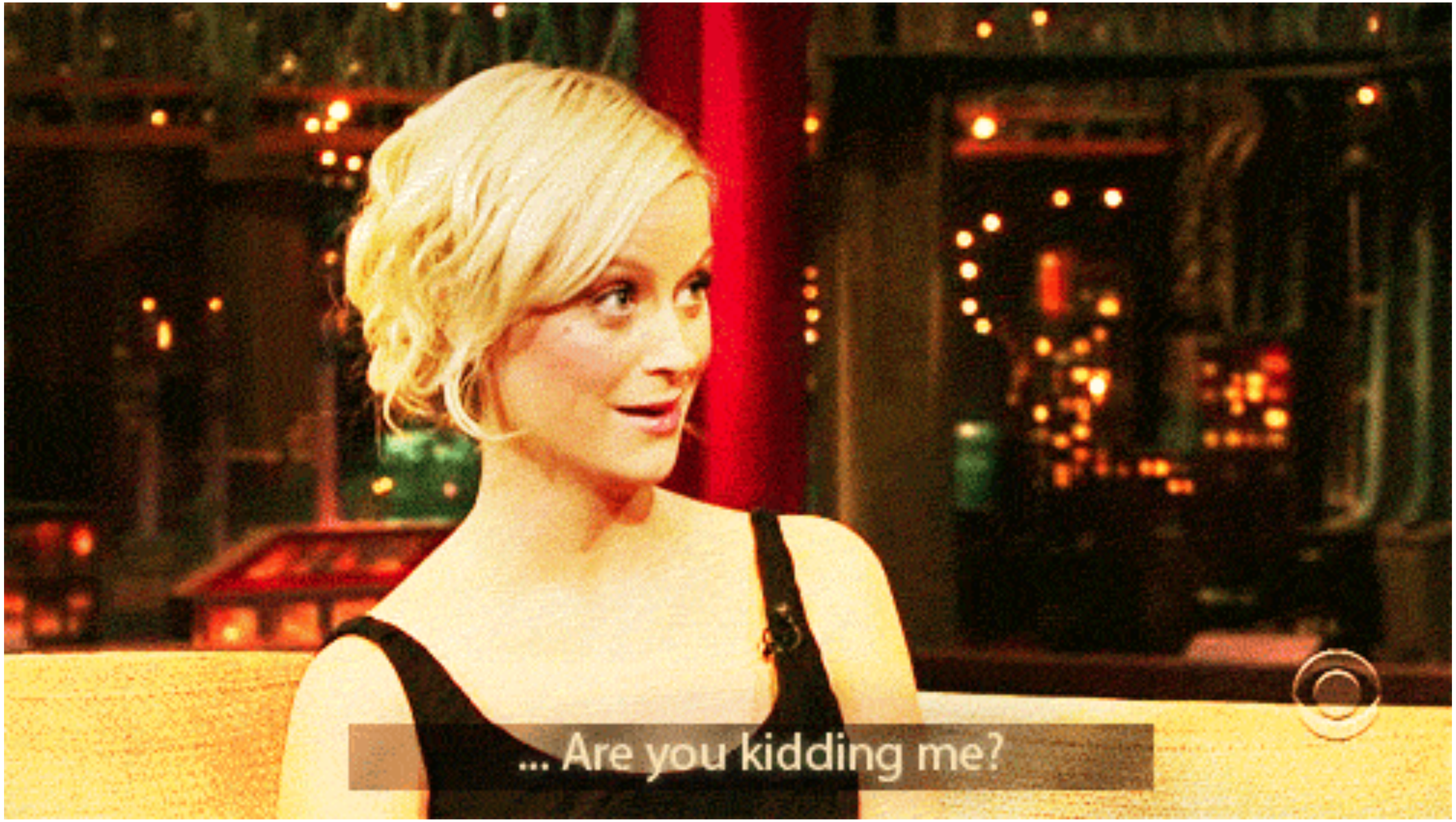
Logging with Elasticsearch, Logstash & Kibana

Bastian Widmer / @dasrecht

But why?

„Can you check the errors from yesterday between 15.02 and 15.07“



Amazee Labs

Visualization > Plaintext


3,761
302
386
90
821
955
1,851
127
65
77
204
5
6,563
2,011
5,837
729
832
312
527
305
Bastian Widmer
GoPro: Red Bull X Alps
15:51
08:51
18:51
YES
118
4.9K
344.3K
79
76
50


VideoWall
FILTERING
an hour ago to a few seconds ago refreshed every 30s
EVENTS OVER TIME
View
amazeeweb1 (319)
amazeeweb2 (2309)
amazeeweb3 (2291)
amazeeweb4 (2185)
amazeeweb5 (2394)
count per 30s | (9498 hits)
140
120
100
80
60
40
20
15:34:00 15:36:00 15:38:00 15:40:00 15:42:00 15:44:00 15:46:00 15:48:00 15:50:00 15:52:00 15:54:00 15:56:00 15:58:00 16:00:00 16:02:00 16:04:00 16:06:00 16:08:00 16:10:00 16:12:00 16:14:00 16:16:00 16:18:00 16:20:00 16:22:00 16:24:00 16:26:00 16:28:00 16:30:00

Who are you?

Bastian Widmer

@dasrecht / bastianwidmer.ch

Switzerland

Development and Operations Engineer

Amazee Labs

Agenda

1. Introduction
2. ELK Stack
3. Architecture
4. Tools!
5. Demo

ELK Stack!


ELK Stack!
Elasticsearch
Logstash
Kibana
Amazee Labs

Elasticsearch

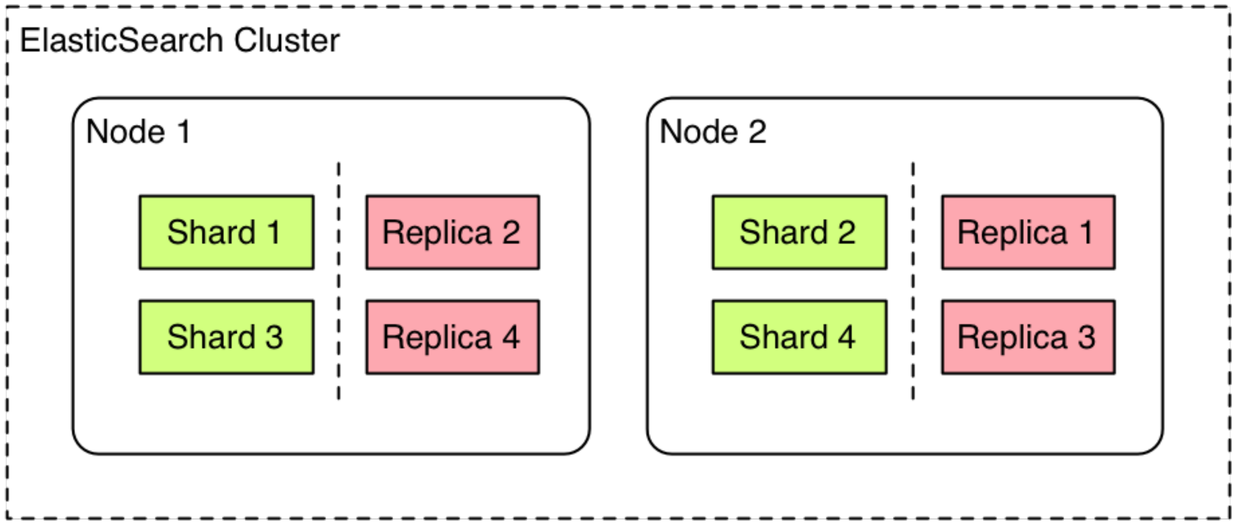
Elasticsearch

- Java
- Search and Index
- Distributed — Copies & Shards
- Clustering
- API — JSON / RESTful
- Apache Lucene

