elastic

Cloud Conf 2016

Awesome Logging Infrastructure Using The Elastic Stack

Alexander Reelsen
@spinscale

© München Tourismus

© München Tourismus

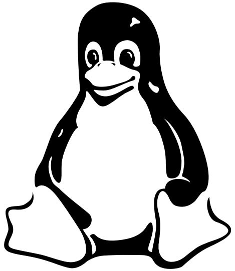
© München Tourismus

elasticsearch

x-pack

elasticsearch

x-pack



About Elastic

elasticsearch

x-pack

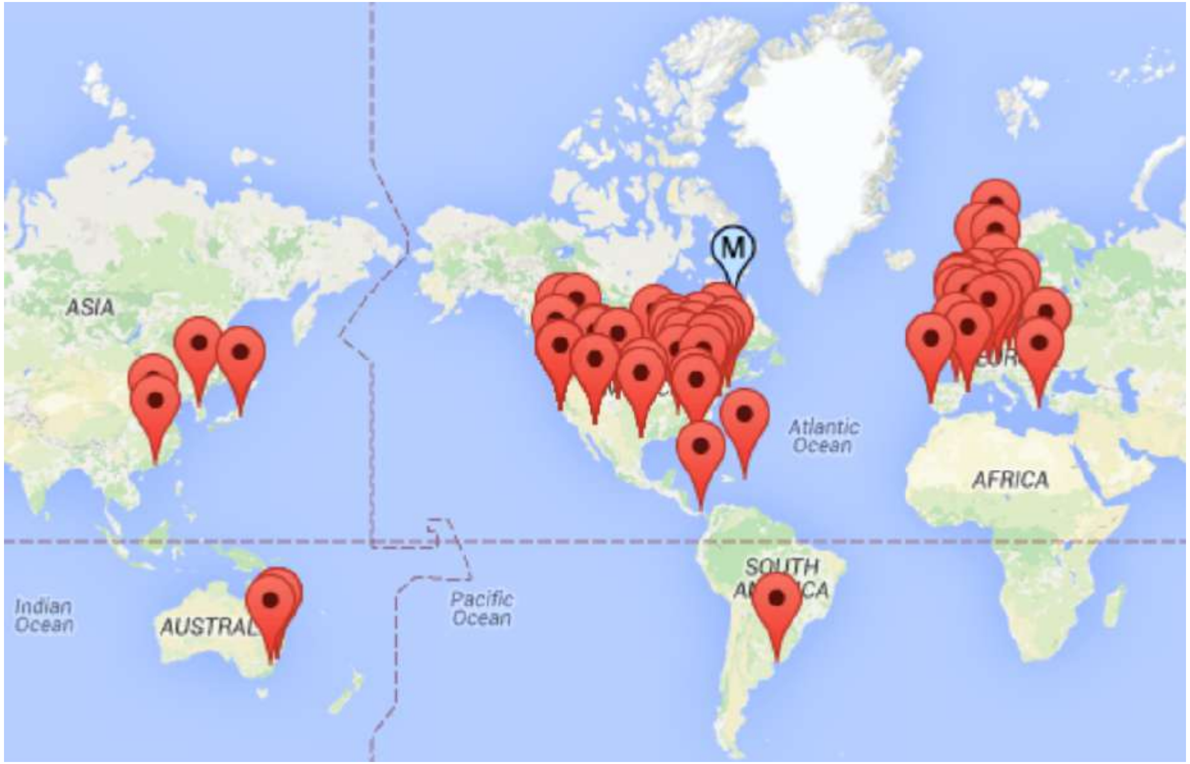
kibana

beats

logstash

cloud

About Elastic - Engineering team



elastic

Why logging?

How many users signed up to our newsletter this week?

Business Analyst

elastic

How successful is our advertising campaign?

Marketing Team

When should we schedule maintenance?

VP Operations

elastic

Why is the database slow?

Sysadmin

elastic

Logging is hard

Required Expertise

elastic

Access Rights

elastic

Unstructured Logging

Unstructured Logging

```
RemoteTransportException[[Anelle][127.0.0.1:9301][indices:data/read/percolate[s]]]; nested: PercolateException[failed to percolate]; nested: PercolateException[failed
to execute]; nested: NullPointerException;
Caused by: PercolateException[failed to percolate]; nested: PercolateException[failed to execute]; nested: NullPointerException;
        at org.elasticsearch.action.percolate.TransportPercolateAction.shardOperation(TransportPercolateAction.java:180)
        at org.elasticsearch.action.percolate.TransportPercolateAction.shardOperation(TransportPercolateAction.java:55)
        at org.elasticsearch.action.support.broadcast.TransportBroadcastAction$ShardTransportHandler.messageReceived(TransportBroadcastAction.java:268)
        at org.elasticsearch.action.support.broadcast.TransportBroadcastAction$ShardTransportHandler.messageReceived(TransportBroadcastAction.java:264)
        at org.elasticsearch.transport.TransportService$4.doRun(TransportService.java:350)
        at org.elasticsearch.common.util.concurrent.AbstractRunnable.run(AbstractRunnable.java:37)
        at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
        at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
        at java.lang.Thread.run(Thread.java:745)
Caused by: PercolateException[failed to execute]; nested: NullPointerException;
        at org.elasticsearch.percolator.PercolatorService$4.doPercolate(PercolatorService.java:583)
        at org.elasticsearch.percolator.PercolatorService.percolate(PercolatorService.java:254)
        at org.elasticsearch.action.percolate.TransportPercolateAction.shardOperation(TransportPercolateAction.java:177)
        ... 8 more
Caused by: java.lang.NullPointerException
        at org.apache.lucene.search.GeoPointTermQueryConstantScoreWrapper$1.getDocIDs(GeoPointTermQueryConstantScoreWrapper.java:86)
        at org.apache.lucene.search.GeoPointTermQueryConstantScoreWrapper$1.scorer(GeoPointTermQueryConstantScoreWrapper.java:126)
        at org.apache.lucene.search.LRUQueryCache$CachingWrapperWeight.scorer(LRUQueryCache.java:628)
        at org.apache.lucene.search.BooleanWeight.scorer(BooleanWeight.java:280)
        at org.apache.lucene.search.LRUQueryCache$CachingWrapperWeight.scorer(LRUQueryCache.java:628)
        at org.apache.lucene.search.BooleanWeight.scorer(BooleanWeight.java:280)
        at org.apache.lucene.search.LRUQueryCache$CachingWrapperWeight.scorer(LRUQueryCache.java:628)
        at org.apache.lucene.search.BooleanWeight.scorer(BooleanWeight.java:280)
        at org.apache.lucene.search.LRUQueryCache$CachingWrapperWeight.scorer(LRUQueryCache.java:628)
        at org.elasticsearch.common.lucene.Lucene.exists(Lucene.java:248)
        at org.elasticsearch.percolator.PercolatorService$4.doPercolate(PercolatorService.java:571)
        ... 10 more
```

Semi-Structured Logging

Semi-Structured Logging

```
Mar  6 10:02:42 my-host mosquitto[18881]: mosquitto version 0.15 (build date
2013-08-23 19:23:43+0000) starting

Mar  7 06:43:06 my-host CRON[28050]: (CRON) info (No MTA installed, discarding
output)

Mar  7 06:45:01 my-host CRON[28325]: (root) CMD (command -v debian-sa1 > /dev/null
&& debian-sa1 1 1)

Mar  7 12:01:40 my-host kernel: [256359.334516] init: meetup-stream main process
(24941) killed by TERM signal
```

Structured Logging

Structured Logging

```
{
   "error": {
      "root_cause": [
         {
            "type": "repository_exception",
            "reason": "[test-6] failed to create repository"
         }
      ],
      "type": "repository_exception",
      "reason": "[test-6] failed to create repository",
      "caused_by": {
         "type": "creation_exception",
         "reason": "Guice creation errors:\n\n1) …",
         "caused_by": {
            "type": "amazon_s3_exception",
            "reason": "The specified location-constraint is not valid (Service:
Amazon S3; Status Code: 400; Error Code: InvalidLocationConstraint; Request ID:
85CFF34E01878232)"
         }
      }
   },
   "status": 500
}
```

Structured Logging

```
1.2.3.4 - - [07/Mar/2016:09:57:02 +0100] "GET /posts/2015-05-04-producing-technical
documentation-an-overview.html HTTP/1.1" 200 11755 "-" "Mozilla/5.0 (Macintosh;
Intel Mac OS X 10_11_3) AppleWebKit/601.4.4 (KHTML, like Gecko)"
```

Timestamps

Timestamps

[29/Apr/2011:07:05:26 +0000]

