

*"Skyline at Night" by Knight Foundation, used under CC BY-SA 2.0 / Converted to black and white*



# Taming logs with Elasticsearch, Logstash, and Kibana

1DEVDAY Detroit 2014  
Dan Grabowski

- Introductions
- Problem
- Solution
- Demo
- Quick Start
- Experiences
- Questions

My favorite programming language is Clojure.

My favorite technical talk is “Are We There Yet?” by Rich Hickey (<http://www.infoq.com/presentations/Are-We-There-Yet-Rich-Hickey>).

The most interesting thing I’ve read recently is Kyle Kingsbury’s Jepsen series of blog posts (<http://aphyr.com/tags/Jepsen>).

Non-technical topics I’ll talk forever about if you get me started include Formula 1 and The Wire.

# Problem

Extracting information from  
large volumes of log data,  
making it accessible, and  
keeping it visible

# Large

- On the order of
  - 10 million log events per day
  - A few dozen applications and services
  - Several dozen servers

# Information

- Details of an occurrence of an issue
- Count and patterns of occurrences for an issue
- Future occurrences of an issue
- Performance trends and variations

# Accessible

- Available to developers, business analysts, system administrators, customer support, etc.
- Can be used effectively by all these groups, perhaps with different levels of sophistication

# Visible

- Keep monitoring in front of people
- ... but don't nag them












# Solution

- Elasticsearch to store and index log events
- Logstash to monitor, parse, and load logs
- Kibana as a UI to search, visualize information, and build dashboards
- Elasticsearch HTTP API to extract data for further analysis










# Elasticsearch (the product)

- Distributed Lucene indexes
- HTTP API
- Runs on JVM
- Apache 2.0 license
- Maintained and supported by Elasticsearch (the company)










# Cluster

   	<b>logstash-2014.10.05</b> ▼ shards: 2 * 2   docs: 18209   size: 9.63MB	<b>logstash-2014.10.06</b> ▼ shards: 2 * 2   docs: 300544   size: 148.01MB	<b>logstash-2014.10.07</b> ▼ shards: 2 * 2   docs: 10010   size: 5.27MB
★ <b>node2</b>  OKE-MACBOOK-MB.local - inet[/1]  <div><div></div>heap<div></div>disk<div></div>cpu</div>	<div>01</div>	<div>01</div>	<div>01</div>
☆ <b>node1</b>  OKE-MACBOOK-MB.local - inet[/1]  <div><div></div>heap<div></div>disk<div></div>cpu</div>	<div>01</div>	<div>01</div>	<div>01</div>
🔍 <b>logstash</b>  OKE-MACBOOK-MB.local - inet[/1] <div><div></div>heap<div></div>disk<div></div>cpu</div>			
❗ <b>unassigned shards</b>			






















# Nodes

<div><div></div><div></div></div>	logstash-2014.10.05 ▼	logstash-2014.10.06 ▼	logstash-2014.10.07 ▼
shards: 2 * 2   docs: 18209   size: 9.63MB		shards: 2 * 2   docs: 300544   size: 148.01MB	shards: 2 * 2   docs: 10010   size: 5.27MB
<div>★ <b>node2</b>  OKE-MACBOOK-MB.local - inet[/10.0.2.15:5044] <div><div></div><div>heap</div><div></div><div>disk</div><div></div><div>cpu</div></div></div> <div>☆ <b>node1</b>  OKE-MACBOOK-MB.local - inet[/10.0.2.15:5044] <div><div></div><div>heap</div><div></div><div>disk</div><div></div><div>cpu</div></div></div> <div>🔍 <b>logstash</b>  OKE-MACBOOK-MB.local - inet[/10.0.2.15:5044] <div><div></div><div>heap</div><div></div><div>disk</div><div></div><div>cpu</div></div></div> <div>🔔 <b>unassigned shards</b></div>	<div><div>0</div><div>1</div></div>	<div><div>0</div><div>1</div></div>	<div><div>0</div><div>1</div></div>
	<div><div>0</div><div>1</div></div>	<div><div>0</div><div>1</div></div>	<div><div>0</div><div>1</div></div>

