

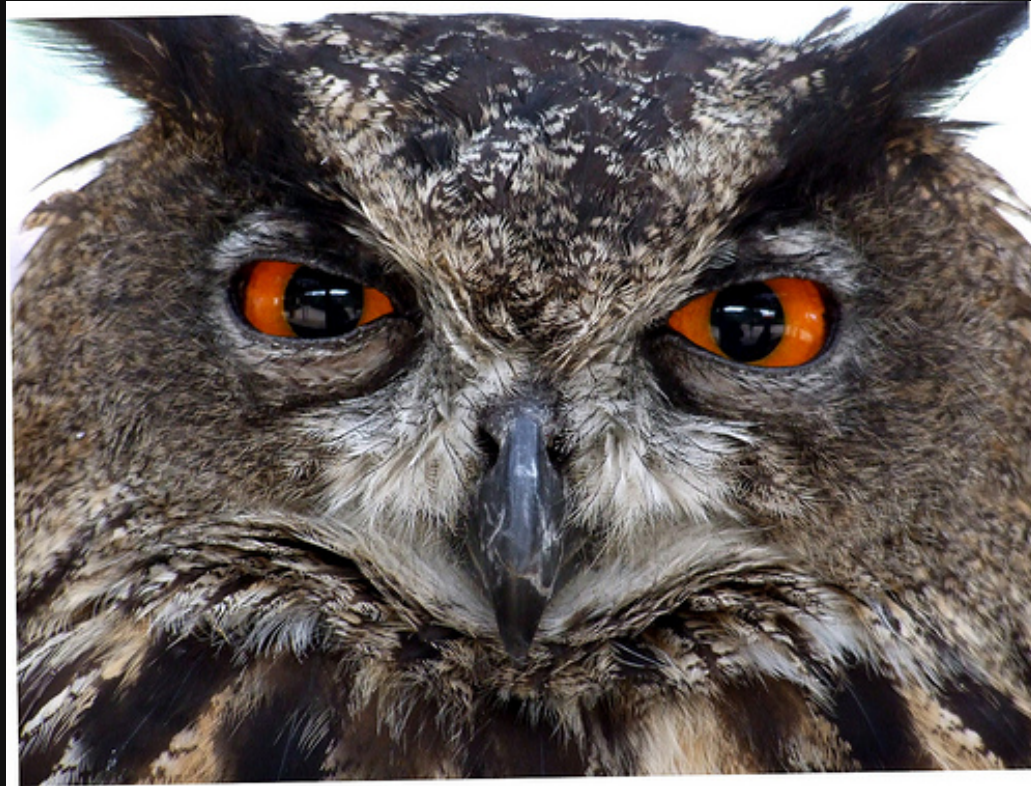
Yes, Logging Can Be Awesome



James Turnbull

@kartar

who



- operations chap
- Puppet chap
- erstwhile Ruby chap
- funny accent

(photo by Jennie Rainsford)

other matters

author

<http://www.jamesturnbull.net>

hack-n-slash developer

<https://github.com/jamtur01>

pontification

<http://www.kartar.net>

books

- **Pro Puppet**
- **Pro Linux System Administration**
- **Pro Nagios 2.0**
- **Hardening Linux**

the logstash book

The **logstash** Book
Log management made easy



James Turnbull

So who are you folks?



so what's a log



(photo by Rick Payette)

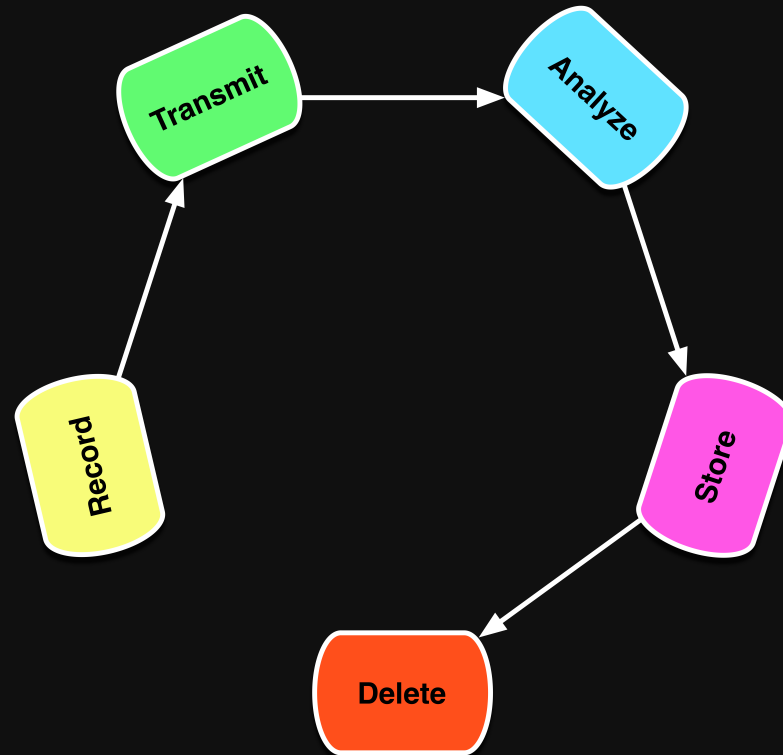

```
May  7 16:07:10 pelin systemd[1]: Starting Command Scheduler...
```

```
May  7 16:07:10      < timestamp
```