Oct 11 20:21:47

130460505

@4000000037c219bf2ef02e94

020805 13:51:24

Enrichment

Centralization

Shipping

elastic

Analytics

elastic

Visualization

elastic

Alerting

elastic

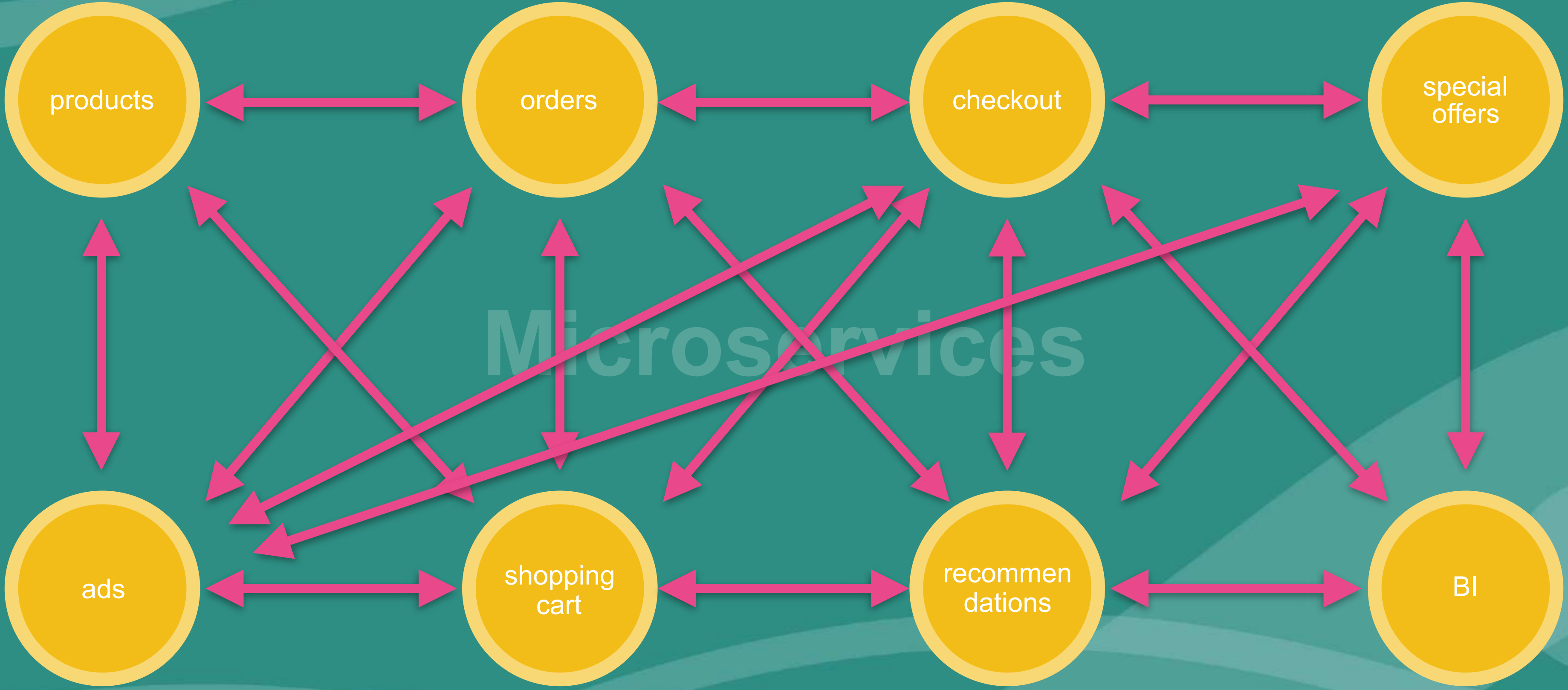
Outages

Peaks

Logging got harder!

Microservices

elastic



products
orders
checkout
special offers
Microservices
ads
shopping cart
recommen dations
BI