Amazee Labs

Elasticsearch

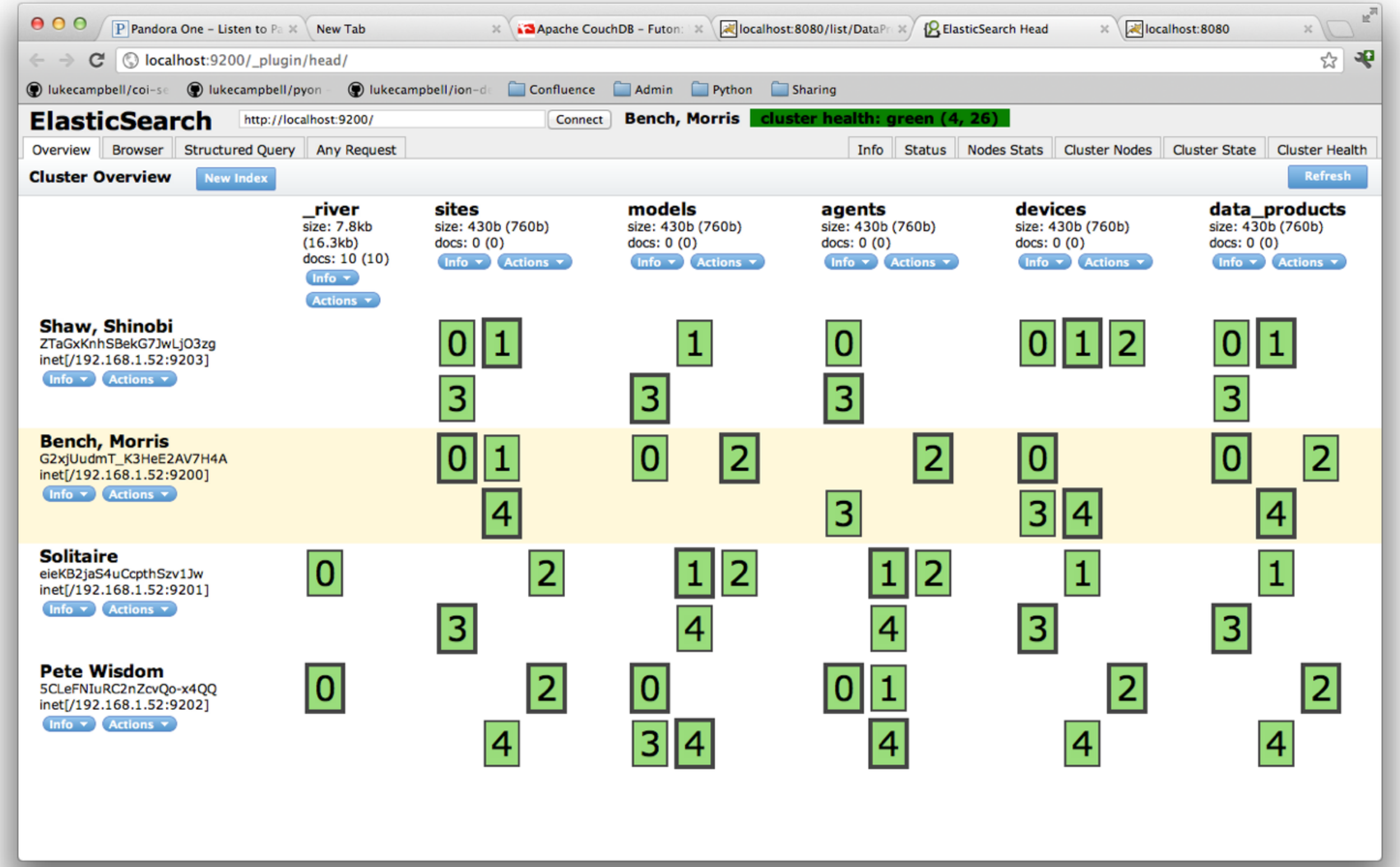
- Index
  like a Database
- Replica
  Copies for Fault Tolerance
- Shard
  Lucene Instance which indexes the Data



see : http://blog.liip.ch/archive/2013/07/19/on-elasticsearch-performance.html

Amazee Labs

Elasticsearch



Logstash

Logstash

- Multiple Input / Multiple Output
- Centralize Logs
  - Collect
  - Parse
  - Store / Forward

Amazee Labs

inputs

- collectd
- drupal_dblog
- elasticsearch
- eventlog
- exec
- file
- ganglia
- gelf
- gemfire
- generator
- graphite
- heroku
- imap
- invalid_input
- irc
- jmx
- log4j
- lumberjack
- pipe
- puppet_facter
- rabbitmq
- redis
- relp
- s3
- snmptrap
- sqlite
- sqs
- stdin
- stomp
- syslog
- tcp
- twitter
- udp
- unix
- varnishlog
- websocket
- wmi

codecs

- cloudtrail
- compress_spooler
- dots
- edn
- edn_lines
- fluent
- graphite
- json
- json_lines
- json_spooler
- line
- msgpack
- multiline
- netflow
- noop
- oldlogstashjson
- plain
- rubydebug
- spool

filters

- advisor
- alter
- anonymize
- checksum
- cidr
- cipher
- clone
- collate
- csv
- date
- dns
- drop
- elapsed
- elasticsearch
- environment
- extractnumbers
- fingerprint
- gelfify
- geoip
- grep
- grok
- grokdiscovery
- i18n
- json
- json_encode
- kv
- metaevent
- metrics
- multiline
- mutate
- noop
- prune
- punct
- railsparallelrequest
- range
- ruby
- sleep

outputs

- boundary
- circonus
- cloudwatch
- csv
- datadog
- datadog_metrics
- elasticsearch
- elasticsearch_http
- elasticsearch_river
- email
- exec
- file
- ganglia
- gelf
- gemfire
- google_bigquery
- google_cloud_storage
- graphite
- graphtastic
- hipchat
- http
- irc
- jira
- juggernaut
- librato
- loggly
- lumberjack
- metriccatcher
- mongodb
- nagios
- nagios_nsca
- null
- opentsdb
- pagerduty
- pipe
- rabbitmq
- redis

Amazee Labs

The life of an event

- Input
- Filters
- Output
- Codecs

Amazee Labs

Logstash

- JRuby*
- >1.4.0 - FlatJAR Release is gone
- Instead of running „java -jar logstash.jar“ — „bin/logstash“
- Contrib Plugins
- Daily Indices

* see https://gist.github.com/jordansissel/978956

Amazee Labs

Input

- File
- Syslog
- Redis
- logstash-forwarder (former Lumberjack)

Amazee Labs

Filters

- Grok
- Mutate
- Drop
- Clone
- GeoIP (!!!)

Amazee Labs

Outputs

- Elasticsearch
- File
- Graphite
- StatsD

Amazee Labs

Logstash

```
1 input {
2   stdin { }
3 }
4
5 output {
6   stdout {
7     codec => rubydebug
8   }
9 }
```

Logstash

```
1 vagrant@precise64$ ./logstash agent -f 1_simpleconfig.cfg
2 very important log message!
3 {
4         "message" => "very important log message!",
5        "@version" => "1",
6      "@timestamp" => "2014-04-21T16:18:02.952Z",
7            "host" => "precise64"
8 }
```

Logstash

```
1  input {
2    stdin { }
3  }
4  output {
5    elasticsearch{
6      host => "127.0.0.1"
7    }
8    stdout {
9      codec => rubydebug
10   }
11 }
```

