



Logstash and friends

Julien Pivotto

Techies Teach Techies
September 2, 2013

① Introduction

② Logstash

- Missions
- Inputs
- Filters
- Output

③ Kibana

④ Conclusion



Logging

- Recording of events
- Voice of your systems and applications
- It tells you almost everything
- It is a source of knowledge



Logging is useful

- Understanding outages



Logging is useful

- Understanding outages
- not only when it's wrong
- you can extract metrics
- no logs means something
- it tells you what, why, who, when



Logging in the wild

- Syslog
- `|tee /var/log/myapp.log`
- Cron + MAILTO=
- `&>/dev/null`



Logging in the past

- Logging to files on each server
- Using syslog protocol
- Decentralized
- Reading requires SSH access
- Not developer friendly



The tools nowadays

- Jenkins, Icinga, Graphite, Foreman
- Nice web interfaces
- Centralized
- Easy to use



Requirements

- Scalable tools
- Configured by text files
- Playing with existing tools
- Scalable
- Following the Unix philosophy



3 separate tools

- Elasticsearch, distributed search & analytics engine
- Logstash, logs management
- Kibana, very nice webui to ES and Logstash



Logstash



Shipping the logs

- Some applications can only write to files
- But you need them on the main logstash server
- Logstash can act as a daemon to ship the logs
- Destinations can be syslog, redis, ...
- Then you can act on your logs



Collecting the logs

- You can plug logstash to a lot of data sources
- It can be passive or active
- Listening on a UDP port vs checking mails
- All your logs are managed by one application
- It creates fields from the logs



Filtering the logs

- Making sense of a log message
- Finding what is important
- Adding and removing fields



Storing the logs

- Output to Elasticsearch
- Sending information to statsd
- Sending to your inbox, to icinga or files



<http://www.flickr.com/photos/quinnanya/7237788632/>



UDP and TCP input

- Compatible with rsyslog protocol
- Each syslog talks with logstash directly
- Allow you to use the syslog toolchains: logger, rsyslog
- UDP is shoot and forget



UDP and TCP input

Logstash configuration

```
input {  
  udp {  
    type => syslog  
    port => 5544  
  }  
  tcp {  
    type => syslog  
    port => 5544  
  }  
}
```



UDP and TCP input

Rsyslog configuration

```
*.* @logstash.example.com:5544
```

- In /etc/rsyslog.conf
- That line will forward all the logs to logstash
- Logstash will make useful fields out of it: priority, severity, program...



File

- Enable you to use logstash with every application
- Useful to ship the logs
- Acts as a `tail -n 0 -F`
- It works even if you use logrotate



File

```
input {  
  file {  
    path => "/var/log/legacyapp.log"  
    type => "legacylog"  
  }  
}
```



Grok

- Extract fields from text
- Useful to read messages
- A lot of pre-existing patterns
- Uses Regex to find out fields



Grok

Input text

Invalid user oracle from 85.249.144.18

Grok pattern

Invalid user %{USERNAME:login} from %{IP:ip}

Result

```
{
  "login": [
    [ "oracle" ]
  ],
  "ip": [
    [ "85.249.144.18" ]
  ]
}
```



Grok

```
filter {  
  grok {  
    type      => "syslog"  
    pattern => ["(?m)<{%{POSINT:syslog_pri}>..."  
    add_field => [ "received_at", "%{@timestamp}" ]  
    add_field => [ "received_from", "%{@source_host}" ]  
    add_tag   => "syslog-%{syslog_program}"  
  }  
}
```



Grep

- Allows you to grep interesting messages
- Useful to count



Grep

```
filter {  
  grep {  
    add_field => ["outputirc", "A puppet package  
                  has been deployed"]  
  
    add_tag => "outputirc"  
    drop => false  
    match => [ "syslog_program", "yum" ]  
    match => [ "@source_host", "puppetmaster" ]  
    match => [ "@message", "puppet-tree" ]  
  }  
}
```



Geoip

```
filter{  
  geoip {  
    tags    => ["syslog-httpd"]  
    source => ["client"]  
  }  
}
```

- Transform ip address into geo data
- Useful to filter by country/map the data



Elasticsearch

- Version of elasticsearch <=> version of logstash
- Unless you use the elasticsearch_http output

```
output {  
  elasticsearch {  
  }  
}
```



IRC

```
output {  
  irc {  
    channels => ["#example"]  
    host => "chat.freenode.net"  
    nick => "loggy"  
    port => 6667  
    tags => "outputirc"  
    user => "loggy"  
    format => "%{outputirc}"  
  }  
}
```



statsd

```
output {  
  statsd {  
    host => '127.0.0.1'  
    sender => "logstash"  
    increment => [ "httpd.%{http_host}.r.%{response}",  
                  "httpd.response.%{response}" ]  
    count => ["apache.%{http_host}.bytes", "%{bytes}" ]  
    timing => ["apache.%{http_host}", "%{duration_msec}"]  
    tags => 'grokked-apache'  
  }  
}
```



Kibana

- Kibana is a web interface for Logstash/ES
- Kibana 1 was written in PHP
- Kibana 2 was written in Ruby
- Kibana 3 is written in AngularJS



Kibana 3

- Everything happens in the browser
- The browser is connected to Elasticsearch
- You can save dashboards into ES
- You can write/template dashboards to files



