

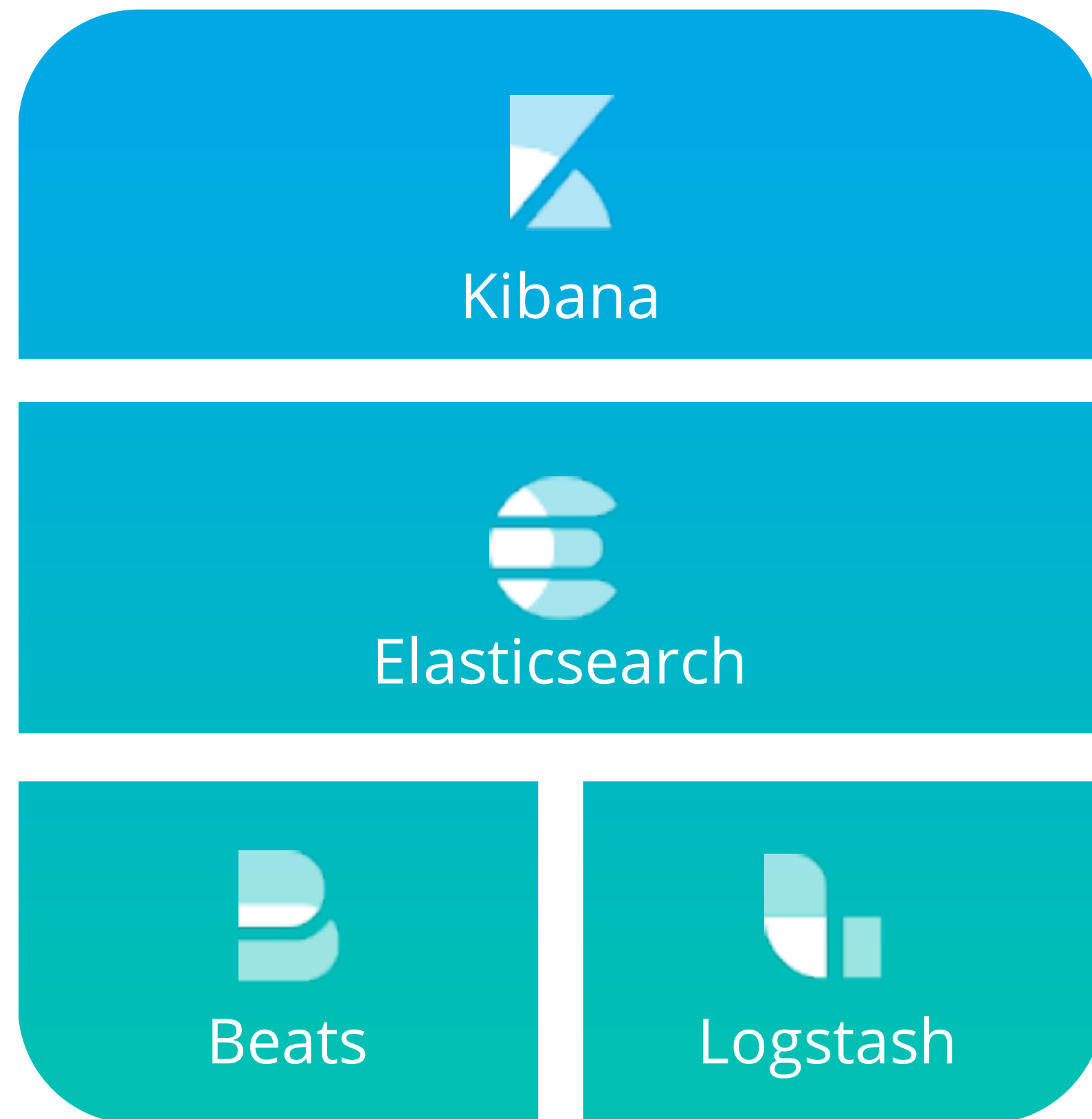


# The State of Elastic Stack

*What's new and what we are working on*

Medcl

# Elastic Stack & X-Pack



**Security**



**Alerting**



**Monitoring**



**Reporting**



**Graph**



**Machine Learning**

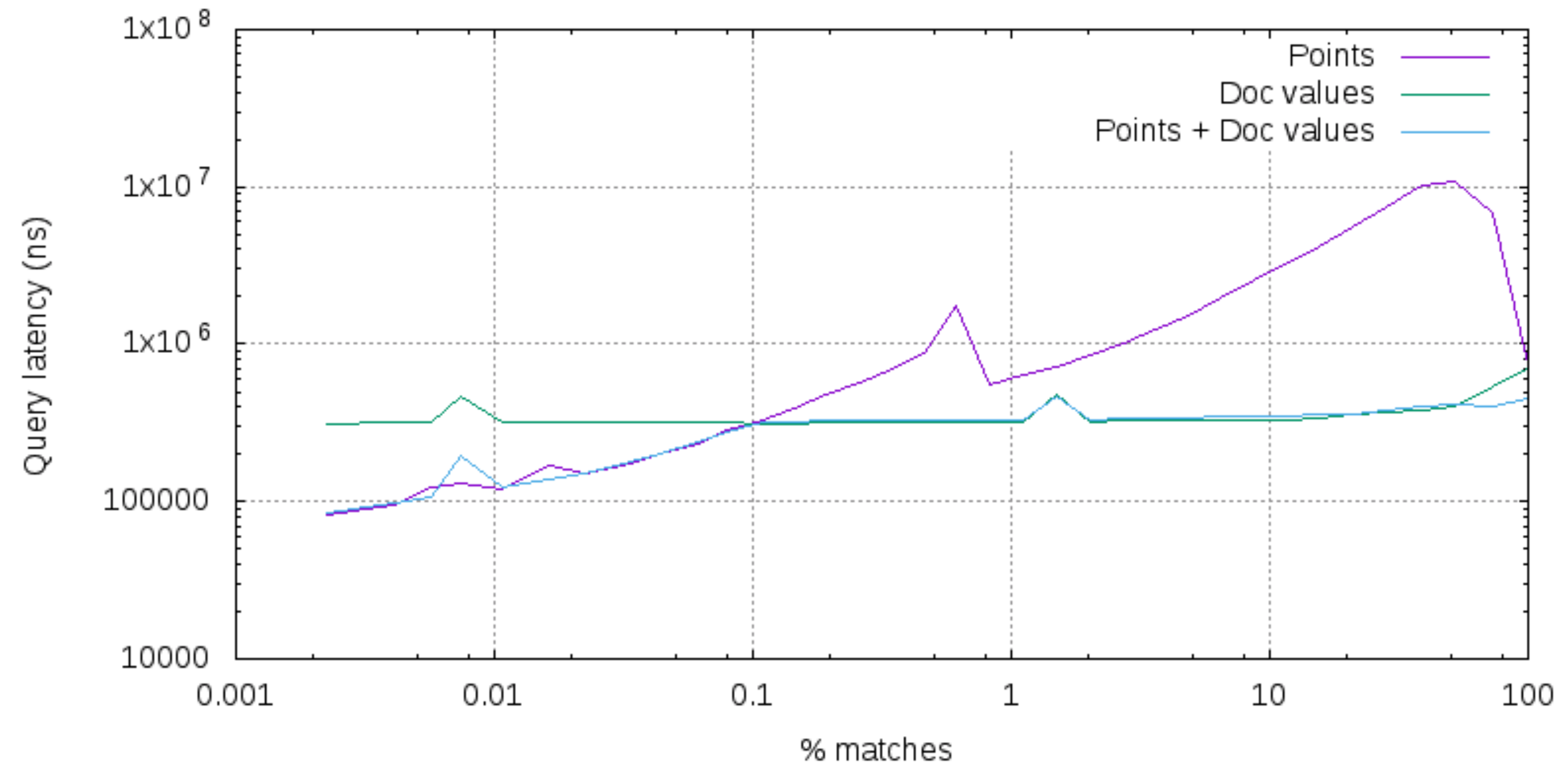


elasticsearch

# Optimized Query Execution

Same queries, just faster.

- range
  - automatically chooses the more efficient of two query modes
  - Additional materials: [blog](#)
- nested
  - Nested mappings / queries get a speed boost
  - Writeup of how/why: [23079](#)



Query latency for 0.1% term and range

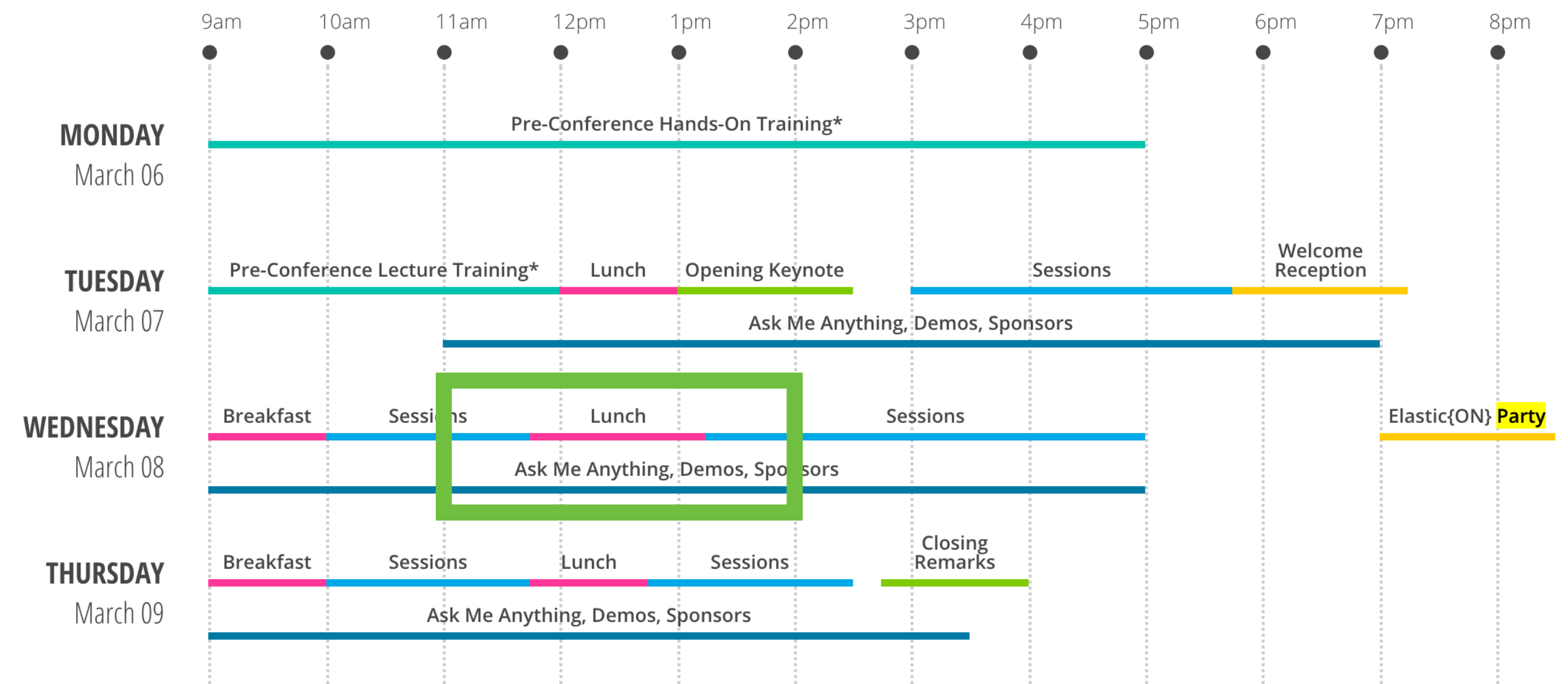


# Numeric & Date Range Fields (5.2)

## Mapping Improvements

- New types for date/number ranges (5.2)  
(*date\_range*, *int\_range*, *float\_range*)

What's happening Wednesday 11am - 2pm



# Terms Aggregation Partitioning (5.2)

## Returning ALL the Terms, in Manageable Chunks

- frequent request
- return all responses from a terms aggs
- Terms can now be broken into partitions and partitions are returned by number

```
{
  "size": 0,
  "aggs": {
    "expired_sessions": {
      "terms": {
        "field": "account_id",
        "include": {
          "partition": 0,
          "num_partitions": 20
        },
        "size": 10000,
        "order": {
          "last_access": "asc"
        }
      },
      "aggs": {
        "last_access": {
          "max": {
            "field": "access_date"
          }
        }
      }
    }
  }
}
```

# Cross Cluster Search (5.3)

Tribe node is dead. Long live Cross-cluster search.

- Minimal viable solution to supersede tribe
- Addresses many of the challenges with tribe node
- Reduces the problem domain to query execution
- Cluster related information is reduced to a namespace

# Field Collapsing (5.3)

One method to rule them all...

