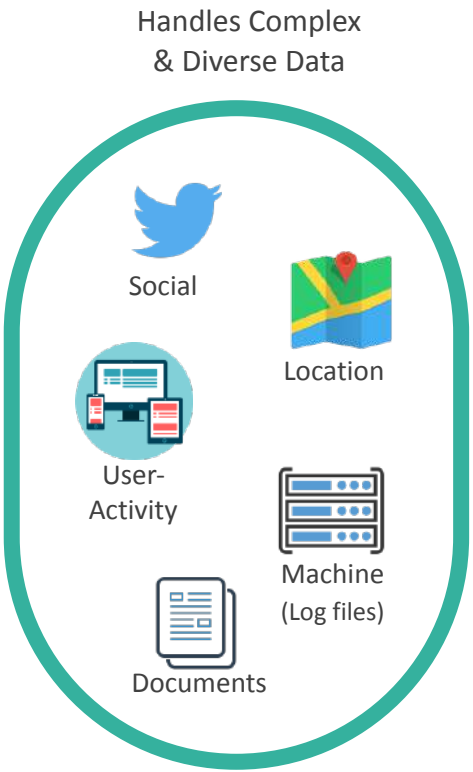
Elastic(search)

Neues aus dem Maschinenraum

Alexander Reelsen
@spinscale
alex@elastic.co

What’s the problem?

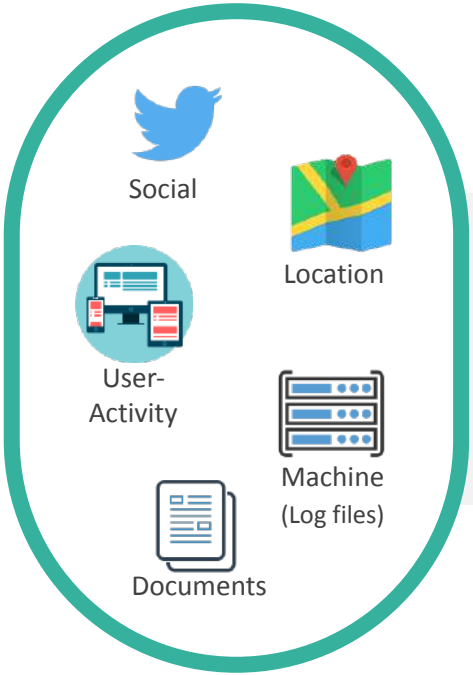
Elastic Makes Building Scalable, Real-Time Systems Simple



elastic

Elastic Makes Building Scalable, Real-Time Systems Simple

Handles Complex
& Diverse Data



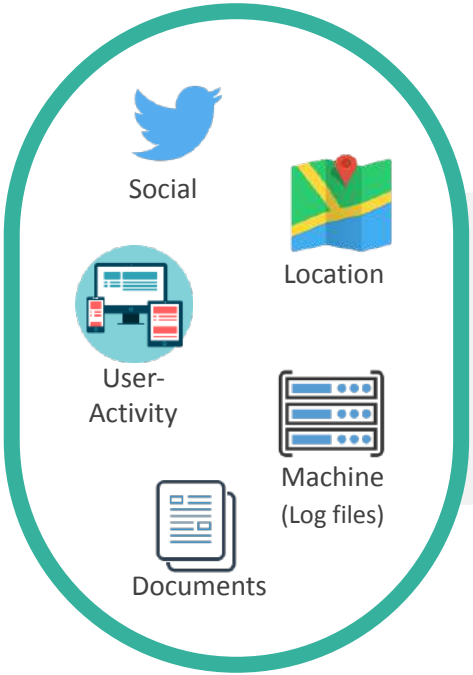
Meets Core
Developer Requirements

- Cross-source insights
- Many users / use cases
- Large data volumes
- Fast data processing
- Data quality & integrity

elastic

Elastic Makes Building Scalable, Real-Time Systems Simple

Handles Complex & Diverse Data



Meets Core Developer Requirements

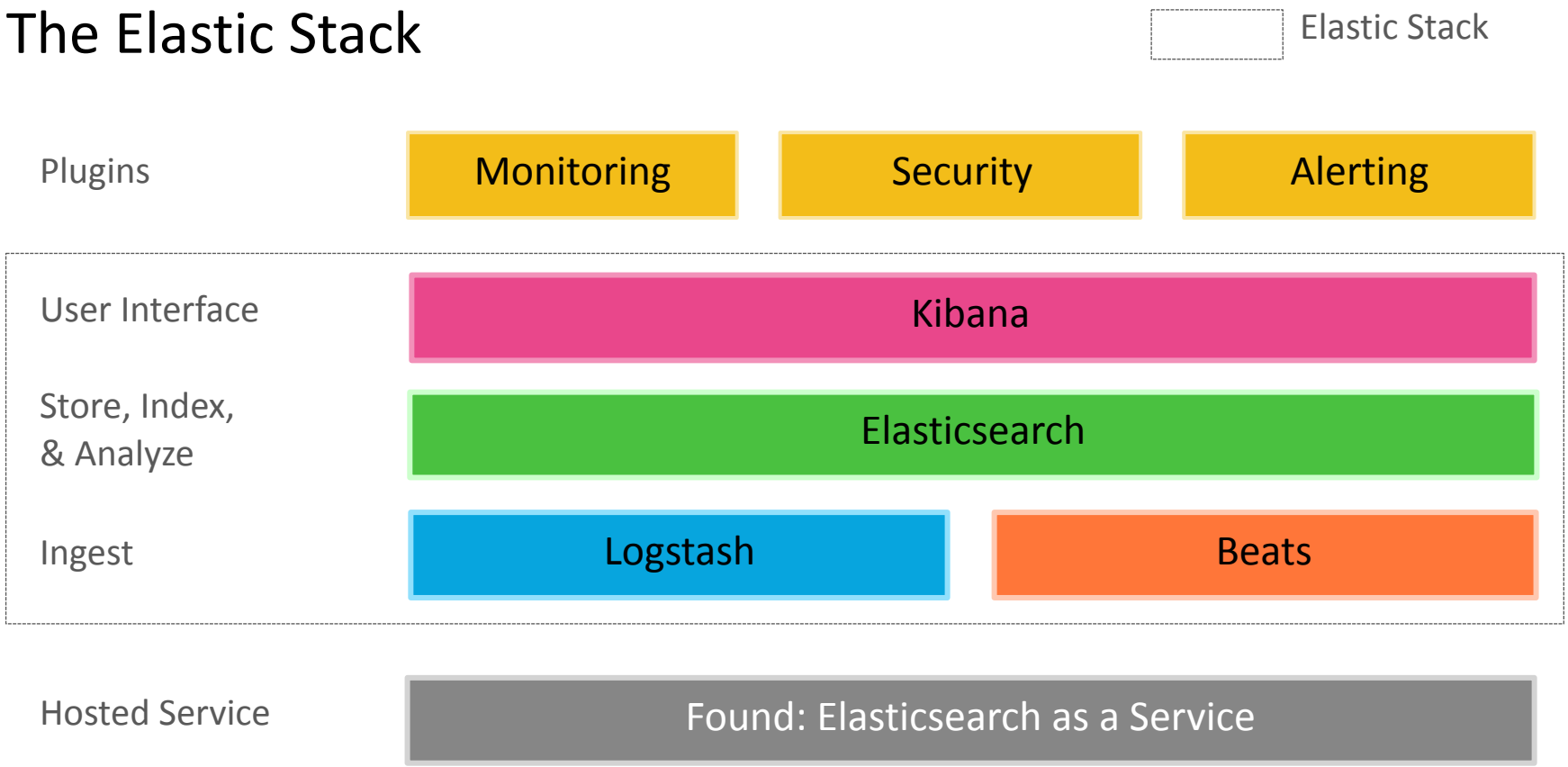
- Cross-source insights
- Many users / use cases
- Large data volumes
- Fast data processing
- Data quality & integrity

Solves Critical Use Cases

- Application Search
- Embedded Search
- Logging
- Security Analytics
- Customer Insights
- More …

elastic

The Elastic Stack



elastic

Elasticsearch

elastic

Elasticsearch: Store, Index, and Analyze

Distributed, scalable, and resilient

Designed for scale-out; high availability

Developer friendly

API-first; schemaless, native JSON & HTTP, client libraries

Real-time Search & Analytics

Real-time aggregations, geospatial, full-text search; query structured and unstructured data

