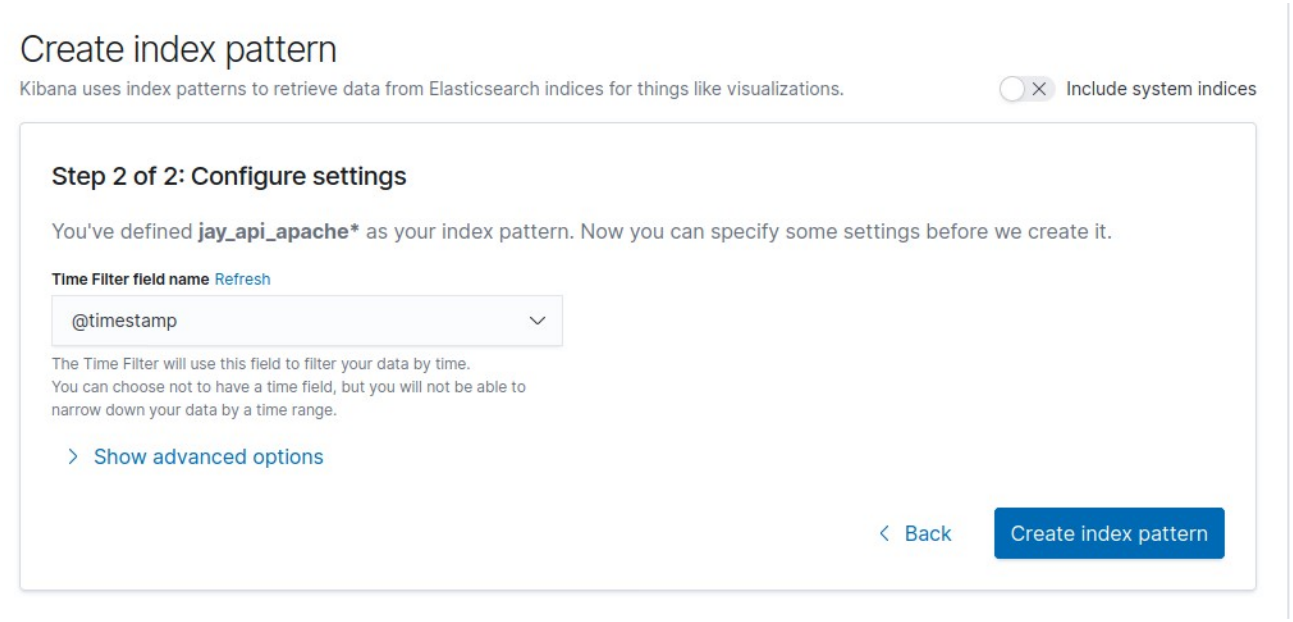

#
LogFormat "%v:%p %h %l %u %t \"%r\" %>s %O \"%{Referer}i\" \"%{User-Agent}i\"" vhost_combined
LogFormat "%h %l %u %t \"%r\" %>s %O \"%{Referer}i\" \"%{User-Agent}i\"" combined
LogFormat "%h %l %u %t \"%r\" %>s %O" common
LogFormat "%{Referer}i -> %U" referer
LogFormat "%{User-agent}i" agent
LogFormat "%h" jay
Include of directories ignores editors' and dpkg's backup files,
see README.Debian for details.
Include generic snippets of statements
IncludeOptional conf-enabled/*.conf
Include the virtual host configurations:

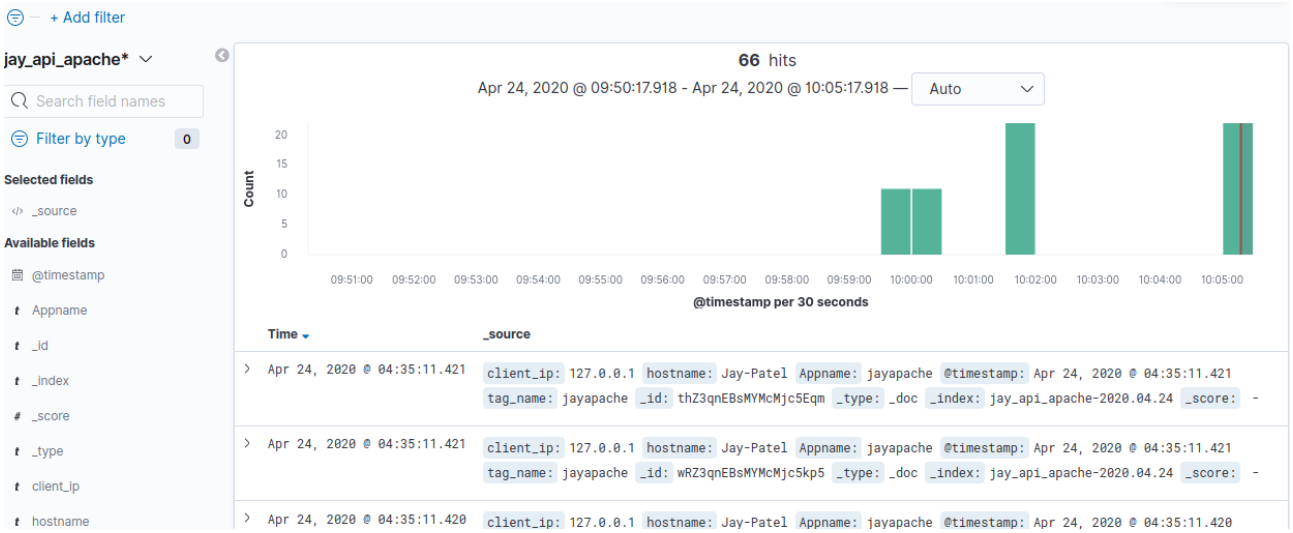
```
ServerAlias www.example.com
DirectoryIndex index.html index.php
DocumentRoot /var/www/wordpress/

CustomLog ${APACHE_LOG_DIR}/access.log jay
<Directory /var/www/wordpress/>
    Options FollowSymLinks
    AllowOverride All
    Require all granted
</Directory>
  RewriteEngine on
  RewriteOptions inherit


  # Catchall redirect to www.example1.com
  RewriteCond %{HTTP_HOST}   !^www.example\.com [NC]
  RewriteCond %{HTTP_HOST}   !^$
  RewriteRule ^/(.*)         https://www.example.com/$1 [L,R
VirtualHost>
```

```
<source>
  @type tail
  read_from_head false
  pos_file /tmp/jayapache.log
  path /var/log/apache2/access.log
  <parse>
  @type regexp
  expression /^(?<client_ip>[0-9.]*)?/
  </parse>
 tag jayapache
</source>
<filter jayapache>
 @type record_transformer
 remove_keys key_id
<record>
   hostname "#{Socket.gethostname}"
   Appname jayapache
</record>
</filter>

<match jayapache>
    @type forest
    subtype elasticsearch
      <template>
        host 192.168.225.53
        port 9200
        logstash_format true
        logstash_prefix jay_api_apache
        include_tag_key true
           tag_key tag_name
          flush_interval 5s
        buffer_type file
       # buffer_path /var/log/td-agent/buffer/${tag_parts[-2]}-buffer
```