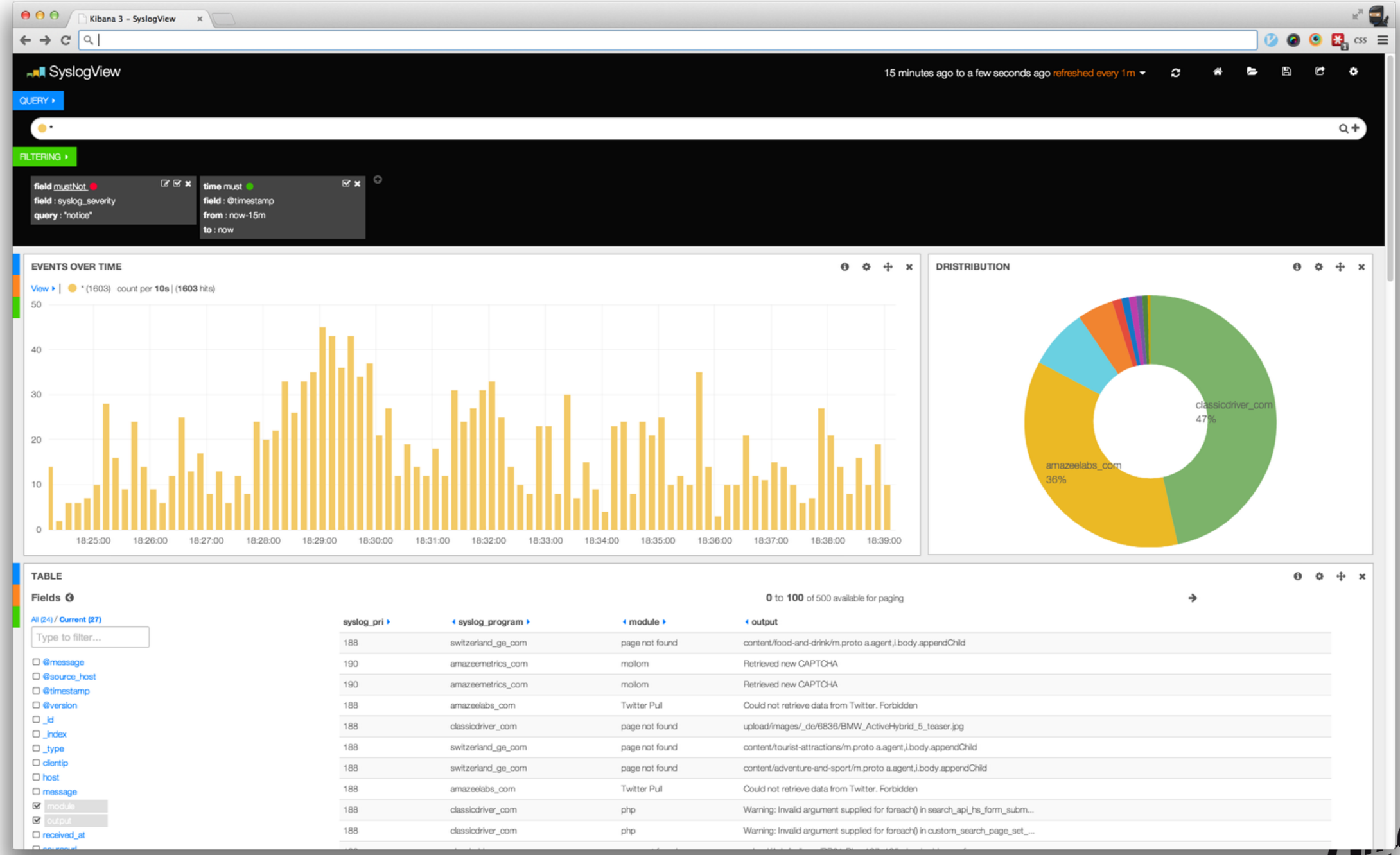
Logstash

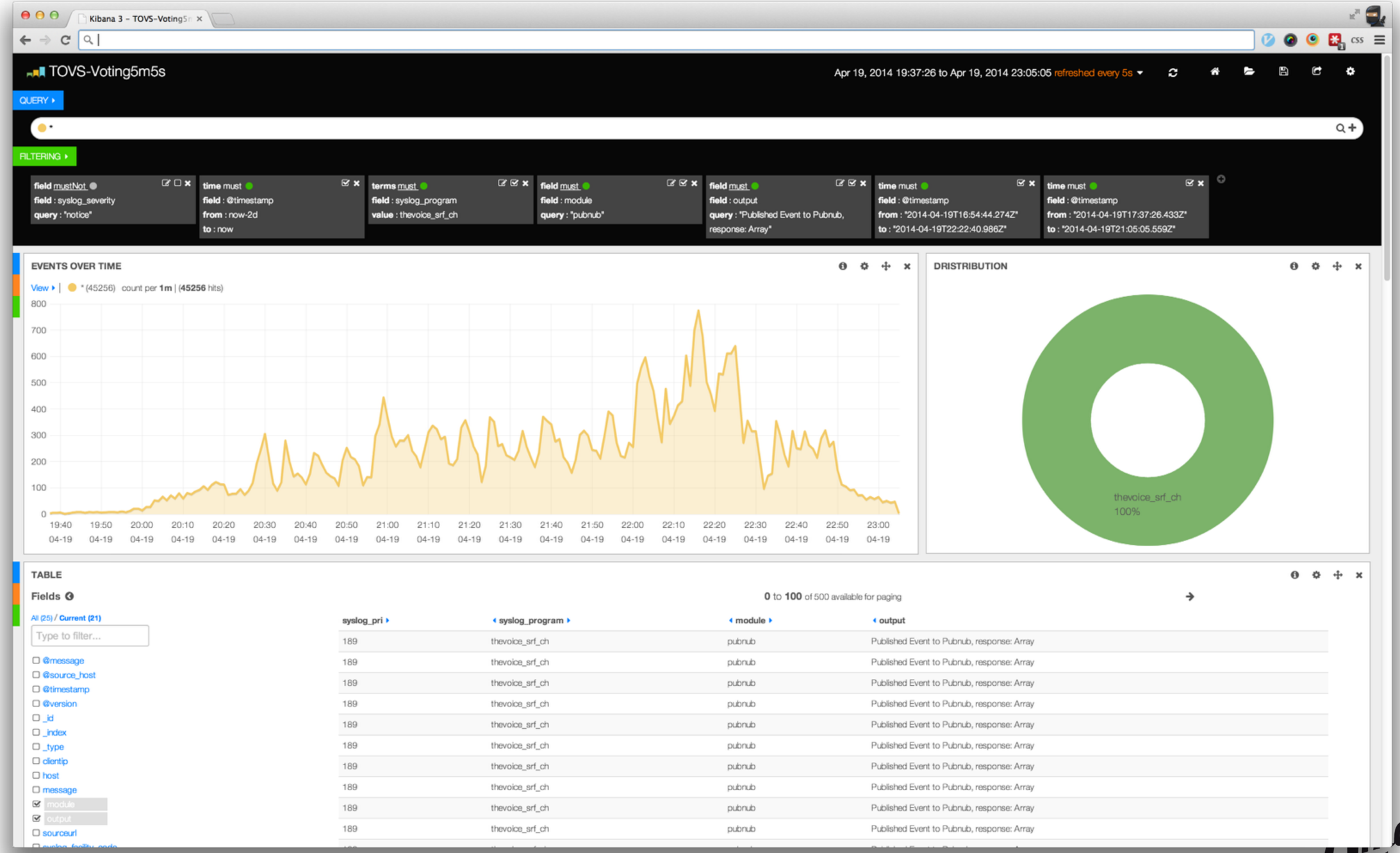
```
input {
    file {
    path => "/var/log/syslog"
    start_position => beginning
  }
}

output {
  stdout {
    codec => rubydebug
  }
    elasticsearch{
    host => "127.0.0.1"
  }
}
```

Logstash

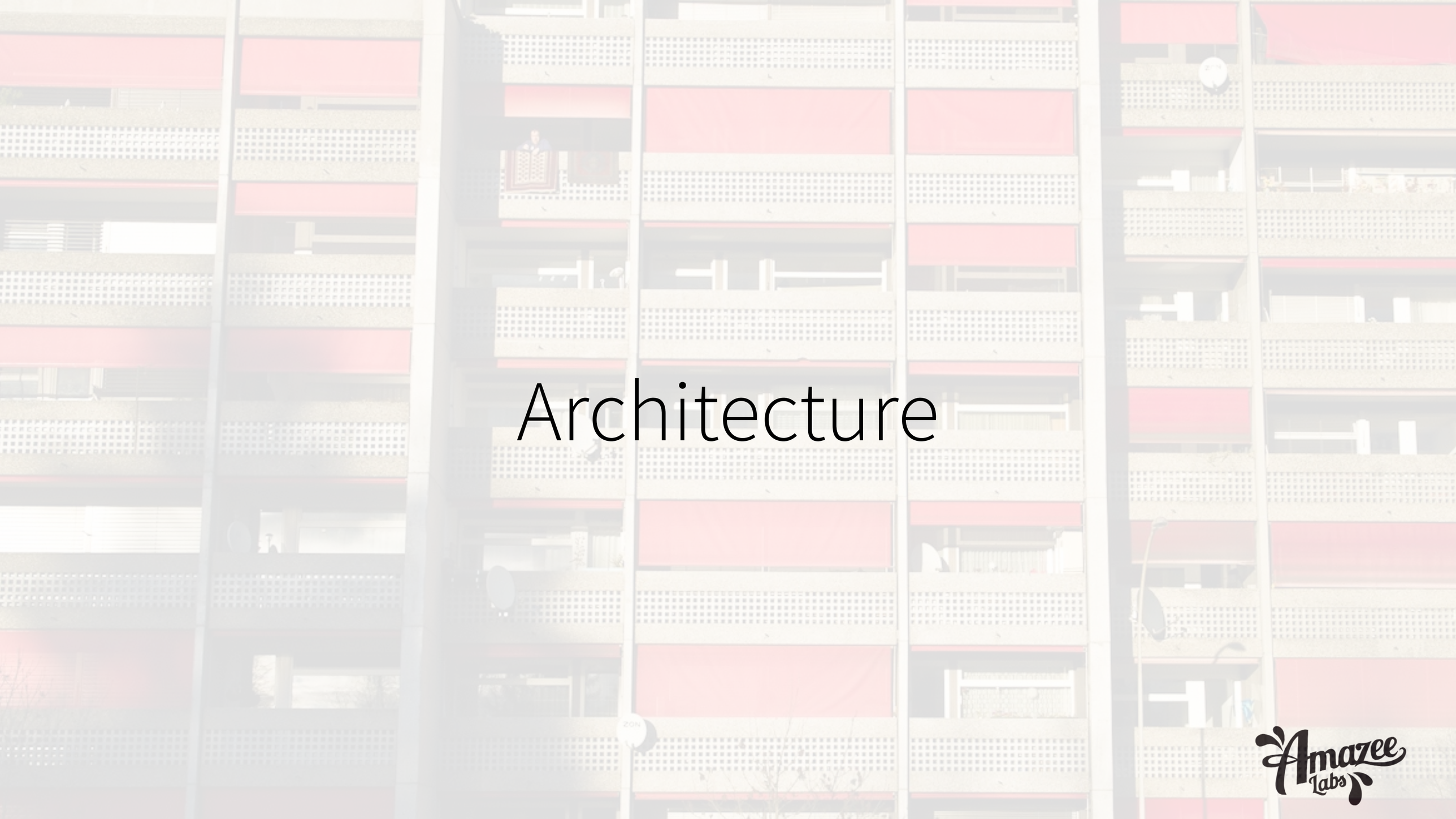
```
Errno::EBADF: Bad file descriptor - Bad file descriptor
```