# Indexes

<div></div> <div><div>★ node2</div><div> OKE-MACBOOK-MB.local - inet[/10.0.2.15:9200]</div><div><div><div></div><div></div><div></div></div><div>heapdiskcpu</div></div></div> <div><div>☆ node1</div><div> OKE-MACBOOK-MB.local - inet[/10.0.2.15:9200]</div><div><div><div></div><div></div><div></div></div><div>heapdiskcpu</div></div></div> <div><div>🔍 logstash</div><div> OKE-MACBOOK-MB.local - inet[/10.0.2.15:5044]</div><div><div></div><div></div><div></div></div><div>heapdiskcpu</div></div> <div><div>❗ unassigned shards</div></div>	<div><div>logstash-2014.10.05 ▾</div><div>shards: 2 * 2   docs: 18209   size: 9.63MB</div></div> <div><div>0</div><div>1</div></div>	<div><div>logstash-2014.10.06 ▾</div><div>shards: 2 * 2   docs: 300544   size: 148.01MB</div></div> <div><div>0</div><div>1</div></div>	<div><div>logstash-2014.10.07 ▾</div><div>shards: 2 * 2   docs: 10010   size: 5.27MB</div></div> <div><div>0</div><div>1</div></div>

# Shards

   	<b>logstash-2014.10.05</b> ▼ shards: 2 * 2   docs: 18209   size: 9.63MB	<b>logstash-2014.10.06</b> ▼ shards: 2 * 2   docs: 300544   size: 148.01MB	<b>logstash-2014.10.07</b> ▼ shards: 2 * 2   docs: 10010   size: 5.27MB
★ <b>node2</b>  OKE-MACBOOK-MB.local - inet[1]  <div><div></div>heap<div></div>disk<div></div>cpu</div>	 	 	 
☆ <b>node1</b>  OKE-MACBOOK-MB.local - inet[1]  <div><div></div>heap<div></div>disk<div></div>cpu</div>	 	 	 
🔍 <b>logstash</b>  OKE-MACBOOK-MB.local - inet[1] <div><div></div>heap<div></div>disk<div></div>cpu</div>			
❗ <b>unassigned shards</b>			