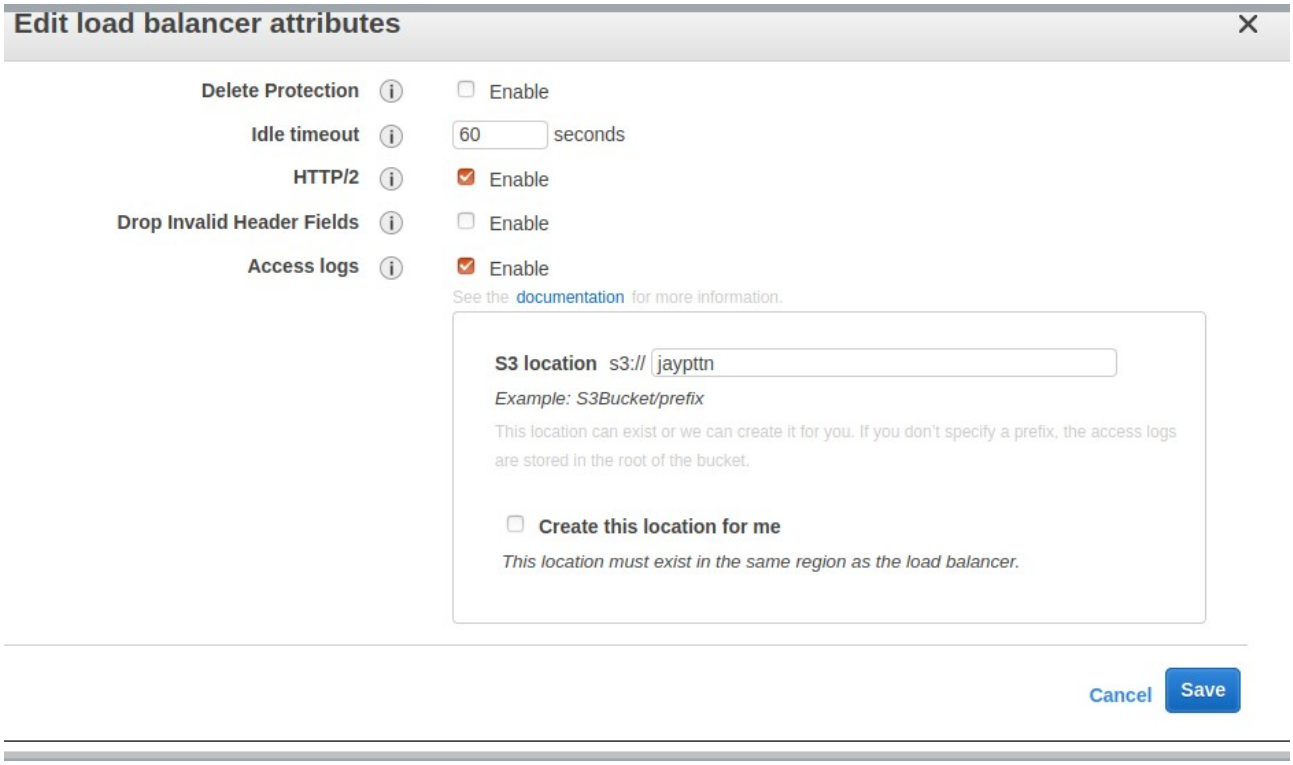
```
</filter>

<match jayapache>
    @type forest
    subtype elasticsearch
      <template>
        host 192.168.225.53
        port 9200
        logstash_format true
        logstash_prefix jay_api_apache
        include_tag_key true
           tag_key tag_name
          flush_interval 5s
        buffer_type file
       # buffer_path /var/log/td-agent/buffer/${tag_parts[-2]}-buffer
        buffer_path /tmp/test-apache.log
        buffer_chunk_limit 100m
        buffer_queue_limit 256
        buffer_queue_full_action drop_oldest_chunk
        retry_wait 15.0
      </template>
</match>
```