Kibana

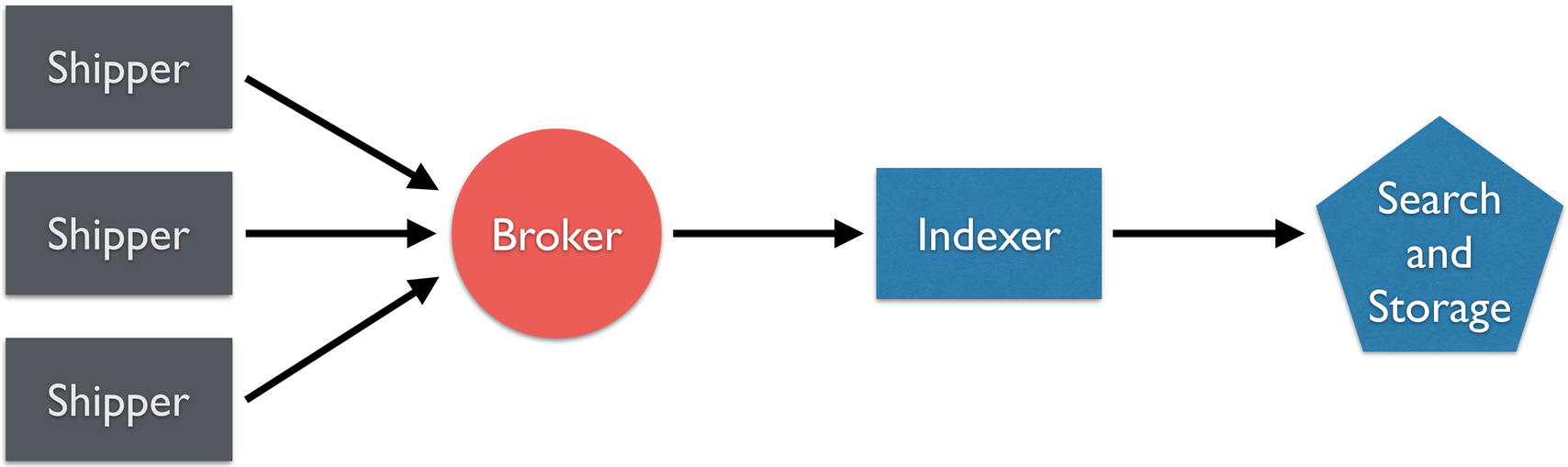

Kibana 3 - SyslogView
SyslogView
15 minutes ago to a few seconds ago refreshed every 1m
QUERY
FILTERING
field mustNot
field : syslog_severity
query : "notice"
time must
field : @timestamp
from : now-15m
to : now
EVENTS OVER TIME
View | * (1603) count per 10s | (1603 hits)
50
40
30
20
10
0
18:25:00
18:26:00
18:27:00
18:28:00
18:29:00
18:30:00
18:31:00
18:32:00
18:33:00
18:34:00
18:35:00
18:36:00
18:37:00
18:38:00
18:39:00
DRISTRIBUTION
classicdriver_com
47%
amazeelabs_com
36%
TABLE
Fields
0 to 100 of 500 available for paging
All (24) / Current (27)
Type to filter...
@message
@source_host
@timestamp
@version
_id
_index
_type
clientip
host
message
module
output
received_at
syslog_pri
syslog_program
module
output
188 switzerland_ge_com page not found content/food-and-drink/m.proto a.agent,i.body.appendChild
190 amazeemetrics_com mollom Retrieved new CAPTCHA
190 amazeemetrics_com mollom Retrieved new CAPTCHA
188 amazeelabs_com Twitter Pull Could not retrieve data from Twitter. Forbidden
188 classicdriver_com page not found upload/images/_de/6836/BMW_ActiveHybrid_5_teaser.jpg
188 switzerland_ge_com page not found content/tourist-attractions/m.proto a.agent,i.body.appendChild
188 switzerland_ge_com page not found content/adventure-and-sport/m.proto a.agent,i.body.appendChild
188 amazeelabs_com Twitter Pull Could not retrieve data from Twitter. Forbidden
188 classicdriver_com php Warning: Invalid argument supplied for foreach() in search_api_hs_form_subm...
188 classicdriver_com php Warning: Invalid argument supplied for foreach() in custom_search_page_set_...


Kibana 3 - TOVS-Voting5m5s
TOVS-Voting5m5s
Apr 19, 2014 19:37:26 to Apr 19, 2014 23:05:05 refreshed every 5s
QUERY
FILTERING
field mustNot
field : syslog_severity
query : "notice"
time must
field : @timestamp
from : now-2d
to : now
terms must
field : syslog_program
value : thevoice_srf_ch
field must
field : module
query : "pubnub"
field must
field : output
query : "Published Event to Pubnub, response: Array"
time must
field : @timestamp
from : "2014-04-19T16:54:44.274Z"
to : "2014-04-19T22:22:40.986Z"
time must
field : @timestamp
from : "2014-04-19T17:37:26.433Z"
to : "2014-04-19T21:05:05.559Z"
EVENTS OVER TIME
View
* (45256) count per 1m | (45256 hits)
800
700
600
500
400
300
200
100
0
19:40 04-19
19:50 04-19
20:00 04-19
20:10 04-19
20:20 04-19
20:30 04-19
20:40 04-19
20:50 04-19
21:00 04-19
21:10 04-19
21:20 04-19
21:30 04-19
21:40 04-19
21:50 04-19
22:00 04-19
22:10 04-19
22:20 04-19
22:30 04-19
22:40 04-19
22:50 04-19
23:00 04-19
DRISTRIBUTION
thevoice_srf_ch
100%
TABLE
Fields
0 to 100 of 500 available for paging
All (25) / Current (21)
Type to filter...
@message
@source_host
@timestamp
@version
_id
_index
_type
clientip
host
message
module
output
sourceurl
syslog_pri
syslog_program
module
output
189 thevoice_srf_ch pubnub Published Event to Pubnub, response: Array