Elasticsearch 2.x

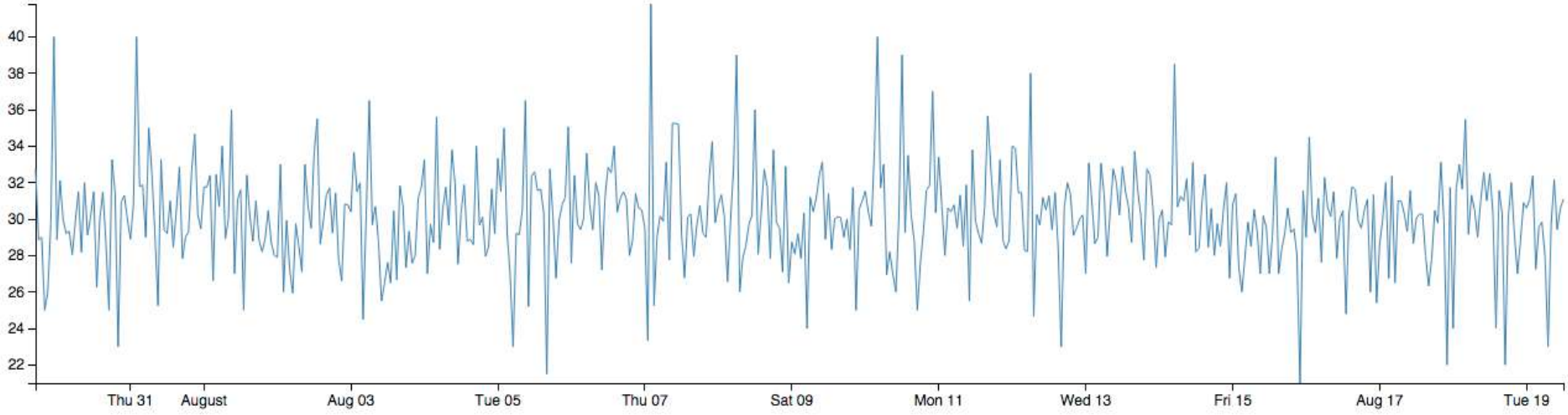
Pipeline Aggregations

Query profiler

Plugins as first class citizen

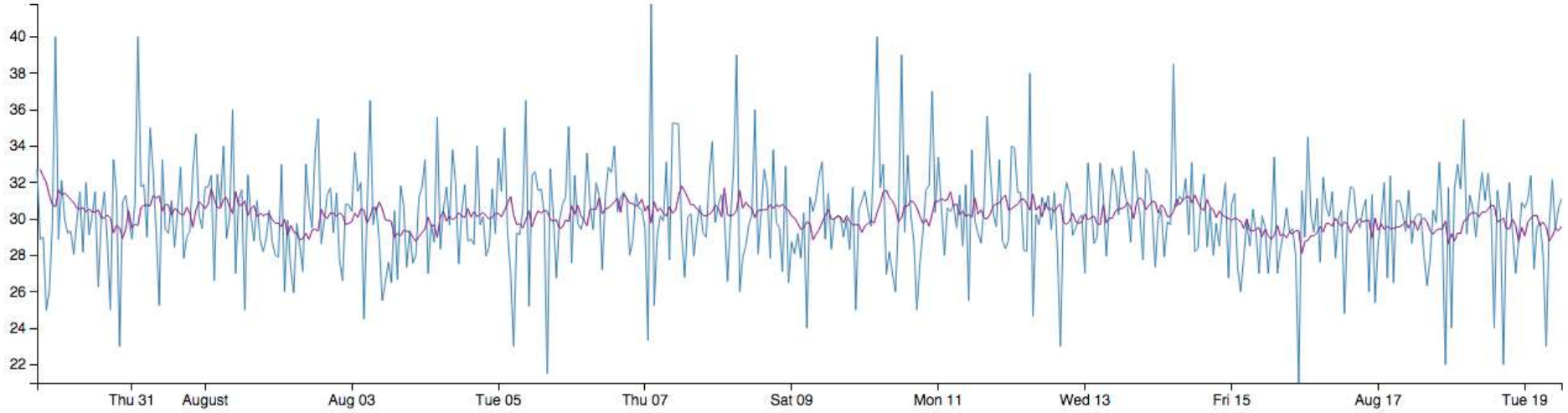
elastic

Pipeline Aggregations



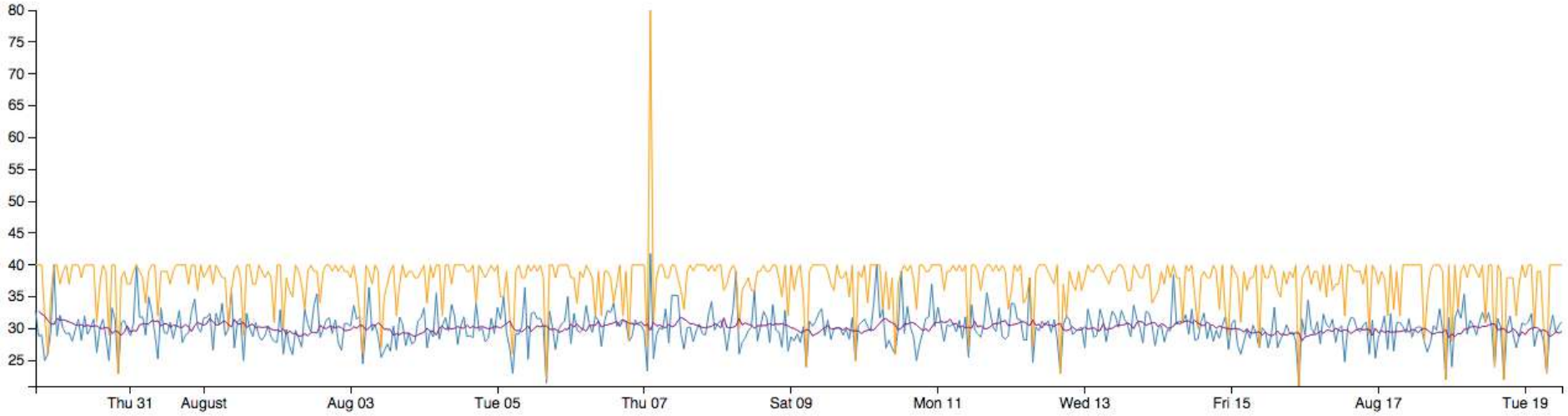
elastic

Pipeline Aggregations



elastic

Pipeline Aggregations



elastic

Query profiler

```
GET /profile/data/_search
{
    "profile": true,
    "query": {
        "match": {
            "foo": "bar baz"
        }
    }
}
```

Query profiler


```
{
  "took": 2, "timed_out": false,
  "_shards": { … },
  "hits": { … },
  "profile": {
    "shards": [
      {
        "id": "[vj4imdlqQOK0Xj_n70xD_A][profile][0]",
       "searches": [
```

Query profiler

```
…
"searches": [ {
  "query": [ {
    "query_type": "BooleanQuery",
    "lucene": "+(foo:bar foo:baz) #ConstantScore(_type:data)",
    "time": "1.056684000ms",
    "breakdown": {
       …
```

elastic

Elasticsearch 2.x

Allocation and recovery

Security Manager

Resilience

elastic

Elasticsearch 2.x

Mapping updates

Two phase query execution

Query/Filter caching

Elasticsearch 3.x

Task Management

Reindex

New scripting language

Elasticsearch 3.x

Strict settings

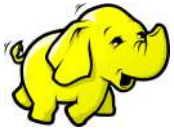
Improved suggester

Percolator

elastic

ES-Hadoop: Integrate with Hadoop, Spark & More

Real-time search on Hadoop data



Standalone, self-contained library on Hadoop

Access ES data bi-directionally

Support for MapReduce, Hive, Pig, Cascading, Spark, and Storm

Leverage HDFS to backup and archive ES data

Security for the Elastic Stack (Shield)

Simply Secure Elasticsearch

Username/password protection

Advanced Security When Needed

LDAP/AD integration

Role-based access control

Field and document level security

Encrypted communications

Audit logging

elastic

Alerting for the Elastic Stack (Watcher)

Alerts based on your data

Flexible Notifications

Wide range of use-cases

Integrations

Slack

Hipchat

Pagerduty

Email

elastic

Monitoring for the Elastic Stack (Marvel)

Monitor Metrics

Track real-time stats and metrics for all clusters and nodes

Diagnose Issues

Analyze historical or real-time data for root cause analyses

Optimize Performance

Utilize in-depth analyses to improve cluster performance

elastic

Elasticsearch as a Service (Found)

The only fully managed and hosted Elasticsearch product supported by the creators of Elasticsearch, Logstash and Kibana

Set up clusters in seconds

Dedicated memory and storage

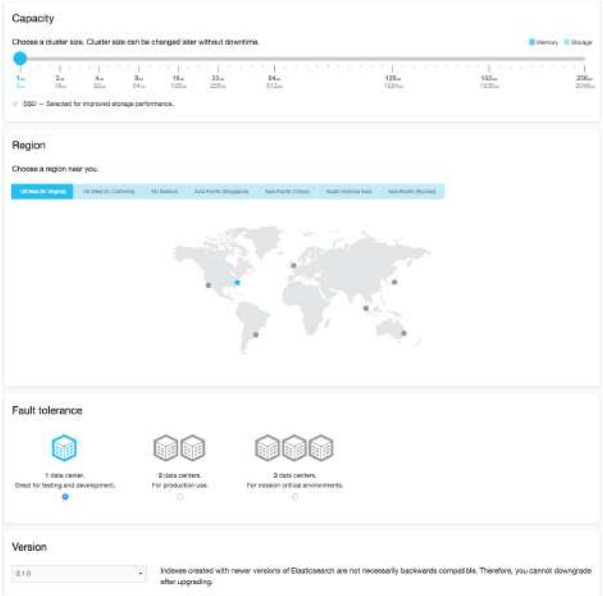
Native, unmodified Elasticsearch endpoint