Installing kibana3

```
git clone https://github.com/elasticsearch/kibana.git  
ssh -NL 9200:127.0.0.1:9200 elasticsearch &  
python -m SimpleHTTPServer
```



Kibana queries

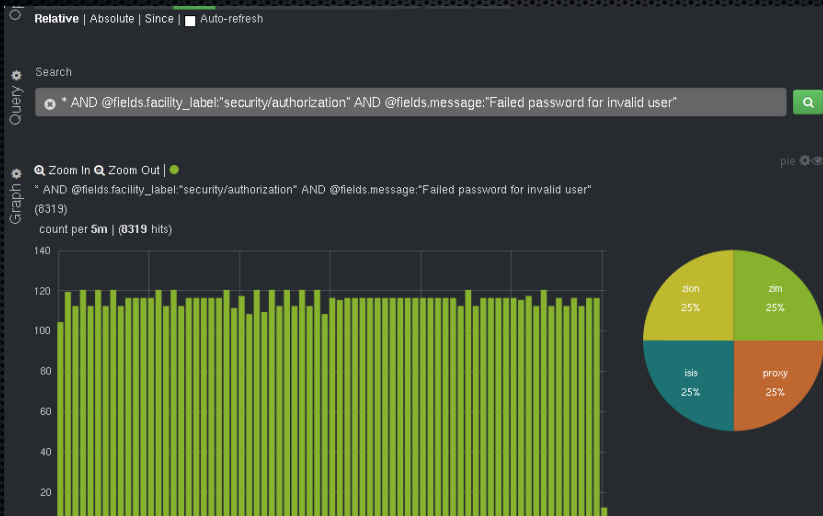
Example of a kibana query

```
@fields.syslog_program:"httpd" AND @fields.http_host:"test.example.com" AND @fields.response:"404"
```

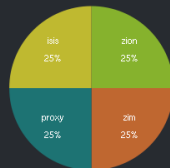
- Lucene query syntax
- Simple and effective
- Point & click web interface



Kibana



Kibana



- ☐ @fields.facility
- ☐ @fields.facility_id
- ☐ @fields.logsource
- ☐ @fields.message
- ☐ @fields.pid
- ☐ @fields.priority
- ☐ @fields.program

0 to 44 of 44 available for paging

table

@timestamp ▾

← @message

```

2013-09-02T04:49:10.000Z <86>Sep 2 06:49:10 proxy sshd[17977]: Failed password for invalid user info from
61.238.156.110 port 48448 ssh2

2013-09-02T04:49:10.000Z <86>Sep 2 06:49:10 zim sshd[6099]: Failed password for invalid user info from 61.238.156.110
port 57343 ssh2

2013-09-02T04:49:10.000Z <86>Sep 2 06:49:10 isis sshd[12951]: Failed password for invalid user info from 61.238.156.110
port 45166 ssh2

2013-09-02T04:49:10.000Z <86>Sep 2 06:49:10 zion sshd[27594]: Failed password for invalid user info from
61.238.156.110 port 43534 ssh2

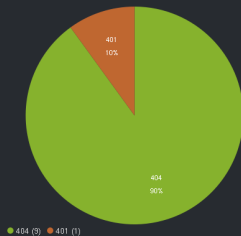
2013-09-02T04:48:58.000Z <86>Sep 2 06:48:58 zim sshd[6866]: Failed password for invalid user info from 61.238.156.110

```

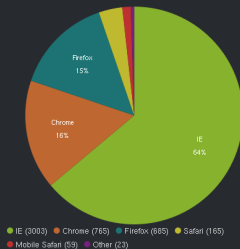


Kibana

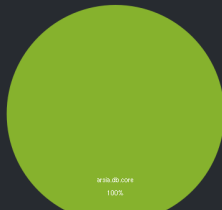
Error codes



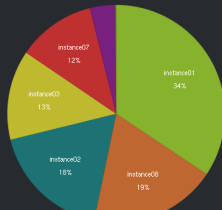
Browsers



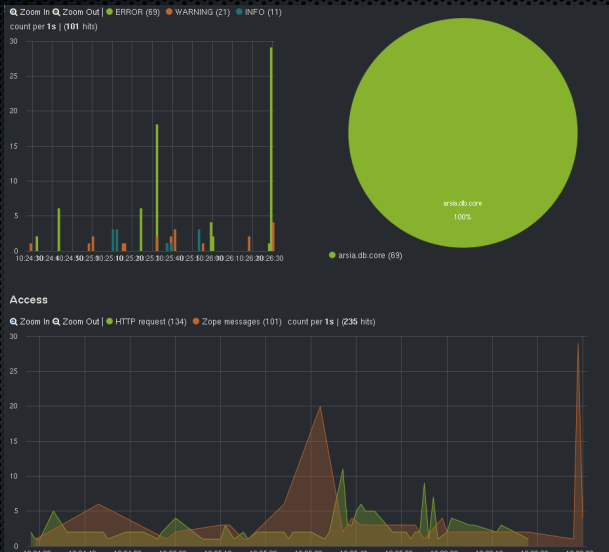
Loggename with errors



Instances



Kibana



Conclusion

- Logstash is a small daemon
- Simple to package & deploy (jar file)
- Scalable thanks to Elasticsearch
- Developer friendly thanks to Kibana