Architecture

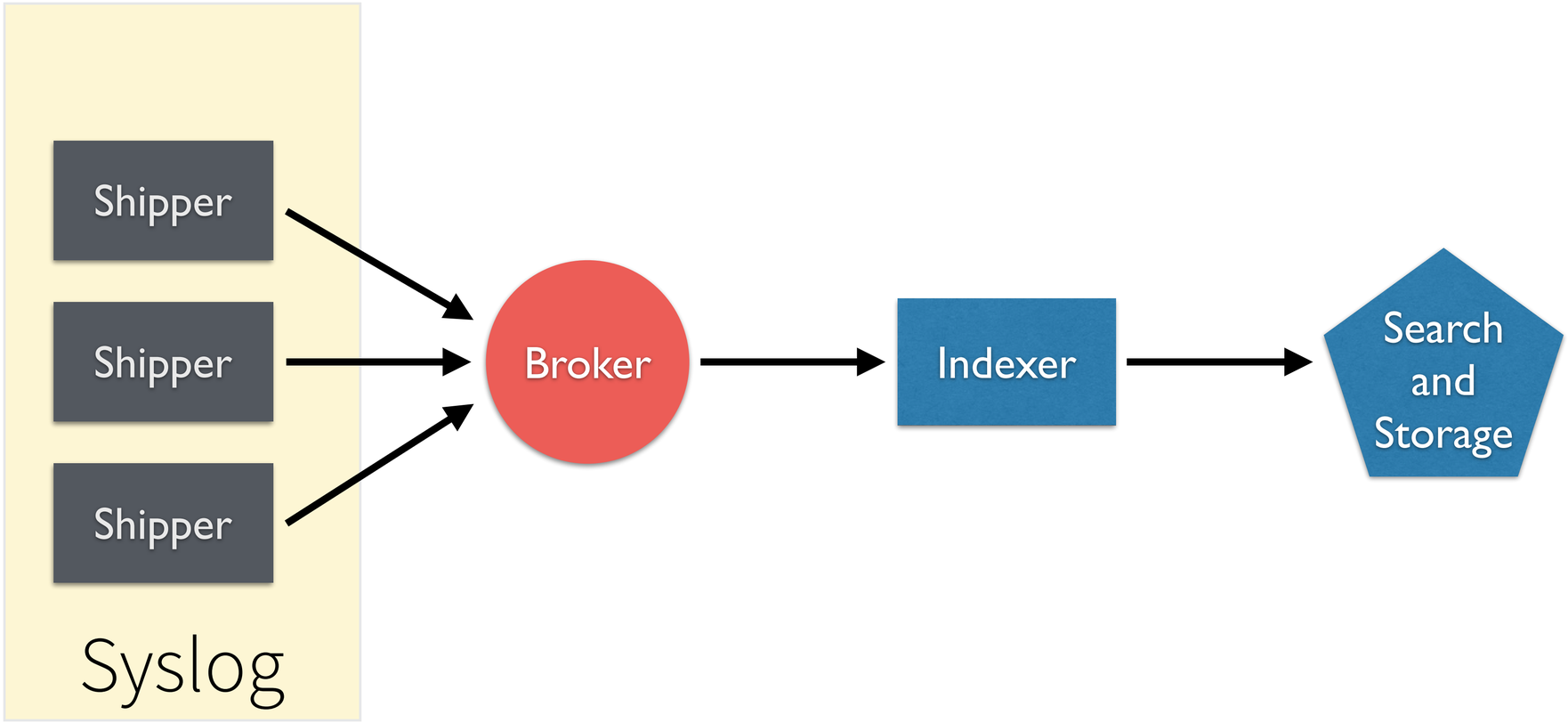


Architecture



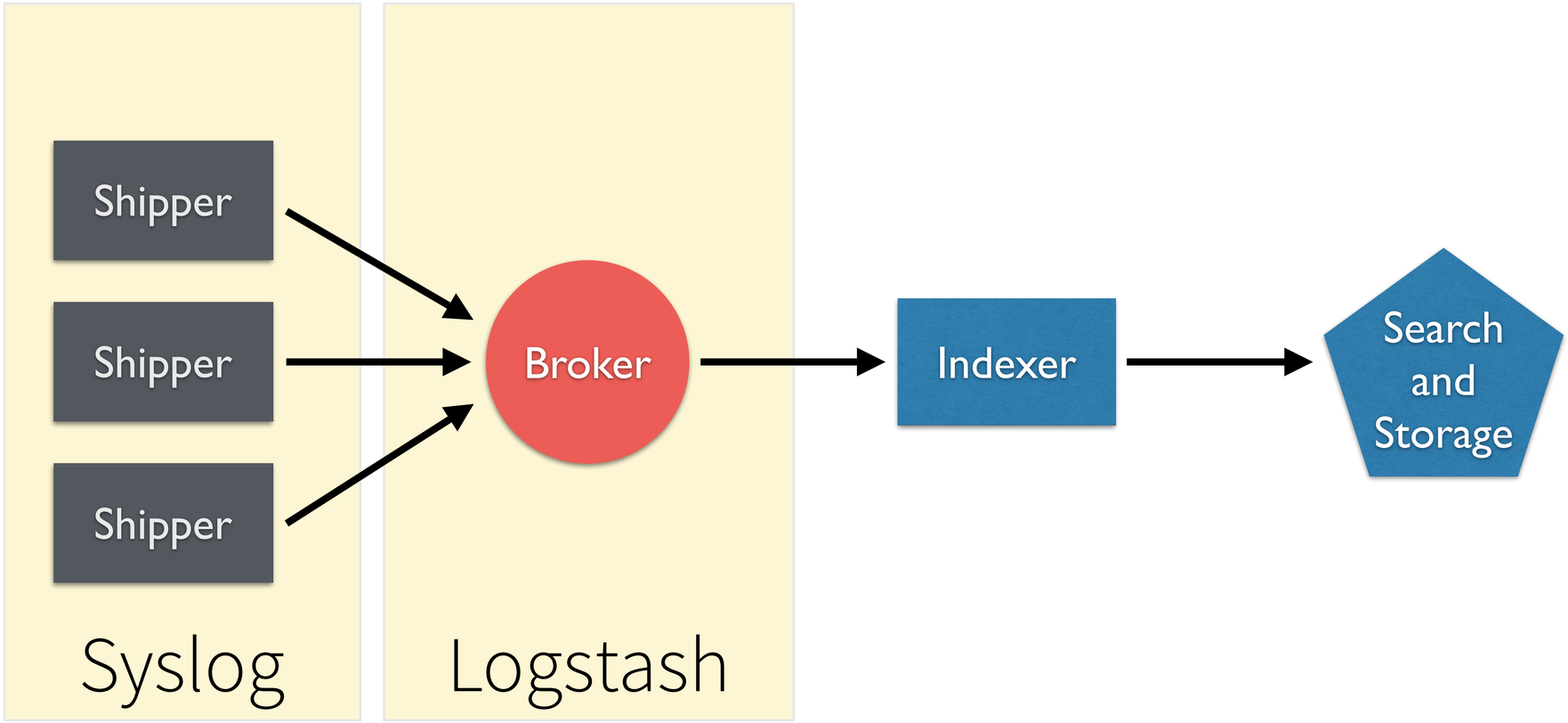
Amazee Labs

Architecture



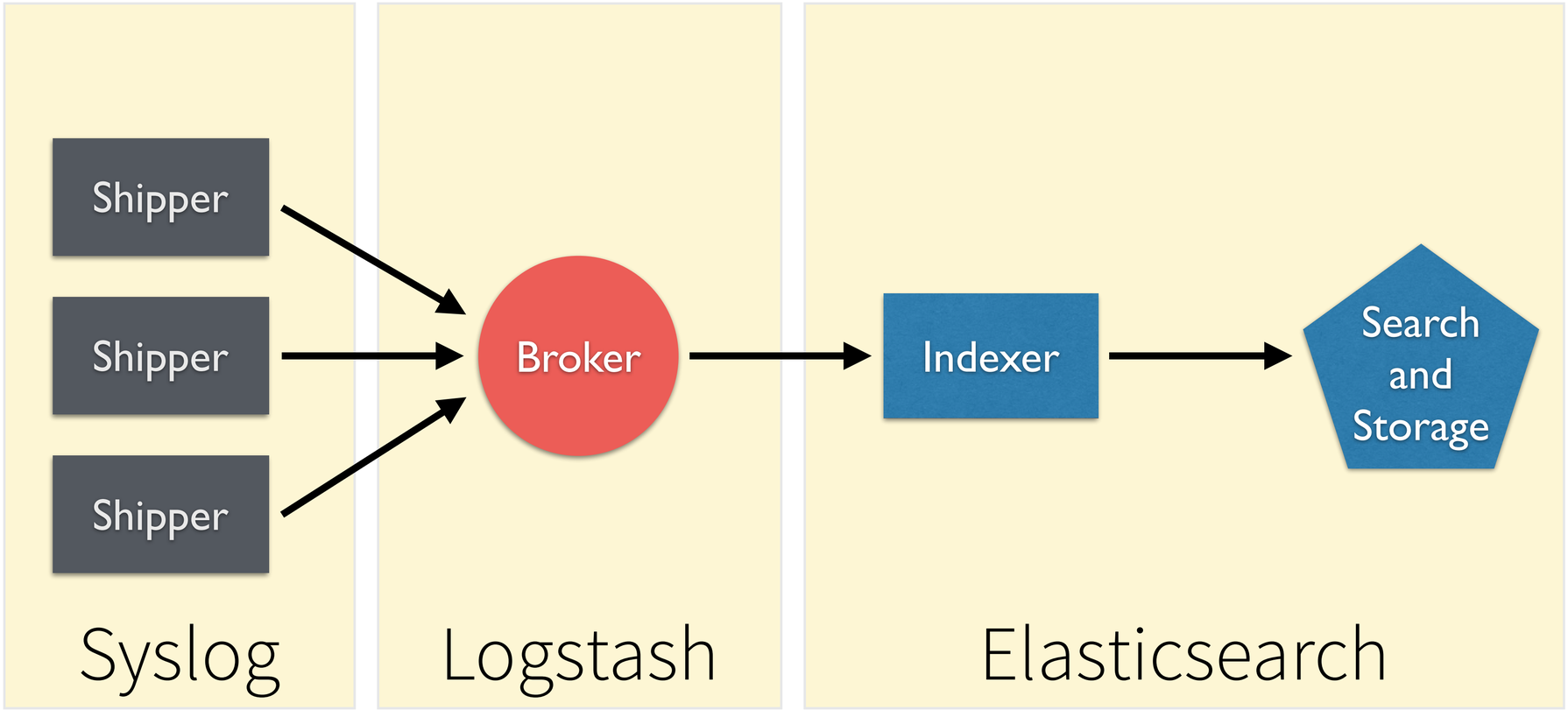
Amazee Labs

Architecture



Amazee Labs

Architecture



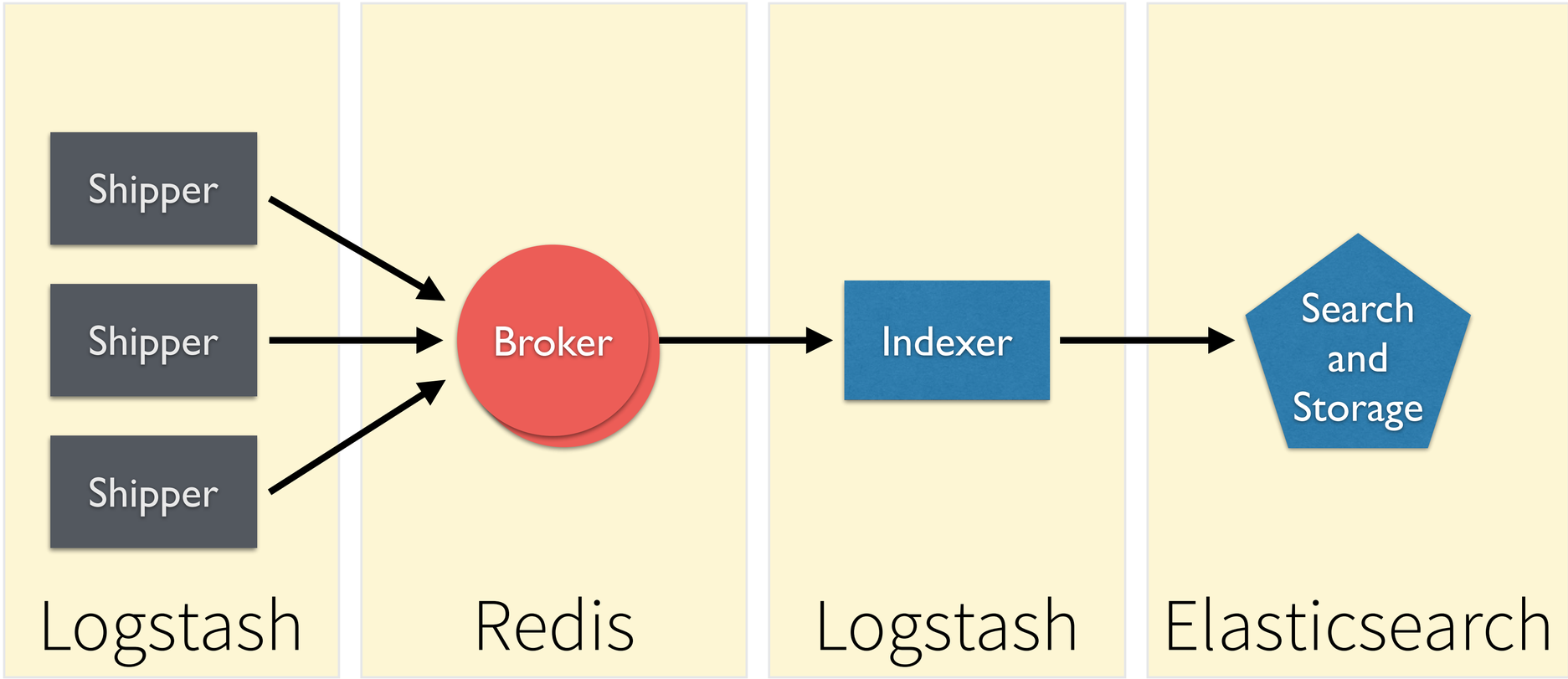
Amazee Labs

Architecture

the real deal!

