grok filter

context

dit document beschrijft de werking van grok.
Deze filter parset raw data in een bruikbaar key/value paar zodat je deze makkelijker kunt doorzoeken

werking

- formaat van logboeken verschilt per toepassing
- voor logboeken worden doorgestuurd naar elaticsearch, wordt die raw data doorzocht op bruikbare info
- aan deze data wordt een key/value paar toegekend
- dit paar wordt meegestuurd naar Elasticsearch
- je kan dan deze keys gaan gebruiken als zoekterm

voorbeeld

- toepassings logboek /var/log/webshop.log bevat onderstaande raw data:

```
Feb 06 16:54:14 67.38.38.23 bestelt theeset Porcelain300
Feb 14 10:09:17 58.23.46.81 aanmelden retour theeset Porcelain300
...
```

- je kan deze data nu gaan opdelen in bruikbare (doorzoekbare) info:
  - tijdstip
  - wie: het ip-adres
  - actie: een bestelling plaatsen, retour aanmelden
  - artikel: aankoop

syntax

- grok kent een tekst patroon (SYNTAX) een veldnaam (SEMANTIC)toe:

```
%{SYNTAX:SEMANTIC}
```

- voorbeeld:

    - `match ⇒ { “message” ⇒ “%{MONTH:maand} %{MONTHDAY:dag} %{TIME:tijdstip}” }`
    - haalt maand, dag en tijdstip uit de entry en slaat die op onder genoemde keys
- voor veelgebruikte velden zijn er al grok filters beschikbaar.
- filters definieer je in /etc/logstash/conf.d/:

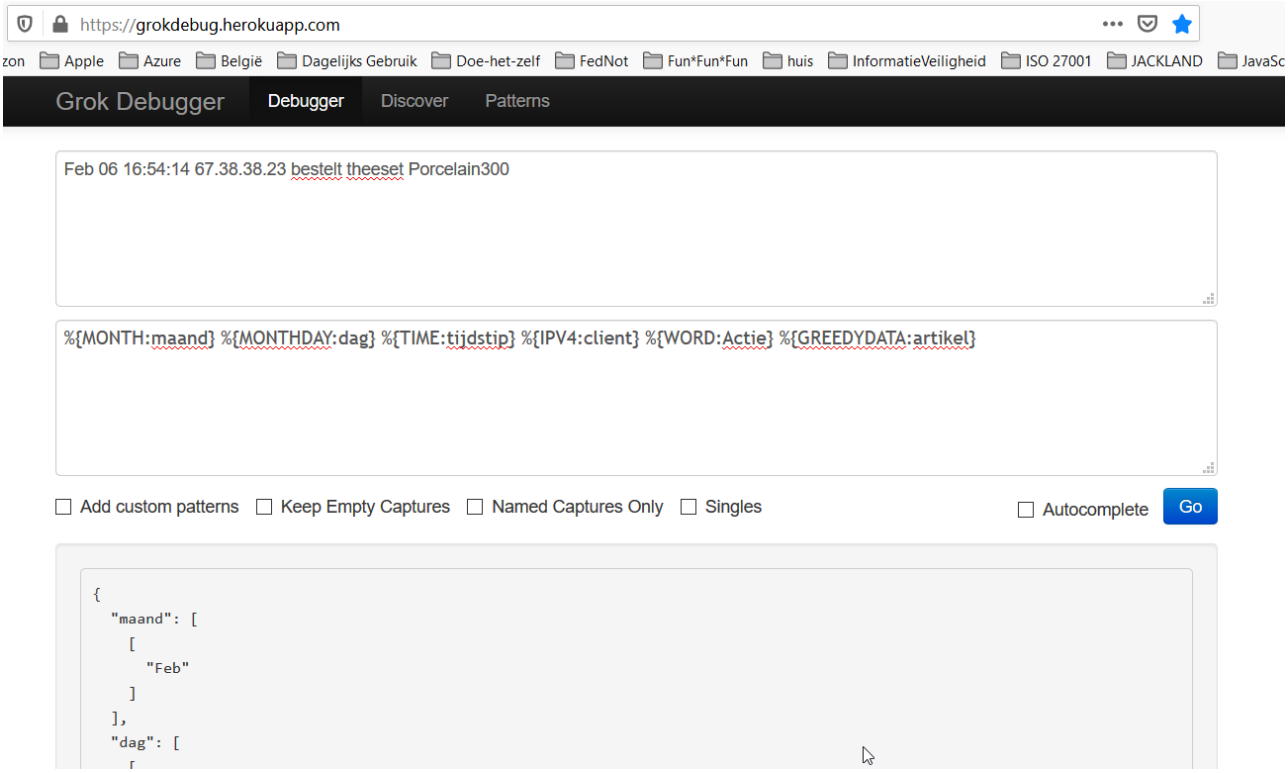
```
filter {
  grok  {
     match => { "message" => "%{MONTH:maand} %{MONTHDAY:dag} %{TIME:tijdstip} %{GREEDYDATA:webshop}" }
  }
}
```

- bij elke aanpassing, dien je logstash te herstarten:

```
sudo systemctl restart logstash
```

testen

1. ga naar Grokdebugger
2. kopieer je log entry in het bovenste veld
3. voeg %{GREEDYDATA} in je filter veld
4. voeg stelselmatig filters toe:



problemen, problemen

logentry wordt niet correct geparset

1. controleer de syntax van je grok filter en herstart logstash:

```
sudo systemctl restart logstash
```

2. controleer `/var/log/logstash/logstash-plain.log`

TER INFO: gefaalde parsing krijgen ook steeds de tag **_grokparsefailure**

fout: mapper_parsing_exception", "reason"=>"object mapping for [client] tried to parse field [client] as object, but found a concrete value"}}}}

- je gebruikt een keynaam die in logstash/filebeat wordt gebruikt. In dit geval: client
- kies een andere keynaam en herstart logstash:

```
sudo systemctl restart logstash
```

meer info

- grok basics

elk stack