- Simple (almost) no setup!
- Great for query-time group/category de-dup

```
GET /twitter/tweet/_search
{
  "query": {
    "match": {
      "message": "elasticsearch"
    }
  },
  "collapse" : {
    "field" : "user", ①
    "inner_hits": {
      "name": "last_tweets", ②
      "size": 5, ③
      "sort": [{ "date": "asc" }] ④
    },
    "max_concurrent_group_searches": 4 ⑤
  },
  "sort": ["likes"]
}
```

# Elasticsearch Keystore

If you like it, you should put it in a keystore.

- Sensitive settings should not be protected by filesystem permissions only.
- Commands feel familiar:
  - `bin/elasticsearch-keystore create`
  - `bin/elasticsearch-keystore list`
  - `bin/elasticsearch-keystore add the.setting.name.to.set`
  - `bin/elasticsearch-keystore remove the.setting.name.to.remove`
- Just the framework/start: sensitive settings to be pulled in

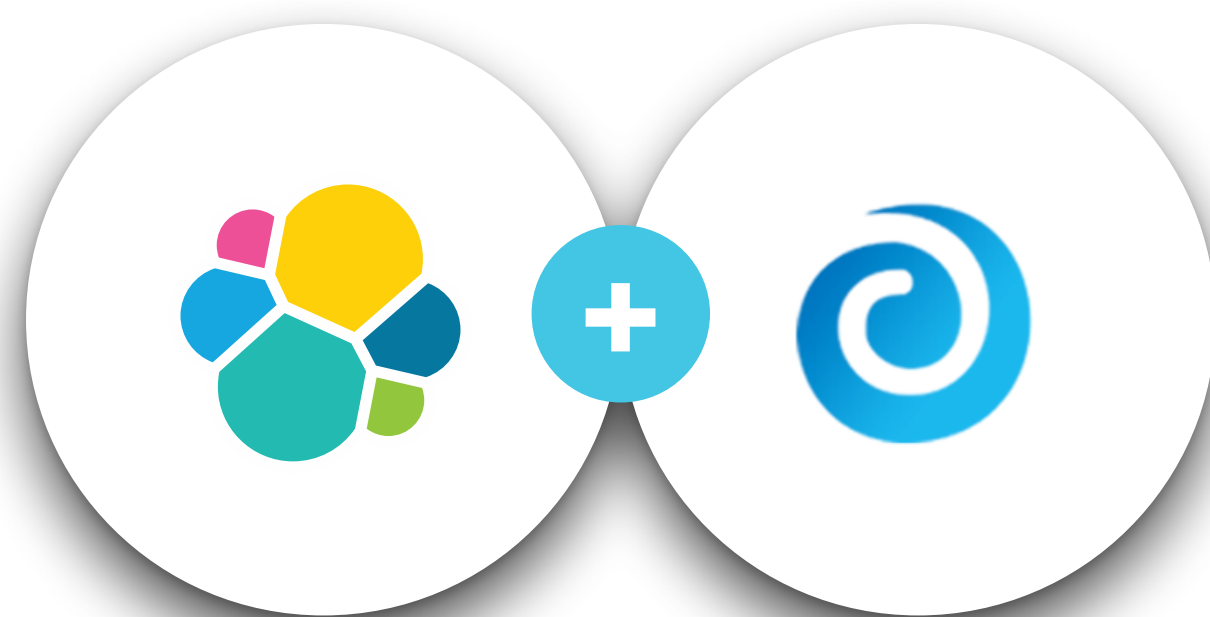
# Elasticsearch 6.0 is coming

- Remove Type
- Sparse Doc Values
- Index Sorting
- Sequence Numbers
- Rolling Upgrades
- ...



Elastic Stack 6.0.0-alpha2 Released





# Machine Learning

## UNSUPERVISED MACHINE LEARNING

- Automatically detect anomalies
- Advanced correlation and categorization
  - Identify root cause(s)
- Expose early warning signs

---

## NEW USE CASES

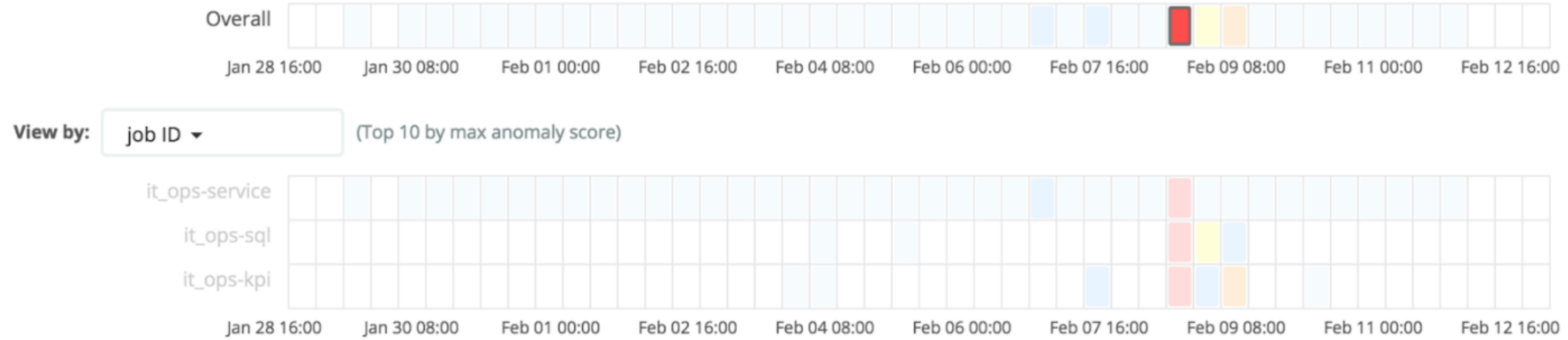
- Analyze time series data
- Expand security, IT Ops, fraud, finance, and many more use cases
  - Currently beta; building a more native integration into the Elastic Stack

Job it\_ops-kpi and 2 others ▼

## Top Influencers

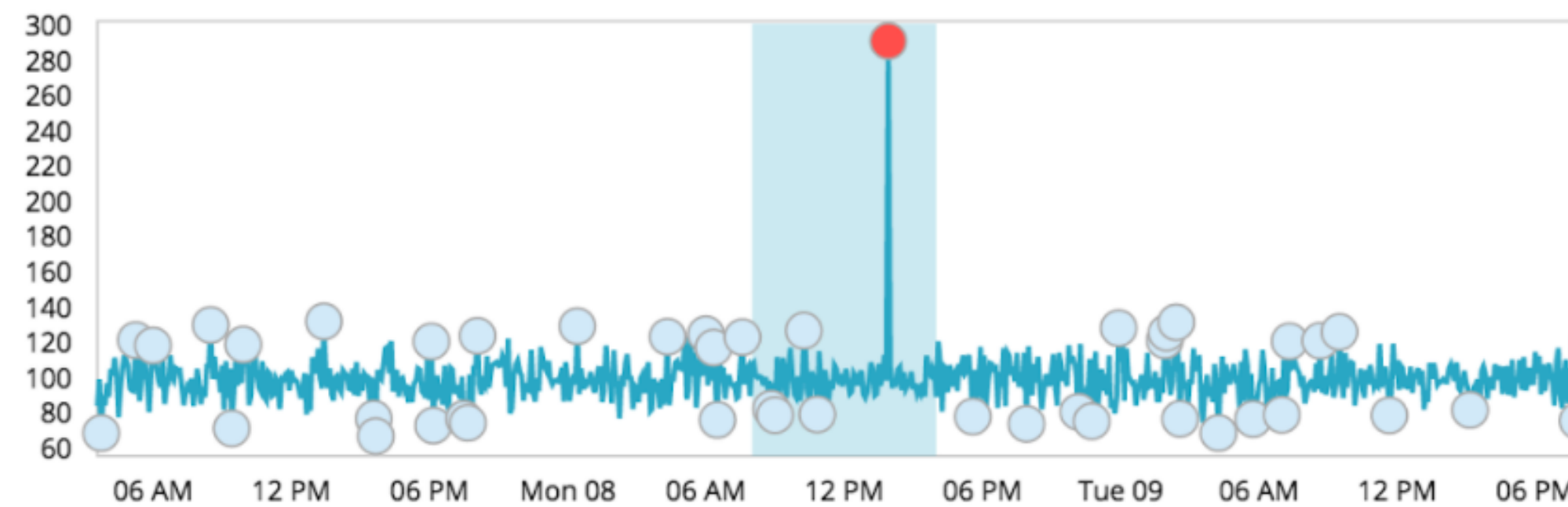


## Anomaly timeline

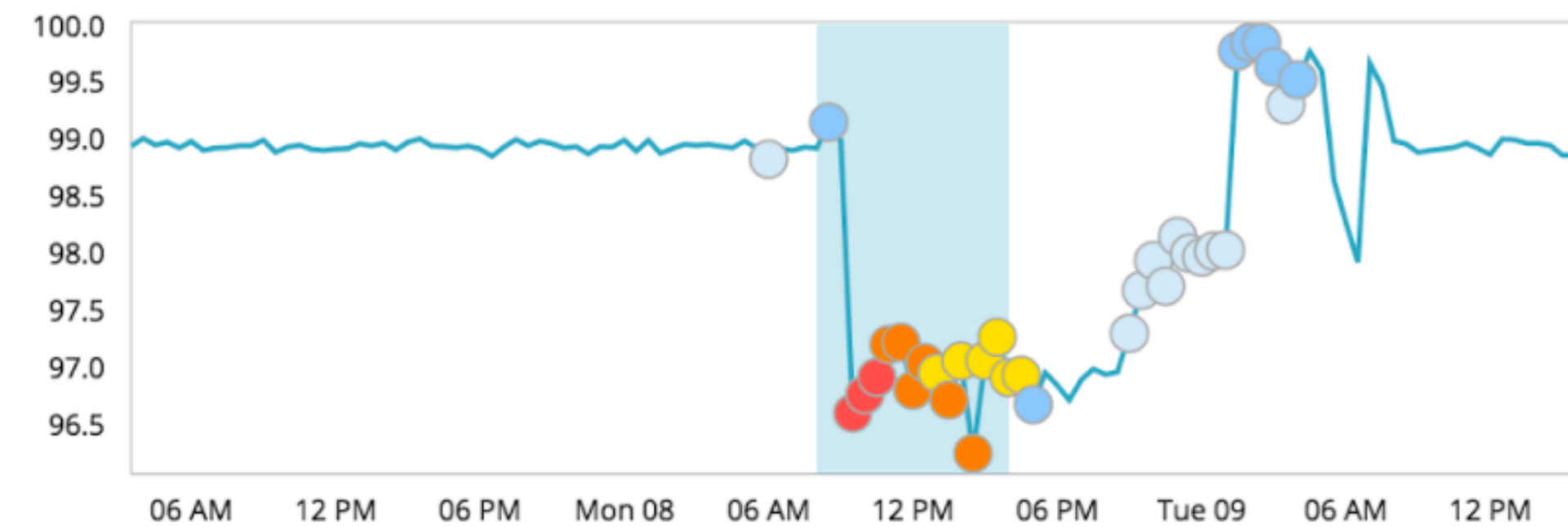


## Anomalies

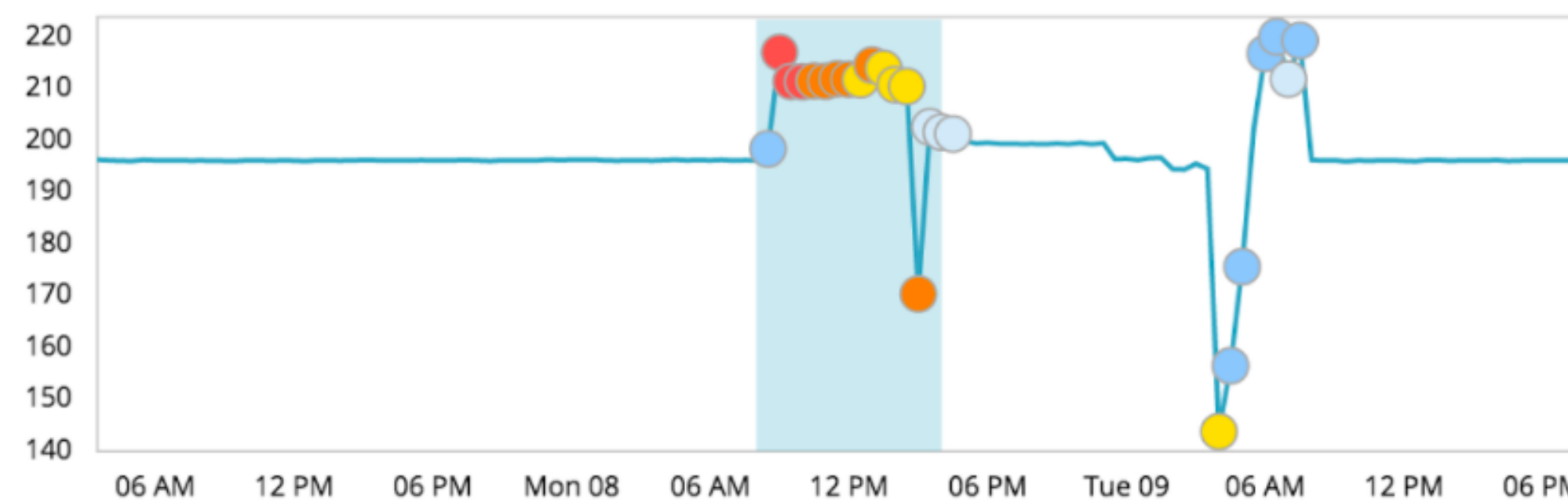
mean responsetime service inventory-us-east-1-34