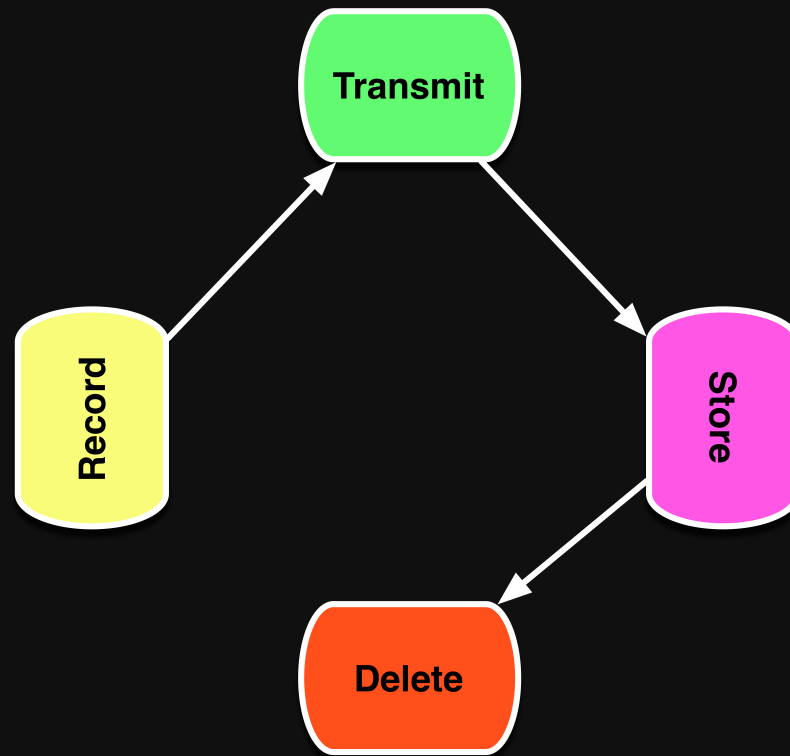
```
pelin systemd[1]: Starting Command  
Scheduler...      < data
```

timestamp + data = log

lifecycle of a log



actual lifecycle of a log



actual actual lifecycle of a log



so why isn't logging awesome?