High availability with replication

Simply scale up or down

Pre-integrated Elastic plugins

Enterprise, SLA-support



elastic

Getting up and running - easy!

```
unzip elasticsearch-2.x.y.zip ; cd elasticsearch 2.x.y
bin/elasticsearch

bin/plugin install analysis-icu
bin/plugin install shield
bin/plugin install watcher
bin/plugin install marvel
```

elastic

Elasticsearch: Further info

https://www.elastic.co/guide/en/elasticsearch/reference/2.2/search-profile.html

https://www.elastic.co/guide/en/elasticsearch/reference/2.2/search-aggregations-pipeline.html

https://www.elastic.co/blog/out-of-this-world-aggregations

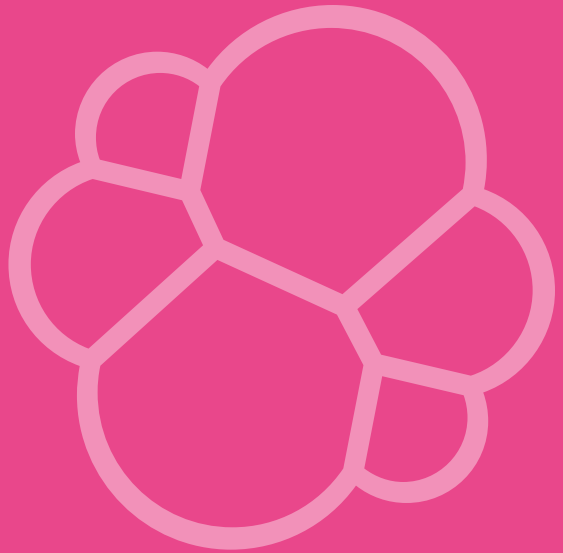
https://www.elastic.co/blog/staying-in-control-with-moving-averages-part-1