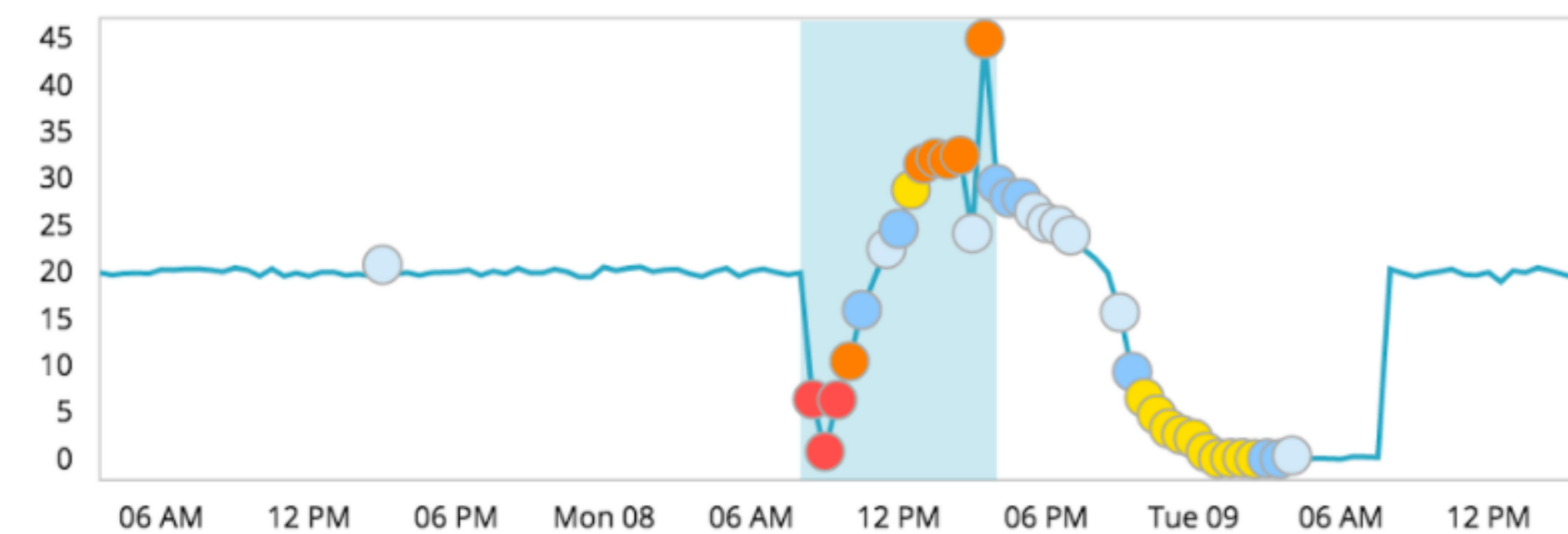
mean SQLServer\_Buffer\_Manager\_Buffer\_cache\_hit\_ratio hostname MSSQL-0783E4076



mean SQLServer\_General\_Statistics\_User\_Connections hostname MSSQL-0783E4076



mean SQLServer\_SQL\_Statistics\_Batch\_Requests\_sec hostname MSSQL-0783E4076



# Elasticsearch-SQL



Elasticsearch-SQL Coming soon!

## CLI

- OS independent
- Quick diagnostics and sanity checks
- Admin focused
- Optimized for efficiency

## JDBC

- Dedicated client (driver) and server component
- JDBC 4.2/Java 8 (downgrade possible)
- Supports `java.sql` and `javax.sql` APIs
- Pays attention to details
  - Timeouts (connect vs read vs network)
  - Logging
- Light, without dependencies



# Elasticsearch-SQL

Dev Tools

Console

```
1 GET /_sql
2 {
3   "query" : "SELECT * FROM emp.emp"
4 }
```

History

Settings

```
1 {
2   "size": 100,
3   "columns": {
4     "age": {
5       "type": "integer"
6     },
7     "emp_no": {
8       "type": "integer"
9     },
10    "first_name": {
```

DbVisualizer Pro 9.5.6 [EVALUATION] - ES/elasticsearch//TABLE/emp.emp

File Edit View Database SQL Commander Tools Window Help

Databases

Connections

Scripts

Favorites

Table: emp.emp

Info Columns Data Row Count Primary Key Indexes Grants Row Id DDL

	birth_date	emp_no	first_name	gender	last_name	salary	tenure
1	1953-09-02	10001	Georgi	M	Facello	52184	31
2	1952-04-19	10009	Sumant	F	Peac	91831	32
3	1963-06-01	10010	Duangkaew	F	Piveteau	91222	28
4	1961-05-02	10016	Kazuhiro	M	Cappelletti	58493	22
5	1952-07-08	10022	Shahaf	M	Famili	51352	22
6	1956-12-13	10029	Otmir	M	Herbst	75996	32
7	1963-07-22	10037	Pradeep	M	Makrudi	65937	27
8	1960-07-23	10046	Lucien	M	Rosenbaum	52903	25
9	1963-07-11	10048	Florian	M	Syrobiuk	54611	32
10	1953-07-28	10051	Hidefumi	M	Caine	76893	25
11	1961-02-26	10052	Heping	M	Nitsch	93008	29
12	1956-06-06	10055	Georgy	M	Dredge	81830	25
13	1963-04-14	10065	Satosi	M	Awdeh	38304	29
14	1953-01-07	10067	Claudi	M	Stavenow	48996	30
15	1955-08-28	10074	Mokhtar	F	Bernatsky	78088	27
16	1964-04-18	10077	Mona	M	Azuma	58183	27
17	1960-05-25	10084	Tuval	M	Kalloufi	100855	22
18	1963-03-21	10089	Sudharsan	F	Flasterstein	86574	31
19	1954-09-16	10096	Jayson	M	Mandell	61522	27
20	1964-06-02	10002	Bezalel	F	Simmel	70000	32
21	1954-05-01	10004	Christian	M	Koblick	43197	31
22	1953-04-20	10006	Anneke	F	Preusig	74702	28
23	1953-11-07	10011	Mary	F	Sluis	93275	27
24	1958-07-06	10017	Cristinel	F	Bouloucos	67904	24
25	1954-06-19	10018	Kazuhide	F	Peha	40575	30
26	1953-01-23	10019	Lillian	M	Haddadi	47152	18
27	1952-12-24	10020	Mayuko	M	Warwick	40969	26
28	1962-07-10	10027	Divier	F	Reistad	102739	28

1: emp [14]

tenure min (+2)

DbVisualizer Pro 9.5.6 [EVALUATION] - Untitled\*

File Edit View Database SQL Commander Tools Window Help

Databases

Connections

Scripts

Favorites

1: Untitled\*

ES emp.emp

Database Connection

Schema

Max Rows

Max Chars

1 SELECT min(salary) min, avg(salary) avg, max(salary) max, tenure FROM emp.emp GROUP BY tenure HAVING min > 30000

2 ORDER BY tenure

Log

1: emp [14]

tenure min (+2)



kibana



# Kibana 5.4!

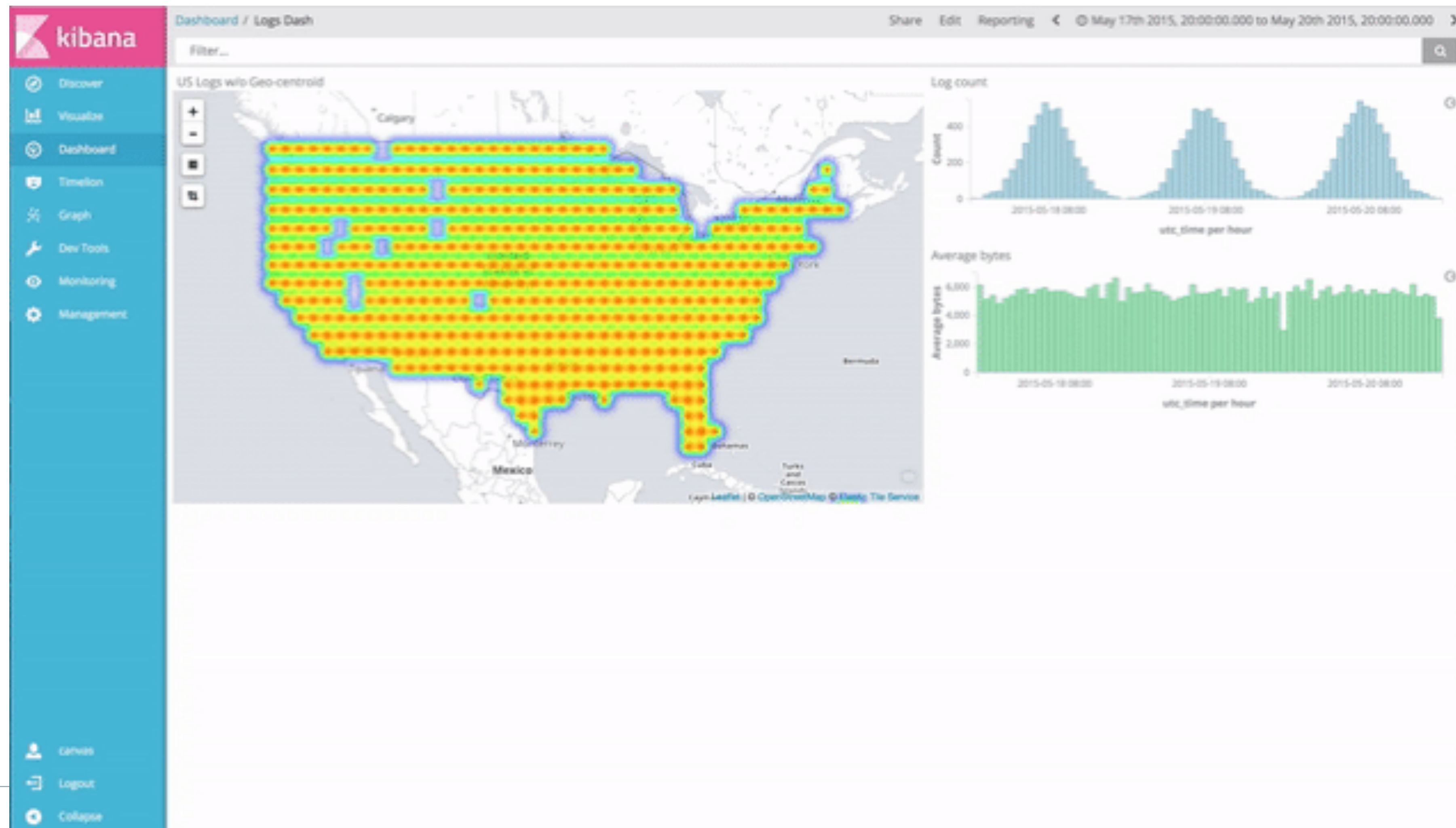
- **20** Kibana enhancements with 5.4
- **10** Visualization customizations
  - Combo chart (2), horizontal bar chart (2), multiple y-axis (4), easily switch between viz (2)
- **5** Pipeline agg support
  - Derivatives (5)
- **2** Timepicker enhancements
  - “to” field in relative, start / end of day
- **1** Dashboard enhancement
  - Allow duplicate dashboard names
- **1** Reporting enhancement
  - Report should reflect visualization query when defined