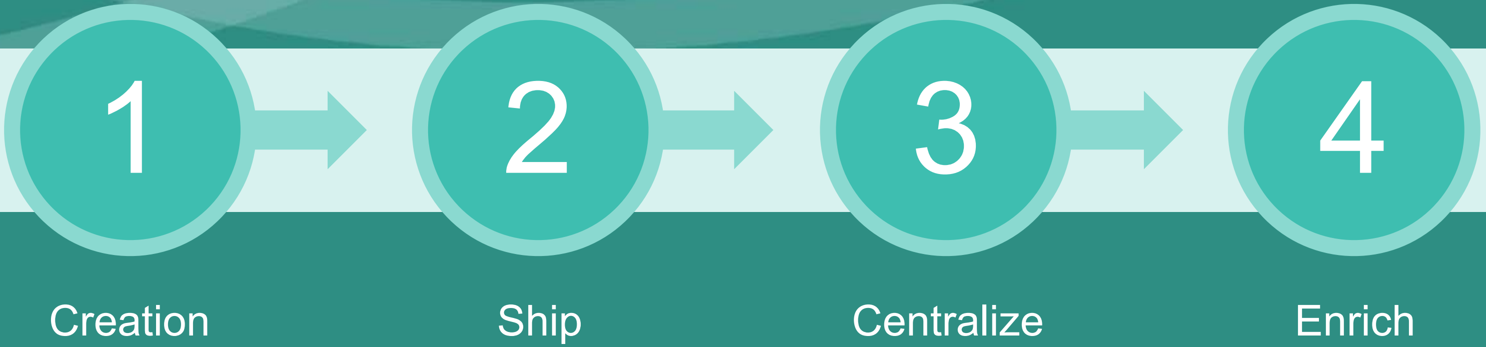
Serverless

Cluster/server/process management platforms

Short lived services

elastic

Lifecycle of a log



1 Creation

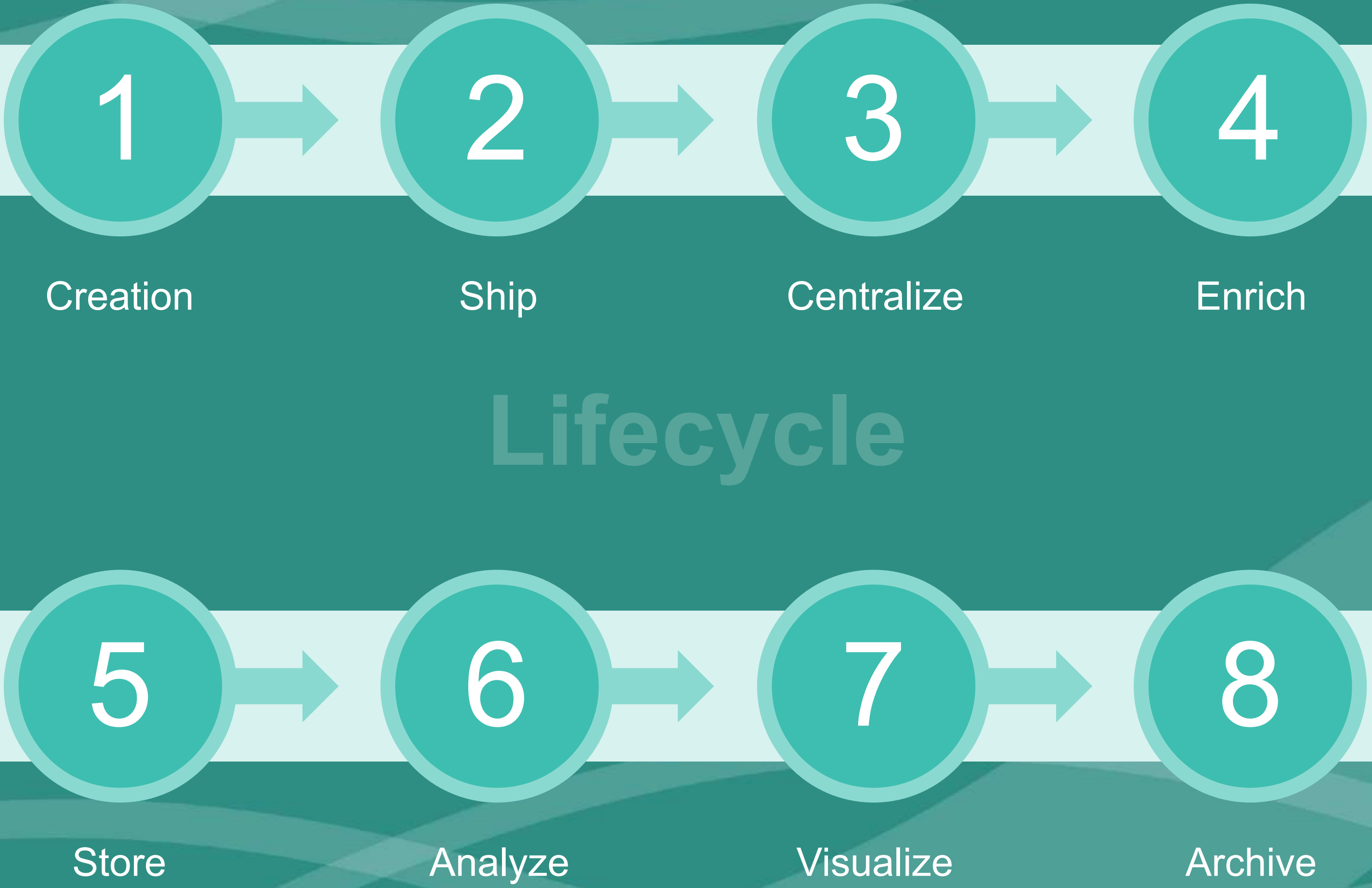
2 Ship

3 Centralize

4 Enrich

Lifecycle

Lifecycle



1 Creation

2 Ship

3 Centralize

4 Enrich

5 Store

6 Analyze

7 Visualize

8 Archive

elastic

Architecture

shipper

Architecture

Architecture

beats

receiver

Architecture

beats

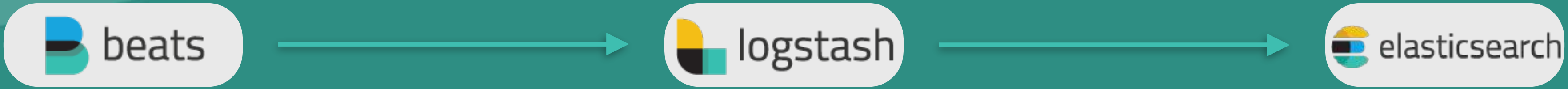
logstash

Architecture

beats

elasticsearch

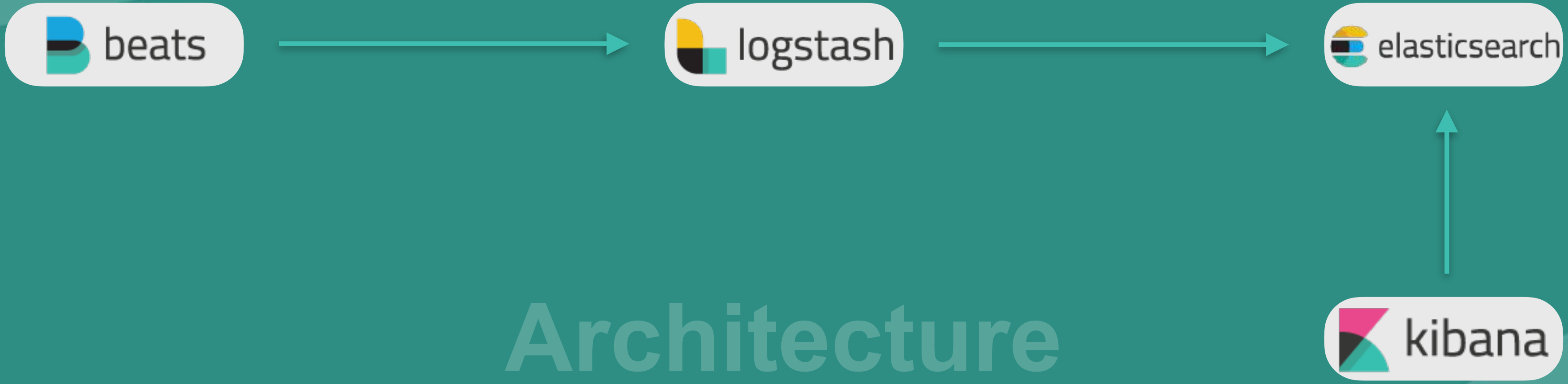
Architecture



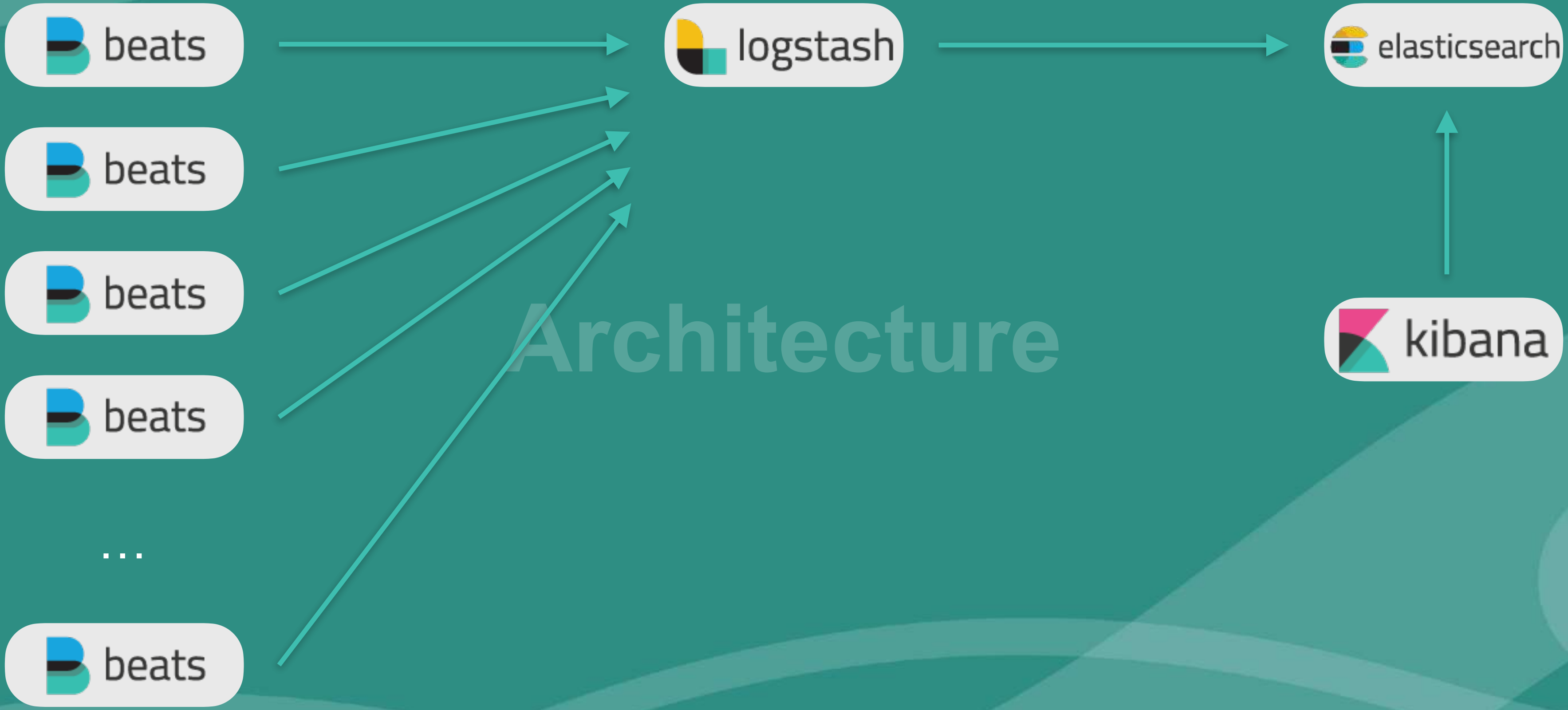
beats → logstash → elasticsearch

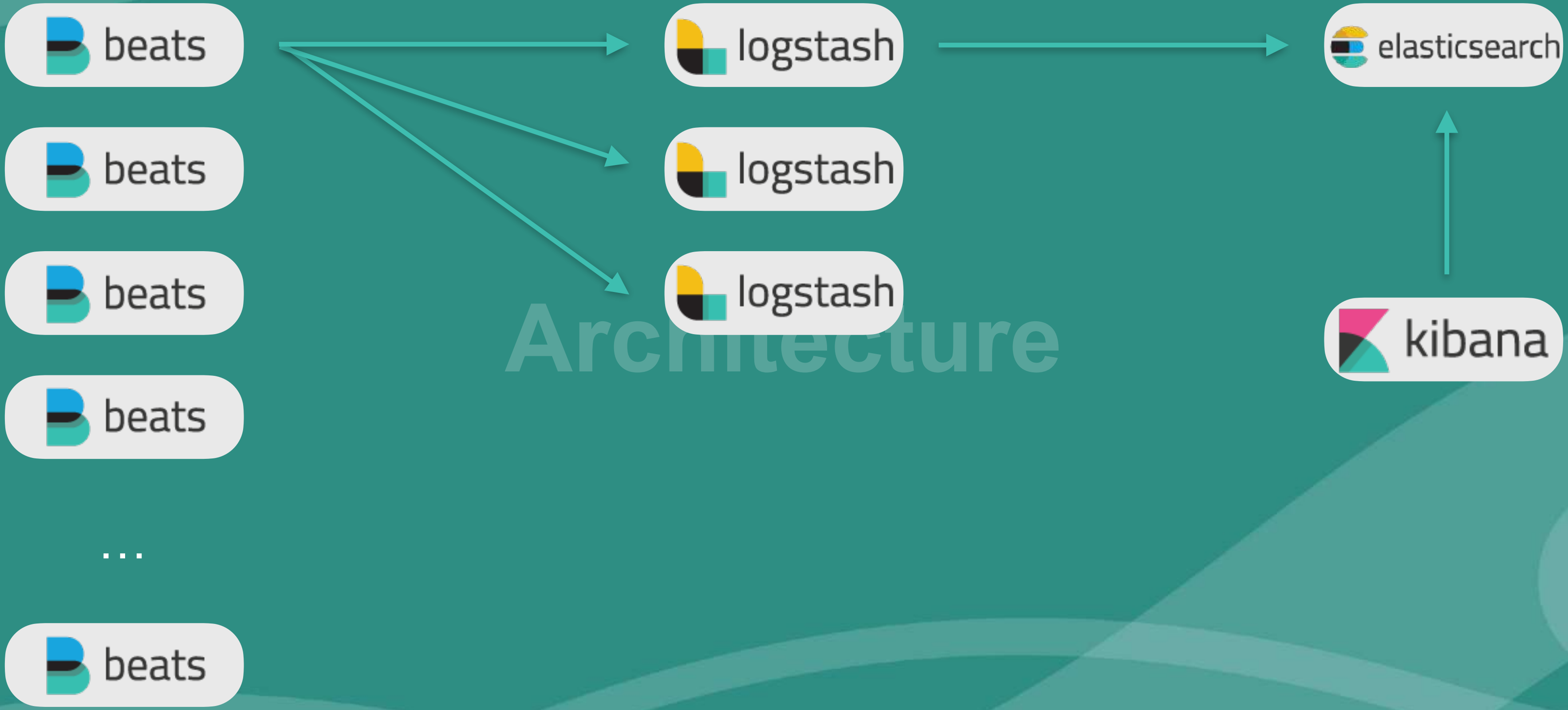
Architecture

Architecture



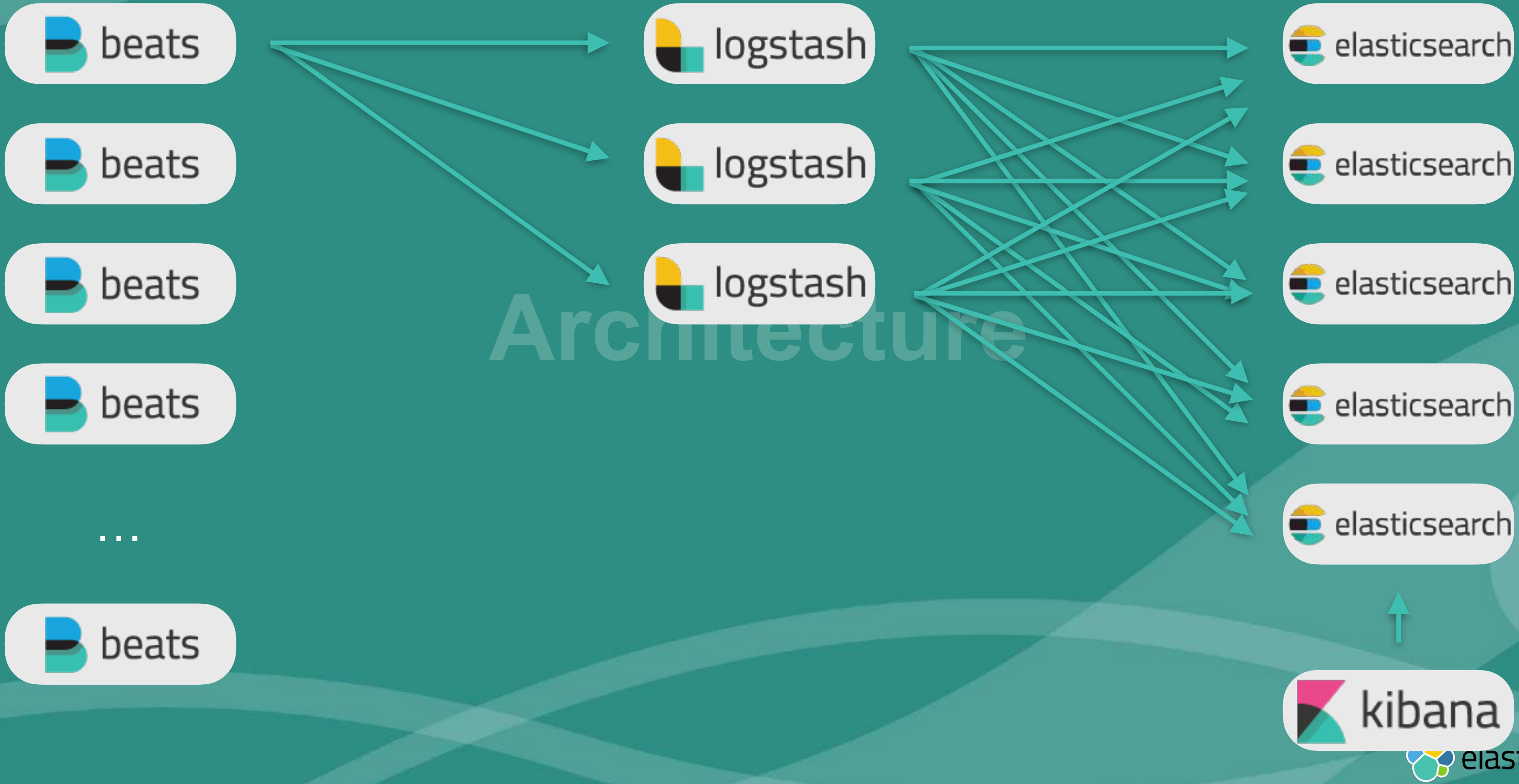
Architecture





Architecture

Architecture

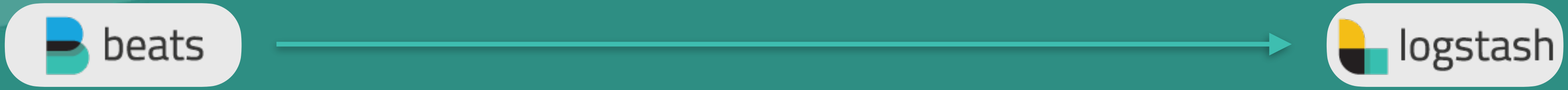