I'll tell you a story



123.151.148.182 - - [11/May/2013:20:48:25 -0400] "GET /2010/08/rag-of-the-week-busted/trackback HTTP/1.1" 302 5 "http://www.stumpdinpx.com/"
"Mozilla/5.0 (compatible; Sosospider/2.0; +http://help.soso.com/web spider.htm)"
123.151.148.182 - - [11/May/2013:20:48:25 -0400] "GET /2010/08/rag-of-the-week-busted/ HTTP/1.1" 200 11678 "http://www.stumpdinpx.com/"
"Mozilla/5.0 (compatible; Sosospider/2.0; +http://help.soso.com/web spider.htm)"
96.126.127.108 - - [11/May/2013:20:48:35 -0400] "POST /wp-cron.php?doing_wp_cron=1368319715.1563251018524169921875 HTTP/1.0" 200 0 "-"
"WordPress/3.5.1; http://www.stumpdinpx.com"
123.151.148.182 - - [11/May/2013:20:48:35 -0400] "GET /2010/08/rag-of-the-week-busted/feed HTTP/1.1" 301 5 "http://www.stumpdinpx.com/"
"Mozilla/5.0 (compatible; Sosospider/2.0; +http://help.soso.com/web spider.htm)"
123.151.148.182 - - [11/May/2013:20:48:35 -0400] "GET /2010/08/rag-of-the-week-busted/feed/ HTTP/1.1" 200 2559 "http://www.stumpdinpx.com/"
"Mozilla/5.0 (compatible; Sosospider/2.0; +http://help.soso.com/web spider.htm)"
107.20.202.46 - - [11/May/2013:20:52:34 -0400] "GET /feed/ HTTP/1.1" 200 135969 "-" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_6; en-US)
AppleWebKit/534.16 (KHTML, like Gecko) Chrome/10.0.648.204 Safari/534.16"
107.20.202.46 - - [11/May/2013:20:52:34 -0400] "GET /feed/ HTTP/1.1" 200 135969 "-" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_6_6; en-US)
AppleWebKit/534.16 (KHTML, like Gecko) Chrome/10.0.648.204 Safari/534.16"
96.126.127.108 - - [11/May/2013:20:54:02 -0400] "POST /wp-cron.php?doing_wp_cron=1368320042.6065499782562255859375 HTTP/1.0" 200 0 "-"
"WordPress/3.5.1; http://www.stumpdinpx.com"
92.64.254.225 - - [11/May/2013:20:54:03 -0400] "POST /wp-login.php HTTP/1.0" 200 4452 "-" "Mozilla/3.0 (compatible; Indy Library)"
209.85.238.233 - - [11/May/2013:21:07:01 -0400] "GET /feed/ HTTP/1.1" 200 46099 "-" "Feedfetcher-Google;
(+http://www.google.com/feedfetcher.html; 48 subscribers; feed-id=5312968832043971344)"
121.219.57.195 - - [11/May/2013:21:08:21 -0400] "GET / HTTP/1.1" 200 6142 "-" "Reeder/1020.09.00 CFNetwork/596.3.3 Darwin/12.3.0 (x86_64)
(MacBookPro8%2C2)"
121.219.57.195 - - [11/May/2013:21:08:21 -0400] "GET / HTTP/1.1" 200 6142 "-" "Reeder/1020.09.00 CFNetwork/596.3.3 Darwin/12.3.0 (x86_64)
(MacBookPro8%2C2)"
96.126.127.108 - - [11/May/2013:21:10:51 -0400] "POST /wp-cron.php?doing_wp_cron=1368321051.2980649471282958984375 HTTP/1.0" 200 0 "-"
"WordPress/3.5.1; http://www.stumpdinpx.com"
94.125.180.90 - - [11/May/2013:21:10:51 -0400] "POST /wp-login.php HTTP/1.0" 200 4452 "-" "Mozilla/3.0 (compatible; Indy Library)"
217.34.181.76 - - [11/May/2013:21:10:51 -0400] "POST /wp-login.php HTTP/1.0" 200 4452 "-" "Mozilla/3.0 (compatible; Indy Library)"
96.126.127.108 - - [11/May/2013:21:12:09 -0400] "POST /wp-cron.php?doing_wp_cron=1368321129.5501360893249511718750 HTTP/1.0" 200 0 "-"
"WordPress/3.5.1; http://www.stumpdinpx.com"
190.199.60.150 - - [11/May/2013:21:12:09 -0400] "POST /wp-login.php HTTP/1.0" 200 4463 "http://www.stumpdinpx.com/wp-login.php" "Mozilla/4.0
(compatible; MSIE 6.0; Windows NT 5.1; SV1)"
184.154.100.20 - - [11/May/2013:21:12:56 -0400] "GET /2012/12/50-things-i-will-miss-about-portland/comment-page-1/ HTTP/1.0" 200 12699
"http://www.stumpdinpx.com/" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; Mozilla/4.0 (compatible; MSIE 6.0; Windows NT
5.1; SV1) ; .NET CLR 3.5.30729)"
96.126.127.108 - - [11/May/2013:21:13:29 -0400] "POST /wp-cron.php?doing_wp_cron=1368321209.4377140998840332031250 HTTP/1.0" 200 0 "-"
"WordPress/3.5.1; http://www.stumpdinpx.com"
217.91.37.3 - - [11/May/2013:21:13:29 -0400] "POST /wp-login.php HTTP/1.0" 200 4452 "-" "Mozilla/3.0 (compatible; Indy Library)"
80.93.213.249 - - [11/May/2013:21:15:32 -0400] "GET /2010/05/food-carts-of-melbourne-all-four-of-them/ HTTP/1.1" 200 16569
"http://www.stumpdinpx.com/" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; FunWebProducts; .NET CLR 1.1.4322; PeoplePal 6.2)"
80.93.213.249 - - [11/May/2013:21:15:33 -0400] "GET /2012/12/50-things-i-will-miss-about-portland/comment-page-1/ HTTP/1.1" 200 12720
"http://www.stumpdinpx.com/" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; FunWebProducts; .NET CLR 1.1.4322; PeoplePal 6.2)"

```
[11-May-2013 14:10:04 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:11:32 UTC] PHP Fatal error: Call to a member function setting() on a non-object in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/feedwordpress.php on line 606
[11-May-2013 15:21:58 UTC] PHP Fatal error: Call to a member function setting() on a non-object in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/feedwordpress.php on line 606
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 15:50:03 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
[11-May-2013 17:10:07 UTC] PHP Warning: Invalid argument supplied for foreach() in /var/www/html/planetdevops/wp-content/plugins/feedwordpress/magpiefromsimplepie.class.php on line 531
```

```
Jun 4, 2011 10:01:06 AM org.apache.coyote.http11.Http11Protocol init
INFO: Initializing Coyote HTTP/1.1 on http-8080
Jun 4, 2011 10:24:48 AM org.apache.catalina.loader.WebappClassLoader clearThreadLocalMap
SEVERE: The web application [] created a ThreadLocal with key of type [null] (value [clojure.lang.Var$1@564ca930]) and a value of type
[clojure.lang.Var.Frame] (value [clojure.lang.Var$Frame@42f7ba93]) but failed to remove it when the web application was stopped. This is very
likely to create a memory leak.
Jun 4, 2011 10:24:48 AM org.apache.catalina.loader.WebappClassLoader clearThreadLocalMap
SEVERE: The web application [] created a ThreadLocal with key of type [java.lang.ThreadLocal] (value [java.lang.ThreadLocal@15fa2b3e]) and a
value of type [clojure.lang.LockingTransaction] (value [clojure.lang.LockingTransaction@5b2cfeb7]) but failed to remove it when the web
application was stopped. This is very likely to create a memory leak.
Jun 4, 2011 10:24:50 AM org.apache.catalina.core.StandardContext resourcesStart
SEVERE: Error starting static Resources
java.lang.IllegalArgumentException: Document base /var/lib/tomcat6/webapps/ROOT does not exist or is not a readable directory
    at org.apache.naming.resources.FileDirContext.setDocBase(FileDirContext.java:142)
    at org.apache.catalina.core.StandardContext.resourcesStart(StandardContext.java:4249)
    at org.apache.catalina.core.StandardContext.start(StandardContext.java:4418)
    at org.apache.catalina.startup.HostConfig.checkResources(HostConfig.java:1244)
    at org.apache.catalina.startup.HostConfig.check(HostConfig.java:1342)
    at org.apache.catalina.startup.HostConfig.lifecycleEvent(HostConfig.java:303)
    at org.apache.catalina.util.LifecycleSupport.fireLifecycleEvent(LifecycleSupport.java:119)
    at org.apache.catalina.core.ContainerBase.backgroundProcess(ContainerBase.java:1337)
    at org.apache.catalina.core.ContainerBase$ContainerBackgroundProcessor.processChildren(ContainerBase.java:1601)
    at org.apache.catalina.core.ContainerBase$ContainerBackgroundProcessor.processChildren(ContainerBase.java:1610)
    at org.apache.catalina.core.ContainerBase$ContainerBackgroundProcessor.run(ContainerBase.java:1590)
    at java.lang.Thread.run(Thread.java:662)
Jun 4, 2011 10:24:50 AM org.apache.catalina.core.StandardContext start
SEVERE: Error in resourceStart()
Jun 4, 2011 10:24:50 AM org.apache.catalina.core.StandardContext start
SEVERE: Error getConfigured
```