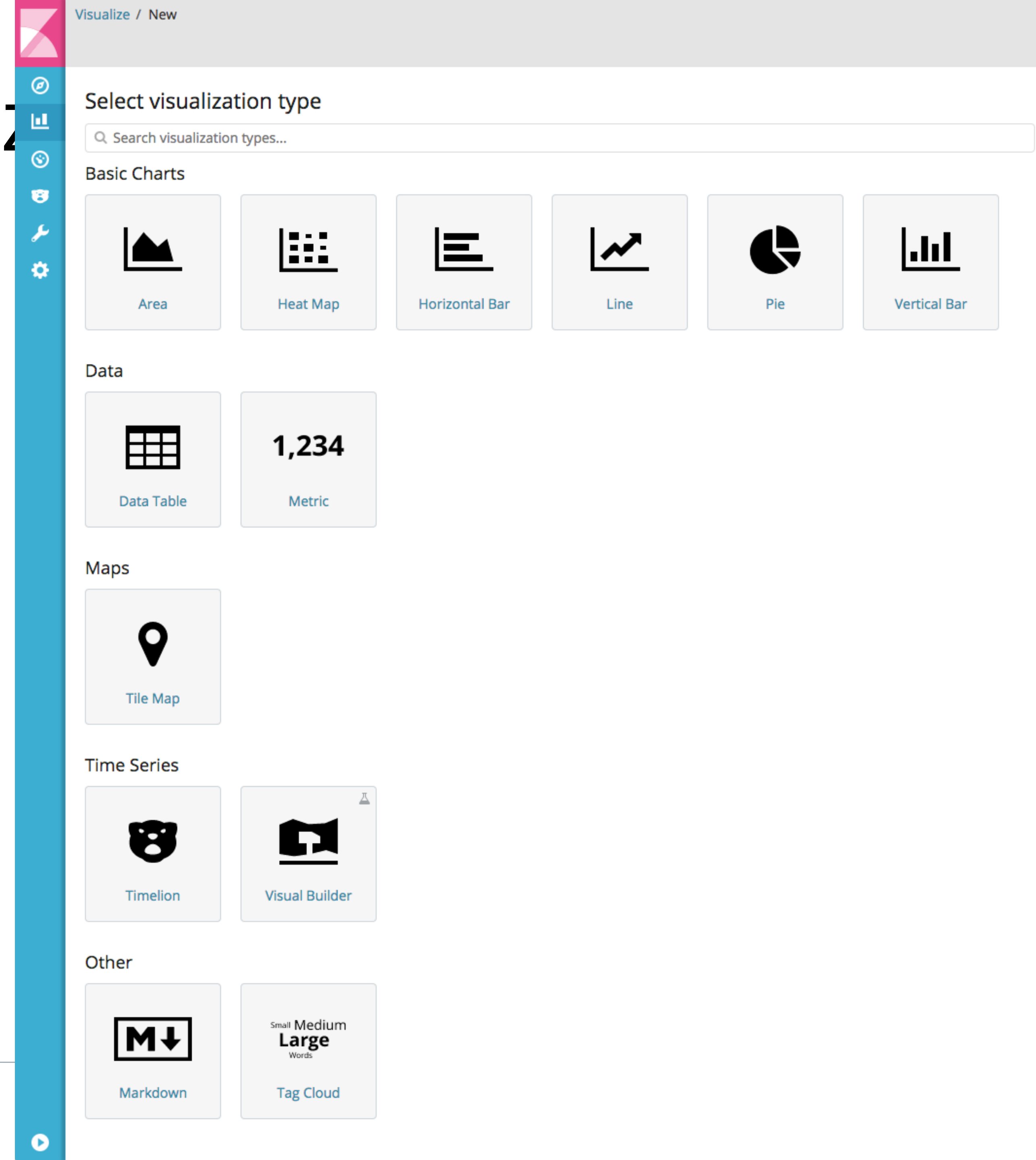
# Dashboard View & Edit Mode

Gone are the days of accidental panel resizing



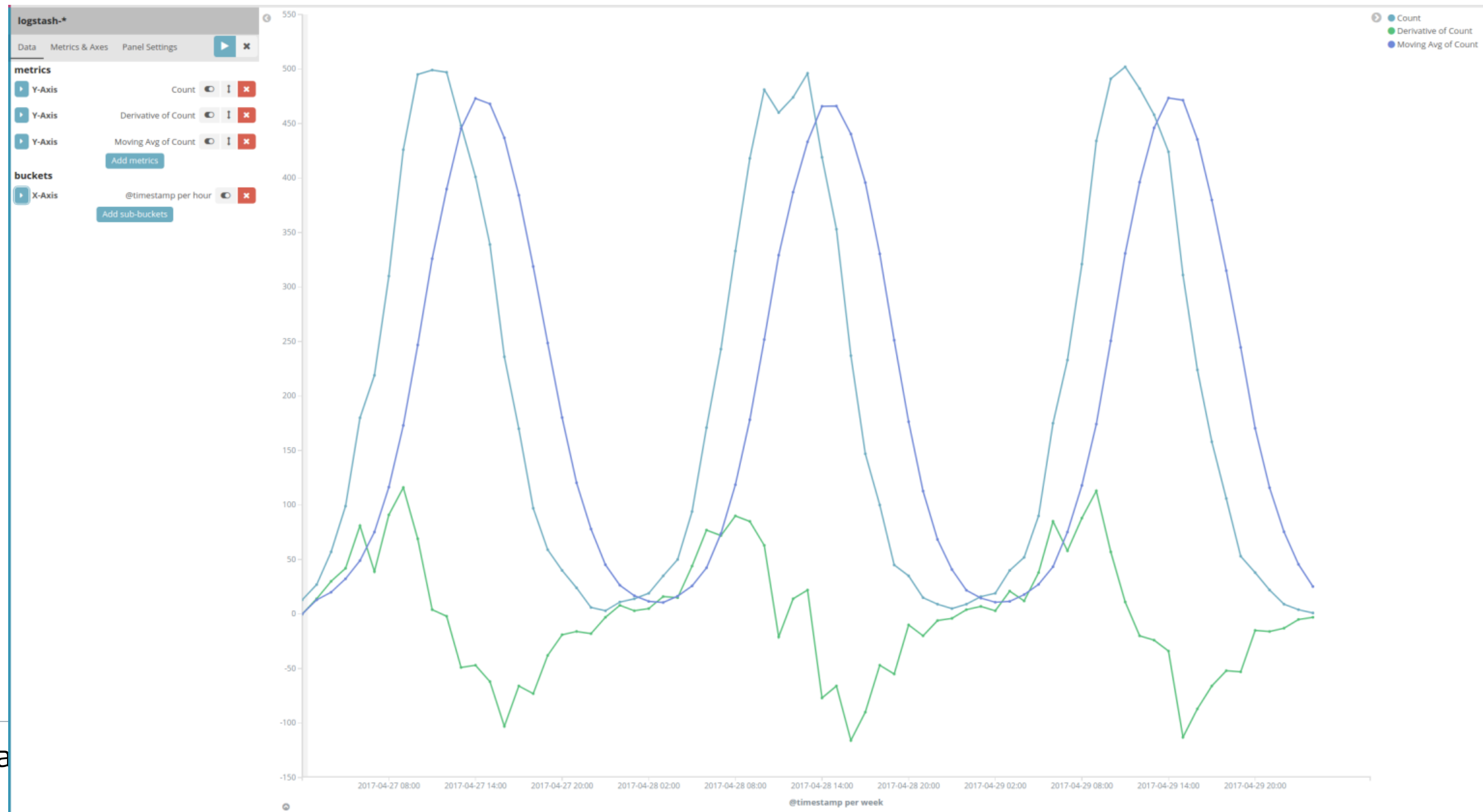
# Create Visualization Wizard

- Creating a new visualization can be daunting
- Group visualization types into buckets with icons for each visualization type



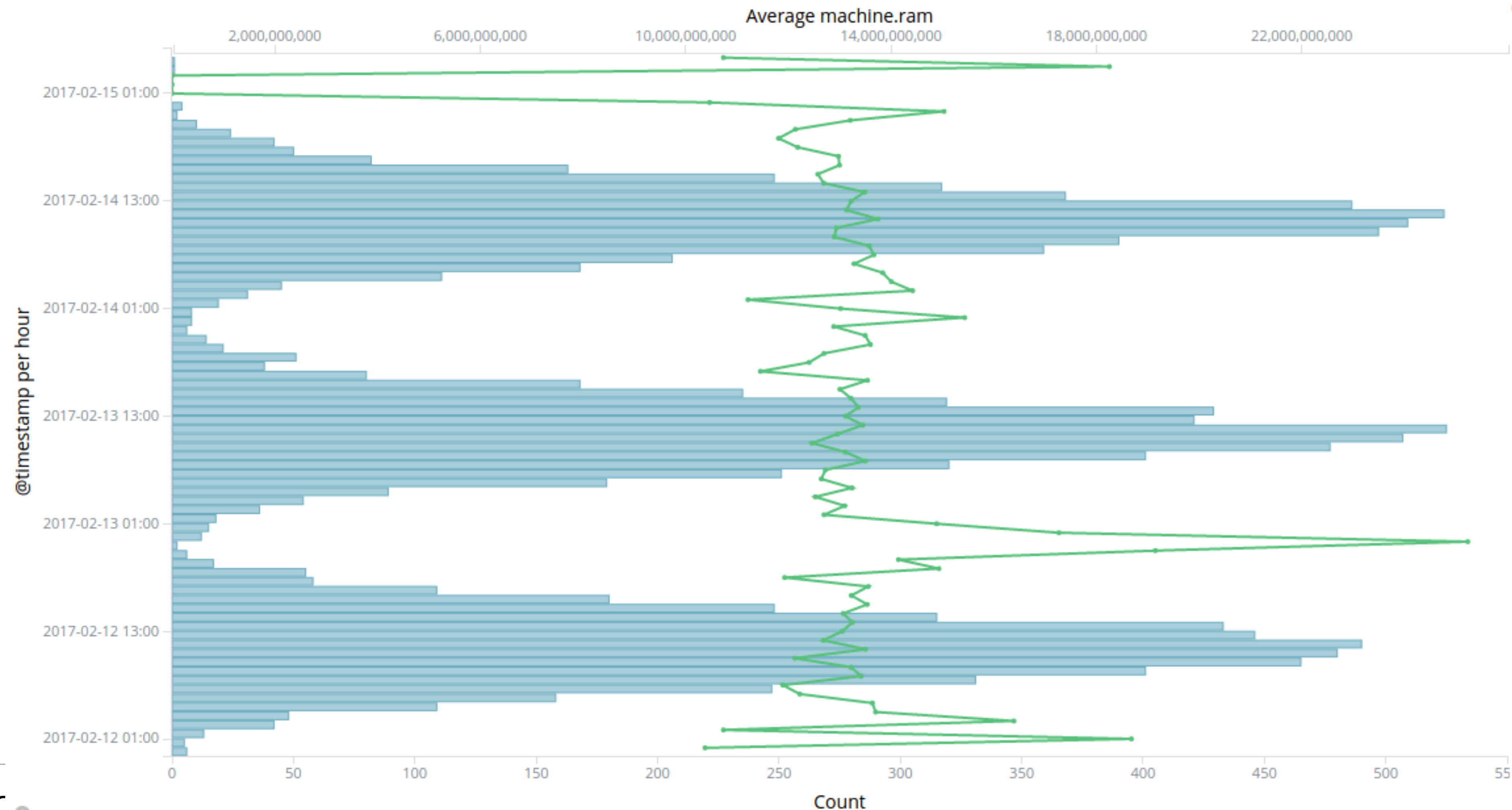
# Pipeline Aggs in Kibana!

We've all wanted it, here it is!



# Visualization Enhancements

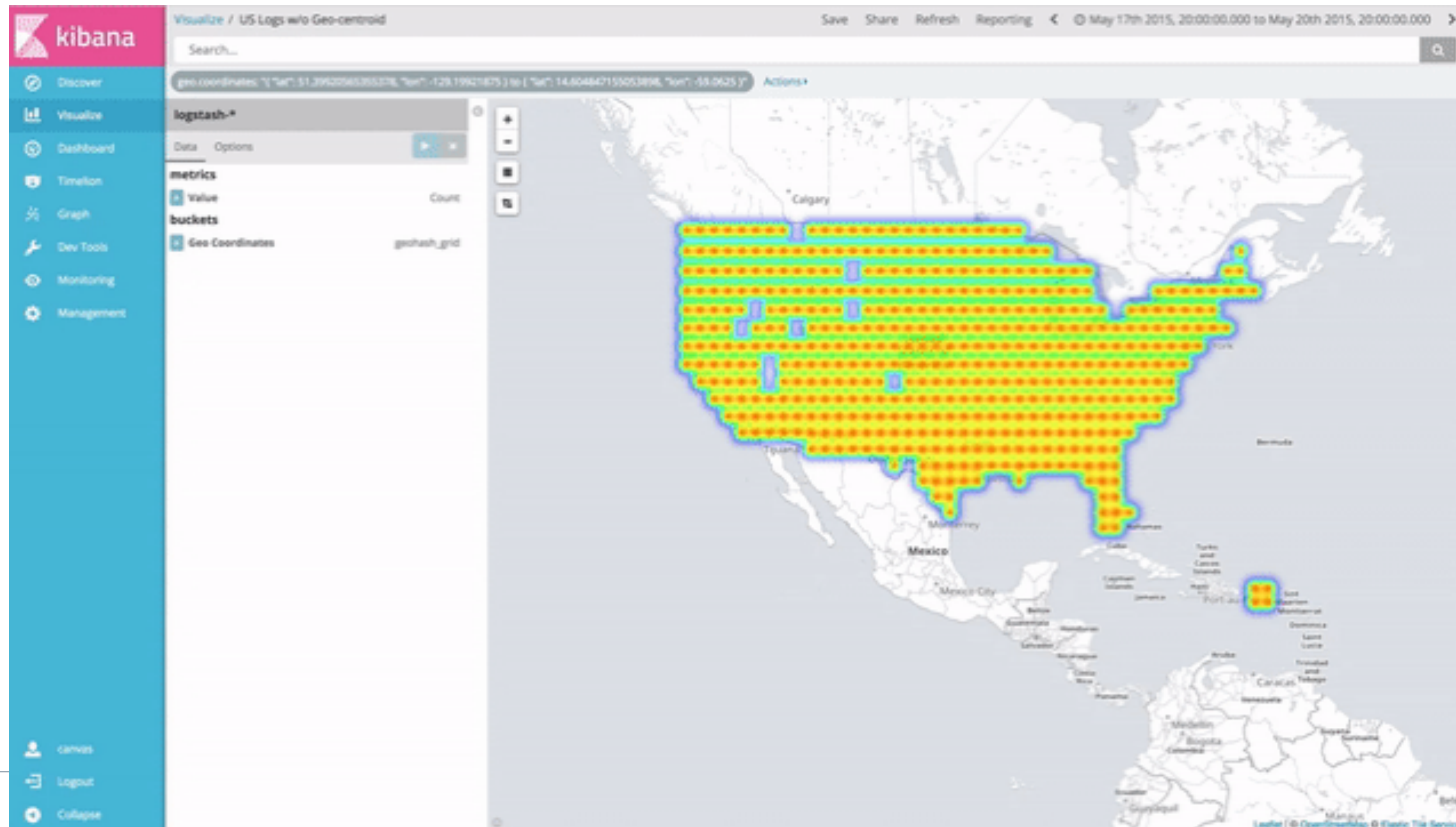
Combination charts, horizontal charts, multiple y-axis & more!