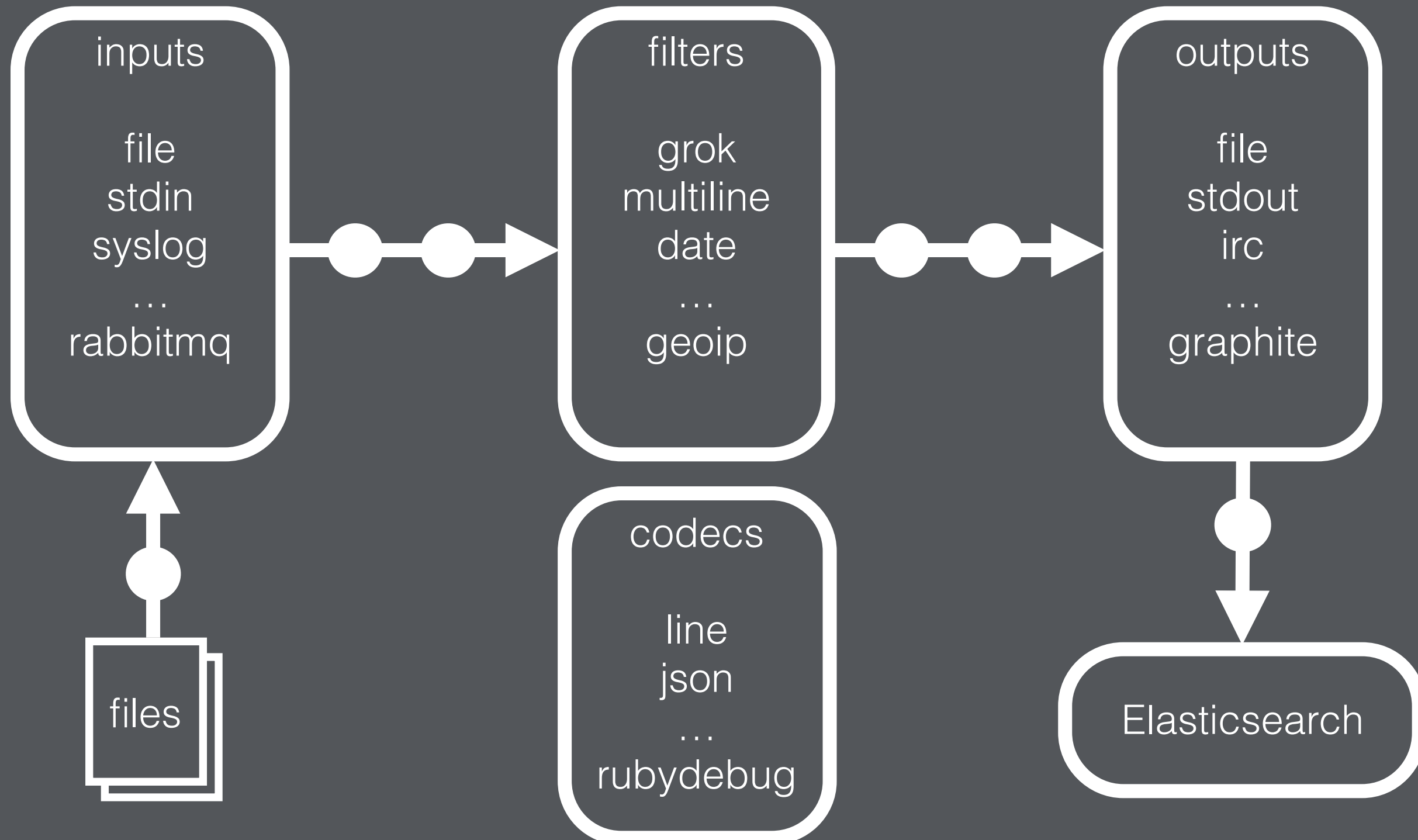
# Logstash

- Log stream processing
- Implemented with JRuby, runs on JVM
- Apache 2.0 license
- Created by Jordan Sissel
- Maintained and supported by Elasticsearch (the company)



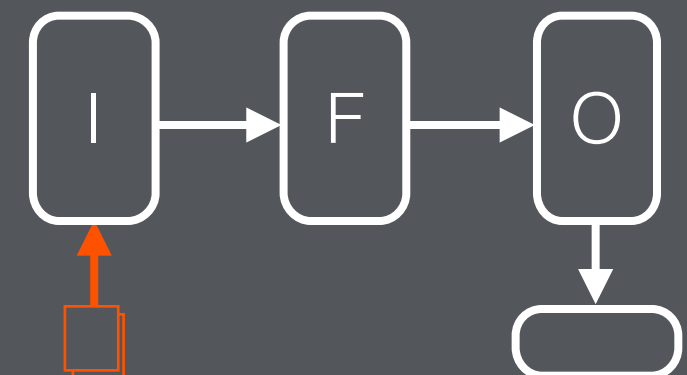
*Logstash logo from <http://logstash.net/images/logstash.png>*