**all of these logs tell us (useful)
stories**

**pretty confusing stories though
eh?**

so what's wrong?

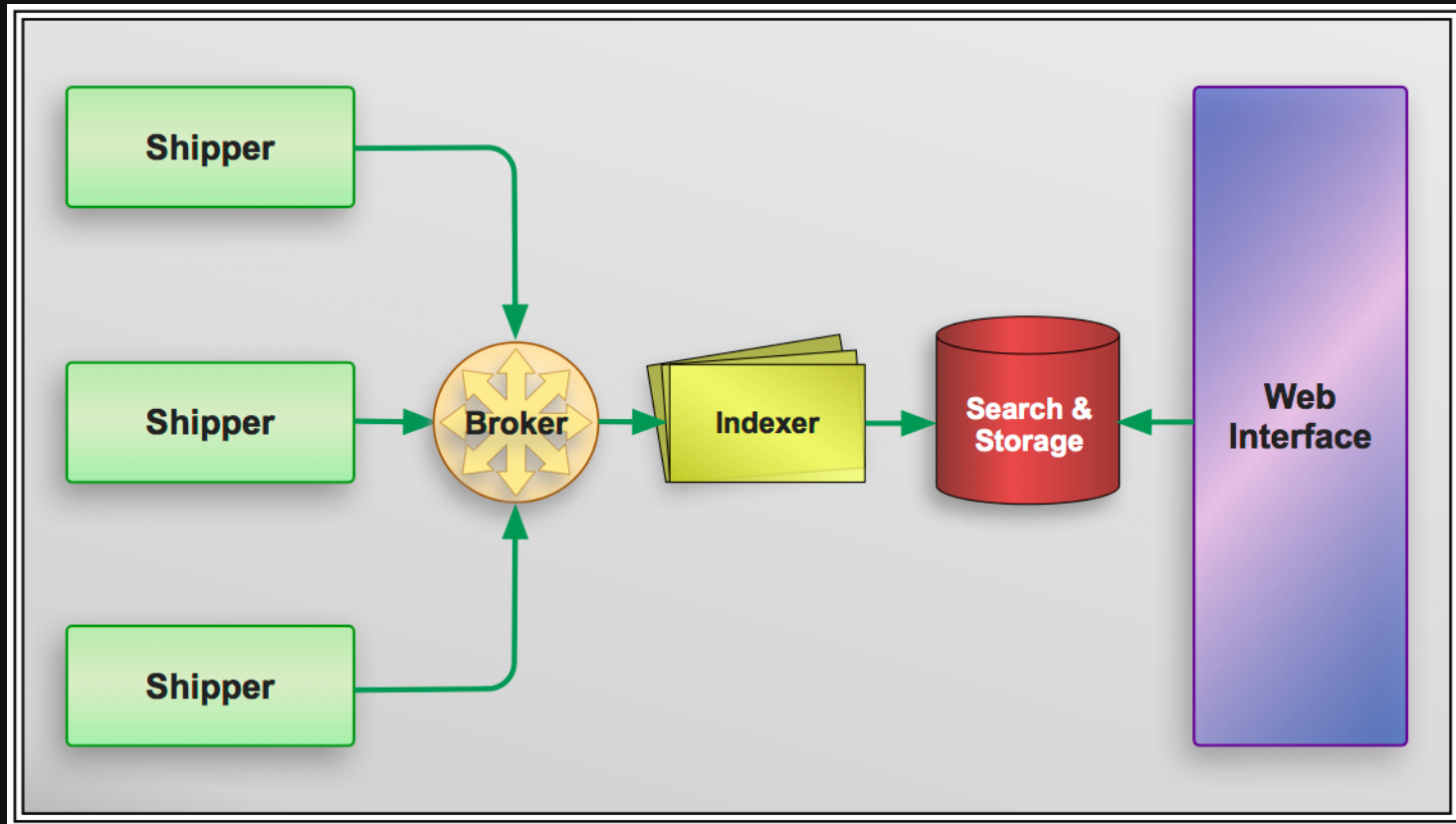
- so many sodding formats
- don't even get me started on timestamps
- no context
- really unhelpful error messages
- doesn't scale

enter logstash, parsing heavily

what?

