链滴

EFK 实战二 - 日志集成

作者：jianzh5

原文链接：https://ld246.com/article/1578558894406

来源网站：链滴

前言

在EFK基础架构中，我们需要在客户端部署Filebeat，通过Filebeat将日志收集并传到LogStash中。在ogStash中对日志进行解析后再将日志传输到ElasticSearch中，最后通过Kibana查看日志。

上文EFK实战一 - 基础环境搭建已经搭建好了EFK的基础环境，本文我们通过真实案例打通三者之间数据传输以及解决EFK在使用过程中的一些常见问题。

首先看一下实际的业务日志

```
2020-01-09 10:03:26,719 INFO ========GetCostCenter Start===============
2020-01-09 10:03:44,267 WARN 成本中心编码少于10位！{"deptId":"D000004345","companyCo
e":"01"}
2020-01-09 10:22:37,193 ERROR java.lang.IllegalStateException: SessionImpl[abcpI7fK-WYnW
nzXrv7w,]: can't call getAttribute() when session is no longer valid.
    at com.caucho.server.session.SessionImpl.getAttribute(SessionImpl.java:283)
    at weaver.filter.PFixFilter.doFilter(PFixFilter.java:73)
    at com.caucho.server.dispatch.FilterFilterChain.doFilter(FilterFilterChain.java:87)
    at weaver.filter.MonitorXFixIPFilter.doFilter(MonitorXFixIPFilter.java:30)
    at weaver.filter.MonitorForbiddenUrlFilter.doFilter(MonitorForbiddenUrlFilter.java:133)
```

日志组成格式为：

时间 日志级别 日志详情

主要任务就是将这段日志正常写入EFK中。

filebeat安装配置

- 下载 filebeat7.5.1
- 将下载后的文件上传至服务器并解压

tar -zxvf filebeat-7.5.1-linux-x86_64.tar.gz

- 修改filebeat.yml,

```
filebeat.inputs:
- type: log
  enabled: true
  paths:
    - /app/weaver/Resin/log/xxx.log
```

此段配置日志输入，指定日志存储路径

```
output.logstash:
  # The Logstash hosts
  hosts: ["172.31.0.207:5044"]
```

此段配置日志输出，指定Logstash存储路径

- 启动filebeat

./filebeat -e -c filebeat.yml

如果需要静默启动，则使用nohup ./filebeat -e -c filebeat.yml & 命令启动即可

logstash配置

logstash的配置主要分为三段input，filter，output。

input用于指定输入，主要是开放端口给Filebeat用于接收日志

filter用于指定过滤，对日志内容进行解析过滤。

output用于指定输出，直接配置ES的地址即可

```
input {
  beats {
    port => 5044
  }
}

output {
  elasticsearch {
    hosts => ["http://172.31.0.127:9200"]
    index => "myindex-%{+YYYY.MM.dd}"
    user => "elastic"
    password => "xxxxxx"
  }
}
```

我们配置好logstash后通过命令重启logstash

docker-compose -f elk.yml restart logstash

经过上述两步配置后应用程序往日志文件写入日志，filebeat会将日志写入logstash。在kibana查看入的日志结果如下：

日志显示有2个问题：

- 由于错误日志堆栈信息有多行，在kibana中展示成了多行，数据查看很乱。需要将堆栈异常整理成行显示。
- 需要对日志进行解析，拆成“时间 日志级别 日志详情”的显示格式。

优化升级

- 在filebeat中设置合并行

filebeat默认是行传输的，但是我们的日志肯定是多行一个日志，我们要把多行合并到一起就要找到志的规律。比如我们的日志格式全都是以时间格式开头，所以我们在filebeat中filebeat.inputs区域添如下几行配置