beats

logstash

Architecture

Architecture



acks

lightweight

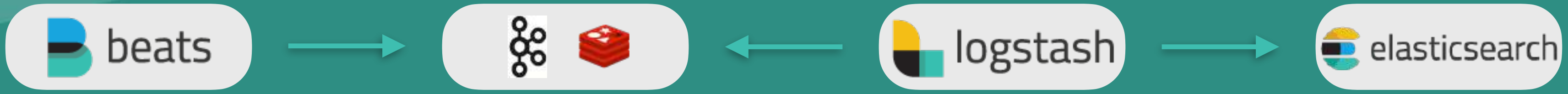
TLS

auth

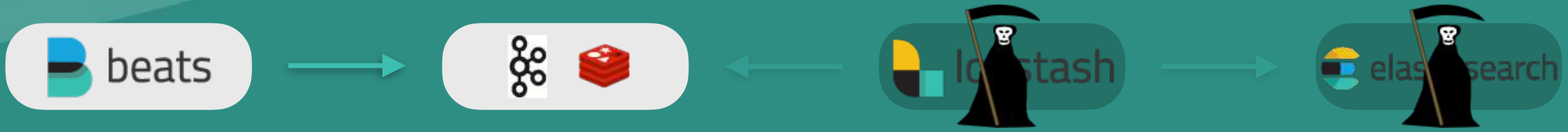
no deps

tags

compression



Architecture

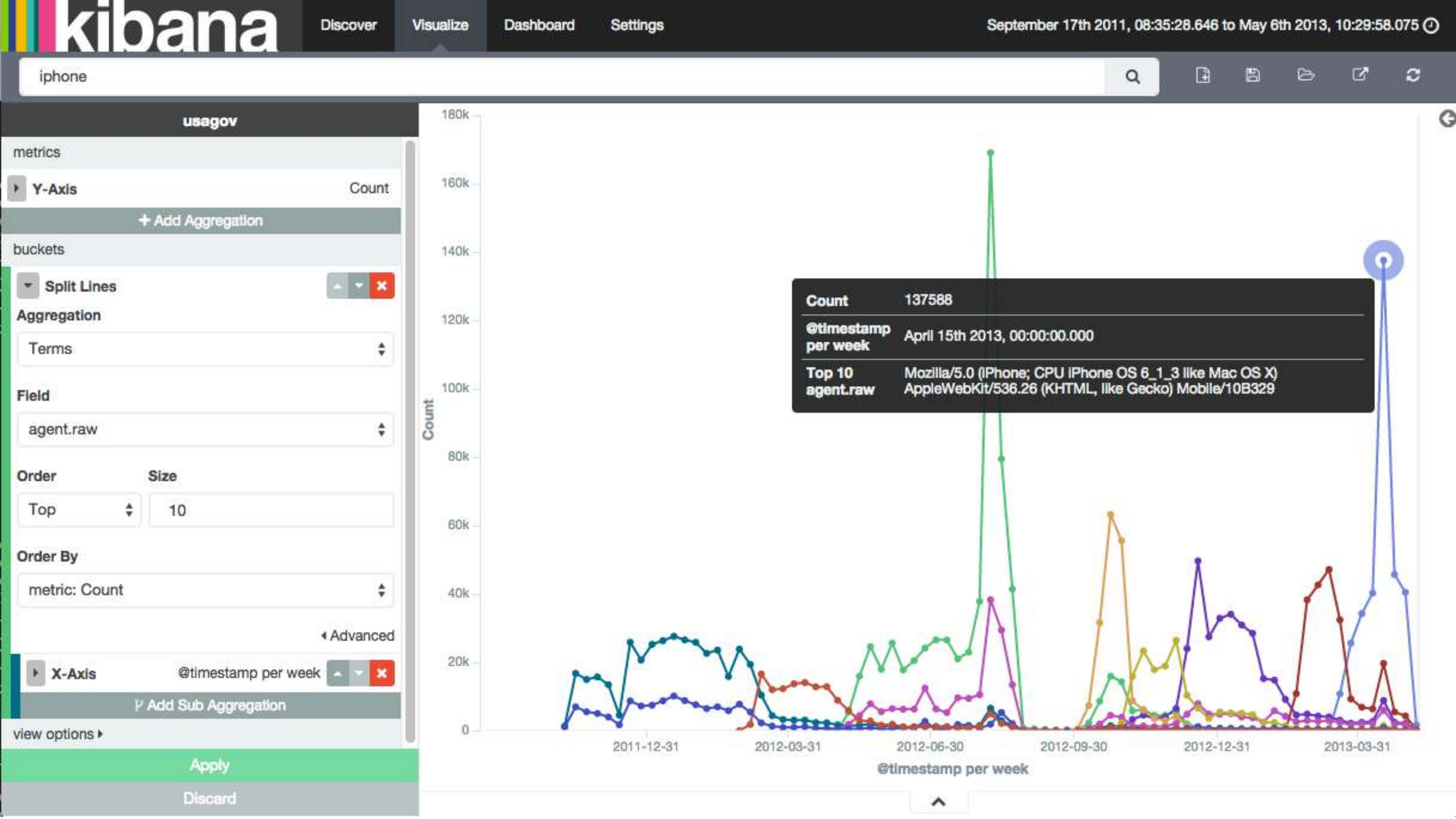


Architecture

elastic

elastic

Visualization



September 17th 2011, 08:35:28.646 to May 6th 2013, 10:29:58.075

iphone

Y-Axis Count

Aggregation: Terms

Field: agent.raw

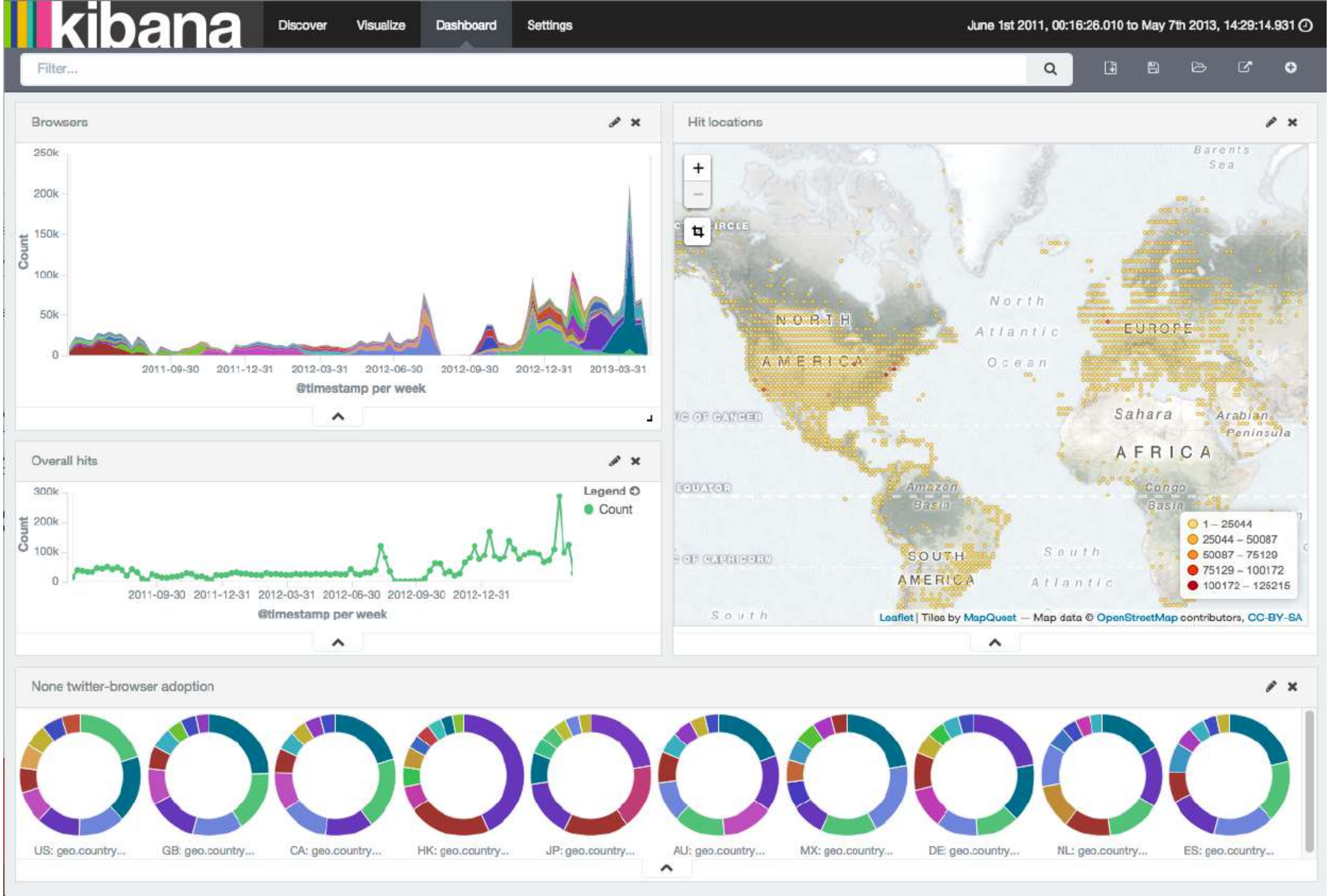
Order: Top

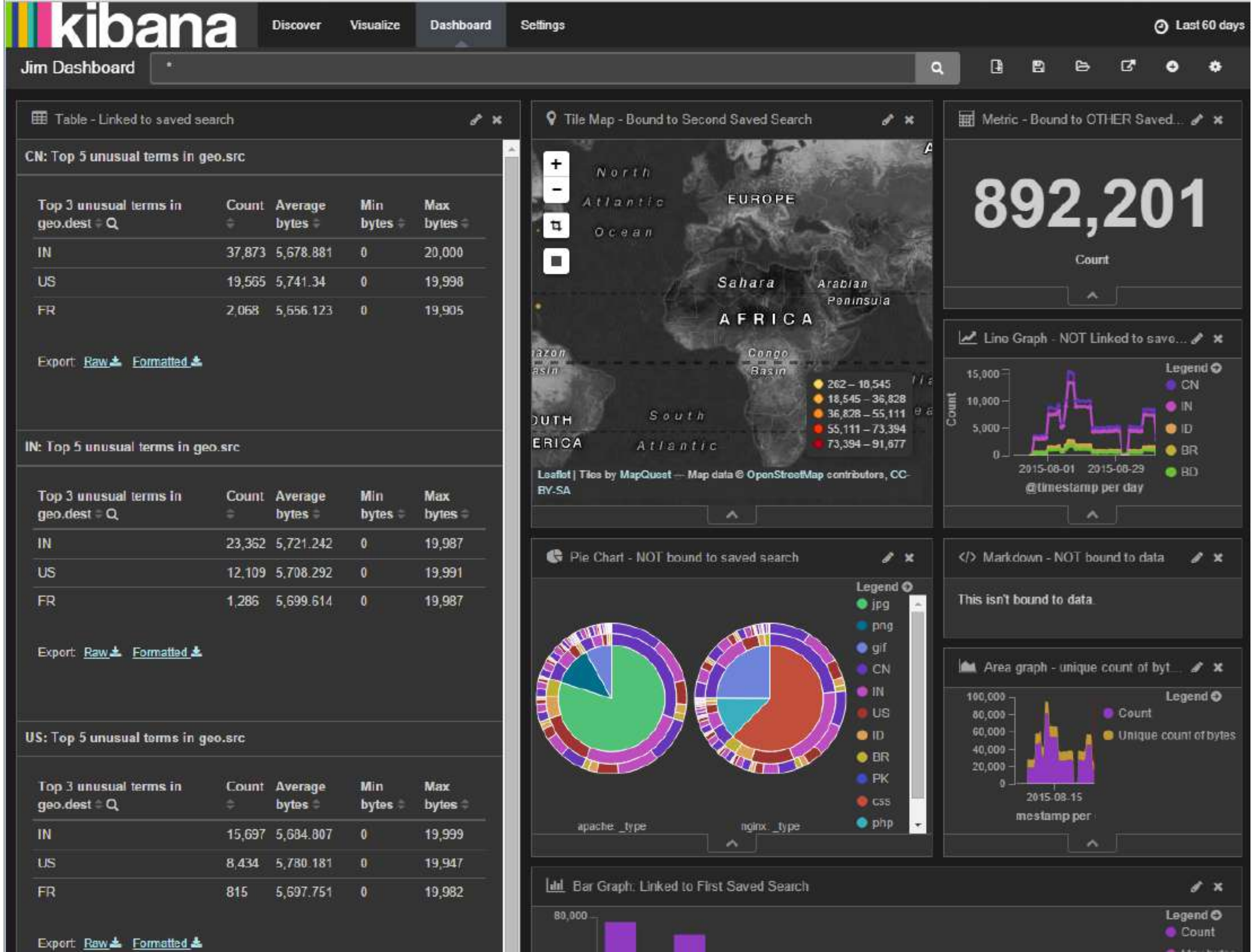
Size: 10

Order By: metric: Count

X-Axis @timestamp per week

Count	137588
@timestamp per week	April 15th 2013, 00:00:00.000
Top 10 agent.raw	Mozilla/5.0 (iPhone; CPU iPhone OS 6_1_3 like Mac OS X) AppleWebKit/536.26 (KHTML, like Gecko) Mobile/10B329


kibana
Discover
Visualize
Dashboard
Settings
June 1st 2011, 00:16:26.010 to May 7th 2013, 14:29:14.931
Filter...
Browsers
250k
200k
150k
100k
50k
0
Count
2011-09-30
2011-12-31
2012-03-31
2012-06-30
2012-09-30
2012-12-31
2013-03-31
@timestamp per week
Overall hits
300k
200k
100k
0
Count
Legend
● Count
2011-09-30
2011-12-31
2012-03-31
2012-06-30
2012-09-30
2012-12-31
@timestamp per week
Hit locations
+
−
Barents Sea
CIRCLE
NORTH AMERICA
North Atlantic Ocean
EUROPE
Sahara
Arabian Peninsula
AFRICA
TROPIC OF CANCER
EQUATOR
Amazon Basin
Congo Basin
TROPIC OF CAPRICORN
SOUTH AMERICA
South Atlantic
South
1 – 25044
25044 – 50087
50087 – 75129
75129 – 100172
100172 – 125215
Leaflet | Tiles by MapQuest — Map data © OpenStreetMap contributors, CC-BY-SA
None twitter-browser adoption
US: geo.country...
GB: geo.country...
CA: geo.country...
HK: geo.country...
JP: geo.country...
AU: geo.country...
MX: geo.country...
DE: geo.country...
NL: geo.country...
ES: geo.country...