# Logstash concepts

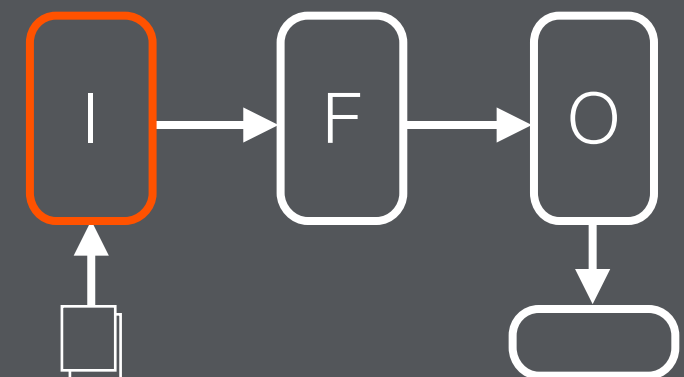




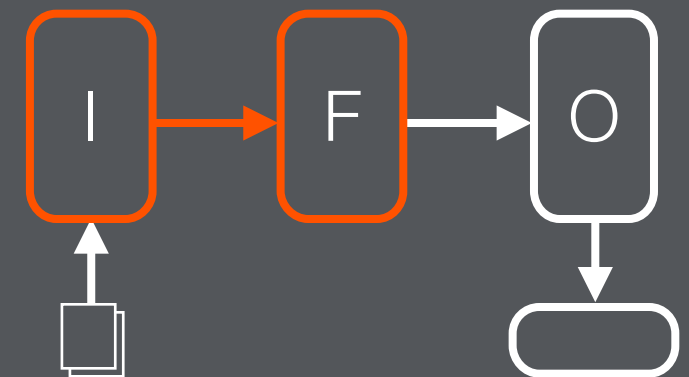
```
2014-10-11 12:52:22 GET /cgi-bin/ 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /default.asp 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /index.jsp 301 0 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /scripts/formmail.html 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /demo/../../%3f.jsp 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /q79w_38jg__.shtml 404 10049 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /displaytable.php 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /default.jsp 404 10049 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /scripts/mailform.html 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /forum/default.asp 404 10049 192.168.0.1 - - - 0.0010
```



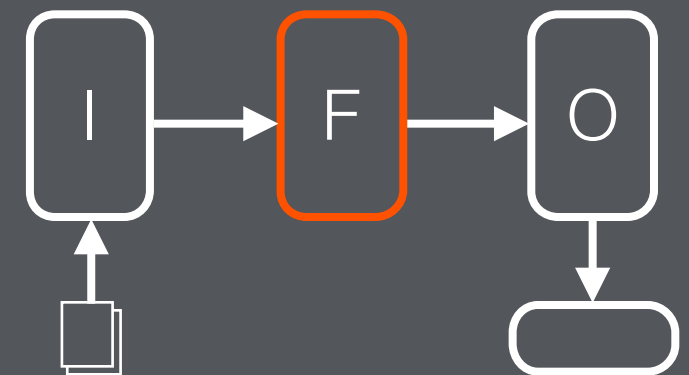
```
2014-10-11 12:52:22 GET /cgi-bin/ 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /index.jsp 301 0 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /demo/../../%3f.jsp 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /q79w_38jg__.shtml 404 10049 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /displaytable.php 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /default.jsp 404 10049 192.168.0.1 - - - 0.0010
2014-10-11 12:52:22 GET /scripts/mailform.html 404 10049 192.168.0.1 - - - 0.0
2014-10-11 12:52:22 GET /forum/default.asp 404 10049 192.168.0.1 - - - 0.0010
```



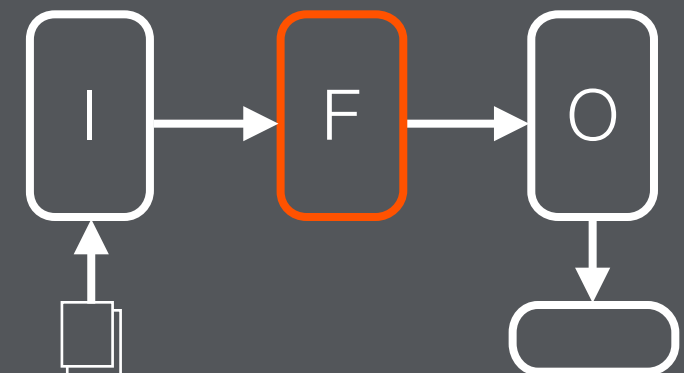
```
{  
    "message" => "2014-10-11 12:52:22 GET /index.jsp 301 0  
192.168.0.1 - - 0.0010",  
    "@version" => "1",  
    "@timestamp" => "2014-11-13T20:39:41.753Z",  
    "host" => "logstash.example.com",  
    "path" => "/logs/access.log"  
}
```