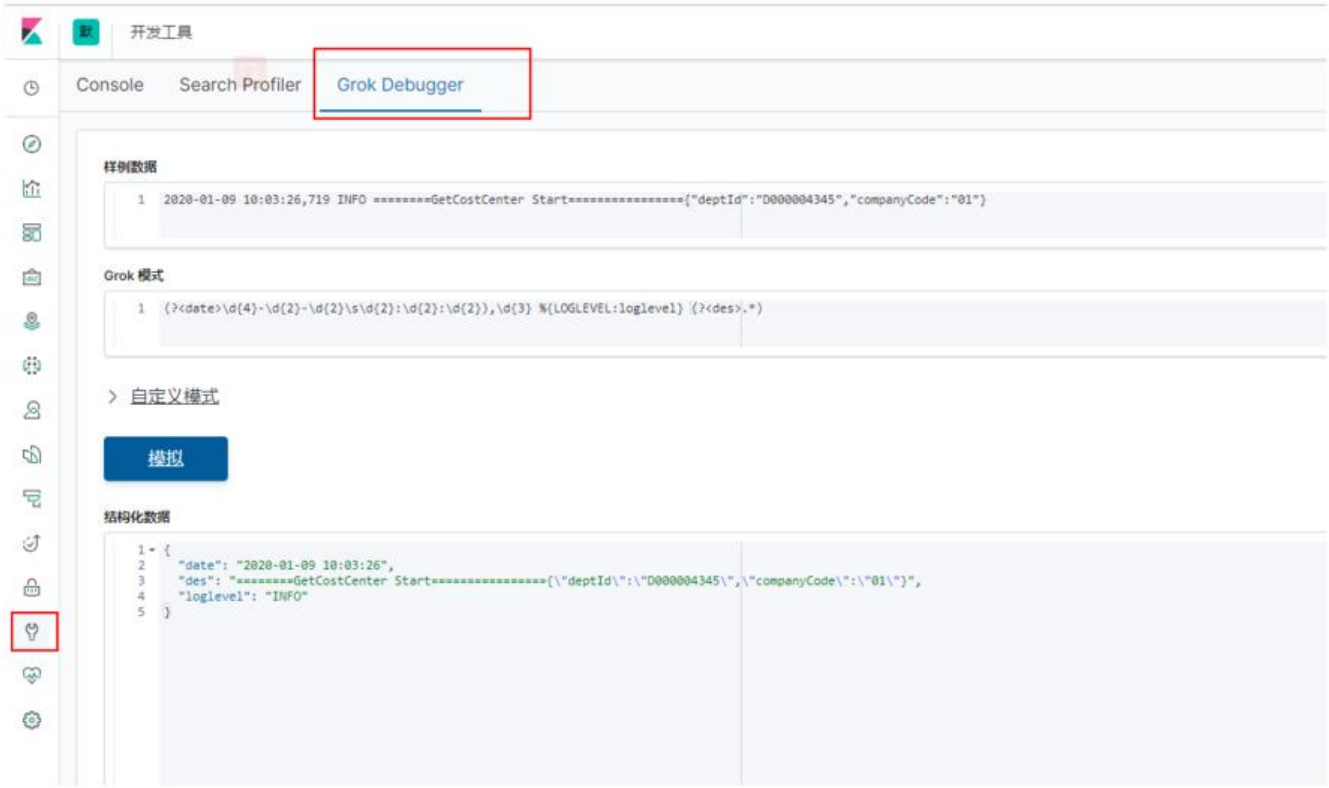
```
# 以日期作为前缀
multiline.pattern: ^\d{4}-\d{1,2}-\d{1,2}
# 开启多行合并
multiline.negate: true
# 合并到上一行之后
multiline.match: after
```

- 在logstash中设置对日志的解析

将日志解析成“时间 日志级别 日志详情”的展示格式，所以我们需要在logstash配置文件中添加filte段

```
filter {
   grok{
      match => {
         "message" => "(?<date>\d{4}-\d{2}-\d{2}\s\d{2}:\d{2}:\d{2}),\d{3} %{LOGLEVEL:loglevel
 (?<des>.*)"
      }
   }
}
```

这里主要是使用grok语法对日志进行解析，通过正则表达式对日志进行过滤。大家可以通过kibana的grok调试工具进行调试



配置完成后我们重新打开kibana Discover界面查看日志，符合预期，完美!

