filebeat+logstash做CDN的ncache日志采集

0、环境说明：

环境搭建：https://download.zhoufengjie.cn/document/loganalysis/elasticsearch-logstash-kibana-install.pdf

filebeat+logstash做CDN的ncache日志采集上传到elasticsearch，通过kibana查看。

filebeat设备ip地址：192.168.0.108

logstash设备ip地址：192.168.0.108

elasticsearch设备ip地址：192.168.0.97

kibana设备ip地址：192.168.0.97

nginx的配置文件log格式定义为：

####

```
log_format filebeat_log '$msec $request_time $remote_addr $upstream_cache_status $status
$body_bytes_sent $request_method $scheme $host $uri $args $remote_user $upstream_addr
$upstream_status $upstream_response_time "$upstream_http_content_type" "$http_referer"
"$http_user_agent" "$http_cookie"';

access_log /var/log/nginx/access.log filebeat_log;
```

####

1、filebeat部署：

编辑vi /etc/filebeat/filebeat.yml输入以下内容：

####

```
filebeat.inputs:
- input_type: log
  paths:
    - /var/log/nginx/access*.log
  exclude_files: ['.gz$']
  tags: ["filebeat-nginx-accesslog"]
  document_type: nginxaccess
- input_type: log
  paths:
    - /var/log/nginx/error*.log
  tags: ["filebeat-nginx-errorlog"]
  exclude_files: ['.gz$']
  document_type: nginxerror
tags: ["filebeat-nginx-logs"]

filebeat.config.modules:
  path: ${path.config}/modules.d/*.yml
  reload.enabled: false

output.logstash:
  hosts: ["127.0.0.1:5044"]

#output:
##  console:
##    pretty: true
```

####

临时测试【测试的时候，可以把output改成console，然后看输出】

/usr/bin/filebeat -e -c /etc/filebeat/filebeat.yml

启动filebeat:

systemctl start filebeat

2、配置logstash:

编辑/etc/logstash/conf.d/nginx.conf配置一个专门采集nginx的logstash配置，用来把filebeat上传上来的日志进行字段过滤和拆解，输入如下内容:

####

```
input {
  beats {
    port => 5044
    host => "127.0.0.1"
  }
}

filter {
  if "filebeat-nginx-logs" in [tags] {
    if "filebeat-nginx-accesslog" in [tags] {
      grok {
        match => { "message" => ["%{DATA:[nginx][access][time]} %{DATA:[nginx][access][request_time]} %{IPORHOST:[nginx
        remove_field => "message"
        add_field => {"logtype"=>"nginxLogs"}
      }
      grok {
        match => {"[nginx][access][url]" =>  "%{URIPATH:api}"}
      }
      mutate {
        add_field => { "read_timestamp" => "%{@timestamp}" }
        convert => ["totaltime","float"]
      }
      #date {
      #  match => [ "[nginx][access][time]", "dd/MMM/YYYY:H:m:s Z" ]
      #  remove_field => "[nginx][access][time]"
      #}
      useragent {
        source => "[nginx][access][agent]"
        target => "[nginx][access][user_agent]"
        remove_field => "[nginx][access][agent]"
      }
      geoip {
        source => "[nginx][access][remote_ip]"
        target => "[nginx][access][geoip]"
      }
    }
    if "filebeat-nginx-errorlog" in [tags] {
      grok {
        match => { "message" => ["%{DATA:[nginx][error][time]} \[%{DATA:[nginx][error][level]}\] %{NUMBER:[nginx][error][pi
        remove_field => "message"
        add_field => {"logtype"=>"nginxLogs"}
      }
      mutate {
        rename => { "@timestamp" => "read_timestamp" }
      }
      date {
        match => [ "[nginx][error][time]", "YYYY/MM/dd H:m:s" ]
        remove_field => "[nginx][error][time]"
      }
    }
```