Amazee Labs

Architecture

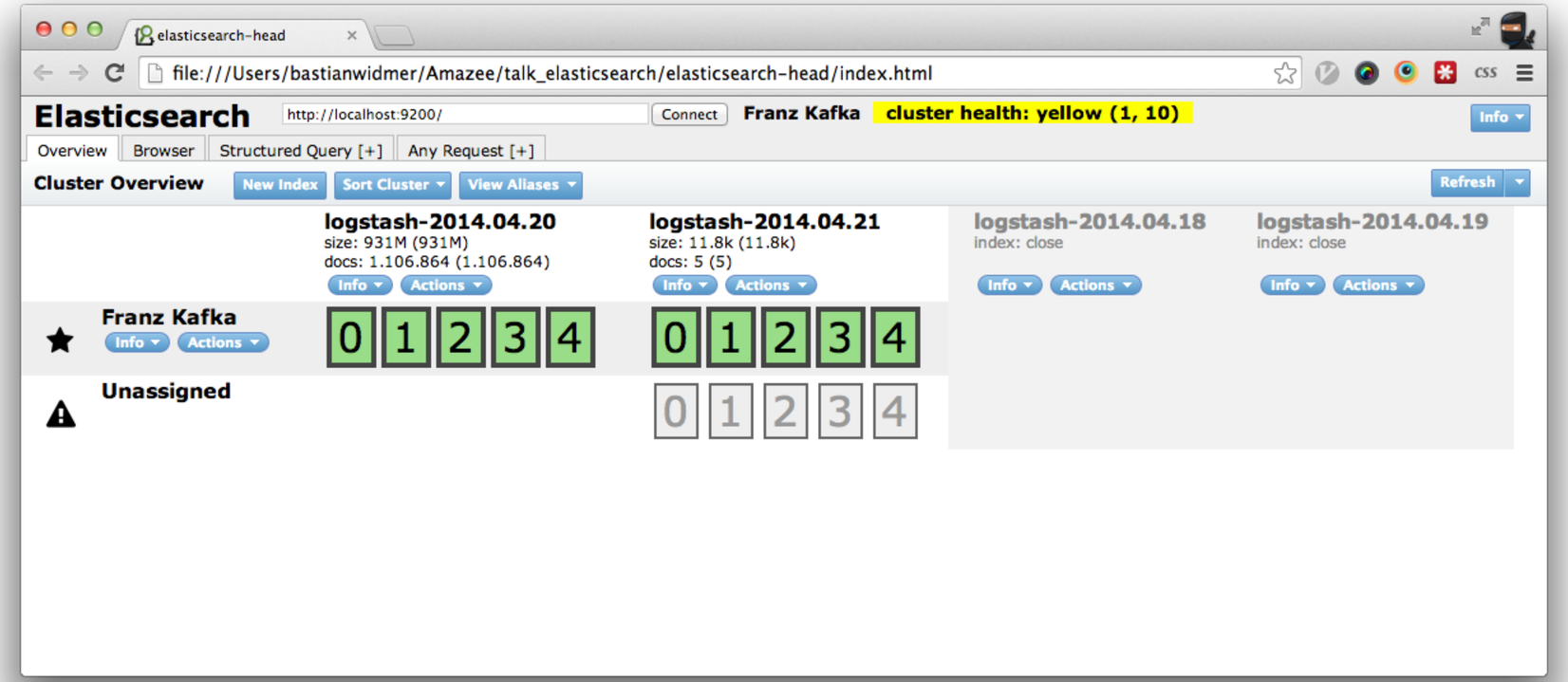


Amazee Labs

Tools!

(because anyone needs a bit help)