# Geo-Centroid Support

More natural looking maps. Less awkward boxes.



# Event Context

## The Needle in the Haystack

Filter a number of events

Understand ‘before’ and ‘after’ events

Enter from Discover through ‘View surrounding documents’ link

Surrounding Documents in logstash-\*

log#AVtdkRwgxIjtxpRZ3Q9K

Load 5 more

5

newer documents

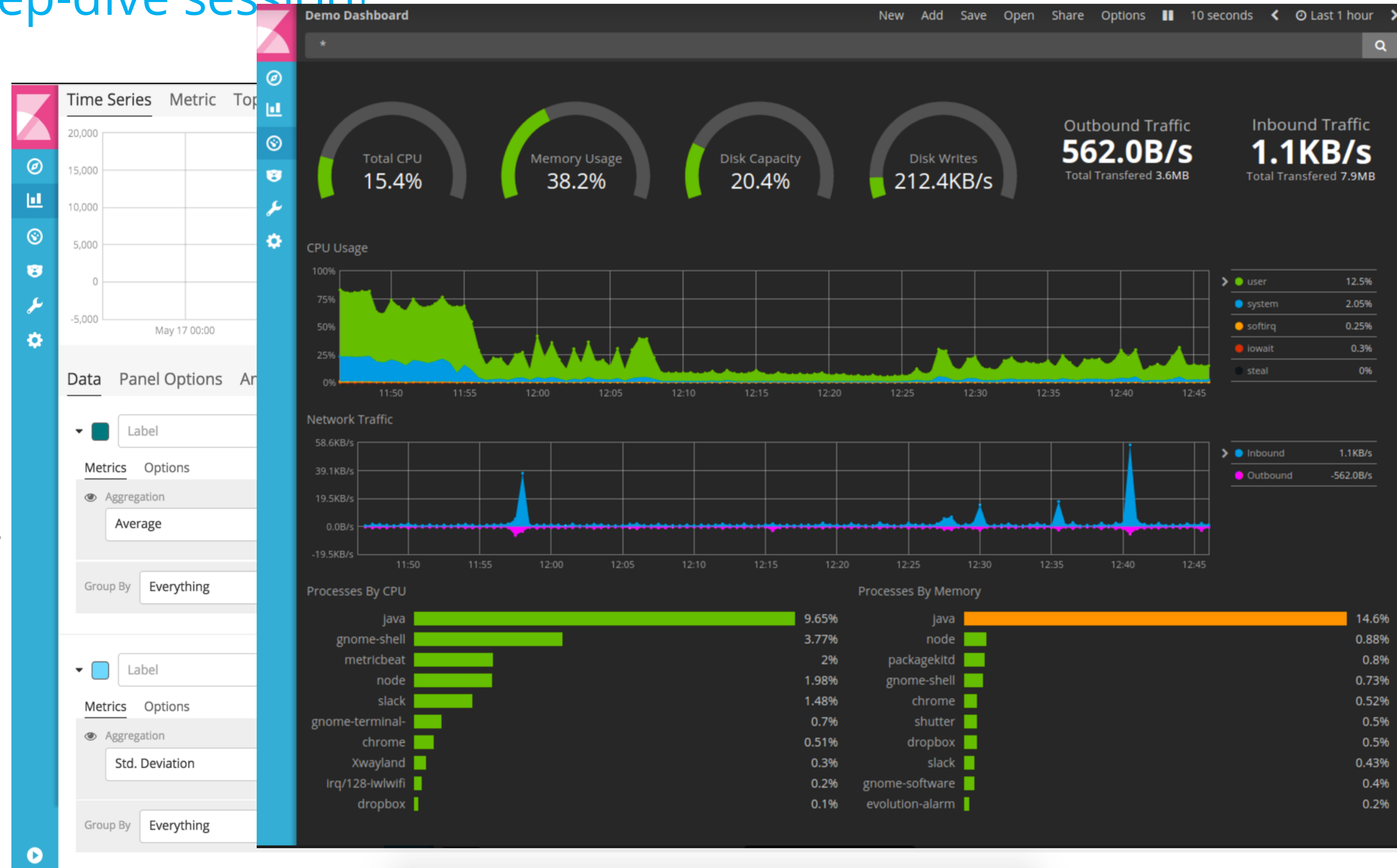
Time	@message
▶ May 20th 2015, 13:41:28.906	189.244.165.222 - - [2015-05-20T20:41:28.906Z] "GET /uploads/ronald-mcnair.jpg HTTP/1.1" 200 3407 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:40:05.434	8.105.89.9 - - [2015-05-20T20:40:05.434Z] "GET /uploads/michael-p-anderson.jpg HTTP/1.1" 200 9179 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:36:52.315	49.151.142.71 - - [2015-05-20T20:36:52.315Z] "GET /people/type:astronauts/name:krasimir-stoyanov/profile HTTP/1.1" 200 4663 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:36:16.940	167.67.215.75 - - [2015-05-20T20:36:16.940Z] "GET /uploads/john-blaha.jpg HTTP/1.1" 200 8458 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:34:26.604	124.254.160.199 - - [2015-05-20T20:34:26.604Z] "GET /uploads/voskhod-1.jpg HTTP/1.1" 200 9979 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:32:14.679	161.107.162.58 - - [2015-05-20T20:32:14.679Z] "GET /styles/ad-blocker.css HTTP/1.1" 200 4825 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)"
▶ May 20th 2015, 13:32:06.085	128.155.230.205 - - [2015-05-20T20:32:06.085Z] "GET /uploads/philip-k-chapman.jpg HTTP/1.1" 200 3088 "-" "Mozilla/5.0 (X11; Linux i686) AppleWebKit/534.24 (KHTML, like Gecko) Chrome/11.0.696.50 Safari/534.24"
▶ May 20th 2015, 13:31:24.323	167.137.188.138 - - [2015-05-20T20:31:24.323Z] "GET /uploads/charles-e-brady-jr-.jpg HTTP/1.1" 200 3959 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:28:49.156	50.11.235.131 - - [2015-05-20T20:28:49.156Z] "GET /uploads/nicole-marie-passonno-stott.jpg HTTP/1.1" 200 3323 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:26:32.000	152.150.148.109 - - [2015-05-20T20:26:32.000Z] "GET /uploads/joseph-m-acaba.png HTTP/1.1" 200 6106 "-" "Mozilla/5.0 (X11; Linux x86_64; rv:6.0a1) Gecko/20110421 Firefox/6.0a1"
▶ May 20th 2015, 13:23:17.483	163.15.240.221 - - [2015-05-20T20:23:17.483Z] "GET /uploads/sergei-ryazanski.jpg HTTP/1.1" 200 3251 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1; .NET CLR 1.1.4322)"



# Time Series Visual Builder

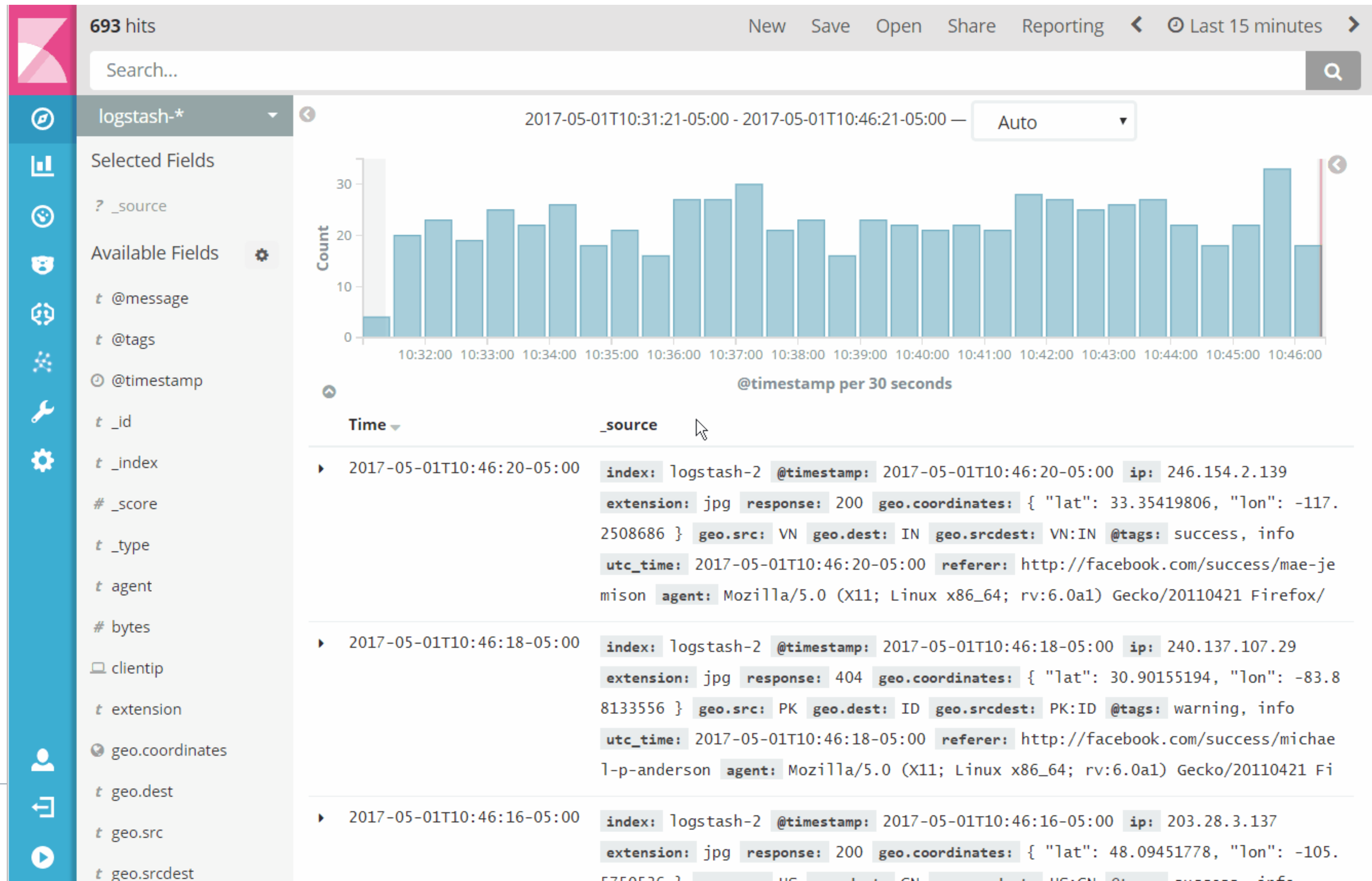
More to come, an entire deep-dive session!

- A curated UI - just for time series data with features such as...
- Chart multiple indices
- Pipeline aggregations
- Complex calculations
- Conditional formatting
- Annotations
- Series offset



# Watcher UI

## Phase 1 - View, manage, and create Watches



# Cluster Alerts

Built-in Watches to help diagnose cluster issues quickly

kibana

Discover

Visualize

Dashboard

Timelion

Machine Learning

Graph

Dev Tools

Monitoring

Management

elastic

Logout

Collapse

Clusters

elasticsearch

Your Trial license will expire on May 24, 2017.

Top Cluster Alerts

Elasticsearch cluster status is red. Allocate missing primary shards and replica shards.

