Filebeat实时收集Nginx日志

JoyLau 2018-05-08

说明

1. Filebeat 版本为 5.3.0
   之所以使用 beats 家族的 Filebeat 来替代 Logstash 是因为 Logstash 实在太消耗资源了（服务器资源充足的土豪请无视）
   在官网下载 Logstash 有 89M，而 Filebeat 才8.4M，由此可见一斑
   Logstash 可以配置 jvm 参数，经过我本身的调试，内存分配小了，启动很慢有时根本起不来，分配大了，其他服务就没有资源了
   所有说对于配置低的服务器，选择 Filebeat 是最好的选择了，而且现在 Filebeat 已经开始替代 Logstash 了
2. 依然需要修改 nginx 的日志格式

nginx.config

更改日志记录的格式

```
log_format json '{ "@timestamp": "$time_iso8601", '
                        '"time": "$time_iso8601", '
                        '"remote_addr": "$remote_addr", '
                        '"remote_user": "$remote_user", '
                        '"body_bytes_sent": "$body_bytes_sent", '
                        '"request_time": "$request_time", '
                        '"status": "$status", '
                        '"host": "$host", '
                        '"request": "$request", '
                        '"request_method": "$request_method", '
                        '"uri": "$uri", '
                        '"http_referrer": "$http_referer", '
                        '"body_bytes_sent":"$body_bytes_sent", '
                        '"http_x_forwarded_for": "$http_x_forwarded_for", '
                        '"http_user_agent": "$http_user_agent" '
                    '}';

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

filebeat.yml

```


filebeat.prospectors:

- input_type: log

```

```
    paths:
      - /var/log/nginx/*access*.log
    json.keys_under_root: true
    json.overwrite_keys: true


output.elasticsearch:

    hosts: ["ip:port","ip:port"]
    index: "filebeat_server_nginx_%{+YYYY-MM}"
```

这里面需要注意的是
json.keys_under_root： 默认这个值是FALSE的，也就是我们的json日志解析后会被放在json键上。设为TRUE，所有的keys就会被放到根节点
json.overwrite_keys: 是否要覆盖原有的key，这是关键配置，将keys_under_root设为TRUE后，再将overwrite_keys也设为TRUE，就能把filebeat默认的key值给覆盖了

还有其他的配置
json.add_error_key：添加json_error key键记录json解析失败错误
json.message_key：指定json日志解析后放到哪个key上，默认是json，你也可以指定为log等。

说白了，差别就是，未配置前elasticsearch的数据是这样的：


```
{
        "_index": "filebeat_server_nginx_2018-05",
        "_type": "log",
        "_id": "AWM9sVOkCcRcg0IPg399",
        "_version": 1,
        "_score": 1,
        "_source": {
                "@timestamp": "2018-05-08T03:00:17.544Z",
                "beat": {
                        "hostname": "VM_252_18_centos",
                        "name": "VM_252_18_centos",
                        "version": "5.3.0"
                },
                "input_type": "log",
                "json": {},
                "message": "{ "@timestamp": "2018-05-08T11:00:11+08:00", "time": "2018-05-08T11:00:11+08:00", "
                "offset": 7633,
                "source": "/var/log/nginx/access.log",
                "type": "log"
        }
}
```


配置后，是这样的：


```
{
        "_index": "filebeat_server_nginx_2018-05",
        "_type": "log",
        "_id": "AWM9rjLd8mVZNgvhdnN9",
        "_version": 1,
```

```
        "_score": 1,
        "_source": {
                "@timestamp": "2018-05-08T02:56:50.000Z",
                "beat": {
                        "hostname": "VM_252_18_centos",
                        "name": "VM_252_18_centos",
                        "version": "5.3.0"
                },
                "body_bytes_sent": "12576",
                "host": "blog.joylau.cn",
                "http_referrer": "http://blog.joylau.cn/",
                "http_user_agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like G
                "http_x_forwarded_for": "-",
                "input_type": "log",
                "offset": 3916,
                "remote_addr": "60.166.12.138",
                "remote_user": "-",
                "request": "GET /2018/03/01/JDK8-Stream-Distinct/ HTTP/1.1",
                "request_method": "GET",
                "request_time": "0.000",
                "source": "/var/log/nginx/access.log",
                "status": "200",
                "time": "2018-05-08T10:56:50+08:00",
                "type": "log",
                "uri": "/2018/03/01/JDK8-Stream-Distinct/index.html"
        }
}
```


这样看起来就很舒服了

启动 FileBeat

进入 Filebeat 目录

```
nohup sudo ./filebeat -e -c filebeat.yml >/dev/null 2>&1 &
```

更新

nginx 的日志里含有中文的话，会将中文转为 Unicode 编码，如果不转的话，加入 `escape=json` 参数就可以了

```
log_format json escape=json '{ "@timestamp": "$time_iso8601", '
                            '"time": "$time_iso8601", '
                            '"remote_addr": "$remote_addr", '
                            '"remote_user": "$remote_user", '
                            '"body_bytes_sent": "$body_bytes_sent", '
                            '"request_time": "$request_time", '
                            '"status": "$status", '
                            '"host": "$host", '
                            '"request": "$request", '
                            '"request_method": "$request_method", '
                            '"uri": "$uri", '
                            '"http_referrer": "$http_referer", '
                            '"body_bytes_sent":"$body_bytes_sent", '
                            '"http_x_forwarded_for": "$http_x_forwarded_for", '
                            '"http_user_agent": "$http_user_agent" '
                        '}';

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