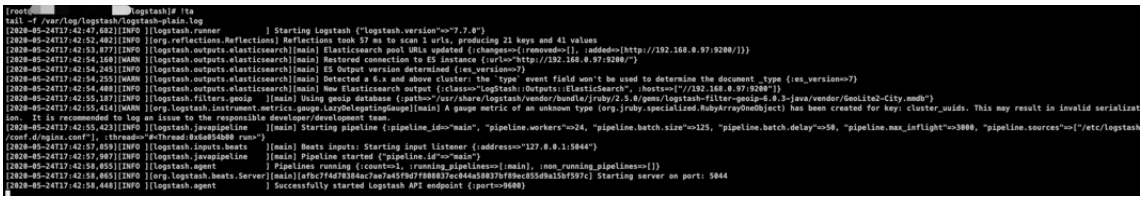
```
  }
}

output {
  if "filebeat-nginx-logs" in [tags] {
    if [logtype] == "nginxLogs" {
      elasticsearch {
        hosts => ["192.168.0.97:9200"]
        manage_template => false
        index => "%{[@metadata][beat]}-%{[@metadata][version]}-%{+YYYY.MM.dd}"
      }
      #stdout {
      #  codec => rubydebug
      #}
    }
  }
}
```

####

重启logstash：

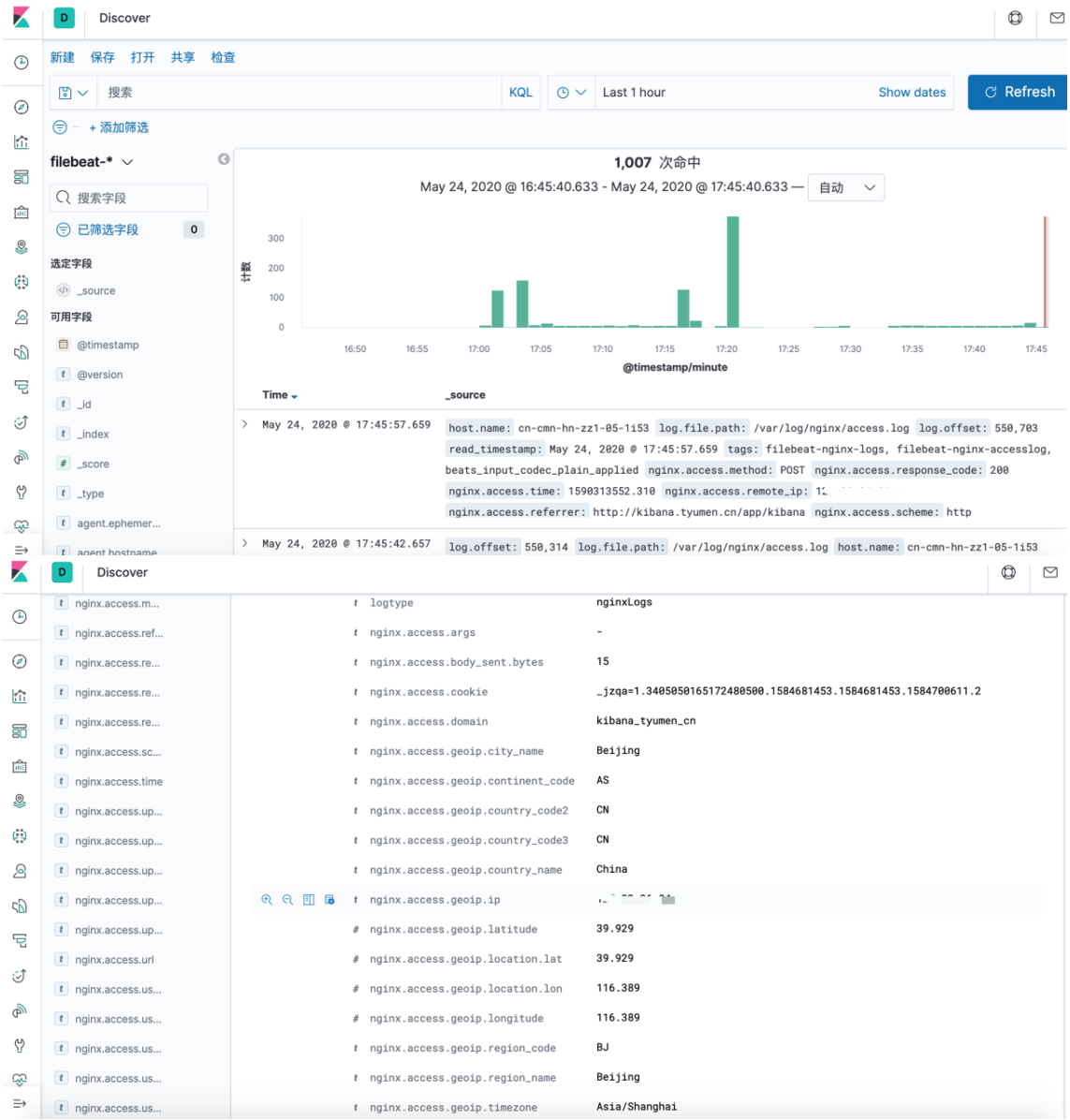
systemctl restart logstash



3、查看kibana日志

管理=>索引模式=>创建索引模式=>filebeat*

然后就可以进一步分析查看了，点击discover可以查看日志状态。



上传到elasticsearch上面的日志，通过kibana查看出来的格式如下：


```
{
  "_index": "filebeat-7.7.0-2020.05.24",
  "_type": "_doc",
  "_id": "hbwTRnIBvMHPzEOk43k7",
  "_version": 1,
  "_score": null,
  "_source": {
    "log": {
      "offset": 553424,
      "file": {
        "path": "/var/log/nginx/access.log"
      }
    },
    "host": {
      "name": "cn-cmn-hn-zz1-05-1i53"
    },
    "read_timestamp": "2020-05-24T09:46:46.666Z",
    "tags": [
      "filebeat-nginx-logs",
      "filebeat-nginx-accesslog",
      "beats_input_codec_plain_applied"
    ],
    "nginx": {
      "access": {
        "method": "POST",
        "response_code": "200",
        "time": "1590313604.749",
        "remote_ip": "125.34.23.95",
        "referrer": "http://kibana.tyumen.cn/app/kibana",
        "scheme": "http",
        "user_agent": {
          "name": "Chrome",
          "os": "Mac OS X",
```

```
          "os_name": "Mac OS X",
          "os_minor": "15",
          "major": "81",
          "os_major": "10",
          "device": "Other",
          "minor": "0",
          "build": "",
          "patch": "4044"