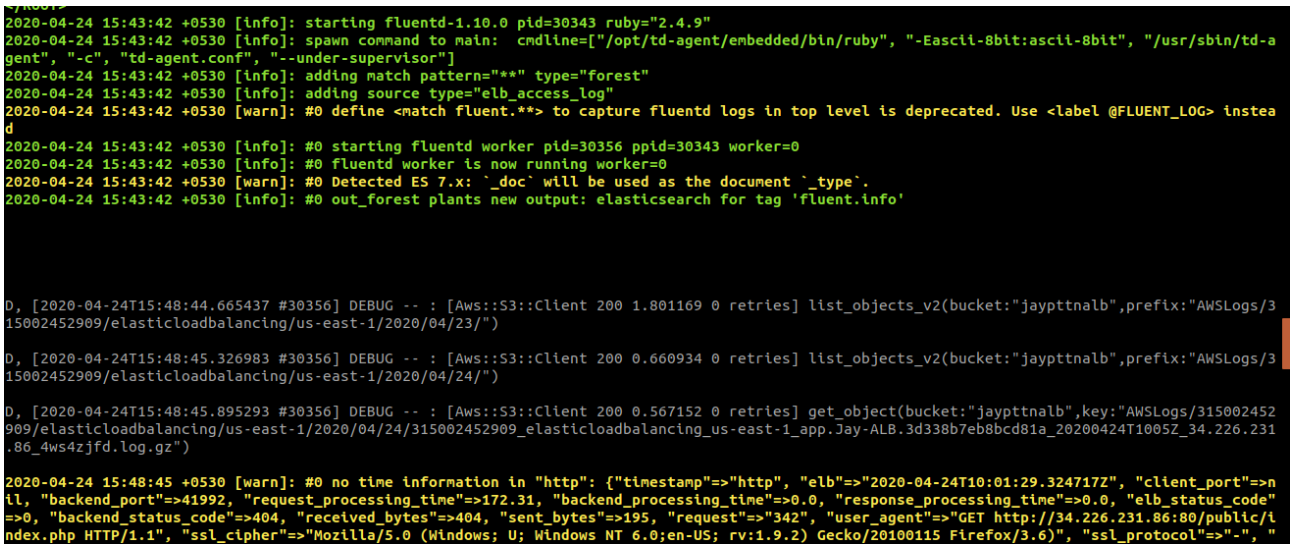


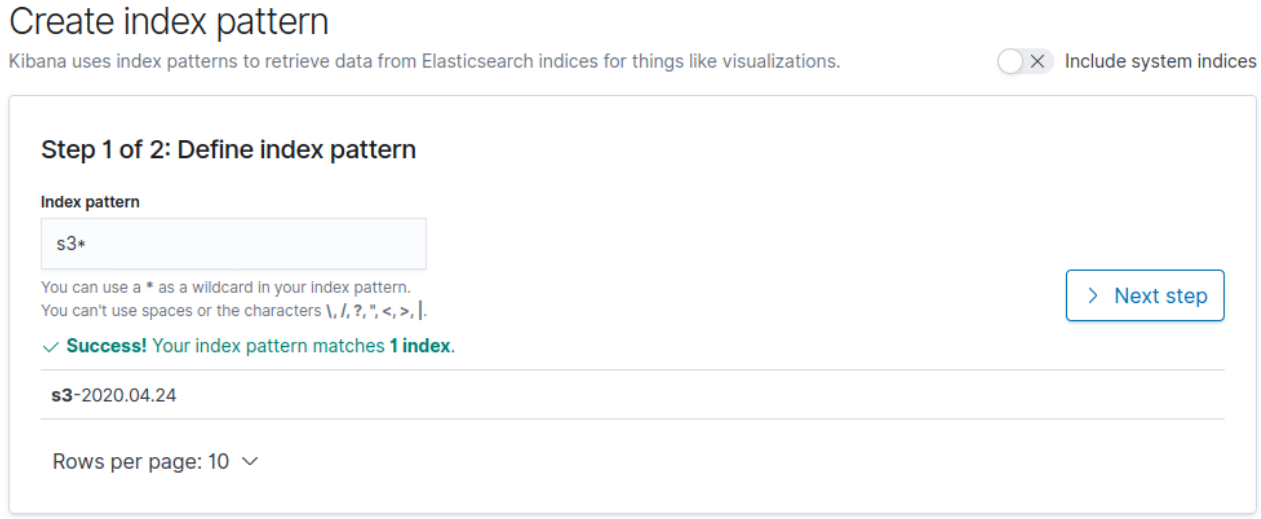


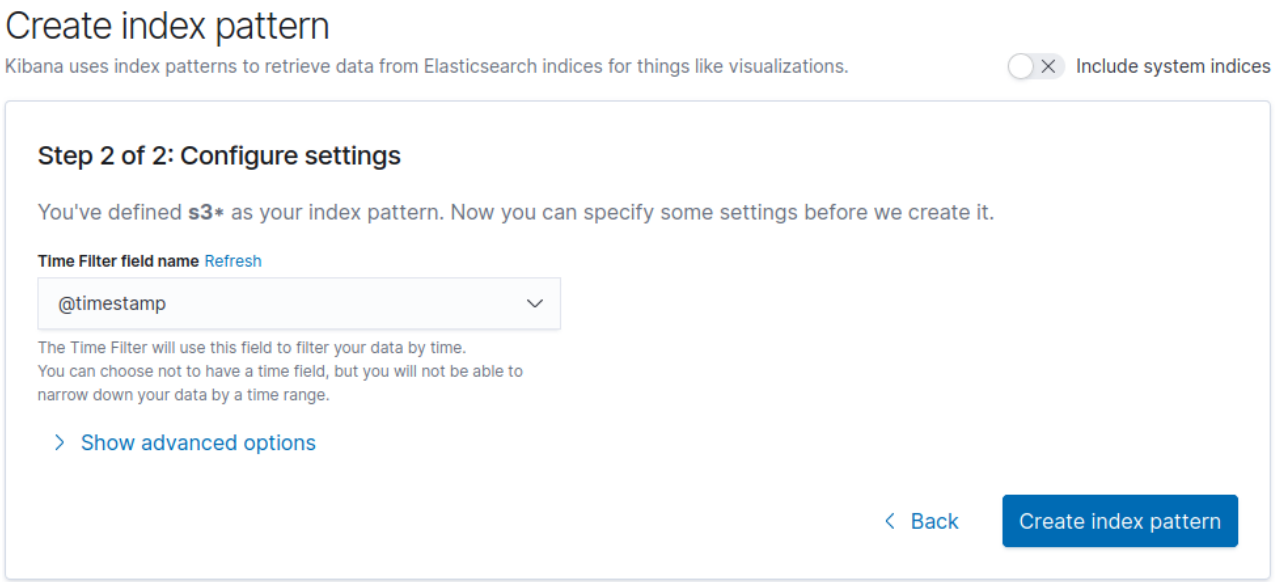
3) Alb logs to s3 then to elk
Enable alb logs on s3
Using td-agent send the logs on s3 to elk (search for this plugin elb-access-log plugin)