- collects, transmits, interprets, stores
- free and open source
- primarily written by Jordan Sissel
- maxim: if a new user has a bad time, it's a bug in logstash
- awesome!

logstash architecture



simple is as simple does

```
input {  
  file {  
    type => "web"  
    path => "/var/log/httpd/access.log" }  
}
```

```
filter {  
  grok {  
    type => "web"  
    pattern => "%{COMBINEDAPACHELOG}" }  
}
```

```
date {  
  type => "web"  
  timestamp => "dd/MMM/yyyy:HH:mm:ss Z" }  
}
```

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

the input

```
input {  
  file {  
    type => "web"  
    path => "/var/log/httpd/access.log" }  
}
```


turns

```
202.46.63.192 - - [21/Jan/2013:16:41:38 -0800] "GET / HTTP/1.1" 200 935 "-"  
"Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.0)"
```

into

```
{"@source"=>"file://pelin.example.com/var/httpd/access.log", "@tags"=>[], "@fields"=>
{}}, {"@timestamp"=>"2013-01-21T16:41:38.030Z", "@source_host"=>"pelin.example.com",
"@source_path"=>"/var/log/httpd/access.log", "@message"=>"202.46.63.192 - -
[21/Jan/2013:16:41:38 -0800] GET / HTTP/1.1 200 935 - Mozilla/4.0 (compatible; MSIE
7.0; Windows NT 6.0)", "@type"=>"web"}
```

**still looks like a
mess eh?**

**but it's now a
structured mess!**

**structured data
for the win!**

the filters

```
grok {  
  type => "web"  
  pattern => "%{COMBINEDAPACHELOG}"  
}
```

use the power of regex



to add context

%{SYNTAX:SEMANTIC}

```
Log: May 12 03:36:31 pelin dhclient[2335]: DHCPACK from 97.107.143.38 (xid=0x6f62572d)
```

```
Grok: %{SYSLOGTIMESTAMP:timestamp} %{HOSTNAME:host} %{SYSLOGPROG:program}: %{DATA:message}
```

```
SYSLOGTIMESTAMP: %{MONTH} +%{MONTHDAY} %{TIME}
```

```
HOSTNAME: \b(?:[0-9A-Za-z][0-9A-Za-z-]{0,62})(?:\.(?:[0-9A-Za-z][0-9A-Za-z-]{0,62}))*(\.|b)
```

```
SYSLOGPROG %{PROG:program}(?:\[ %{POSINT:pid} \])?
```

remember this?

```
{"@source"=>"file:///pelin.example.com/var/httpd/access.log", "@tags"=>[], "@fields"=>
{}}, "@timestamp"=>"2013-01-21T16:41:38.030Z", "@source_host"=>"pelin.example.com",
"@source_path"=>"/var/log/httpd/access.log", "@message"=>"202.46.63.192 - -
[21/Jan/2013:16:41:38 -0800] GET / HTTP/1.1 200 935 - Mozilla/4.0 (compatible; MSIE
7.0; Windows NT 6.0)", "@type"=>"web"}
```

with grok it becomes

```
{ "@source"      => "file:///pelin.example.com/var/httpd/access.log",
  "@tags"        => [],
  "@fields"      => {
    "clientip": [ "202.46.63.192" ],
    "ident":    [ "-" ],
    "auth":     [ "-" ],
    "timestamp": [ "21/Jan/2013:16:41:38 -0800" ],
    "verb":     [ "GET" ],
    "request":  [ "/" ],
    "httpversion": [ "1.1" ],
    "response": [ "200" ],
    "bytes":    [ "935" ],
    "referrer": [ "\"-\"" ],
    "agent":    [ "\"Mozilla/4.0 (compatible; MSIE 7.0; Windows NT
6.0)\"" ] },
  "@timestamp"   => "2013-01-21T16:41:38.030Z",
  "@source_host" => "pelin.example.com",
  "@source_path" => "/var/log/httpd/access.log",
  "@message"     => "202.46.63.192 - - [21/Jan/2013:16:41:38 -0800] GET / HTTP/1.1 200
935 - Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.0)",
  "@type"        => "web" }
```

grok makes better

- over 100 patterns
- numbers, strings, hosts, network addresses, urls, etc
- chain patterns together
- easy to extend, easy to test

you can test your patterns

<http://grokdebug.herokuapp.com/>

**or you can even write tests for
your patterns**

you write tests right?

did I mention time?

```
date {  
  type => "web"  
  timestamp => "dd/MMM/yyyy:HH:mm:ss Z" }  
}
```

problem?

so many fucking time formats

**seriously. stop adding time
formats.**

solution.

standardize with the time filter.