Elasticsearch Head

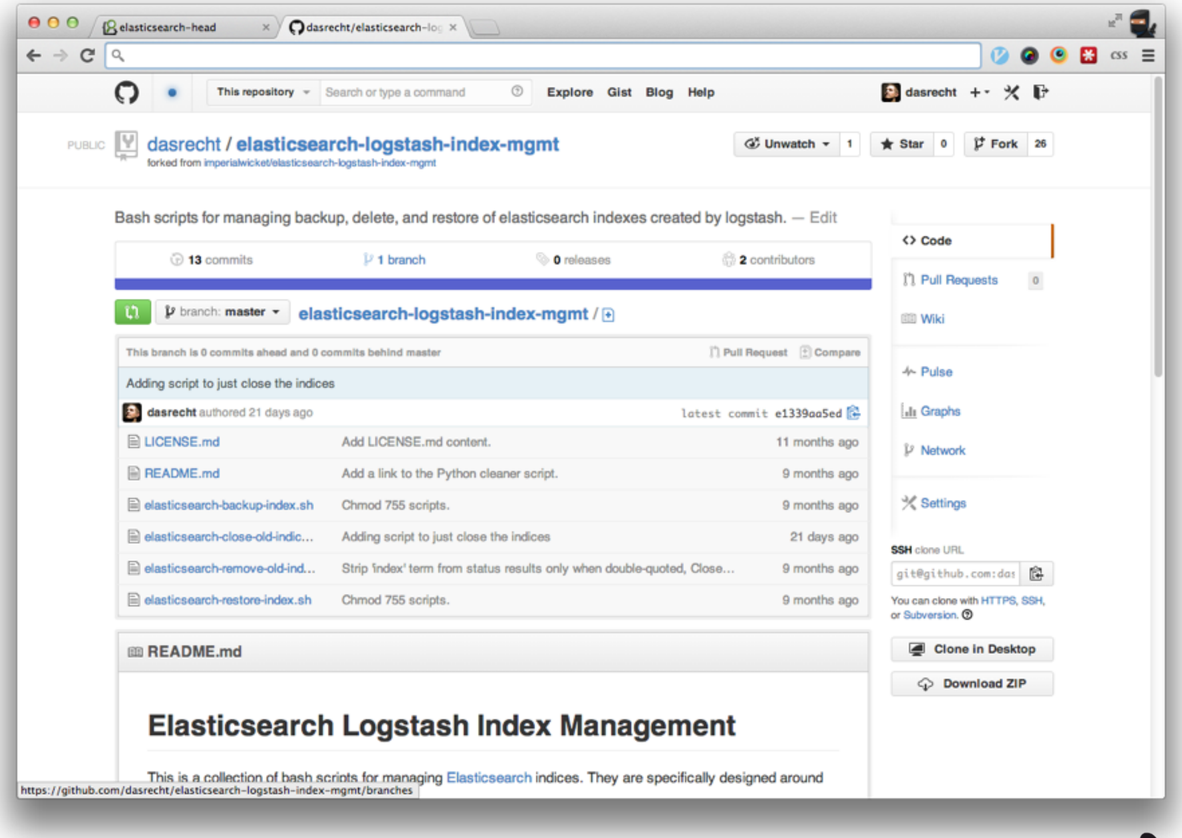


http://mobz.github.io/elasticsearch-head/

Amazee Labs

elasticsearch-index-mgmt

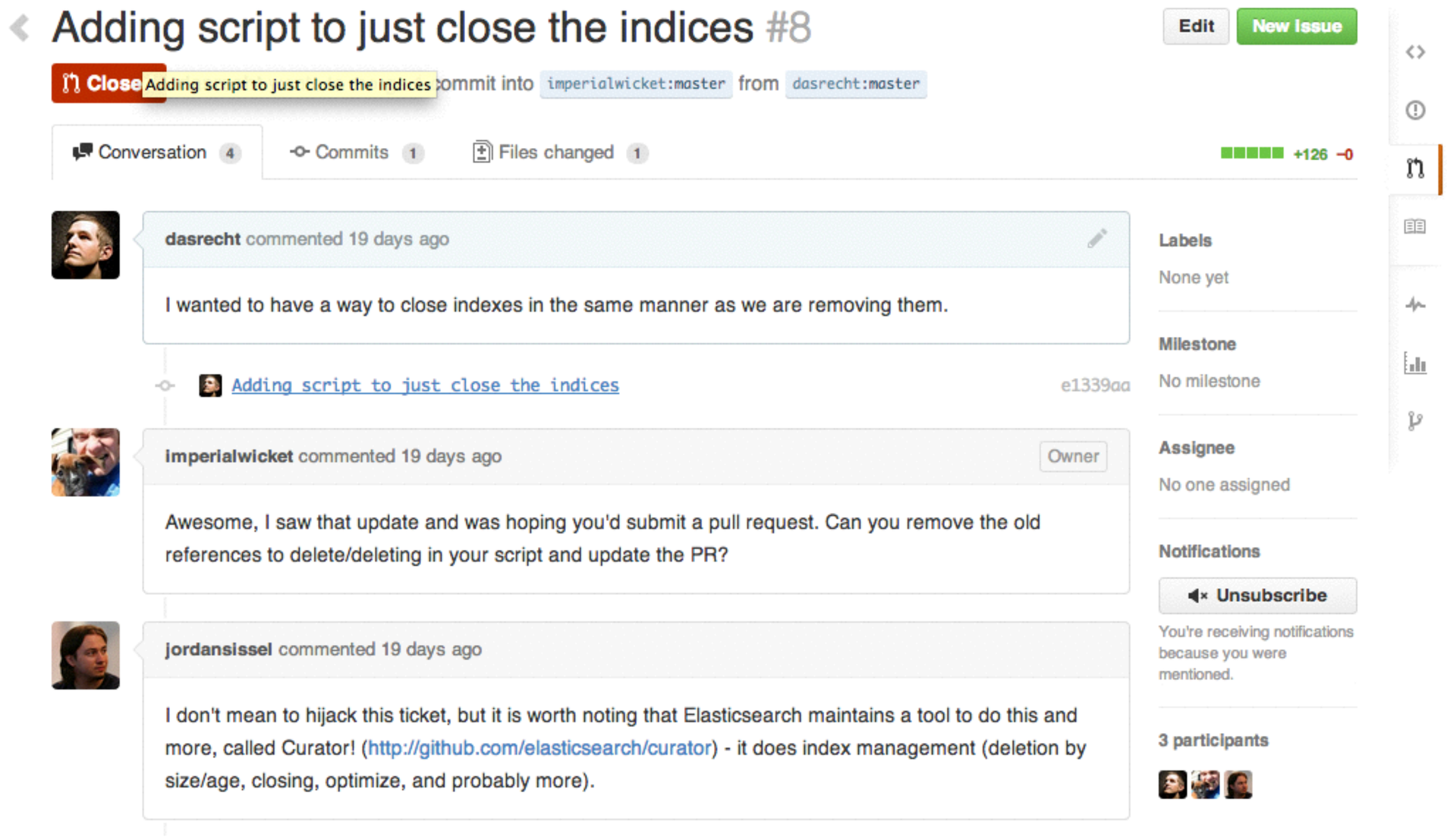
- Close
- Remove
- Backup
- Restore



http://s.nrdy.ch/eee

Amazee Labs

But then…


Adding script to just close the indices #8
Edit
New Issue
Close
Adding script to just close the indices
commit into
imperialwicket:master
from
dasrecht:master
Conversation 4
Commits 1
Files changed 1
+126 −0
dasrecht commented 19 days ago
I wanted to have a way to close indexes in the same manner as we are removing them.
Adding script to just close the indices
e1339aa
imperialwicket commented 19 days ago
Owner
Awesome, I saw that update and was hoping you'd submit a pull request. Can you remove the old references to delete/deleting in your script and update the PR?
jordansissel commented 19 days ago
I don't mean to hijack this ticket, but it is worth noting that Elasticsearch maintains a tool to do this and more, called Curator! (http://github.com/elasticsearch/curator) - it does index management (deletion by size/age, closing, optimize, and probably more).
Labels
None yet
Milestone
No milestone
Assignee
No one assigned
Notifications
Unsubscribe
You're receiving notifications because you were mentioned.
3 participants

Curator

- Time Series Indices? THIS IS THE TOOL!
- Close Indexes
- Delete (by space or time)
- Disable Bloom Filter
- ~~Optimize~~ / ForceMerge
- https://github.com/elasticsearch/curator

Amazee Labs

Curator

- Time Series Indexes? THIS IS THE TOOL!
- Cl
- Re
- Remove by Space Usage
- Disable Bloom Filter
- https://github.com/elasticsearch/curator

Curator
Perfect for Time Series Indexes

Curator

- Close indices older than 14 days, delete indices older than 30 days

```
curator --host my-elasticsearch -d 30 -c 14
```

- Disable bloom filter for indices older than 2 days, close indices older than 14 days, delete indices older than 30 days:

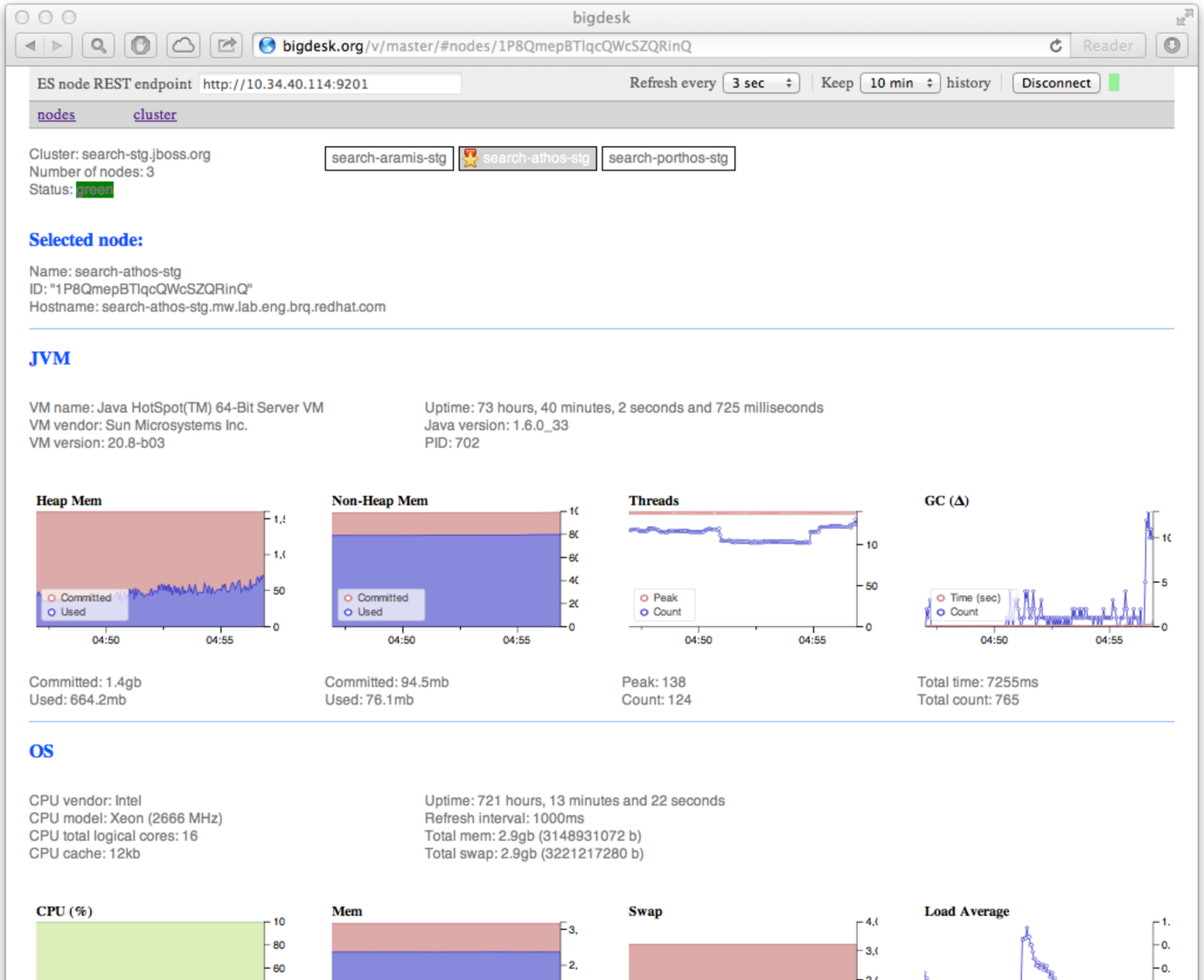
```
curator --host my-elasticsearch -b 2 -c 14 -d 30
```