https://www.elastic.co/blog/staying-in-control-with-moving-averages-part-2

https://www.elastic.co/blog/implementing-a-statistical-anomaly-detector-part-1

https://www.elastic.co/blog/implementing-a-statistical-anomaly-detector-part-2

https://www.elastic.co/blog/implementing-a-statistical-anomaly-detector-part-3



Kibana

elastic

Kibana: A User Interface for All of Your Data

Explore and discover insights

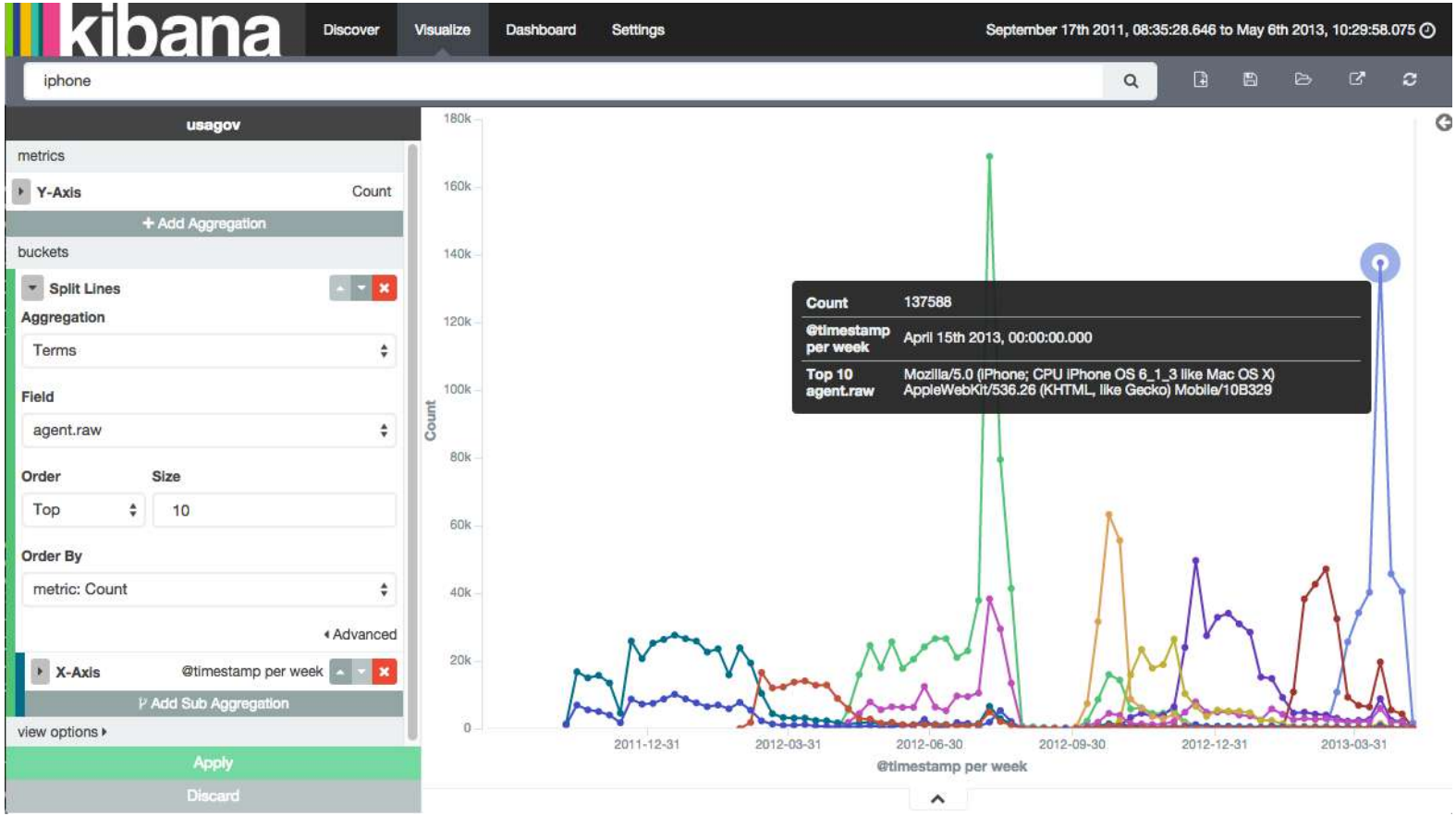
Instant response at any scale

Build interactive dashboards

Share, embed, and integrate

elastic

Single visualizations…



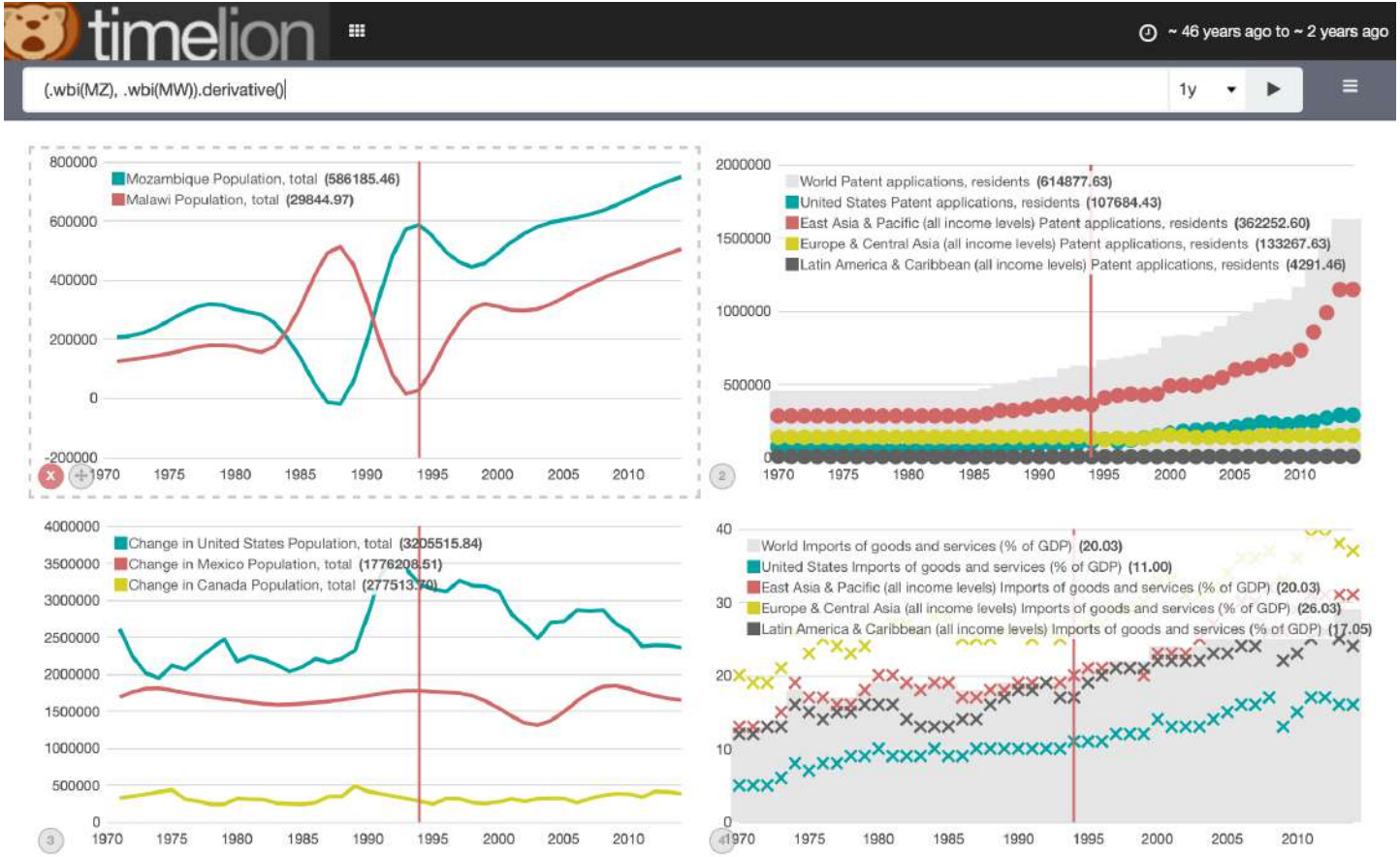
elastic

… form a dashboard

elastic

… the dark side

Timelion - Time series composer



elastic

Timelion - Time series composer

Composable functions

`abs, derivative, cusum, divide, first, max, min, movingaverage, movingstd, multiply, substract, sum`

`bars, color, hide, label, legend, lines, points, precision, yaxis`

`es, graphite, quandl, worldbank, wbi`

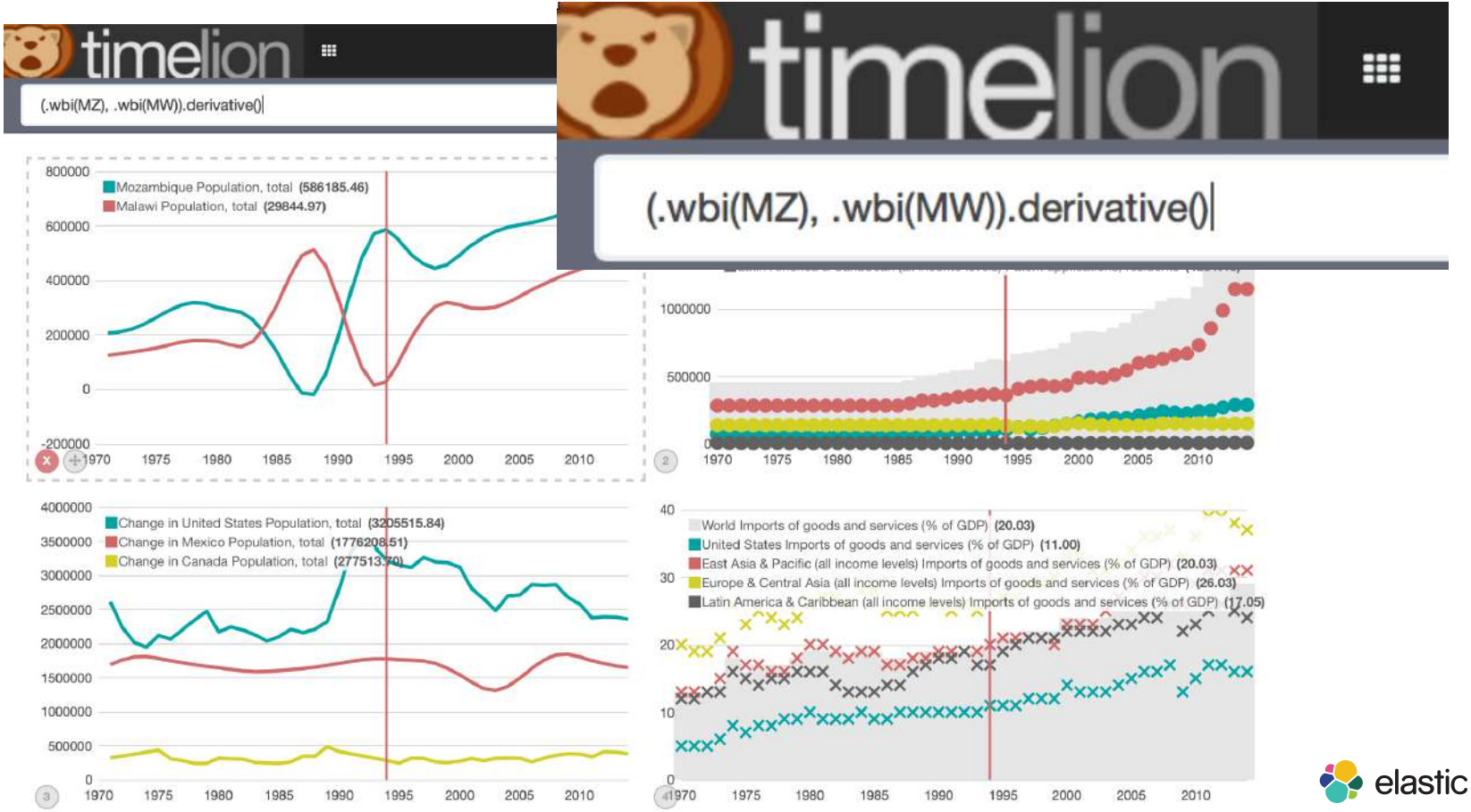
More info

https://www.elastic.co/blog/timelion-timeline

https://www.youtube.com/watch?v=-sgZdW5k7eQ

elastic

Timelion - Time series composer



elastic

Sense - The missing UI

elastic

Sense - Features

Suggestions for all requests

Multiple Requests

Auto-Indent

Copy as cURL

Keyboard shortcuts

History

More info

https://www.elastic.co/blog/sense-2-0-0-beta1

https://www.elastic.co/guide/en/sense/current/index.html

elastic

Getting up and running - easy!

```
tar xvf kibana-….tar.gz ; cd kibana
bin/kibana

./bin/kibana plugin --install elastic/sense
./bin/kibana plugin --install elastic/timelion
```

elastic

Writing own plugins - easy!

```
npm install -g yo

npm install -g generator-kibana-plugin


mkdir my-new-plugin

cd my-new-plugin

yo kibana-plugin
```

Logstash

Logstash: Collect, Enrich, and Transport

Collect data from many sources

- Application
- Social data
- Infra/web/audit logs
- Sensor data
- Message queues
- Documents
- Transaction/wire

Open-source ETL engine with more than 200+ community extensible plugins

Enrich

Parse, transform, clean

Transport

Output to Elasticsearch and other systems

elastic

Logstash 2.x - Changes

Next generation Pipeline in 2.2

Better performance, works in micro batches, automatic worker scaling

Plugins

kafka input/output, JDBC input, HTTP input, WebHDFS output, Salesforce input, HTTP poller

Logstash 3.x

Persistent Queues

Performance

Clustering

Monitoring

elastic

Getting up and running - easy!

```
unzip logstash-2.X.Y.zip ; cd logstash-2.X.Y

bin/logstash -f logstash.conf


bin/plugin install logstash-output-jms
```

Writing own plugins - easy!

```
git clone https://github.com/logstash-plugins/logstash-input-example

git clone https://github.com/logstash-plugins/logstash-output-example

git clone https://github.com/logstash-plugins/logstash-filter-example

git clone https://github.com/logstash-plugins/logstash-codec-example
```

