Filebeat pipeline Dissect doc

● Subject

Dissect instructions and yml Important reminder

● Yml Important reminder

1. Case sensitive
2. Use indentation to represent hierarchical relationships
3. Tabs are not allowed when indenting, only spaces are allowed.
4. The number of indented spaces is not important, as long as elements of the same level are aligned to the left

● Sample pipeline yml

```
- type: log
  paths:
    -   "${PWD}/sampleLogs/FullRecord.log"
  fields_under_root: true
  scan_frequency: 3s
  idle_timeout: 3s
  pipeline: awc_fullrecordapi
  ignore_older: 2h
  close_inactive: 1h
  clean_inactive: 3h
  close_removed: true
  clean_removed: true
  processors:
    - dissect:
        tokenizer: "[%{logdate}][%{class}][%{severity}][%{thread}]
[Status]:%{Status},[Agent_ID]:%{Agent_ID},[Request]:%{Request},[Response]:%{Response}"
        field: "message"
        target_prefix: ""
    - drop_event:
        when:
          has_fields: ["log.flags"]
```

● Paths in pipeline yml

If there are multiple paths in Paths, it is recommended to separate different yml for future debugging and maintenance . The following is an example

Multiple paths

```
$ vi awc_fullrecordapi.yml
- type: log
  paths:
    -   /var/data/log/awc_fullrecordapi/*.log
    -   /var/data/log1/awc_fullrecordapi/*.log
  fields_under_root: true
  scan_frequency: 3s
  idle_timeout: 3s
  pipeline: awc_fullrecordapi
```

Separate yml

```
$ vi awc_fullrecordapi.yml
- type: log
  paths:
    -   /var/data/log/awc_fullrecordapi/*.log
  fields_under_root: true
  scan_frequency: 3s
  idle_timeout: 3s
  pipeline: awc_fullrecordapi
$ vi awc_fullrecordapi1.yml
- type: log
  paths:
    -   /var/data/log1/awc_fullrecordapi/*.log
  fields_under_root: true
  scan_frequency: 3s
  idle_timeout: 3s
  pipeline: awc_fullrecordapi
```

● Dissect

1. One of the Processors used by Filebeat to cut logs
2. Dissect mainly cuts out the key through% {key_name}, and the corresponding content is the value of this key
3. Tips for cutting the log: do not need to cut the text or special characters in the log, please write it into the dissect processor

The following demonstrates a log dissect processor

Log

```
[2020/02/11  10:00:00.175][BaseService.java:2566][ERROR][http-nio-1213-exec-306]  [Status]:{"status":"1028"},[Agent_ID]:002,[Request]:{"Method":"POST","IP":"10.10.80.2"},[Response]:{"status":"1028"}
```

Dissect processor

```
[%{logdate}][%{class}][%{severity}][%{thread}]
[Status]:%{status},[Agent_ID]:%{agentId},[Request]:%{request},[Response]:%{response}
```

● Debugging procedure

First comment out the drop event of test pipeline yml

```
- type: log
   paths:
      -   "${PWD}/sampleLogs/FullRecord.log"
   fields_under_root: true
   scan_frequency: 3s
   idle_timeout: 3s
   pipeline: awc_fullrecordapi
   ignore_older: 2h
   close_inactive: 1h
   clean_inactive: 3h
   close_removed: true
   clean_removed: true
   processors:
      - dissect:
            tokenizer: "[%{logdate}][%{class}][%{severity}][%{thread}]
[Status]:%{Status},[Agent_ID]:%{Agent_ID},[Request]:%{Request},[Response]:%{Response}"
            field: "message"
            target_prefix: ""
#       - drop_event:
#             when:
#                has_fields: ["log.flags"]
```

Secondly, rewrite the path in test.yml to the pipeline yml path to be tested

```
filebeat.config.inputs:
  enabled: true
  path: /home/elk/filebeat/CustomConfig/inputs.d/test.yml
output.console:
  pretty: true
```

Finally use command ***./filebeat -c CustomConfig/test.yml*** to see debug message
if you see the message like following. it is mean your dissect not success cut log

```
{
  "@timestamp": "2020-04-10T12:28:54.081Z",
  "@metadata": {
    "beat": "filebeat",
    "type": "_doc",
    "version": "7.3.2",
    "pipeline": "ckf_manager"
  },
  "ecs": {
    "version": "1.0.1"
  },
  "message": "kqpowkepowqkepoqkwepo12po21po,qwpd[2019/12/13 23:22:20.969][Match
BO.java.closeToSettled:693][INFO][QuartzScheduler_Worker-17] [Function]:Change Match St
atus-Settle Match(Match ID=865445),[Action]:Edit,[Website]:null,[Previous Value]:96,[Curren
t Value]:288,[Update By]:System",
  "log": {
    "offset": 0,
    "file": {
      "path": "/var/data/log/ckf_manager/2.log"
    },
    "flags": [
      "dissect_parsing_error"
    ]
  },
  "input": {
    "type": "log"
  },
```

Version control

In order to enhance the synchronization of the Filebeat settings of each project and the Elasticsearch pipeline of the ITIG team. In the future, Filebeat will add the version number.

```
- type: log
  paths:
    -   "${PWD}/sampleLogs/FullRecord.log"
  fields:
    version: "1.0"
  fields_under_root: true
  scan_frequency: 3s
  idle_timeout: 3s
  pipeline: awc_fullrecordapi
  ignore_older: 2h
  close_inactive: 1h
  clean_inactive: 3h
  close_removed: true
  clean_removed: true
  processors:
    - dissect:
          tokenizer: "[%{logdate}][%{class}][%{severity}][%{thread}]
[Status]:%{Status},[Agent_ID]:%{Agent_ID},[Request]:%{Request},[Response]:%{Response}"
          field: "message"
          target_prefix: ""
    - drop_event:
```

If each project needs to modify the filebeat config. Please be sure to inform the ITIG team to modify the version number to avoid the log from failing to parse

Log add field process

If project team need to add new fields to the log , please notify Polo first and then Polo will open the Jira demand list. After the ITIG team assessed that it was feasible. The log will be tested according to the environment dev-> stg-> prod order

1. The project team has the need to add new fields in the log
2. Make a request to Polo and evaluate it
3. After Polo assessment is feasible, open a demand list through Jira to the ITIG team for evaluation
4. After evaluation, the ITIG team will start testing with the project leader in the dev environment
5. After the test is completed, it will be updated to the prod environment

Notice ：

1. Any spaces or special characters will affect whether the log can be cut correctly, please make sure the log format is unified
2. Please use the lower camel case to name the key