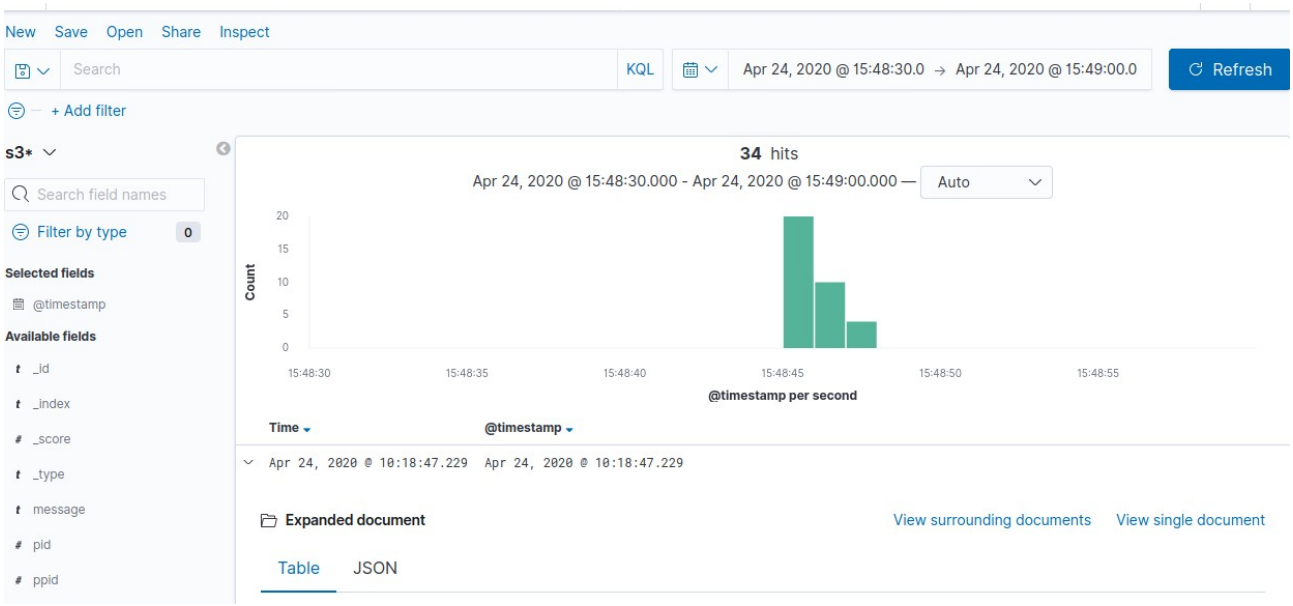



```
<source>
 @type elb_access_log
 account_id 315002452909 # required
 region us-east-1
 s3_bucket jaypttnalb
 tag elb.access_log
 debug true
</source>
<match **>
  @type forest
  subtype elasticsearch
    <template>
      host 192.168.225.53
      port 9200
      logstash_format true
      logstash_prefix s3
      include_tag_key true
            tag_key tag_name
    </template>
</match>
```