Beats

elastic

Beats: Lightweight Data Shippers

Libbeat
Library for forwarding host-based metrics to Elasticsearch

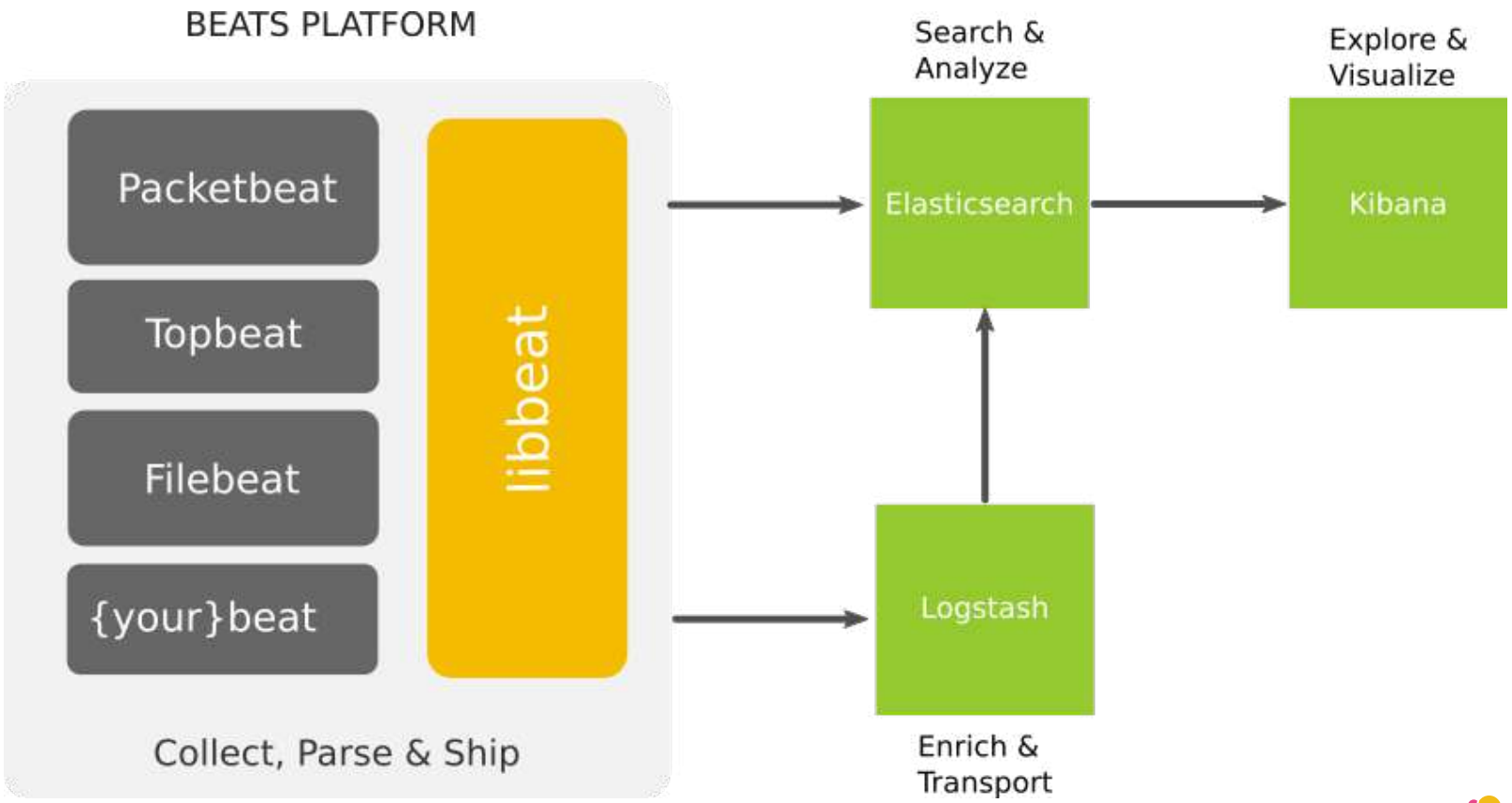
Packetbeat
Real-time network packet analytics for web, database, and any network protocols

Topbeat
Gather resource utilization data such as CPU, memory, etc and ship it to Elasticsearch to analyze .

Filebeat
Next-generation Logstash forwarder to collect, pre-process, and forward log files.

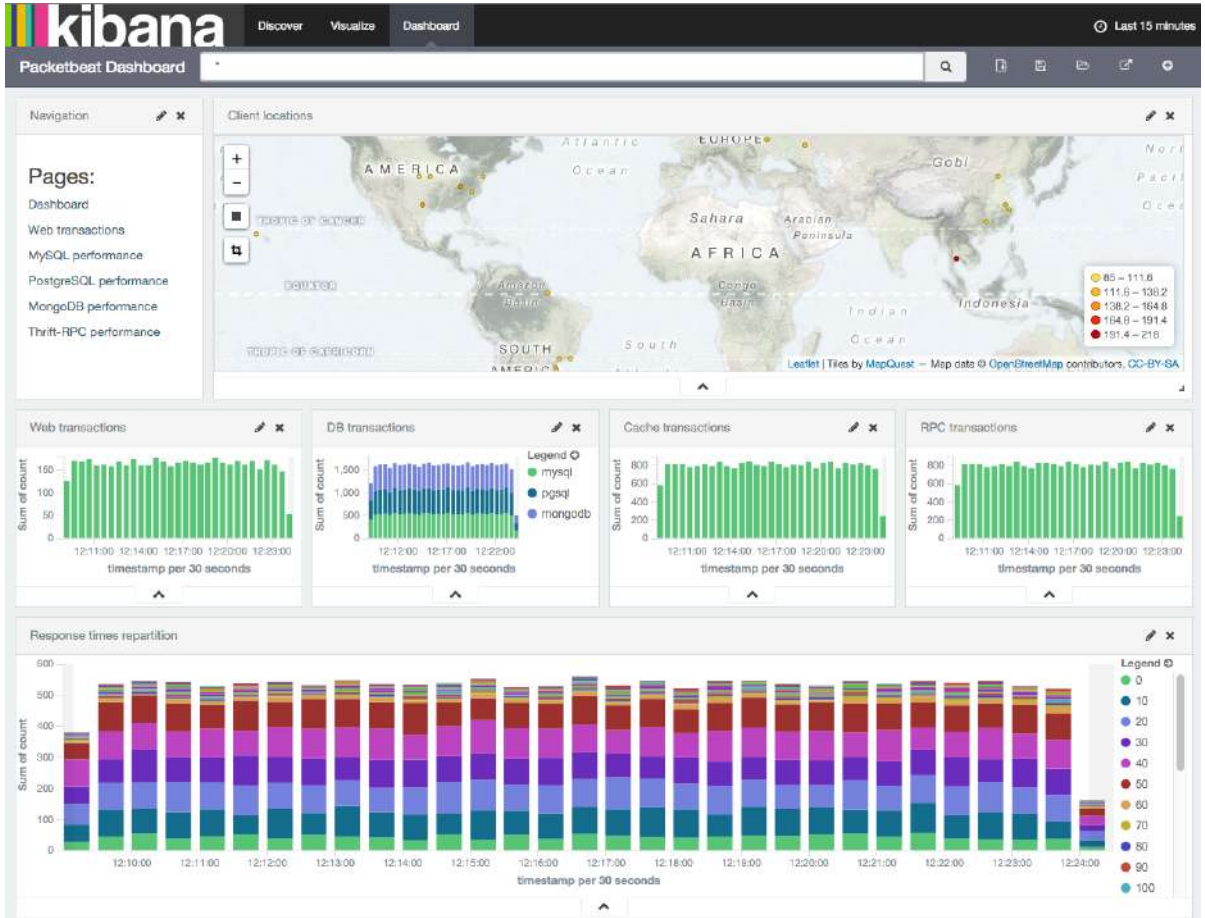
elastic

Beats: Lightweight Data Shippers



elastic

Packetbeat


kibana
Discover
Visualize
Dashboard
Last 15 minutes
Packetbeat Dashboard
Navigation
Pages:
Dashboard
Web transactions
MySQL performance
PostgreSQL performance
MongoDB performance
Thrift-RPC performance
Client locations
AMERICA
EUROPE
AFRICA
Sahara
Gobi
Indonesia
SOUTH AMERICA
Leaflet | Tiles by MapQuest — Map data © OpenStreetMap contributors, CC-BY-SA
Web transactions
DB transactions
Cache transactions
RPC transactions
Sum of count
timestamp per 30 seconds
Legend
mysql
pgsql
mongodb
Response times repartition

Packetbeat

Protocols

ICMP (v4 and v6), DNS, HTTP, Mysql, PostgreSQL, Redis, Thrift-RPC, MongoDB, Memcache