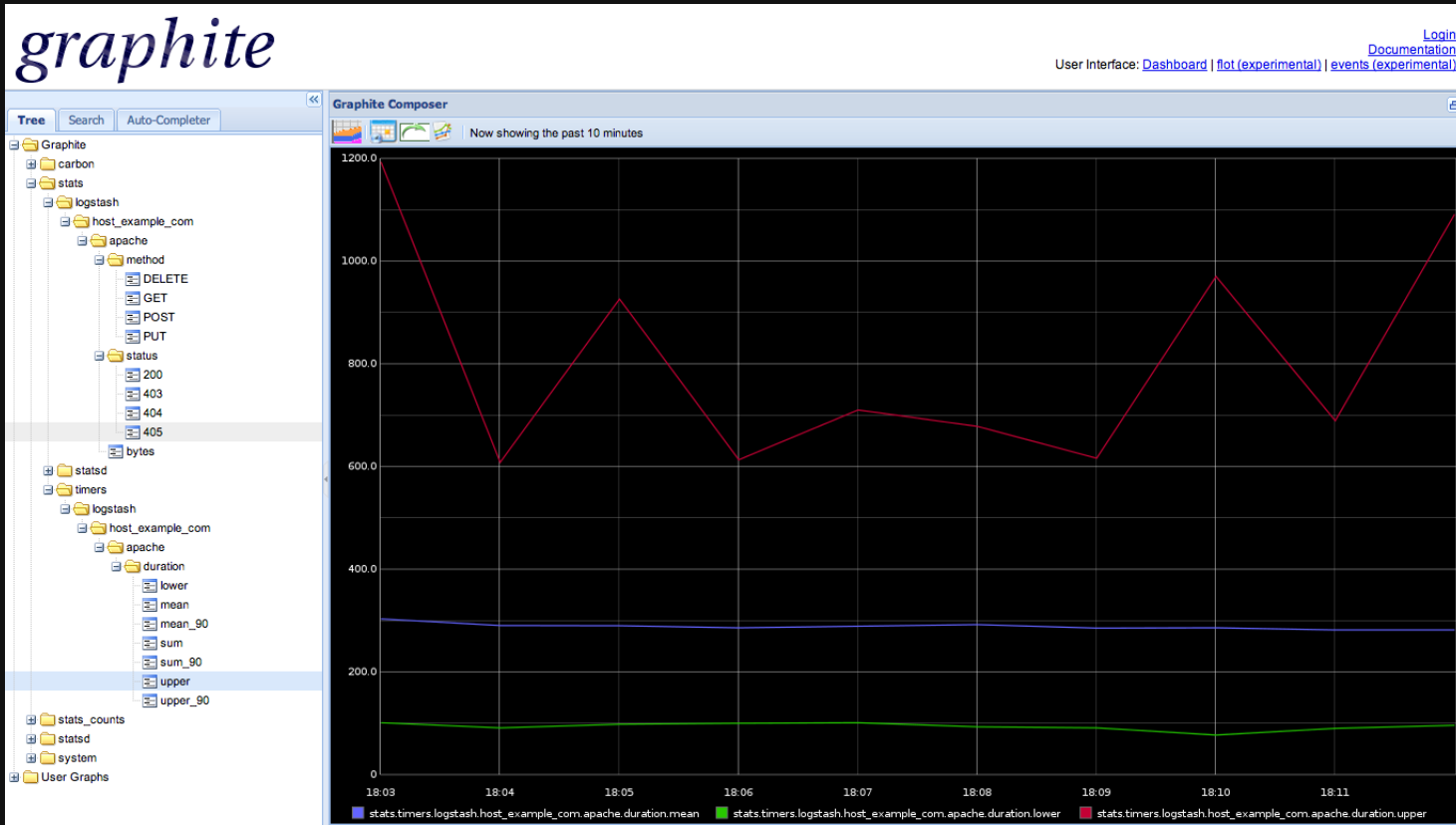
filters rock

- 30+ filters
- munge, mangle, mutate
- lookup, research, aggregate

filters turn abstract information like

```
202.46.63.192 - - [21/Jan/2013:16:41:38 -0800] "GET / HTTP/1.1" 200 935 "-"  
"Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.0)"
```