Amazee Labs

Curator

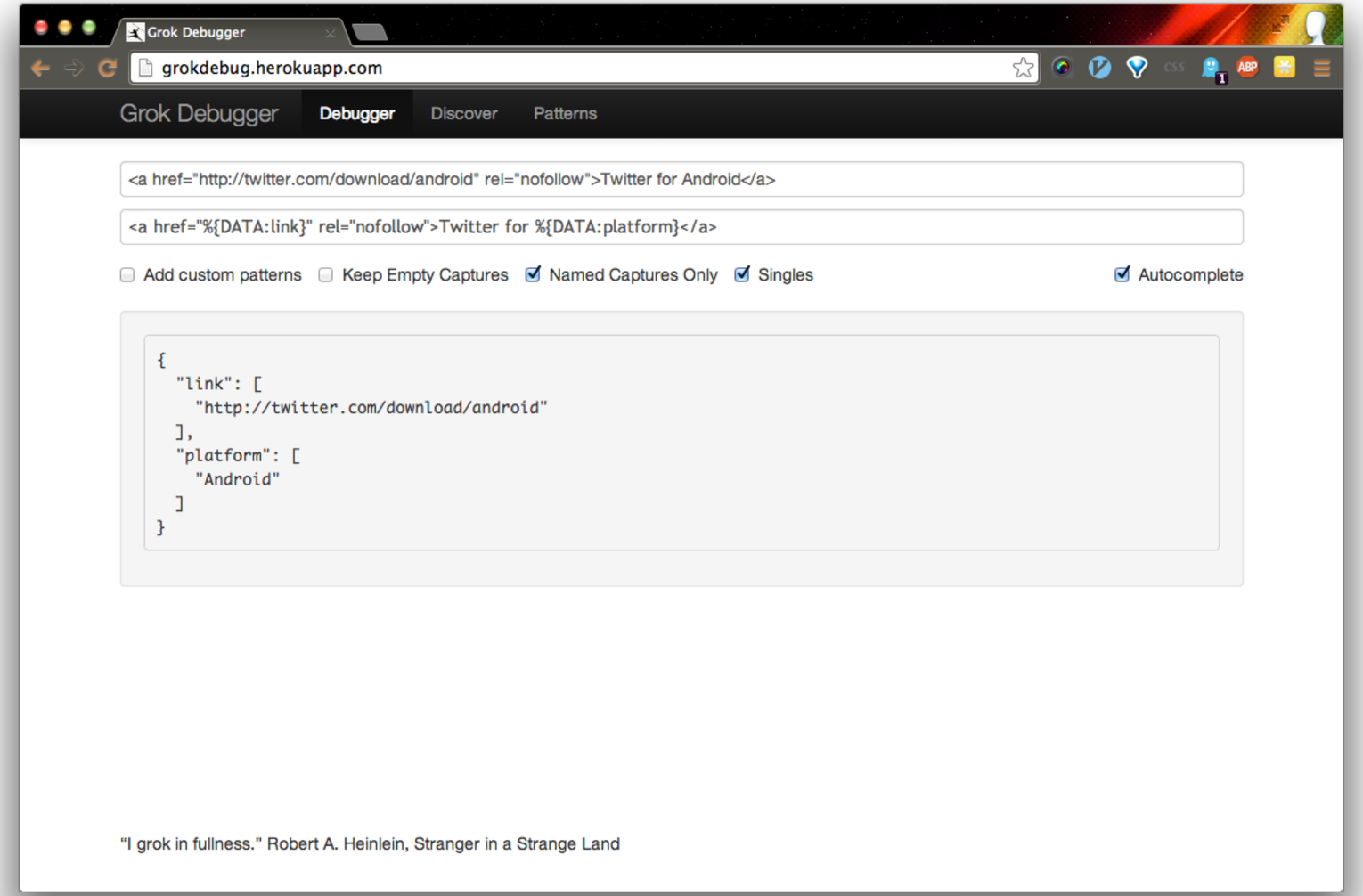
```
root@precise64:/home/vagrant# curator -c 7 -b 2 -d 10
2014-04-21T17:57:19.419 INFO                          main:333  Job starting...
2014-04-21T17:57:19.420 INFO                      _new_conn:180  Starting new HTTP connection (1): localhost
2014-04-21T17:57:19.422 INFO          log_request_success:49   GET http://localhost:9200/ [status:200 request:0.002s]
2014-04-21T17:57:19.423 INFO                          main:359  Deleting indices older than 10 days...
2014-04-21T17:57:19.430 INFO          log_request_success:49   GET http://localhost:9200/logstash-*/_settings?
expand_wildcards=closed [status:200 request:0.007s]
2014-04-21T17:57:19.433 INFO         find_expired_indices:209  logstash-2014.04.21 is 10 days, 0:00:00 above the cutoff.
2014-04-21T17:57:19.433 INFO                    index_loop:309  DELETE index operations completed.
2014-04-21T17:57:19.433 INFO                          main:364  Closing indices older than 7 days...
2014-04-21T17:57:19.434 INFO          log_request_success:49   GET http://localhost:9200/logstash-*/_settings?
expand_wildcards=closed [status:200 request:0.001s]
2014-04-21T17:57:19.435 INFO         find_expired_indices:209  logstash-2014.04.21 is 7 days, 0:00:00 above the cutoff.
2014-04-21T17:57:19.435 INFO                    index_loop:309  CLOSE index operations completed.
2014-04-21T17:57:19.435 INFO                          main:369  Disabling bloom filter on indices older than 2 days...
2014-04-21T17:57:19.437 INFO          log_request_success:49   GET http://localhost:9200/logstash-*/_settings?
expand_wildcards=closed [status:200 request:0.002s]
2014-04-21T17:57:19.438 INFO         find_expired_indices:209  logstash-2014.04.21 is 2 days, 0:00:00 above the cutoff.
2014-04-21T17:57:19.438 INFO                    index_loop:309  DISABLE BLOOM FILTER FOR index operations completed.
2014-04-21T17:57:19.438 INFO                          main:379  Done in 0:00:00.020348.
```

BigDesk



bigdesk.org

Grok Debugger



grokdebug.herokuapp.com

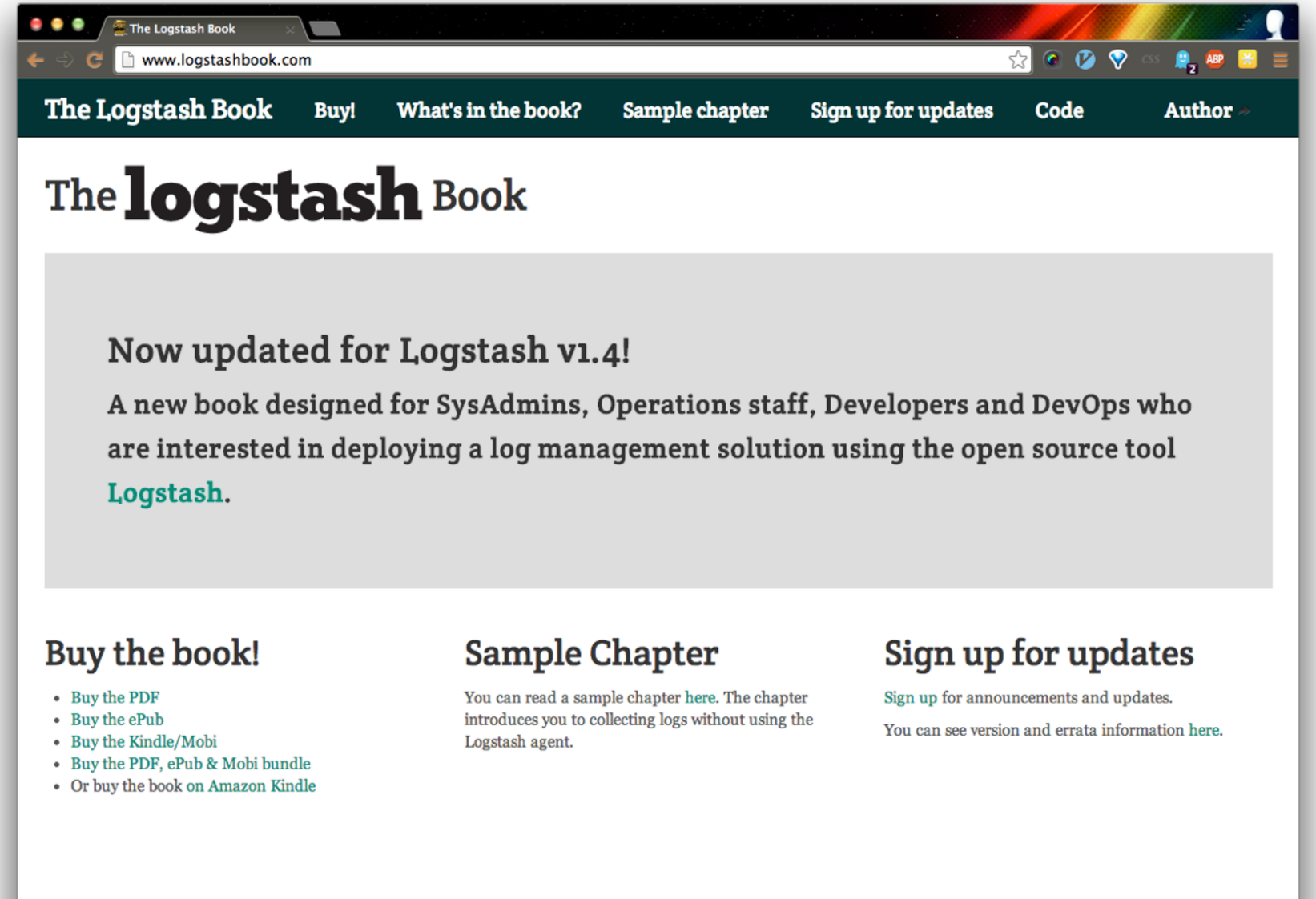
Amazee Labs

Logstash Cookbook

cookbook.logstash.net

Amazee Labs

The Logstash Book



logstashbook.com

DEMO!

~~Logfiles~~

Logstash

Elasticsearch

Kibana

Take Home

- Centralized Logging saves time
- Is fun with the ELK Stack
- Gives you Graphs to Interpret
- „can you check the errors from yesterday between 15.02 and 15.07“ get’s A LOT easier
- Start here tomorrow: http://logstash.net/docs/1.4.0/tutorials/getting-started-with-logstash

Amazee Labs

Thank you for having me here!

Slides : http://s.nrdy.ch/campus-logging

Images Used

- Elk : https://www.flickr.com/photos/ucumari/353839518/
- Paper Stash : https://www.flickr.com/photos/shehan365/8394630603/
- Architecture : https://www.flickr.com/photos/dasrecht/6743411525/

Amazee Labs