CN: Top 5 unusual terms in geo.src

Top 3 unusual terms in geo.dest	Count	Average bytes	Min bytes	Max bytes
IN	37,873	5,678.881	0	20,000
US	19,565	5,741.34	0	19,998
FR	2,068	5,656.123	0	19,905

Export: Raw Formatted

IN: Top 5 unusual terms in geo.src

Top 3 unusual terms in geo.dest	Count	Average bytes	Min bytes	Max bytes
IN	23,362	5,721.242	0	19,987
US	12,109	5,708.292	0	19,991
FR	1,286	5,699.614	0	19,987

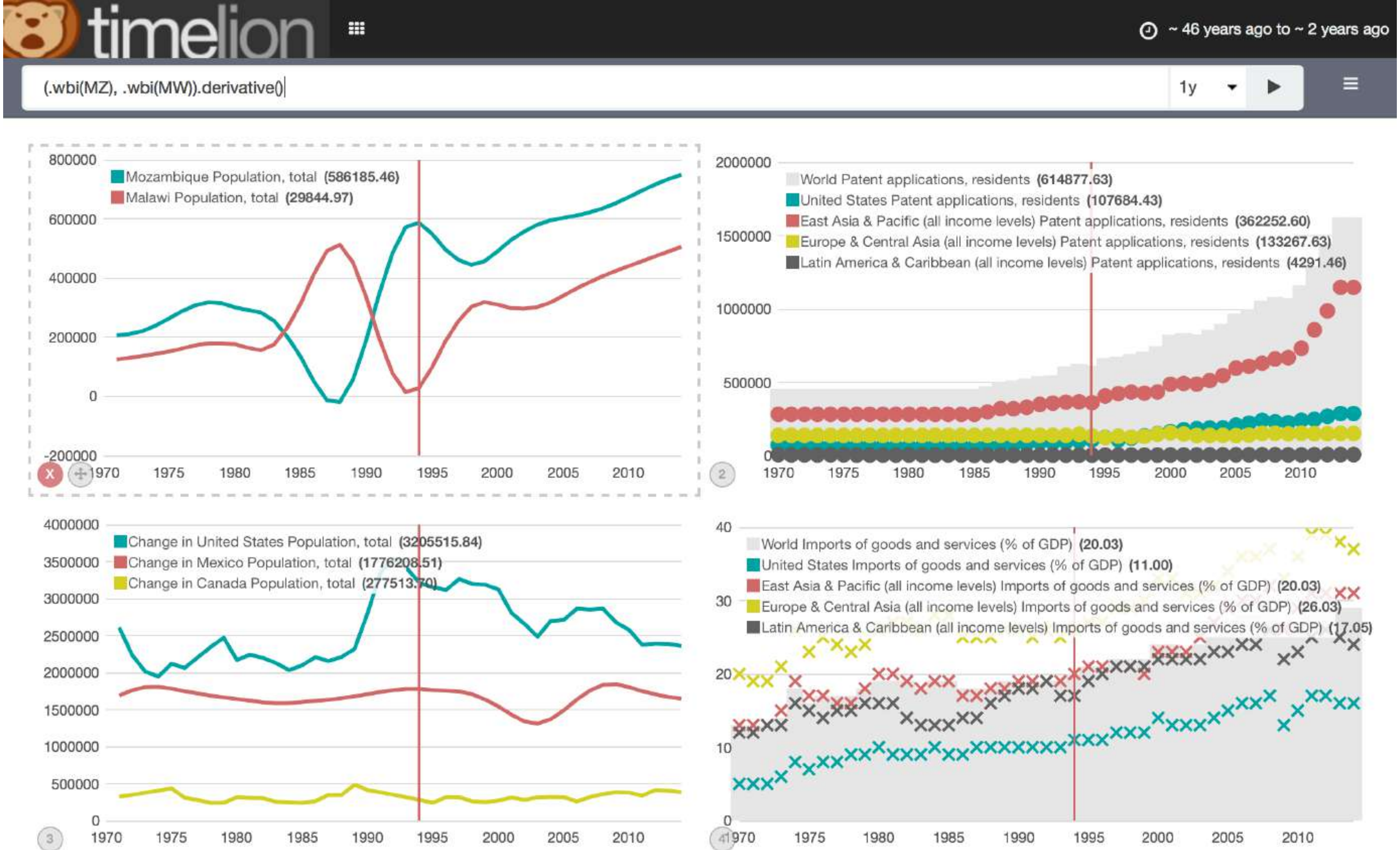
Export: Raw Formatted

US: Top 5 unusual terms in geo.src

Top 3 unusual terms in geo.dest	Count	Average bytes	Min bytes	Max bytes
IN	15,697	5,684.807	0	19,999
US	8,434	5,780.181	0	19,947
FR	815	5,697.751	0	19,982

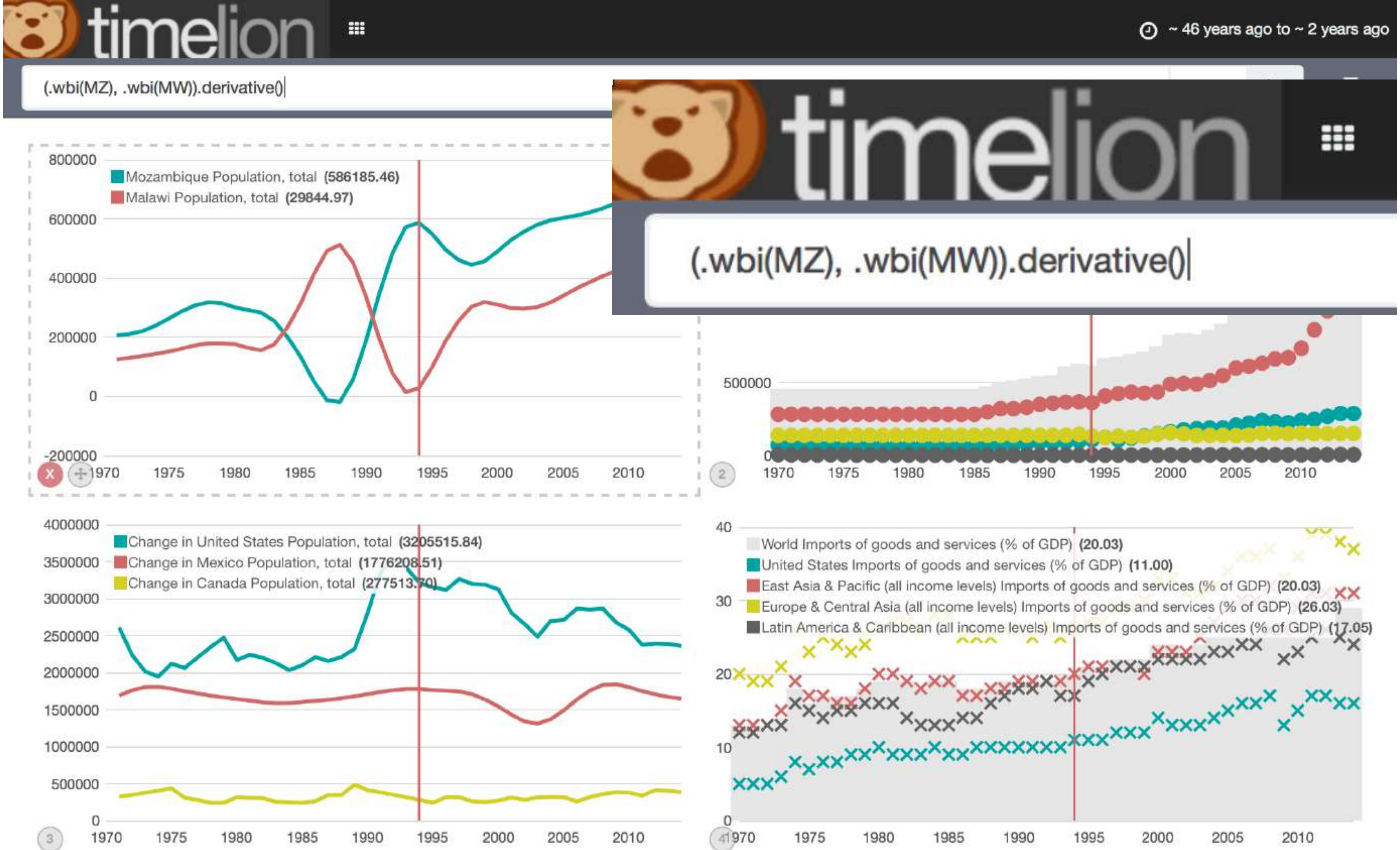
Export: Raw Formatted

This isn't bound to data.