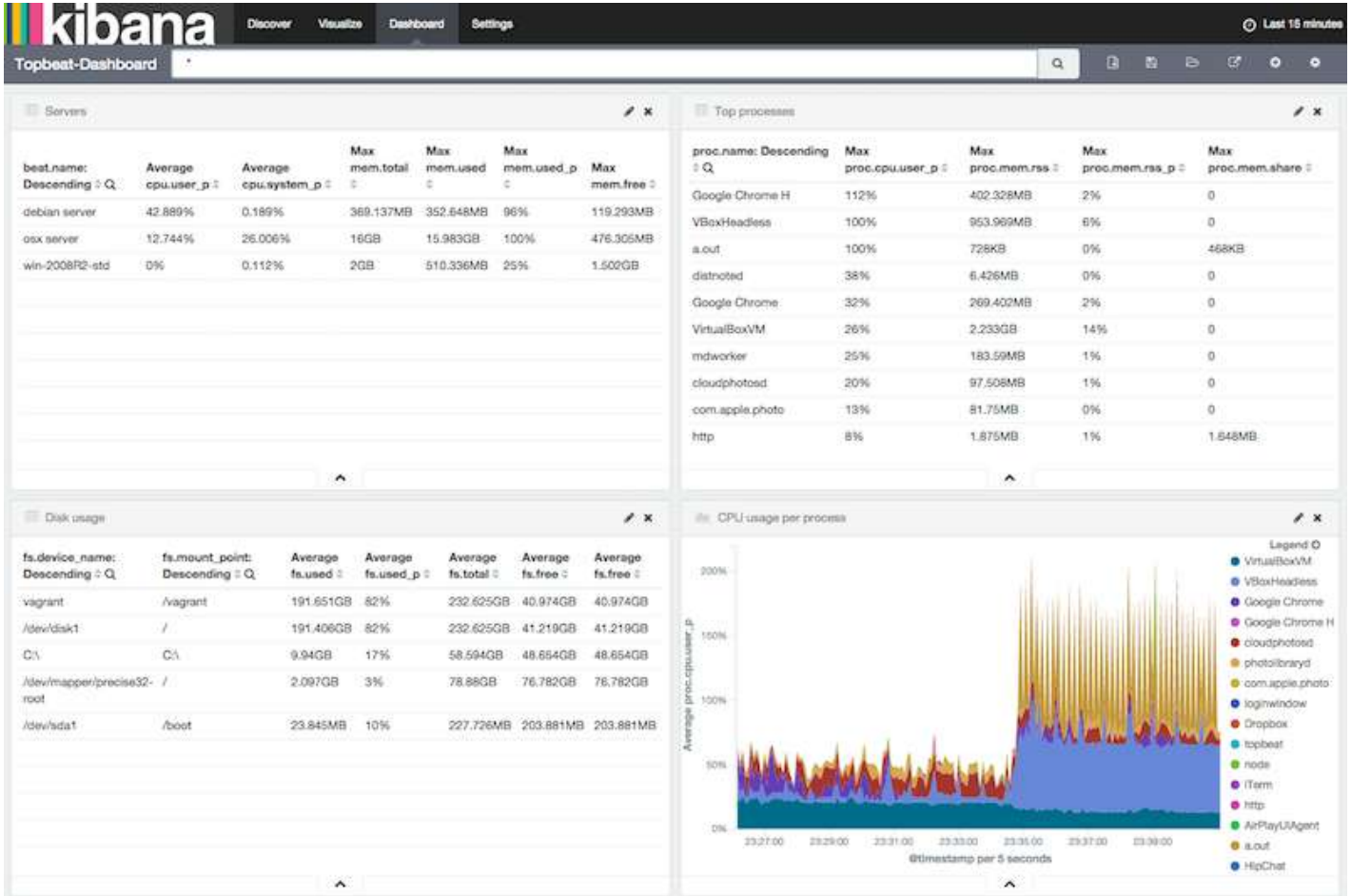
Output

Elasticsearch, Logstash, File, console

Extensibility

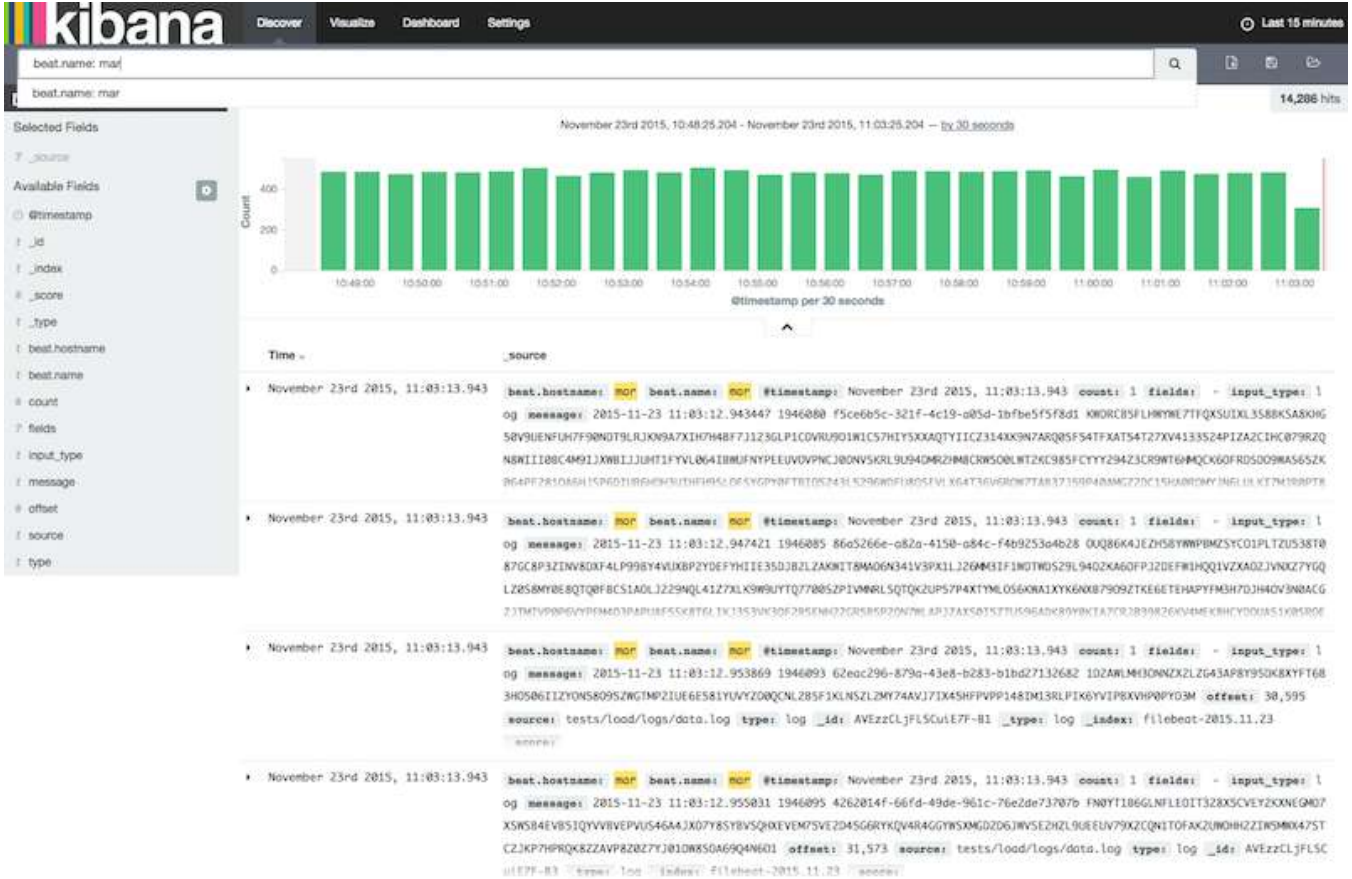
protocols can be added easily

Topbeat



elastic

Filebeat - logstash forwarder as beat



elastic

Beats

Metricbeat, Winlogbeat

Beat stats

Multiline filtering

elastic

Getting up and running - easy!

```
tar zxvf filebeat-1.X.Y-darwin.tgz ; cd filebeat-1.Y.Z
./filebeat -c filebeat.yml


./topbeat -c topbeat.yml
./packetbeat -c packetbeat.yml
```

Community Beats

apachebeat

elasticbeat

httpbeat

dockerbeat

redisbeat

phpfpmbeat

uwsgibeat

execbeat

unifiedbeat

pingbeat

nginxbeat

factbeat

hsbeat

nagioscheckbeat

elastic

Beats: Further info

https://www.elastic.co/guide/en/beats/libbeat/current/getting-started.html

https://www.elastic.co/guide/en/beats/packetbeat/current/index.html

https://www.elastic.co/guide/en/beats/filebeat/current/index.html

https://www.elastic.co/guide/en/beats/topbeat/current/index.html

https://www.elastic.co/guide/en/beats/winlogbeat/current/index.html

https://speakerdeck.com/tsg/get-real-time-insights-from-your-application-with-packetbeat-and-elasticsearch

Elasticsearch

elastic

Ingest node

Document enrichment before indexing

Simple document editing

Processors

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

Dead letter queue

failure handlers to change field or destination index

Ingest node

master only

Document enrichment before indexing

Simple document editing

Processors

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

Dead letter queue

failure handlers to change field or destination index

Ingest node - Configure pipeline

master only

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

elastic

Ingest node - Grok Processor

master only

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

elastic

Ingest node - Convert Processor

master only

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

elastic

Ingest node - Convert Processor

master only

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

elastic

Ingest node - Date Processor

master only

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

elastic

Ingest node - GeoIP Processor

master only

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

elastic

Ingest node - Index document

master only

```
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
=&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\""

}
```

elastic

Ingest node - Indexed document

master only


```
{
  "_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 }
    },
```


elastic

Ingest node - Indexed document

master only


```
…
    "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"
  }
}
```


elastic

All pluggable

master only

```
bin/plugin install ingest-geoip
```

elastic

Community
&
Documentation

elastic

The Elastic Community

40,000

Community members

1,500

Global subscription customers

35,000

Commits against Elastic stack to-date








elastic

ElasticSearch Meetups

Find out what's happening in ElasticSearch Meetup groups around the world and start meeting up with the ones near you.