        },
        "cookie": "_jzqa=1.3405050165172480500.1584681453.1584681453.1584700611.2",
        "body_sent": {
          "bytes": "15"
        },
        "url": "/api/ui_metric/report",
        "domain": "elk.test.tyumen.cn",
        "user_name": "ngaa",
        "request_time": "0.815",
        "upstream": {
          "cache_status": "-",
          "response_time": "0.798",
          "response_code": "200",
          "upstream_ip": "192.168.0.97:5601",
          "content_type": "application/json; charset=utf-8"
        },
        "args": "-",
        "geoip": {
          "country_code3": "CN",
          "region_name": "Beijing",
          "continent_code": "AS",
          "longitude": 116.3889,
          "location": {
            "lat": 39.9288,
            "lon": 116.3889
          },
          "latitude": 39.9288,
          "timezone": "Asia/Shanghai",
          "country_code2": "CN",
          "region_code": "BJ",
          "ip": "125.33.26.94",
          "country_name": "China",
          "city_name": "Beijing"
        }
      }
    },
    "@version": "1",
    "ecs": {
      "version": "1.5.0"
    },
    "logtype": "nginxLogs",
    "@timestamp": "2020-05-24T09:46:46.666Z",
    "agent": {
      "ephemeral_id": "83799e29-f72d-47f7-979e-0776efddf80a",
      "type": "filebeat",
      "hostname": "cn-cmn-hn-zz1-05-1i53",
      "id": "39d4d5cf-f746-4cd8-b4ed-3c1d63d64edd",
      "version": "7.7.0"
    },
    "api": "/api/ui_metric/report"
  },
  "fields": {
    "read_timestamp": [
      "2020-05-24T09:46:46.666Z"
    ],
    "@timestamp": [
      "2020-05-24T09:46:46.666Z"
    ]
  },
  "sort": [
    1590313606666
  ]
}
```