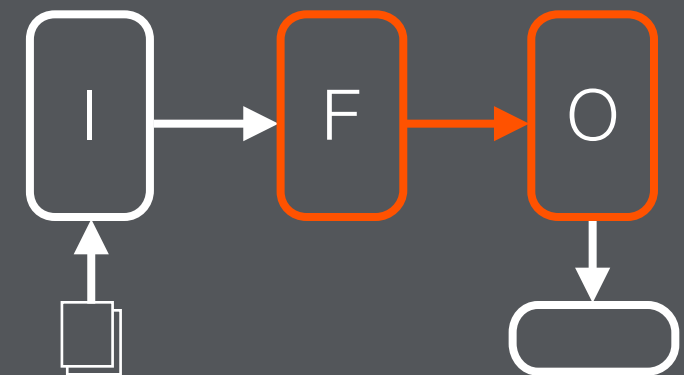
```
{
  "message" => "2014-10-11 12:52:22 GET /index.jsp 301 0
192.168.0.1 - - 0.0010",
  "@version" => "1",
  "@timestamp" => "2014-11-13T20:39:41.753Z",
  "host" => "logstash.example.com",
  "path" => "/logs/access.log",
  "date" => "2014-10-11",
  "time" => "12:52:22",
  "method" => "GET",
  "uri-path" => "/index.jsp",
  "status" => 301,
  "bytes" => 0,
  "ip-address" => "192.168.0.1",
  "query" => "-",
  "referrer" => "-",
  "user-agent" => "-",
  "elapsed_s" => 0.001
}
```



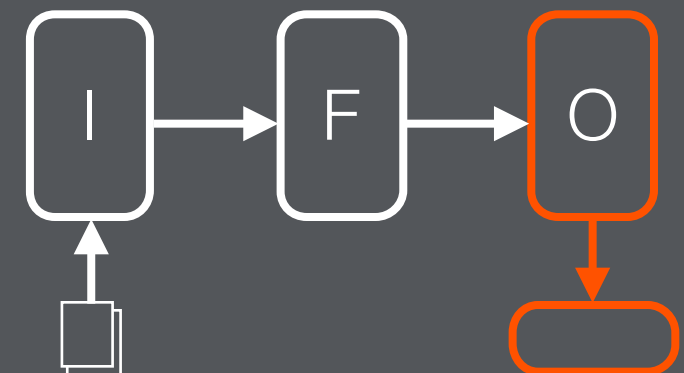
```
{  
  "message" => "2014-10-11 12:52:22 GET /index.jsp 301 0  
192.168.0.1 - - 0.0010",  
  "@version" => "1",  
  "@timestamp" => "2014-10-11T17:52:22.000Z",  
  "host" => "logstash.example.com",  
  "path" => "/logs/access.log",  
  "method" => "GET",  
  "uri-path" => "/index.jsp",  
  "status" => 301,  
  "bytes" => 0,  
  "ip-address" => "192.168.0.1",  
  "query" => "-",  
  "referrer" => "-",  
  "user-agent" => "-",  
  "elapsed_s" => 0.001  
}
```



```
{
  "message" => "2014-10-11 12:52:22 GET /index.jsp 301 0
192.168.0.1 - - 0.0010",
  "@version" => "1",
  "@timestamp" => "2014-10-11T17:52:22.000Z",
  "host" => "server1.example.com",
  "path" => "/logs/access.log",
  "method" => "GET",
  "uri-path" => "/index.jsp",
  "status" => 301,
  "bytes" => 0,
  "ip-address" => "192.168.0.1",
  "query" => "-",
  "referrer" => "-",
  "user-agent" => "-",
  "elapsed_s" => 0.001
}
```



```
{
  "message": "2014-10-11 12:52:22 GET /index.jsp 301 0
192.168.0.1 - - 0.0010",
  "@version": "1",
  "@timestamp": "2014-10-11T17:52:22.000Z",
  "host": "server1.example.com",
  "path": "/logs/access.log",
  "method": "GET",
  "uri-path": "/index.jsp",
  "status": 301,
  "bytes": 0,
  "ip-address": "192.168.0.1",
  "query": "-",
  "referrer": "-",
  "user-agent": "-",
  "elapsed_s": 0.001
}
```