Groups	Members	Interested	Cities	Countries
127	44,785	2,440	86	31

elastic

elastic

Products Subscriptions Learn Community Use Cases Blog

Learn | Docs

Docs

Elasticsearch: Store, Search, and Analyze

- Elasticsearch Reference [2.1] — other versions
- Elasticsearch - The Definitive Guide
- Resiliency status
- Plugins and Integrations [2.1] — other versions
- Elasticsearch Clients
- Sense Editor

Marvel: Monitoring for Elasticsearch

- Marvel Reference [2.1] — other versions

Shield: Security for Elasticsearch

- Shield Reference [2.1] — other versions

Watcher: Alerts for Elasticsearch

- Watcher Reference [2.1] — other versions

Found: Elasticsearch as a Service

- Found Reference

Kibana: Explore, Visualize, and Share

- Kibana Reference [4.3] — other versions

Logstash: Collect, Enrich, and Transport

- Logstash Reference [2.1] — other versions
- Logstash Roadmap

Beats: Collect, Parse, and Ship

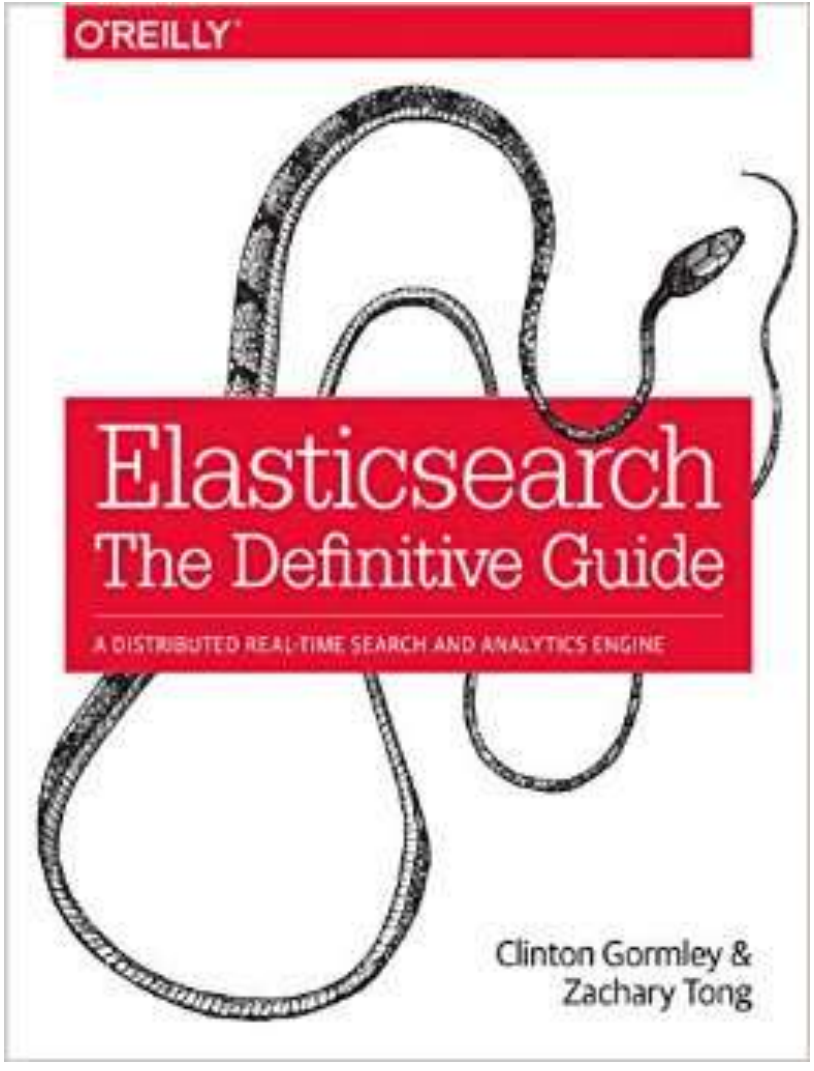
- Beats Platform Reference [1.0.0] — other versions
- Packetbeat Reference [1.0.0] — other versions
- Topbeat Reference [1.0.0] — other versions
- Filebeat Reference [1.0.0] — other versions

Elasticsearch for Apache Hadoop

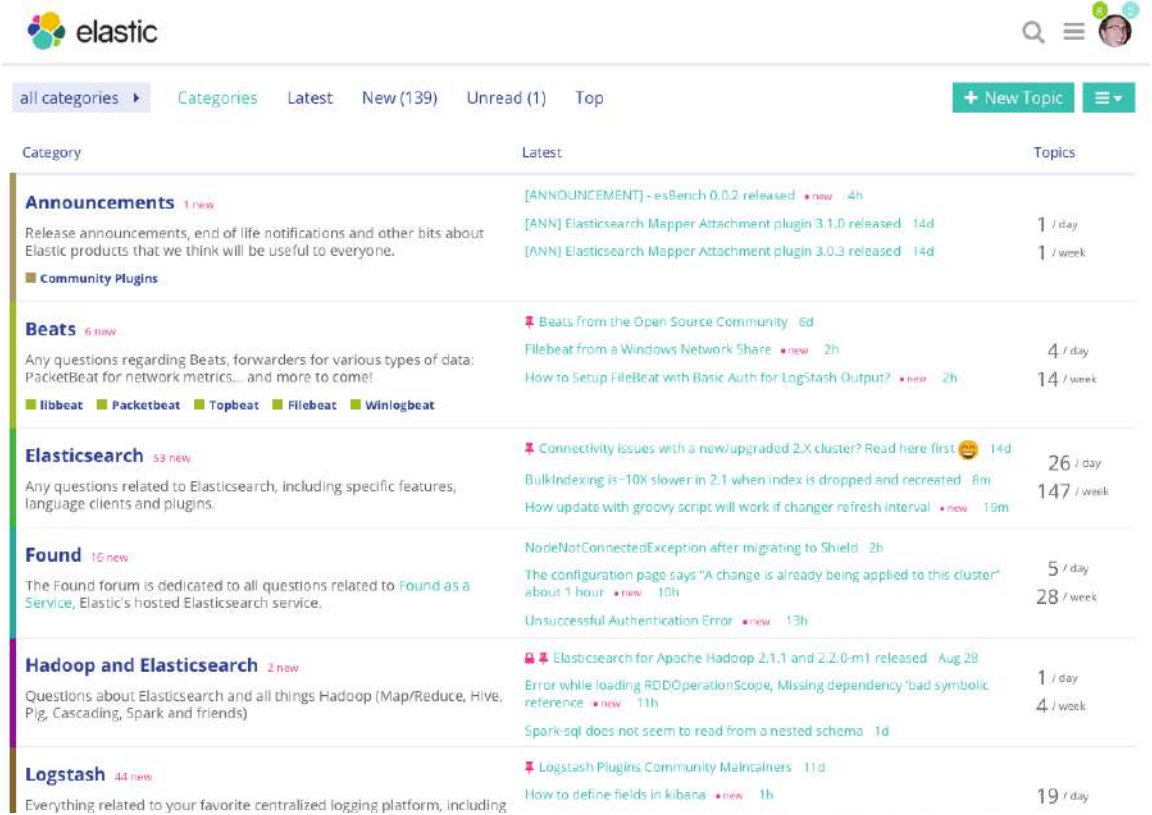
- es-hadoop reference [2.1] — other versions

Utilities

- Curator Index Management [3.4] — other versions
- Rivers [2.0] — other versions