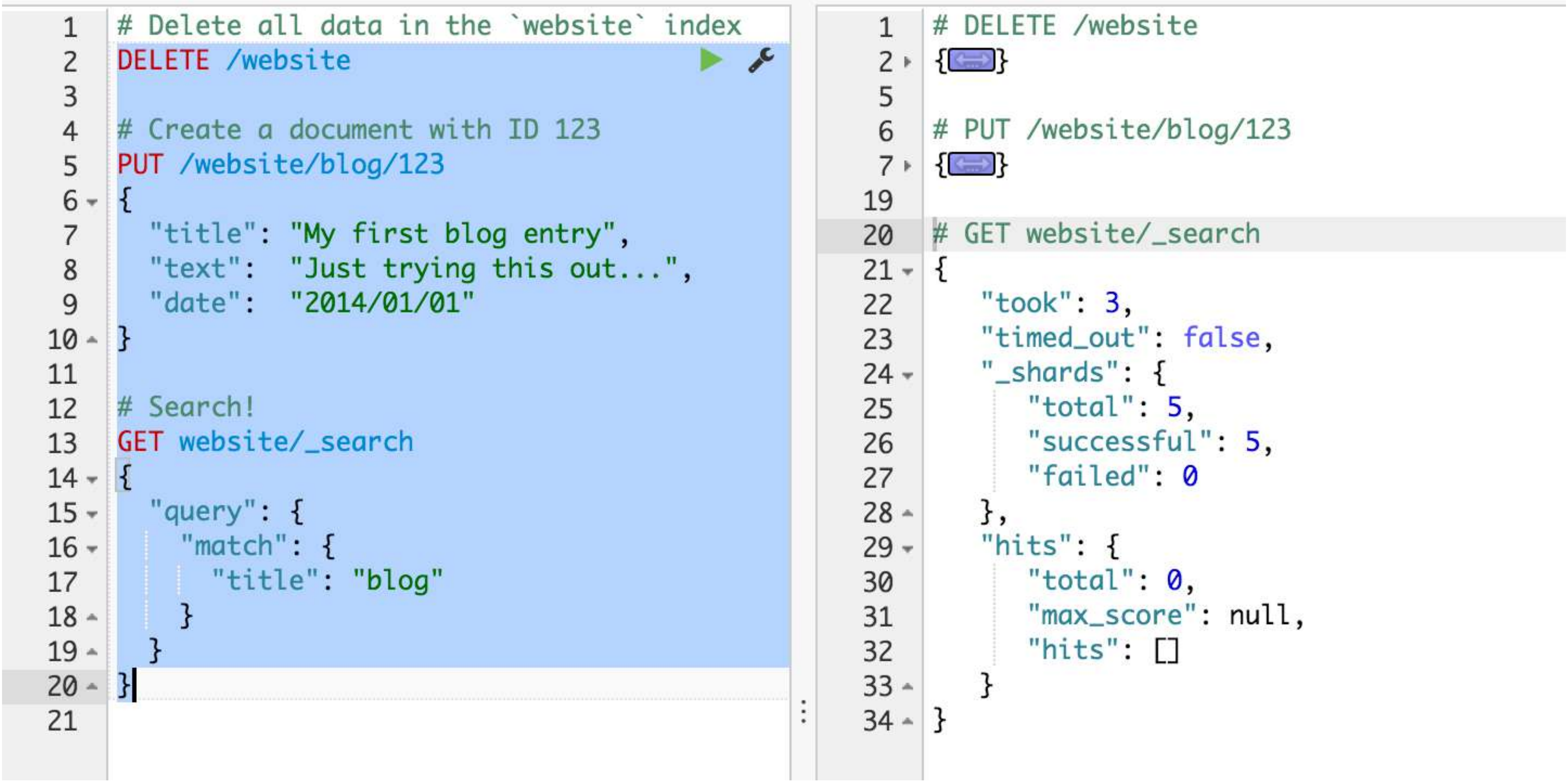


```
(.wbi(MZ), .wbi(MW)).derivative()
```

timelion
~ 46 years ago to ~ 2 years ago
1y
800000
600000
400000
200000
0
-200000
Mozambique Population, total (586185.46)
Malawi Population, total (29844.97)
1970
1975
1980
1985
1990
1995
2000
2005
2010
X
2000000
1500000
1000000
500000
0
World Patent applications, residents (614877.63)
United States Patent applications, residents (107684.43)
East Asia & Pacific (all income levels) Patent applications, residents (362252.60)
Europe & Central Asia (all income levels) Patent applications, residents (133267.63)
Latin America & Caribbean (all income levels) Patent applications, residents (4291.46)
2
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
Change in United States Population, total (3205515.84)
Change in Mexico Population, total (1776208.51)
Change in Canada Population, total (277513.70)
3
40
30
20
10
0
World Imports of goods and services (% of GDP) (20.03)
United States Imports of goods and services (% of GDP) (11.00)
East Asia & Pacific (all income levels) Imports of goods and services (% of GDP) (20.03)
Europe & Central Asia (all income levels) Imports of goods and services (% of GDP) (26.03)
Latin America & Caribbean (all income levels) Imports of goods and services (% of GDP) (17.05)
4


timelion
~ 46 years ago to ~ 2 years ago
800000
600000
400000
200000
0
-200000
Mozambique Population, total (586185.46)
Malawi Population, total (29844.97)
X
1970
1975
1980
1985
1990
1995
2000
2005
2010
500000
0
2
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
Change in United States Population, total (3205515.84)
Change in Mexico Population, total (1776208.51)
Change in Canada Population, total (277513.70)
3
40
30
20
10
0
World Imports of goods and services (% of GDP) (20.03)
United States Imports of goods and services (% of GDP) (11.00)
East Asia & Pacific (all income levels) Imports of goods and services (% of GDP) (20.03)
Europe & Central Asia (all income levels) Imports of goods and services (% of GDP) (26.03)
Latin America & Caribbean (all income levels) Imports of goods and services (% of GDP) (17.05)
4

```
(.wbi(MZ), .wbi(MW)).derivative()
```



```
# Delete all data in the `website` index
DELETE /website

# Create a document with ID 123
PUT /website/blog/123
{
  "title": "My first blog entry",
  "text":  "Just trying this out...",
  "date":  "2014/01/01"
}

# Search!
GET website/_search
{
  "query": {
    "match": {
      "title": "blog"
    }
  }
}
```

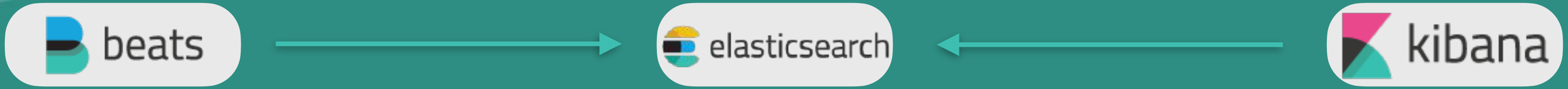
```
# DELETE /website
{⟺}

# PUT /website/blog/123
{⟺}

# GET website/_search
{
   "took": 3,
   "timed_out": false,
   "_shards": {
      "total": 5,
      "successful": 5,
      "failed": 0
   },
   "hits": {
      "total": 0,
      "max_score": null,
      "hits": []
   }
}
```

> **But I just want Apache Logs in Kibana, this is all too complex!**

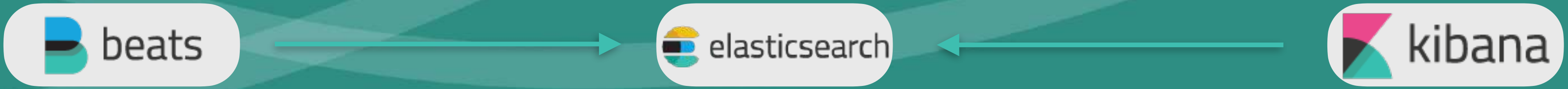
Everyone, ever

elasticsearch
kibana

Ingest pipeline



Ingest pipeline



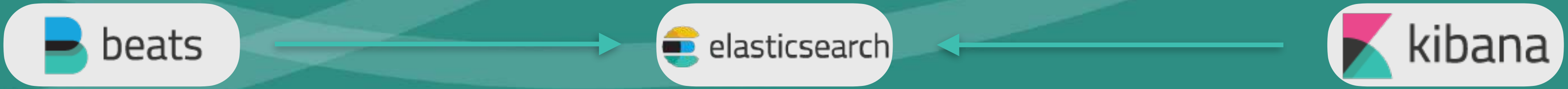
Document enrichment before indexing

failure handlers to change field or destination index on error

Ingest pipeline

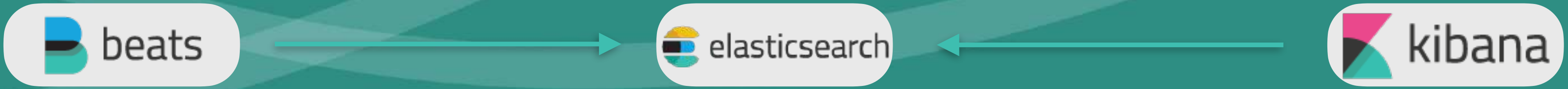
Processors

```
set, append, remove, rename, convert, gsub, join, split,
lowercase, uppercase, trim, grok, date, fail
```



Ingest pipeline

```
PUT/_ingest/pipeline/access-log-pipeline
{
  "description" : "Apache Logs Pipeline",
  "processors" : [
    { "grok" : { … } },
    { "convert" : { … } },
    { "convert" : { … } },
    { "date" : { … } },
    { "geoip" : { … } },
 ]
}
```


beats → elasticsearch ← kibana

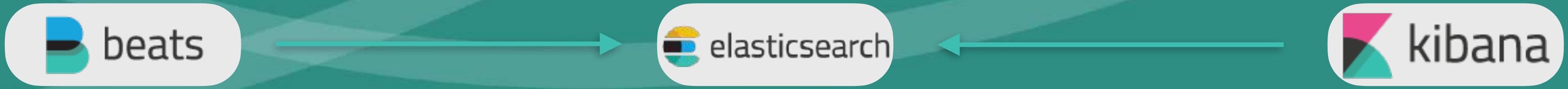
Ingest pipeline

```
…
  {
    "grok" : {
      "field" : "message",
      "pattern" : "%{COMBINEDAPACHELOG}"
    }
  },
…
```

beats → elasticsearch ← kibana

```
…
  {
    "convert" : {
      "field": "response",
      "type": "integer"
    }
  },
…
```

Ingest pipeline



beats

elasticsearch

kibana

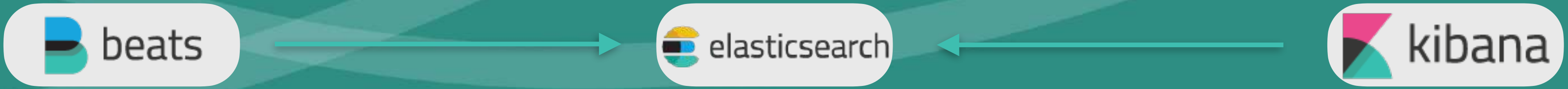
Ingest pipeline

```
…
  {
    "convert" : {
      "field": "bytes",
      "type": "integer"
    }
  },
…
```

beats
elasticsearch
kibana

Ingest pipeline

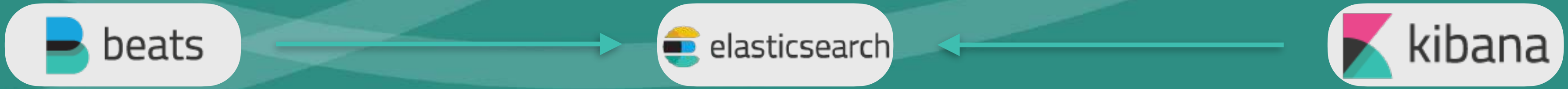
```
…
  {
    "date" : {
      "match_field": "timestamp",
      "match_formats" : [ "dd/MMM/YYYY:HH:mm:ss Z" ]
    }
  },
…
```



```
…
  {
    "geoip" : {
      "source_field" : "clientip"
    }
  }
…
```