# Kibana

- Javascript application that interacts with Elasticsearch HTTP API
- Provides search, visualization, and dashboard capabilities
- Apache 2.0 license
- Maintained and supported by Elasticsearch (the company)



*Kibana logo from <https://github.com/elasticsearch/kibana/blob/3.0/src/img/kibana.png>*





QUERY ▸

host:"server1" AND method:"GET" AND uri-path.raw:/customersV.\*/

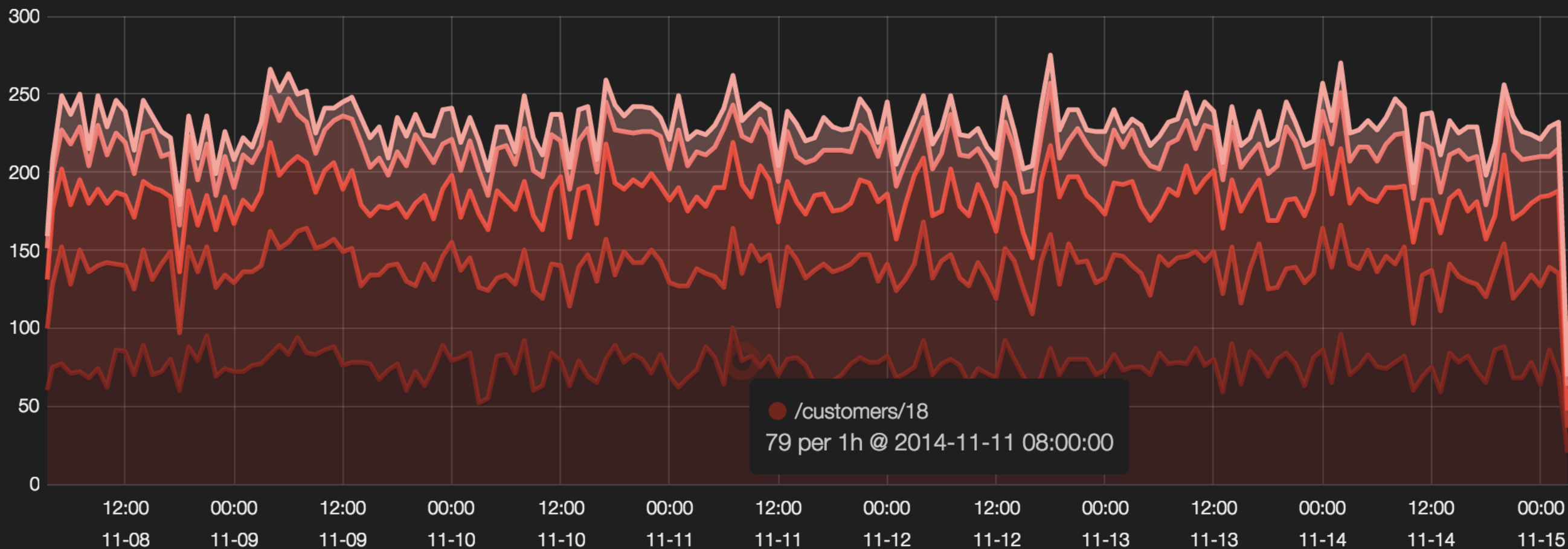


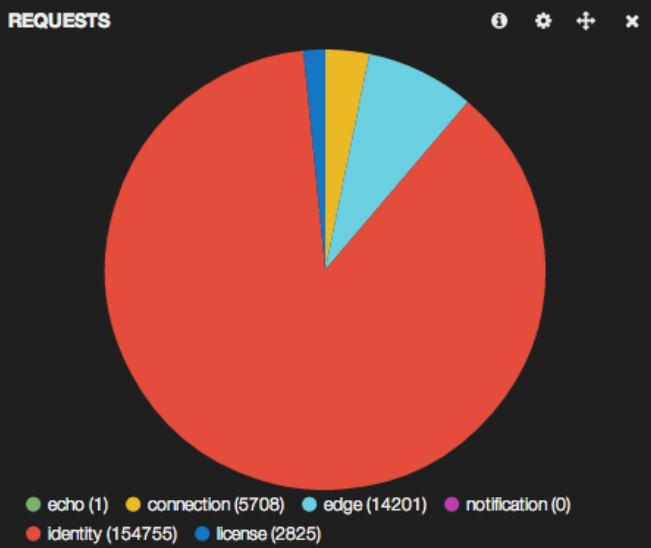
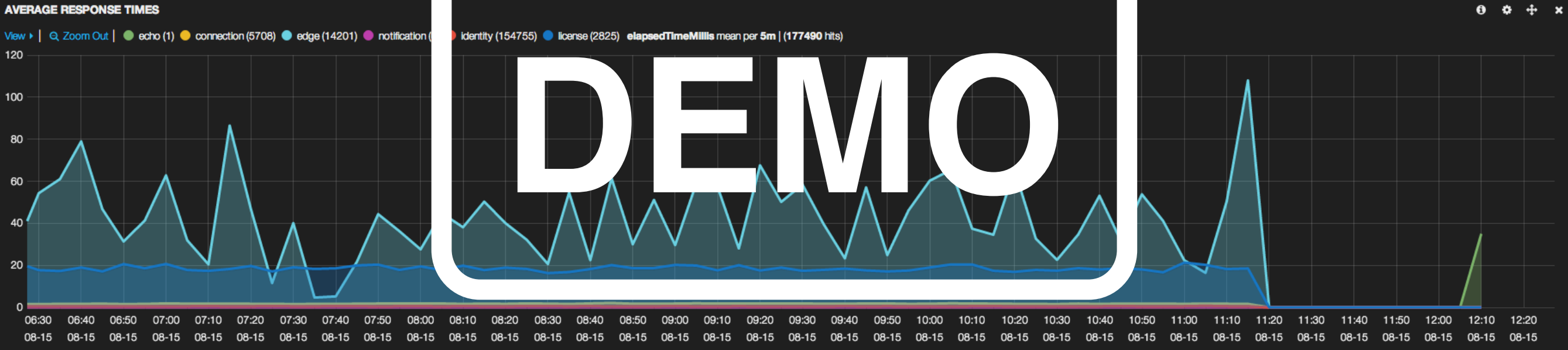
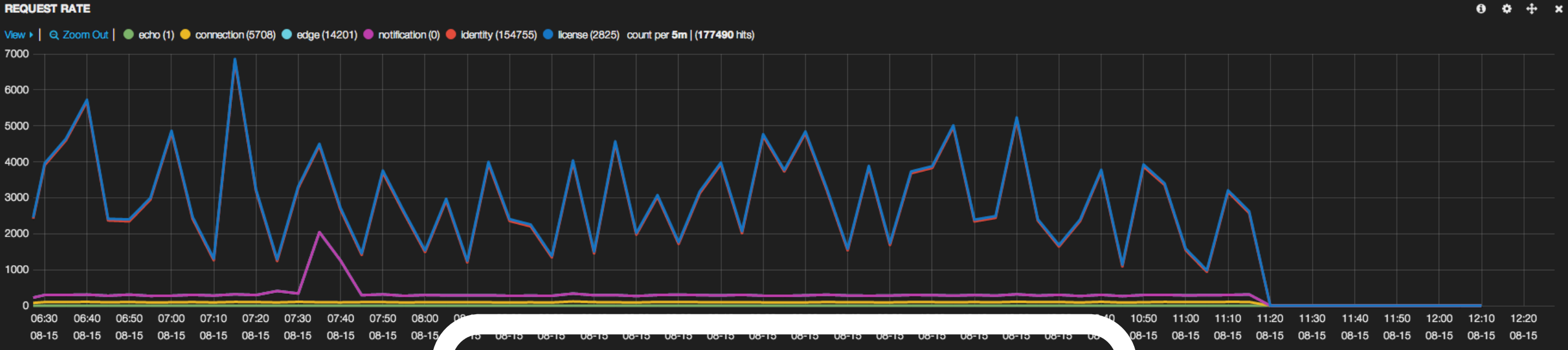
FILTERING ◀