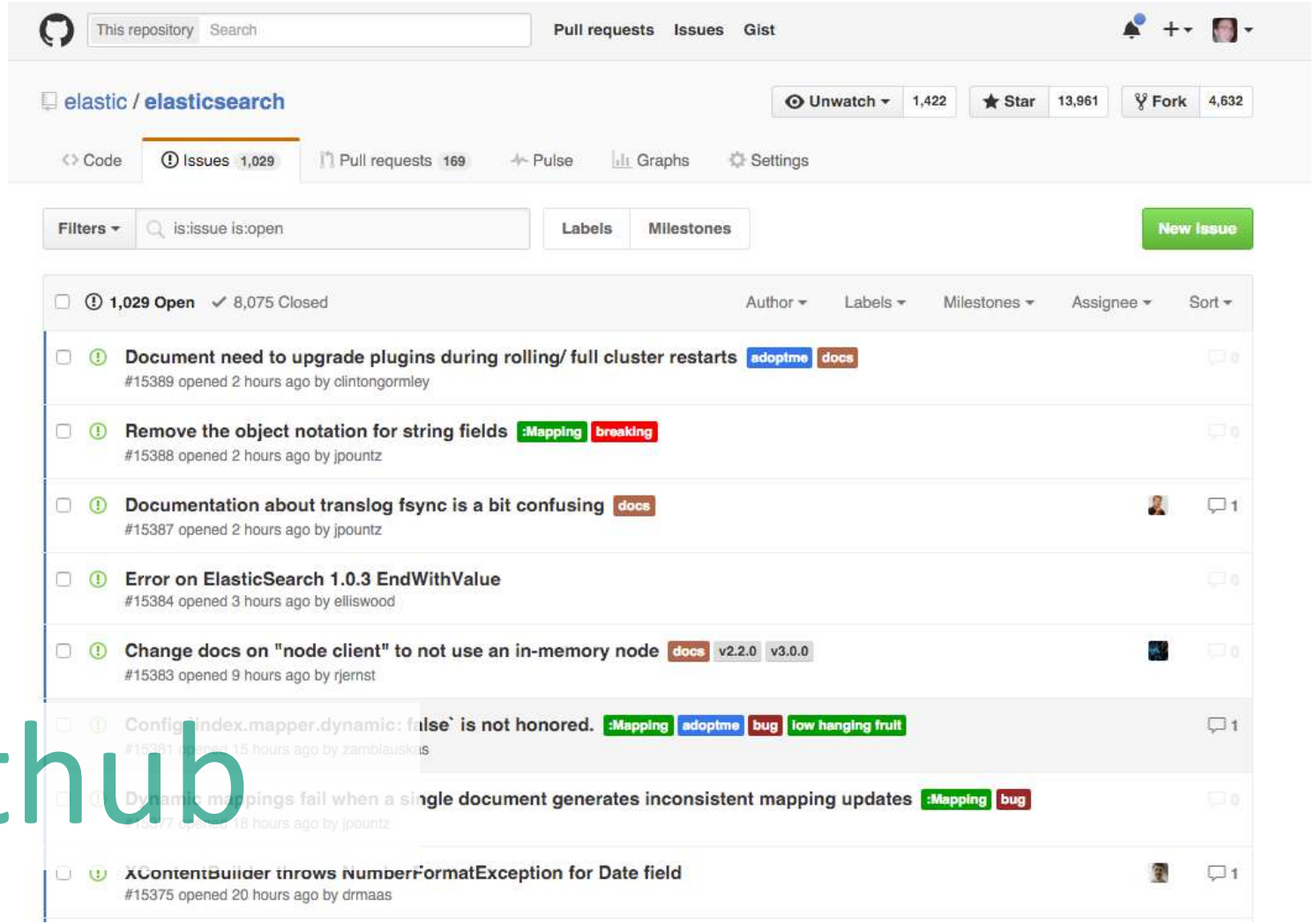

O'REILLY
Elasticsearch
The Definitive Guide
A DISTRIBUTED REAL-TIME SEARCH AND ANALYTICS ENGINE
Clinton Gormley &
Zachary Tong

elastic



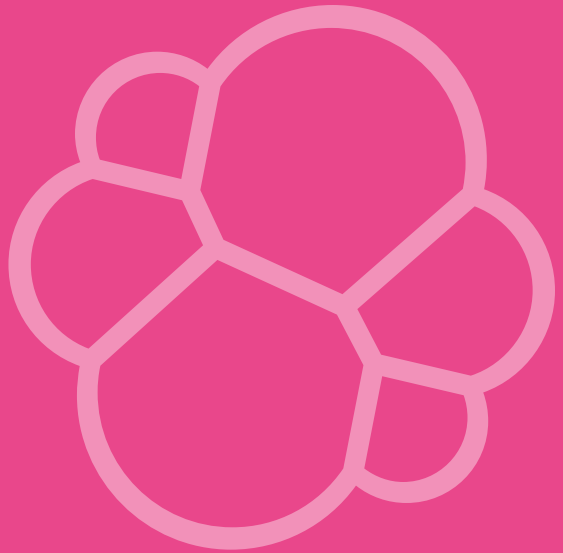
https://discuss.elastic.co/

elastic



github

elastic



Customers

elastic

Global Customer Base

Hi-Tech	     
Finance	     
Telco	     
Retail	     

elastic

"Elasticsearch is the backbone across all of Wikimedia's sites, powering billions of real-time user prefix and full-text searches every day."

Use Case	Search, Logging, Analytics
Products	Elasticsearch, Logstash, Kibana

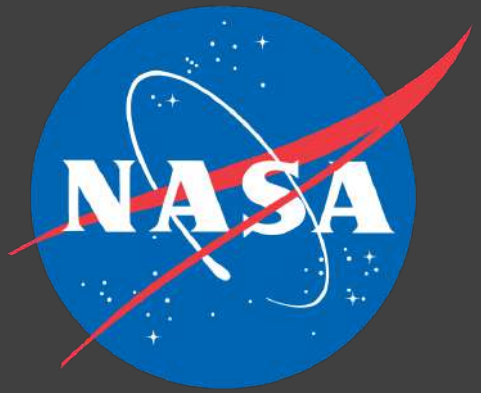
Chad Horohoe
Software Engineering

mozilla

"Elasticsearch, Logstash, and Kibana allow for real-time indexing, search, and analytics for over 300 million events per day. This protects our network, services, and systems from security threats."

Use Case	Search, Logging, Analytics, Security
Products	Elasticsearch, Logstash, Kibana

Jeff Bryner
Security Engineer



Use Case	Search, Logging, Analytics
Products	Elasticsearch, Logstash, Kibana

"With the Elastic Stack, we log more than 30K messages and 100K documents four times every day from the Mars Rover to optimize our space missions."

Dan Isla
Data Scientist

"Using Elasticsearch, we index more than 500 billion documents for real-time logging and analytics for our mission critical applications."

Use Case	Logging, Analytics
Products	Elasticsearch, Logstash

Bhaskar Karambelkar
Sr. Security Data Scientist

Roundup

elastic

Ease-of-use

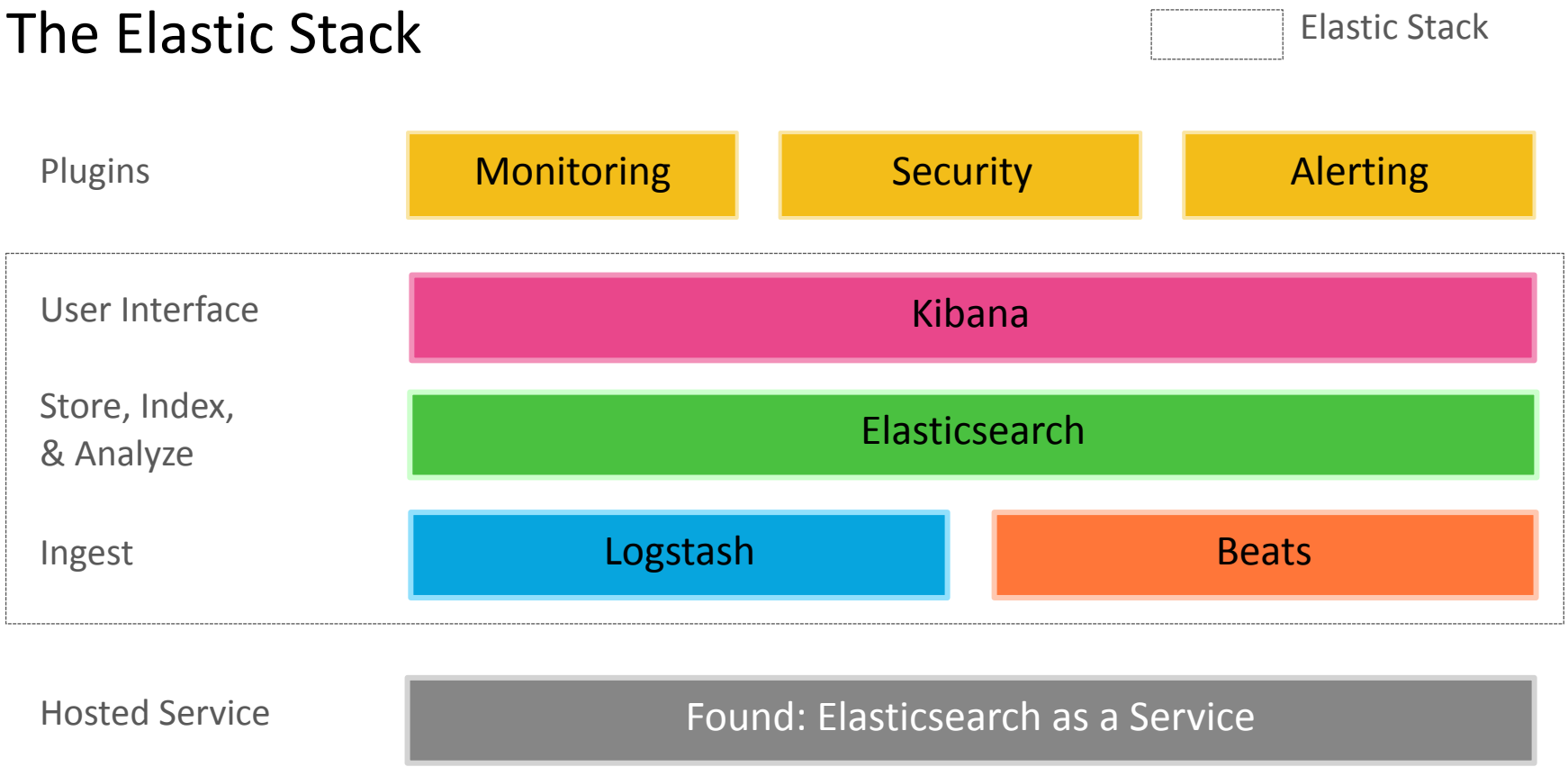
Minimal dependencies

Extensibility

Consistency

Flexibility

The Elastic Stack



elastic

Thank You

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

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