April 28, 2017 11:13:12 AM

[View all alerts](#)

Elasticsearch

Health: Red

Overview

Version: 5.4.0  
Uptime: 38 minutes

Nodes: 1

Disk Available: 208GB / 465GB (44.78%)  
JVM Heap: 32.67% (647MB / 2GB)

Indices: 26

Documents: 1,821,745  
Disk Usage: 746MB  
Primary Shards: 43  
Replica Shards: 0

Kibana

Health: Green

Overview

Requests: 2  
Max. Response Time: 31 ms

Instances: 1

Connections: 4  
Memory Usage: 8.25% (118MB / 1GB)


Logstash

Overview

Events Received: 3  
Events Emitted: 3

Nodes: 1

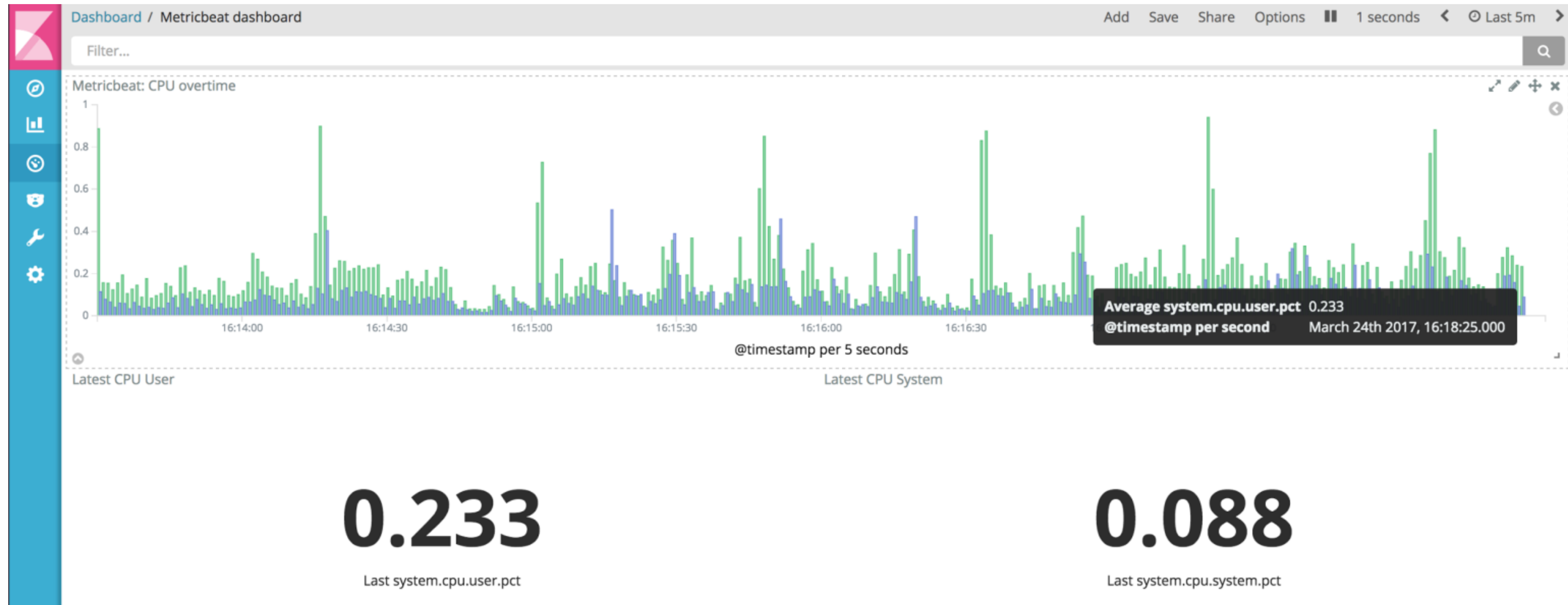
Uptime: 25 minutes  
JVM Heap: 7.41% (73MB / 990MB)

 elastic



# New Analytics

## Top Hits Aggregation (5.3) - Visualize the 'latest' metric



# Profile your Search Queries

## Search Profiler (5.1) - Detect and visualize bottlenecks in your query

Dev Tools

Console

Search Profiler

\_all

Index

Type

1

{

"query": {

"bool": {

"should": [

"match": {

"metric": "5"

}

},

"term": {

"node": {

"value": "1"

}

},

"terms": {

"query": [0,1,2]

}

},

"match": {

"title": "Quick brown

}

},

"match": {

"title": {

"query": "Quick bro

"fuzziness": 2

}

},

"bool": {

"should": [

"range": {

"hour": {

"lte": "2

}

},

"match": {

"title": "Fas

}

},

}

}

}

}

Profile

Query Profile

Aggregation Profile

Index: data

Cumulative Time: 30.290s

> [94Dq9uKuQSiITRnIYwYHKA][2]

6.176s

Type	Self Time	Total Time	% Time
BooleanQuery metric:[5 TO 5] node:[1 TO 1] query:{0 1...	3.0s	6.2s	100.00%
BooleanQuery hour:[-9223372036854775808 TO 9223372036...	1.7s	2.7s	42.99%
hour:[-9223372036854775808 TO 9223372036...	949.0ms	949.0ms	15.37%
BooleanQuery title:fast title:jumping title:spider ti...	0.1ms	1.6ms	0.03%
hour:[-9223372036854775808 TO 9223372036...	395.8ms	395.8ms	6.41%
metric:[5 TO 5]	75.6ms	75.6ms	1.22%
node:[1 TO 1]	49.5ms	49.5ms	0.80%
query:{0 1 2}	22.5ms	22.5ms	0.36%
BooleanQuery title:quick title:brown title:fox	0.2ms	3.1ms	0.05%
TermQuery title:quick	2.4ms	2.4ms	0.04%
TermQuery title:brown	0.3ms	0.3ms	0.00%
TermQuery title:fox	0.3ms	0.3ms	0.00%
BooleanQuery MatchNoDocsQuery("empty BooleanQuery") M...	0.1ms	0.1ms	0.00%

> [94Dq9uKuQSiITRnIYwYHKA][0]

6.164s

Type	Self Time	Total Time	% Time
BooleanQuery metric:[5 TO 5] node:[1 TO 1] query:{0 1...	2.9s	6.2s	100.00%
BooleanQuery hour:[-9223372036854775808 TO 9223372036...	1.8s	2.7s	44.09%
hour:[-9223372036854775808 TO 9223372036...	965.4ms	965.4ms	15.66%

data

[94Dq9uKuQSiITRnIYwYHKA][2]

Type

BooleanQuery

Description

hour:[-9223372036854775808 TO 9223372036854775807] (title:fast title:jumping title:spider title:eats title:small title:mice)

Total Time

2.655s

Self Time

1.705s

Timing Breakdown

advance	1.3s	50.4%
score	1.3s	49.5%
create_weight	1.6ms	0.1%
build_scorer	374.8µs	0.0%
next_doc	0.0ns	0.0%
match	0.0ns	0.0%

\* requires X-Pack (Basic)



# Grok Debugger

[illegible]

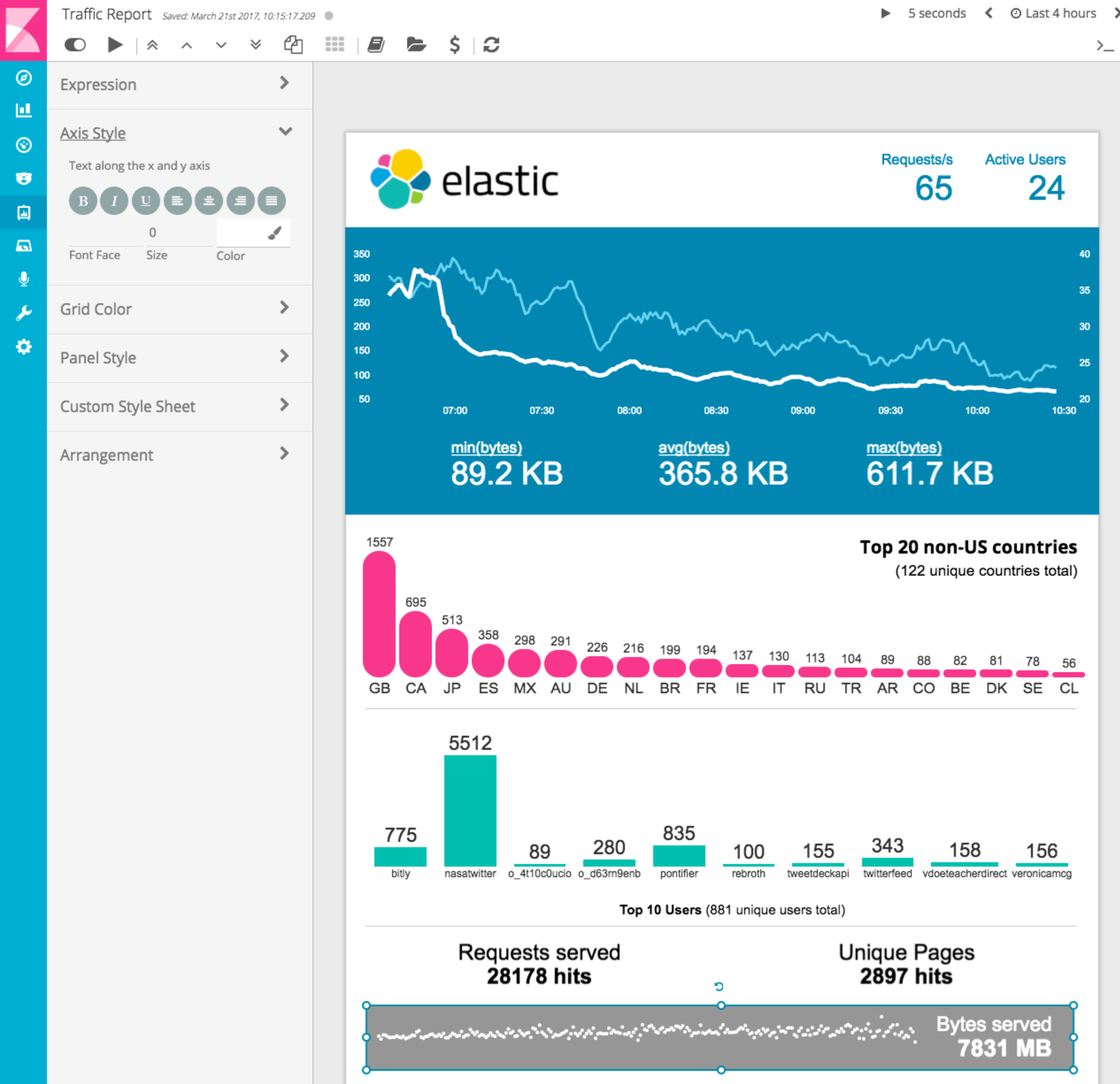


# Internationalization Support

I18N, phase 1 complete (5.2)

- Adheres to browser preference for language
- Translations as plugins
- Thanks IBM!





# Kibana Canvas

- New visualization application on top of Elasticsearch data
- Use Case:
  - live infographics
  - presentations with live data feeds
  - highly customized reports
- Currently, in the prototyping phase
- Release date: TBD



logstash

# Logstash 5.4 Features

## 1. Persistent Queues GA 🥳

## 1. GeoIP2 ISP & ASN Lookup Support

- Lookup ISP and ASN data with the Maxmind GeoIP2 ISP Database (separate license)
- GeoIP filter v4.1.1

## 1. Log4j2 Support (Log4j input plugin deprecated) - two options:

- Log4j2 SocketAppender sends JSON to Logstash TCP input
- Log4j2 logs to file for Filebeat to collect