into



**the truth will set you free
... or at least wake you up.**

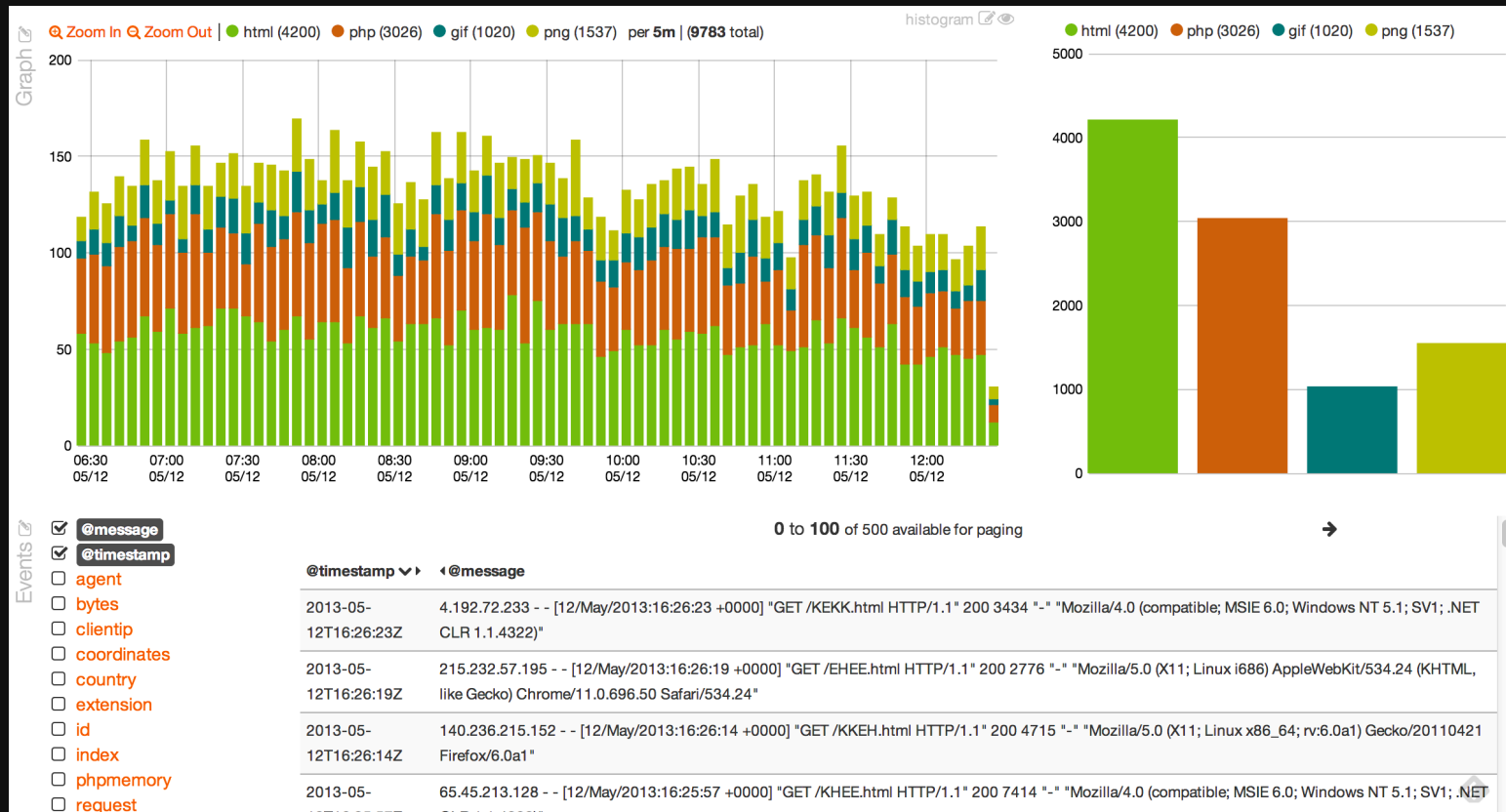
outputs

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

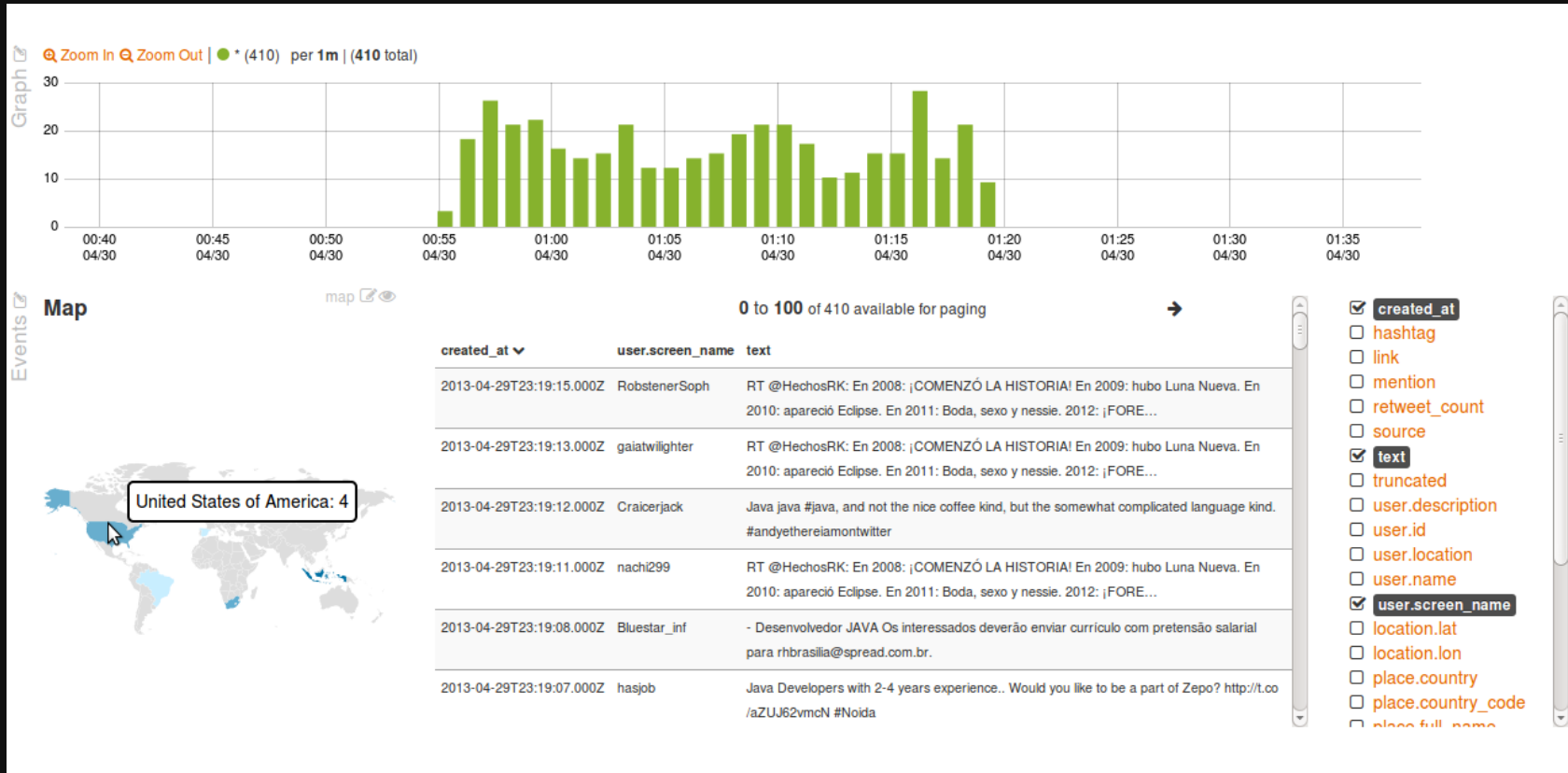
outputs

- 50+ outputs
- search, store, transit
- email, irc, alert
- graph, aggregate, execute