常见问题

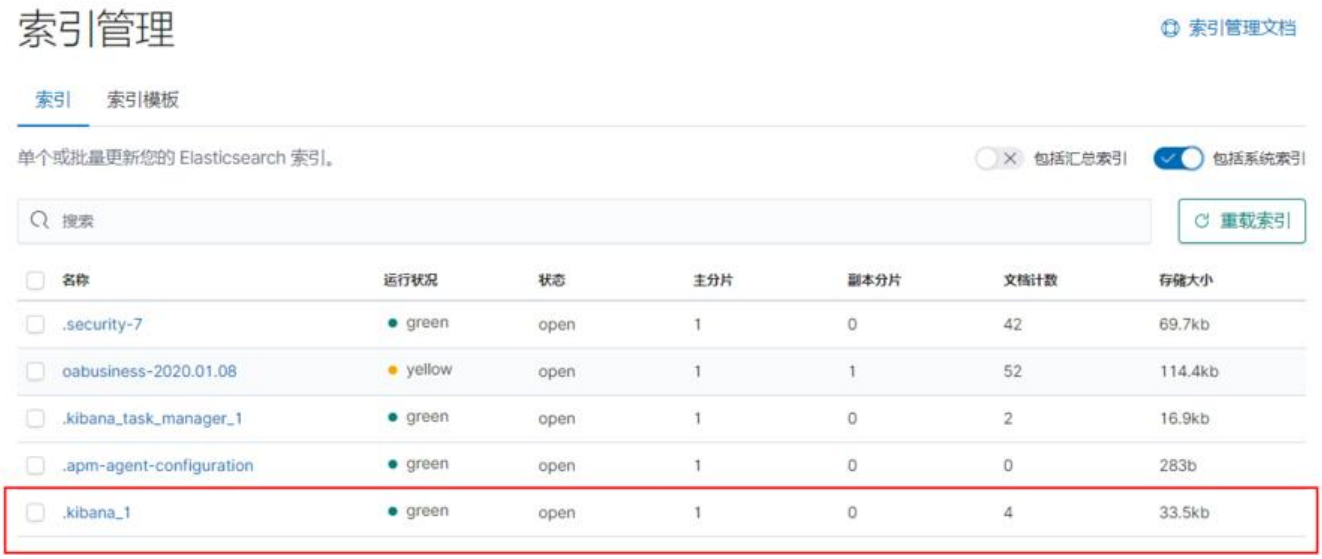
kibana 乱码

这个主要原因还是客户端日志文件格式有问题，大家可以通过file xxx.log查看日志文件的编码格式，果是ISO8859的编码基本都会乱码，我们可以在filebeat配置文件中通过encoding指定日志编码进行输。

```
filebeat.inputs:
- type: log
  enabled: true
  paths:
    - /app/weaver/Resin/log/xxx.log
  encoding: GB2312
```

kibana 提取字段出错

如上所示，打开kibana Discover面板时出现此异常，大家只要删除ES中的.kibana_1索引然后重新访Kibana即可。



查看周围文件

我们在终端查看日志某关键字时一般会查上下文信息便于排查问题，如经常用到的指令cat xxx.log | grep -C50 keyword，那么在Kibana中如何实现这功能呢。

在Kibana中搜索关键字，然后找到具体日志记录，点击左边向下箭头，然后再点击“查看周围文档”可实现。

原文链接：EFK 实战二 - 日志集成

动态索引

我们日志平台可能需要对接多个业务系统，需要根据业务系统建立不同的索引。

- 在filebeat中给日志打上标记

```
- type: log
  ......
  fields:
    logType: oabusiness
```

- 在logstash中根据标记生成索引

```
input {
  beats {
    port => 5044
  }
}
filter {
  if [fields][logType] == "oabusiness" {
        grok{
            match => {
                "message" => "(?<date>\d{4}-\d{2}-\d{2}\s\d{2}:\d{2}:\d{2}),\d{3} %{LOGLEVEL:logle
el} (?<des>.*)"
            }
        }
    }
}
output {
    elasticsearch {
        hosts => ["http://172.31.0.207:9200"]
        index => "%{[fields][logType]}-%{+YYYY.MM.dd}"
        user => "elastic"
        password => "elastic"
    }
}
```

好了，各位朋友们，本期的内容到此就全部结束啦，能看到这里的同学都是优秀的同学，下一个升职薪的就是你了！

如果觉得这篇文章对你有所帮助的话请扫描下面二维码加个关注。"转发" 加 "在看"，养成好习惯！咱下期再见！