# Persistent Queues

## Features

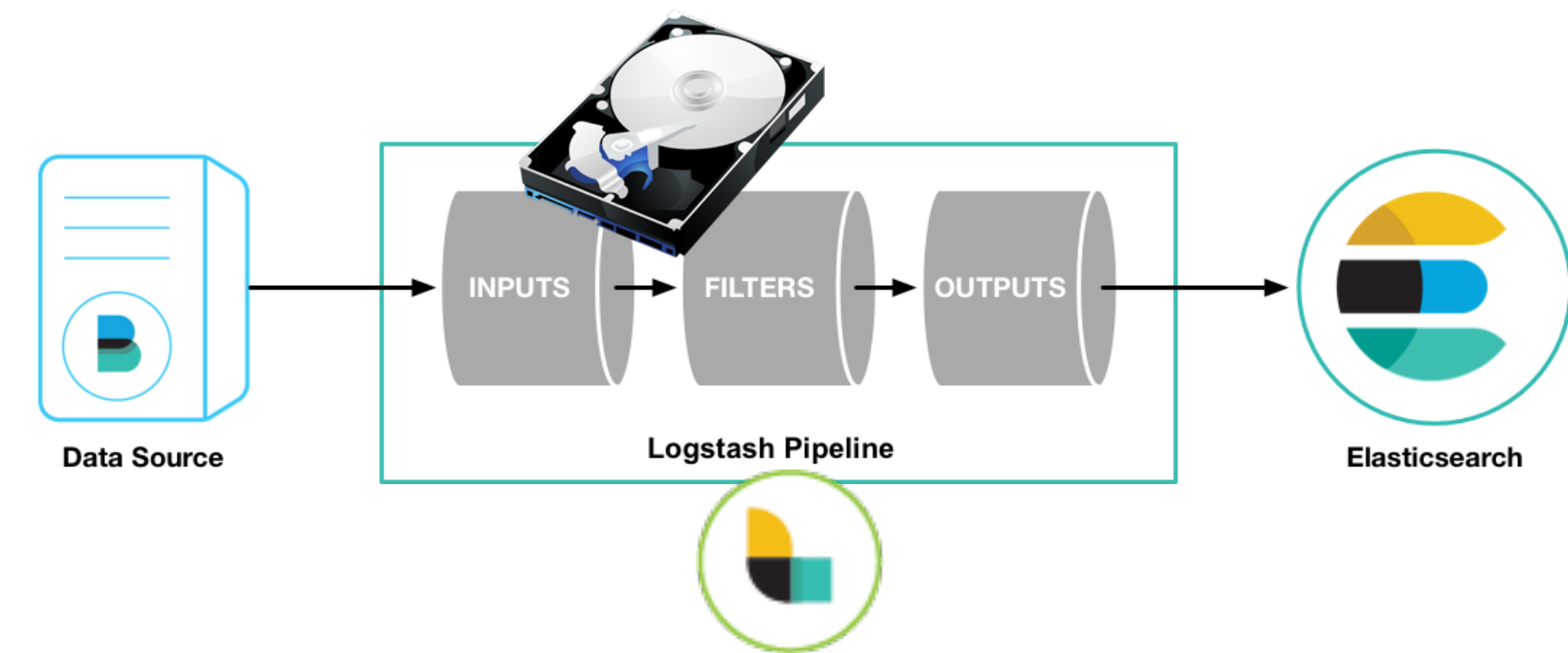
Disk-Based Queuing - GA in 5.4

### Resiliency

- Durability across node failures
- **At-least-once** delivery guarantees

### Adaptive buffering

- External queuing layer no longer required to absorb throughput



Control max disk usage

Limited impact on performance

### Monitoring UI integration

- Queue type and queue lag
- *Future: disk usage and disk IO*

Opt-in feature



# Persistent Queues

## Resiliency

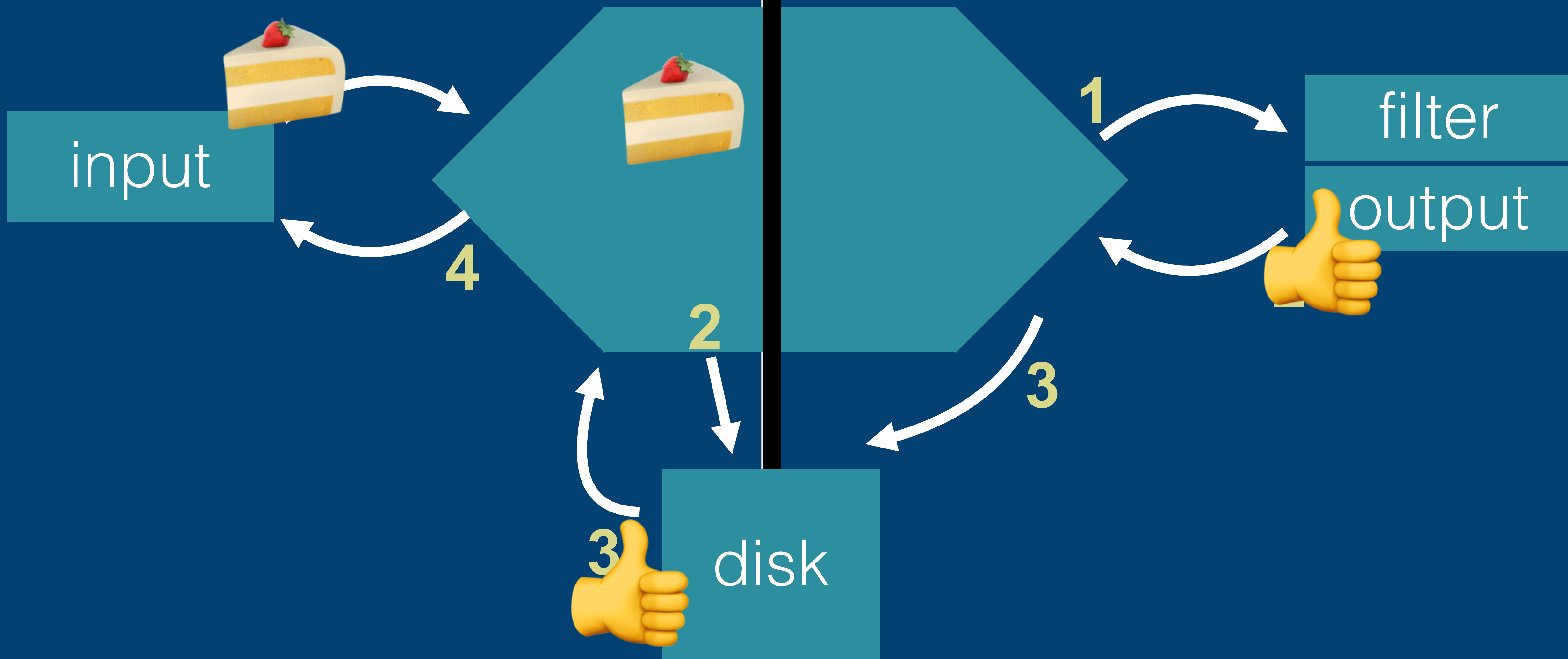
- Make sure `queue.checkpoint.writes` = 1 for at-least-once delivery guarantees
- PQs are resilient across node failures. For protection against disk-level corruption or failure, disk redundancy is recommended:
  - **On-Prem** – RAID
  - **AWS** – EBS (i.e. t2, m4, c4, p2)
  - **GCP** – Persistent Disk
  - **Azure** – Premium / Standard Managed Disks (ZRS or higher for multi-DC redundancy)
- Shared network filesystems (CEPH, GlusterFS, NFS, etc.) are **not** recommended.
- Inputs that do not support acknowledgements (i.e. TCP, UDP) can still lose data.



# Persistent Queues

## Things To Know

- Understand your hardware
  - Know your available disk capacity needs
  - Monitoring UI can help monitor PQs
  - SSDs are **not** required
- PQs should be enabled when ingesting from Kafka
  - Resilient transport from Kafka to ES
  - PQs mitigate the need for expensive reprocessing on recovery
  - Recommend default 1GB PQ max disk size
- For ephemeral storage, multi-DC queue replication, or centralized data replay:
  - Add or leverage an HA Kafka cluster



# Centralised Management

- Elasticsearch as a remote config store
- Manage configurations via UI
- Group multiple Logstash under roles
- Simple alternative to puppet, chef

# Offline Plugin Management (5.2)

## Air-gapped Networks and Offline Environments

### Prepare and Pack Plugins on Staging Box

```
$ bin/logstash-plugin prepare-offline-pack logstash-filter-* logstash-input-beats
```

### Move Offline Pack to Offline Boxes

- Default pack location: `/LOGSTASH_HOME/logstash-offline-plugins-5.2.0.zip`
- Change pack location using `--output /path/to/pack` parameter

### Install or Update Plugins

```
$ bin/logstash-plugin install file:///path/to/logstash-offline-plugins-{logstash_version}.zip
```



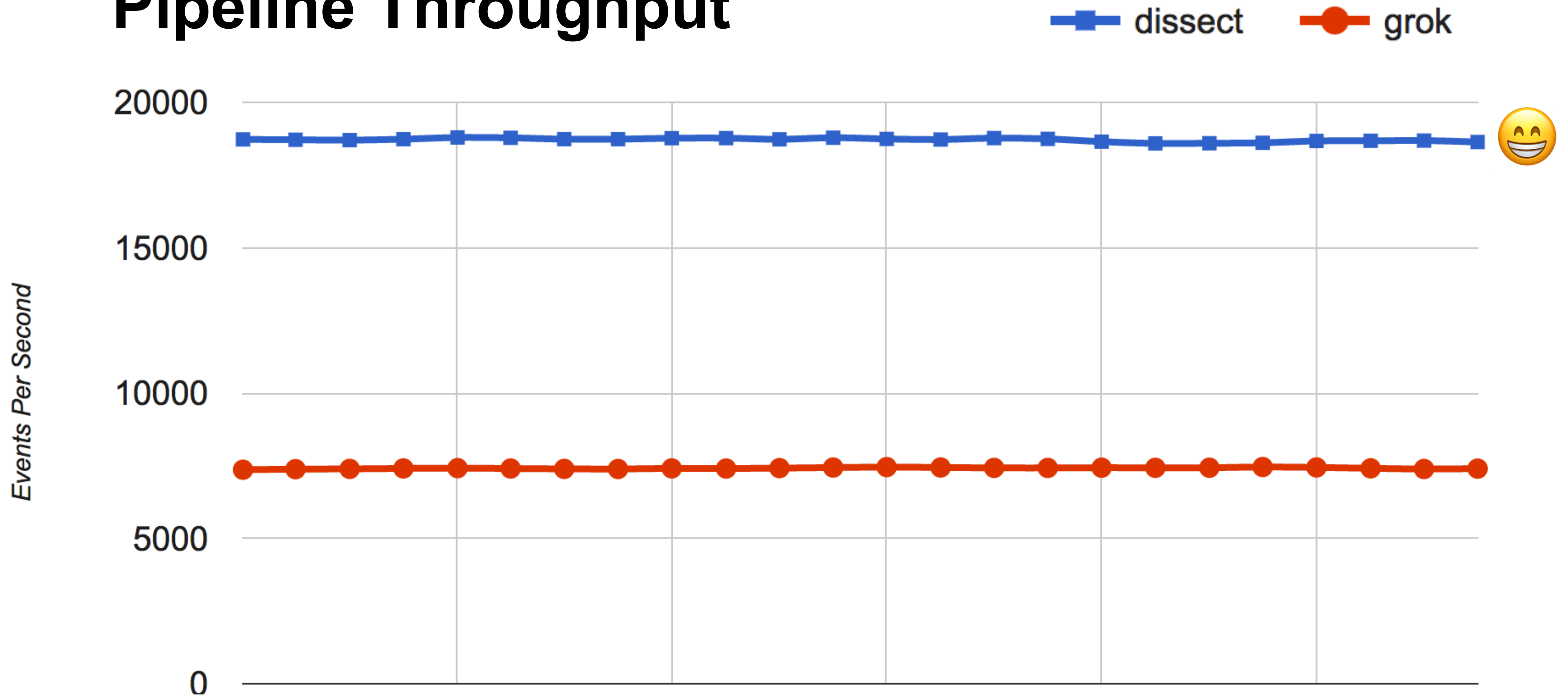
# Database Lookup Enrichment

## JDBC\_streaming filter (5.3)

- Enrich Logstash events with DB data (streaming joins)
- Executes JDBC lookup queries per event (add one or more fields)

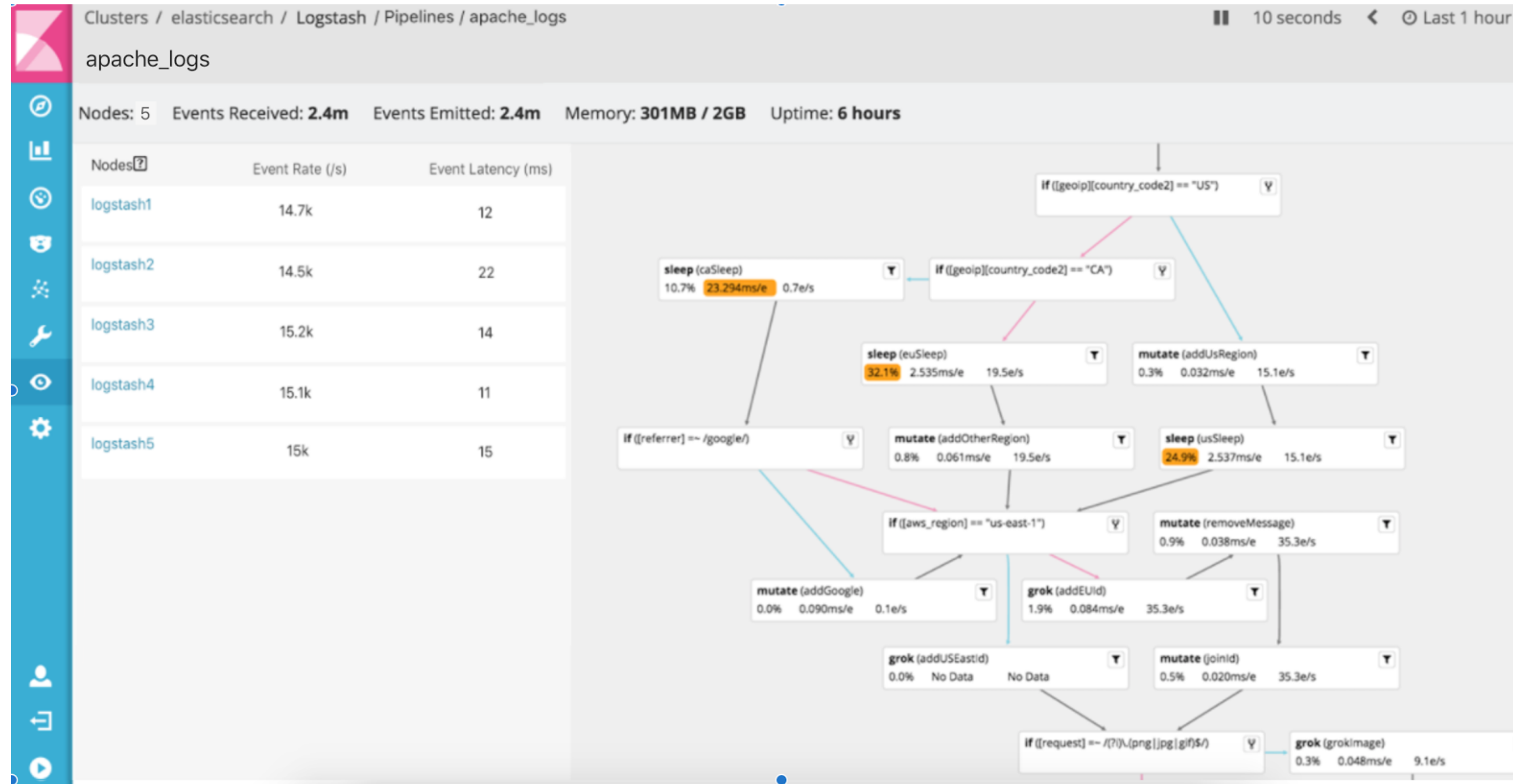
```
jdbc_streaming {  
  jdbc_connection_string => "jdbc:mysql://localhost:3306/mydatabase"  
  statement => "select * from PRODUCTS.FRUIT WHERE SKU = :sku"  
  parameters => { "sku" => "sku_code"  
  }  
}
```