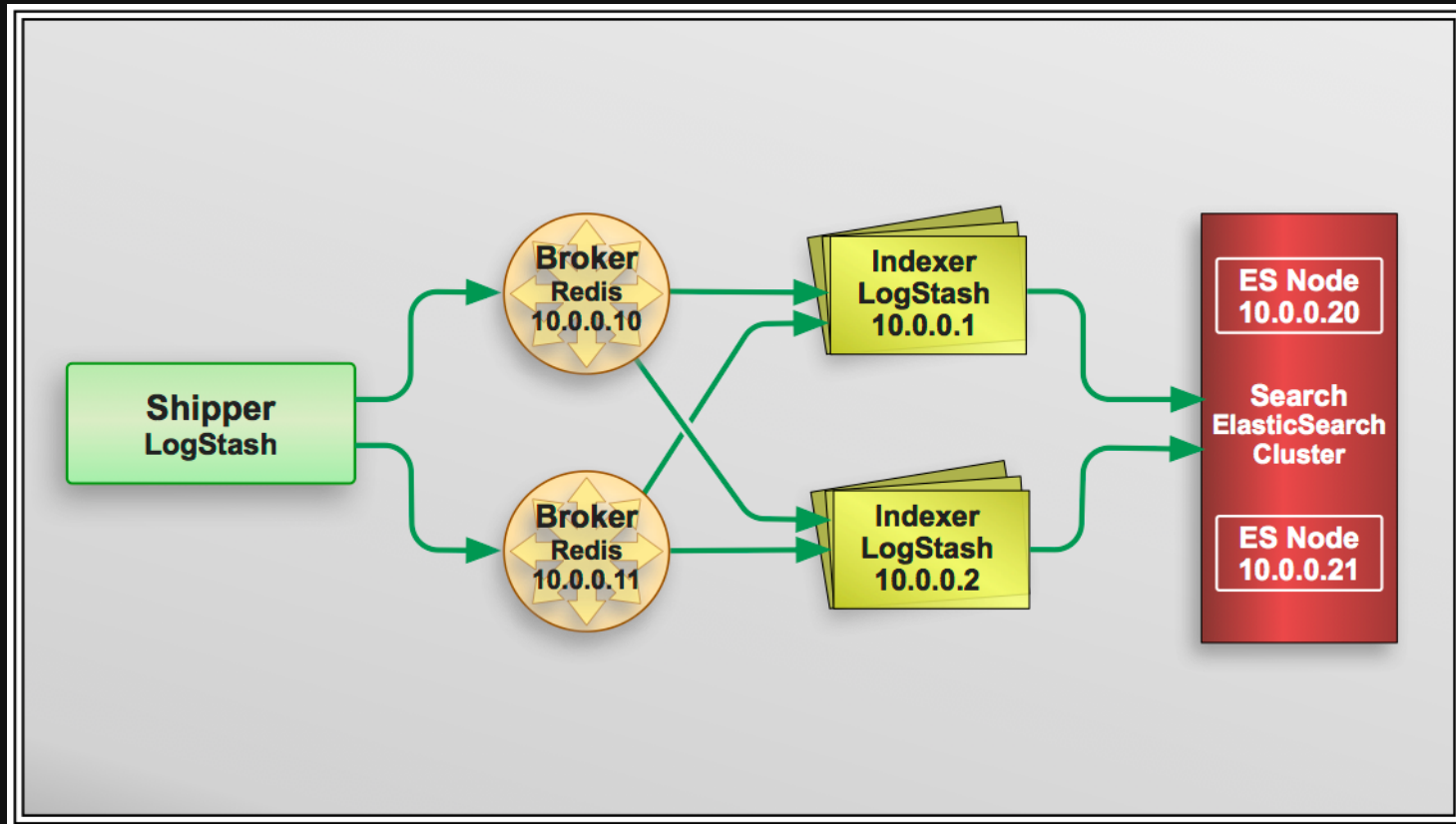
all of the pretty things



all of the pretty things



scales like a mofo



all of the logstashes

- logstash.net
- logstash-users@googlegroups.com
- #logstash on freenode irc
- logstash.jira.com

Questions?

references

- Doctor Who © BBC
- He-Man © Mattel