2020-04-24 15:43:42 +0530 [info]: starting fluentd-1.10.0 pid=30343 ruby="2.4.9"
2020-04-24 15:43:42 +0530 [info]: spawn command to main: cmdline=["/opt/td-agent/embedded/bin/ruby", "-Eascii-8bit:ascii-8bit", "/usr/sbin/td-agent", "-c", "td-agent.conf", "--under-supervisor"]
2020-04-24 15:43:42 +0530 [info]: adding match pattern="**" type="forest"
2020-04-24 15:43:42 +0530 [info]: adding source type="elb_access_log"
2020-04-24 15:43:42 +0530 [warn]: #0 define <match fluent.**> to capture fluentd logs in top level is deprecated. Use <label @FLUENT_LOG> instead
2020-04-24 15:43:42 +0530 [info]: #0 starting fluentd worker pid=30356 ppid=30343 worker=0
2020-04-24 15:43:42 +0530 [info]: #0 fluentd worker is now running worker=0
2020-04-24 15:43:42 +0530 [warn]: #0 Detected ES 7.x: `_doc` will be used as the document `_type`.
2020-04-24 15:43:42 +0530 [info]: #0 out_forest plants new output: elasticsearch for tag 'fluent.info'
D, [2020-04-24T15:48:44.665437 #30356] DEBUG -- : [Aws::S3::Client 200 1.801169 0 retries] list_objects_v2(bucket:"jaypttnalb",prefix:"AWSLogs/315002452909/elasticloadbalancing/us-east-1/2020/04/23/")
D, [2020-04-24T15:48:45.326983 #30356] DEBUG -- : [Aws::S3::Client 200 0.660934 0 retries] list_objects_v2(bucket:"jaypttnalb",prefix:"AWSLogs/315002452909/elasticloadbalancing/us-east-1/2020/04/24/")
D, [2020-04-24T15:48:45.895293 #30356] DEBUG -- : [Aws::S3::Client 200 0.567152 0 retries] get_object(bucket:"jaypttnalb",key:"AWSLogs/315002452909/elasticloadbalancing/us-east-1/2020/04/24/315002452909_elasticloadbalancing_us-east-1_app.Jay-ALB.3d338b7eb8bcd81a_20200424T1005Z_34.226.231.86_4ws4zjfd.log.gz")
2020-04-24 15:48:45 +0530 [warn]: #0 no time information in "http": {"timestamp"=>"http", "elb"=>"2020-04-24T10:01:29.324717Z", "client_port"=>nil, "backend_port"=>41992, "request_processing_time"=>172.31, "backend_processing_time"=>0.0, "response_processing_time"=>0.0, "elb_status_code"=>0, "backend_status_code"=>404, "received_bytes"=>404, "sent_bytes"=>195, "request"=>"342", "user_agent"=>"GET http://34.226.231.86:80/public/index.php HTTP/1.1", "ssl_cipher"=>"Mozilla/5.0 (Windows; U; Windows NT 6.0;en-US; rv:1.9.2) Gecko/20100115 Firefox/3.6)", "ssl_protocol"=>"-", "


Create index pattern
Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.
Include system indices
Step 1 of 2: Define index pattern
Index pattern
s3*
You can use a * as a wildcard in your index pattern.
You can't use spaces or the characters \, /, ?, ", <, >, |.
Next step
Success! Your index pattern matches 1 index.
s3-2020.04.24
Rows per page: 10


Create index pattern
Kibana uses index patterns to retrieve data from Elasticsearch indices for things like visualizations.
Include system indices
Step 2 of 2: Configure settings
You've defined s3* as your index pattern. Now you can specify some settings before we create it.
Time Filter field name Refresh
@timestamp
The Time Filter will use this field to filter your data by time.
You can choose not to have a time field, but you will not be able to narrow down your data by a time range.
Show advanced options
Back
Create index pattern


New Save Open Share Inspect
Search
KQL
Apr 24, 2020 @ 15:48:30.0 → Apr 24, 2020 @ 15:49:00.0
Refresh
+ Add filter
s3*
Search field names
Filter by type
0
Selected fields
@timestamp
Available fields
_id
_index
_score
_type
message
pid
ppid
34 hits
Apr 24, 2020 @ 15:48:30.000 - Apr 24, 2020 @ 15:49:00.000 —
Auto
Count
20
15
10
5
0
15:48:30
15:48:35
15:48:40
15:48:45
15:48:50
15:48:55
@timestamp per second
Time
@timestamp
Apr 24, 2020 @ 10:18:47.229
Apr 24, 2020 @ 10:18:47.229
Expanded document
View surrounding documents
View single document
Table
JSON