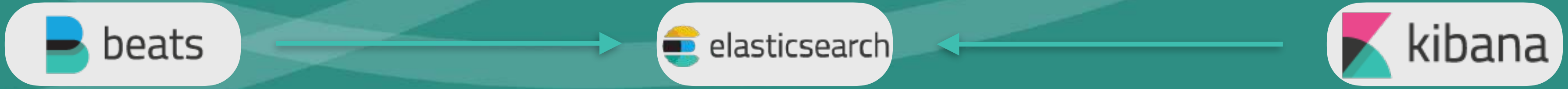
Ingest pipeline

elastic

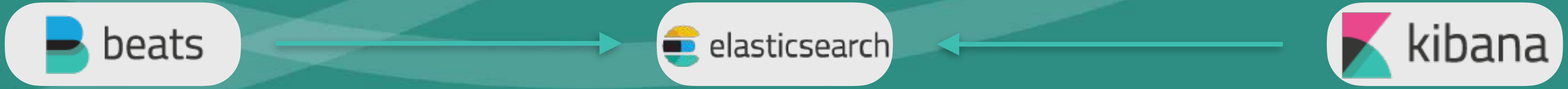


```
POST logs/log?pipeline=access-log-pipeline

{

  "message" : "70.193.17.92 - - [08/Sep/2014:02:54:42 +0000]
\"GET /presentations/logstash-scale11x/images/
ahhh___rage_face_by_samusmmx-d5g5zap.png HTTP/1.1\" 200
175208 \"http://mobile.rivals.com/board_posts.asp?
SID=880&mid=198829575&fid=2208&tid=198829575&Team=&TeamId=&Si
teId=\" \"Mozilla/5.0 (Linux; Android 4.2.2; VS980 4G Build/
JDQ39B) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/
33.0.1750.135 Mobile Safari/537.36\""

}
```

Ingest pipeline



```
{
  "_index": "logs", "_type": "log", "_id": "AVKiNsYu-Si4NcOnCP5b",
  "_version": 1, "found": true,
  "_source": {
    "request": "/presentations/logstash-scale11x/images/
ahhh___rage_face_by_samusmmx-d5g5zap.png",
    agent: "\"Mozilla/5.0 (Linux; Android 4.2.2; VS980 4G Build/JDQ39B)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/33.0.1750.135 Mobile Safari/
537.36\"",
    "geoip": {
      "continent_name": "North America",
      "city_name": "Charlotte",
      "country_iso_code": "US",
      "region_name": "North Carolina",
      "location": { "lon": -80.8431, "lat": 35.2271 }
    },
```



beats

elasticsearch

kibana

```
…
     "auth": "-", "ident": "-", "verb": "GET", "httpversion": "1.1",
     message: "70.193.17.92 - - [08/Sep/2014:02:54:42 +0000] \"GET /
presentations/logstash-scale11x/images/ahhh___rage_face_by_samusmmx-d5g5zap.png
HTTP/1.1\" 200 175208 \"http://mobile.rivals.com/board_posts.asp?
SID=880&mid=198829575&fid=2208&tid=198829575&Team=&TeamId=&SiteId=\" \"Mozilla/
5.0 (Linux; Android 4.2.2; VS980 4G Build/JDQ39B) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/33.0.1750.135 Mobile Safari/537.36\"",
     "referrer": "\"http://mobile.rivals.com/board_posts.asp?
SID=880&mid=198829575&fid=2208&tid=198829575&Team=&TeamId=&SiteId=\"",
     "response": 200, bytes: 175208,
     "clientip": "70.193.17.92",
     "rawrequest": null,
     "@timestamp": "2014-09-08T02:54:42.000Z"
  }
}
```

Ingest pipeline

elastic

Summary

Ease of use

elastic

Minimal dependencies

Extensibility

Flexibility

elasticsearch

x-pack

kibana

Awesome logging infrastructure

beats

logstash

cloud

elastic

Links, Links, Links…

https://www.elastic.co/guide/index.html

https://www.elastic.co/guide/en/beats/filebeat/master/elasticsearch-output.html

https://www.elastic.co/elasticon/conf/2016/sf/whats-evolving-in-elasticsearch

https://www.elastic.co/elasticon/conf/2016/sf/whats-brewing-in-beats

https://www.elastic.co/elasticon/conf/2016/sf/whats-cookin-in-kibana

https://www.elastic.co/elasticon/conf/2016/sf/whats-the-latest-in-logstash

https://www.elastic.co/elasticon/conf/2016/sf/ingest-node-enriching-documents-within-elasticsearch

https://www.elastic.co/elasticon/conf/2016/sf/all-about-elasticsearch-algorithms-and-data-structures

https://www.elastic.co/elasticon/conf/2016/sf/b-b-b-b-b-beats-how-to-build-your-own

https://www.elastic.co/elasticon/conf/2016/sf/grid-monitoring-at-cern-with-the-elastic-stack

https://www.elastic.co/elasticon/conf/2016/sf/quit-yammering-away-analyzing-log-data-microsoft

https://www.elastic.co/elasticon/conf/2016/sf/unleashing-elasticsearch-taking-the-reins-off-at-atlassian

Source: Gray Arial 10pt

Links, Links, Links…

https://www.elastic.co/elasticon/conf/2016/sf

https://www.elastic.co/blog/beats-beta4-filebeat-lightweight-log-forwarding

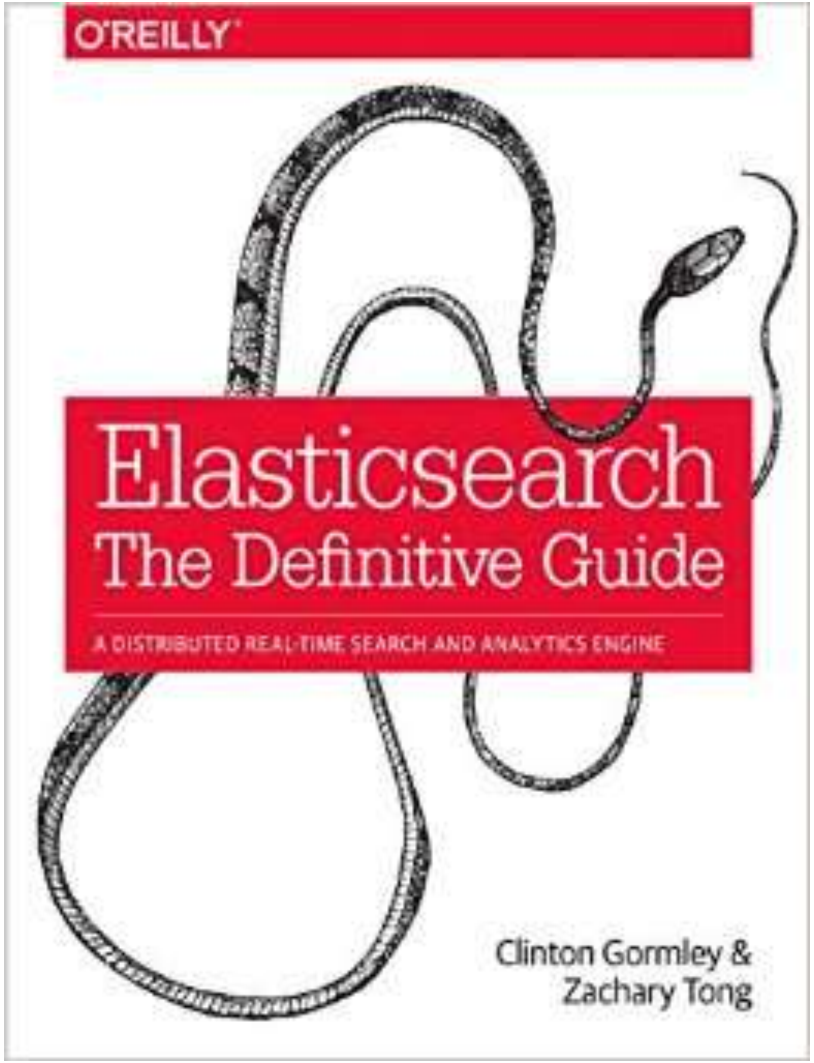
https://www.elastic.co/blog/elasticsearch-command-line-debugging-with-cat

https://www.elastic.co/blog/store-compression-in-lucene-and-elasticsearch

https://discuss.elastic.co/

https://discuss.elastic.co/c/annoucements

Source: Gray Arial 10pt



O'REILLY

Elasticsearch The Definitive Guide

A DISTRIBUTED REAL-TIME SEARCH AND ANALYTICS ENGINE

Clinton Gormley &
Zachary Tong

Thank you for listening!

Alexander Reelsen
alex@elastic.co
@spinscale

We’re hiring
https://www.elastic.co/about/careers

We’re helping
https://www.elastic.co/subscriptions

Images used

https://commons.wikimedia.org/wiki/File:Munich_skyline.jpg

https://commons.wikimedia.org/wiki/File:Skyline_munchen.png

https://commons.wikimedia.org/wiki/File:Olympiapark_M%C3%BCnchen.jpg

https://commons.wikimedia.org/wiki/File:BIER_IM_EG.jpg

Source: Gray Arial 10pt