## EVENTS OVER TIME



View ▸ | [Zoom Out](#) | ● /customers/18 (12747) ● /customers/18/orders (10484) ● /customers/5 (7724) ● /customers/5/orders (5173) ● /customers/32 (2602) count per 1h | (38730 hits)

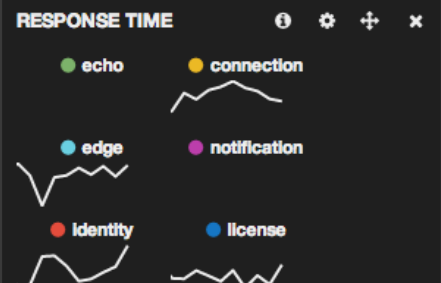
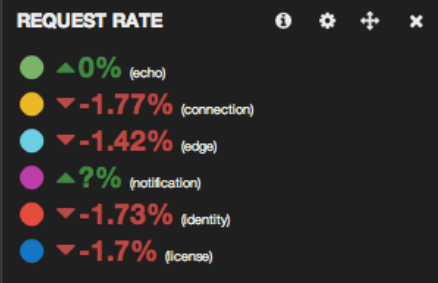




### RESPONSE SIZE (BYTES)

453 (mean)

Query	count	min	max	mean	total	variance	std_deviation
echo	1	6	6	6	6	0	0
connection	5,708	106	263,551	2,814	16,062,867	219,314,862	14,809
edge	14,201	0	642,733	3,003	42,648,398	536,104,447	23,154
notification	0	Infinity	-Infinity	0	0	0	0
Identity	154,755	102	221	124	19,163,573	167	13
license	2,825	307	15,807	904	2,554,249	3,986,432	1,997



# Demo/Quick Start Project

- Downloads, installs, and configures Elasticsearch, Logstash, Kibana, and kopf plugin
- Provides scripts for starting, stopping, and resetting data
- Generates log data to index
- Prerequisites
  - JDK (1.7 or 1.8 should work)
  - node.js
- Caveats
  - Shell scripts require OS X or Linux

# Setup Steps

1. <https://github.com/dgrabows/elk-demo>
2. Clone or download zip
3. Run `./install.sh`
4. Run `./start-all.sh`
5. Kibana: <http://localhost/9200>
6. Admin UI (kopf): [http://localhost:9200/\\_plugin/kopf](http://localhost:9200/_plugin/kopf)

# Experiences

- Don't underestimate power of rsync, find, grep, awk, sed, wc, cron, etc.
- Elasticsearch and Kibana very effective for exploring log data
- Elasticsearch has additional capabilities not exposed through Kibana (e.g. percentile aggregates)

# Experiences (cont.)

- One effective pattern
  - Analyze aggregate metrics with Elasticsearch, node.js, and spreadsheets
  - Drilled down into problem areas with Kibana and Elasticsearch
- You could get a lot done with a 4-8 core/32 GB RAM/1 TB disk server

Questions?

# References

- <http://www.elasticsearch.org>
- <http://logstash.net/>
- <https://github.com/elasticsearch> (Elasticsearch, Logstash, Kibana repos)
- <https://github.com/lmenezes/elasticsearch-kopf> (kopf plugin)



# Attributions

*Full attributions provided inline, where practical*

## ***Title Slide***

*"Skyline at Night" (<https://www.flickr.com/photos/knightfoundation/12119756375>) by Knight Foundation (<https://www.flickr.com/photos/knightfoundation/>), used under CC BY-SA 2.0 (<https://creativecommons.org/licenses/by-sa/2.0/>) / Converted to black and white*

Dan Grabowski

[dan.grabowski@gmail.com](mailto:dan.grabowski@gmail.com)



Except where otherwise noted, this work licensed under

<http://creativecommons.org/licenses/by-sa/4.0/>