# Pipeline Throughput



# More

- Logstash modules (5.5)
- Dead letter queue
- Multi Logstash pipelines
- Logstash Intermediate Representation (LIR)



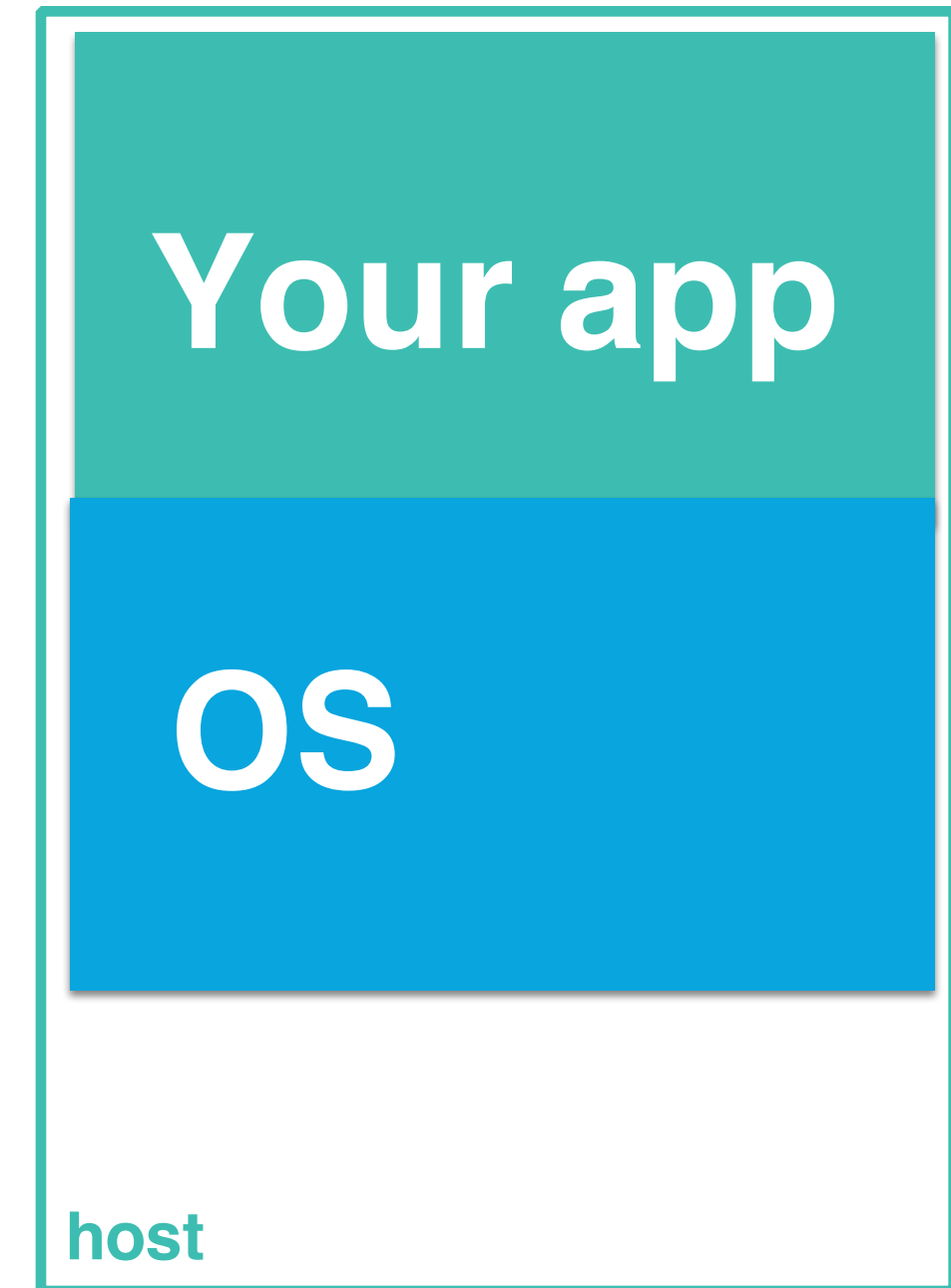
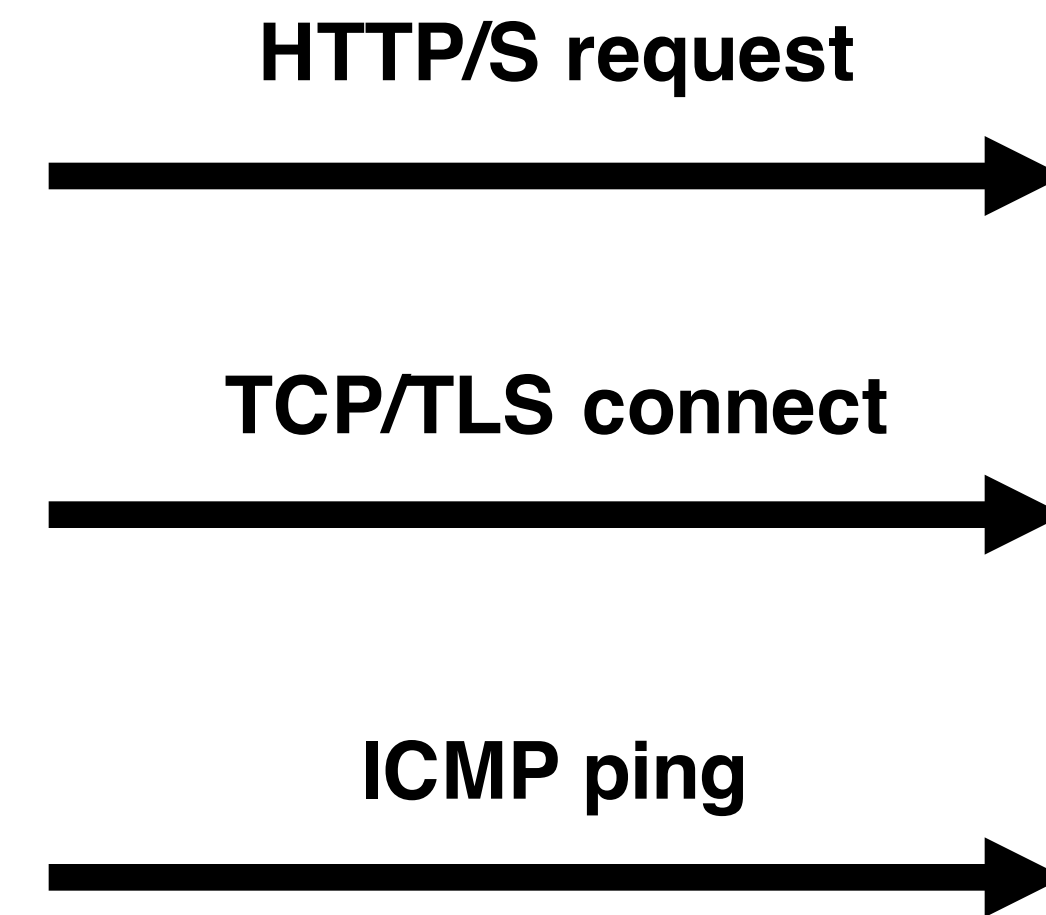


beats



# New Beat: Heartbeat (beta in 5.2)

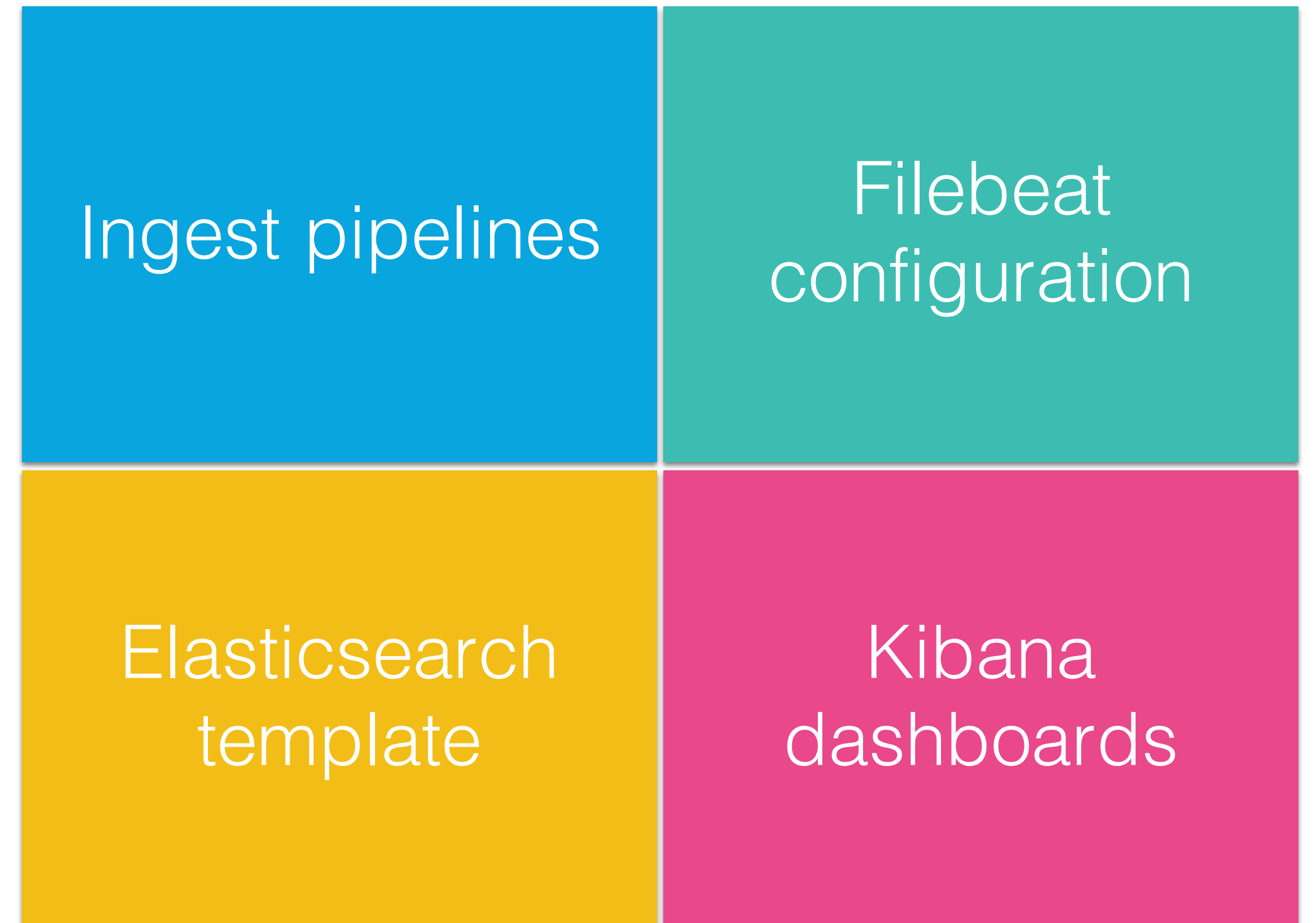
- Ping all the things
- Gather round trip metrics
- Many to many
- Ping IPs behind load balancers



# Filebeat Modules (5.3)

They are like Metricbeat modules...only different

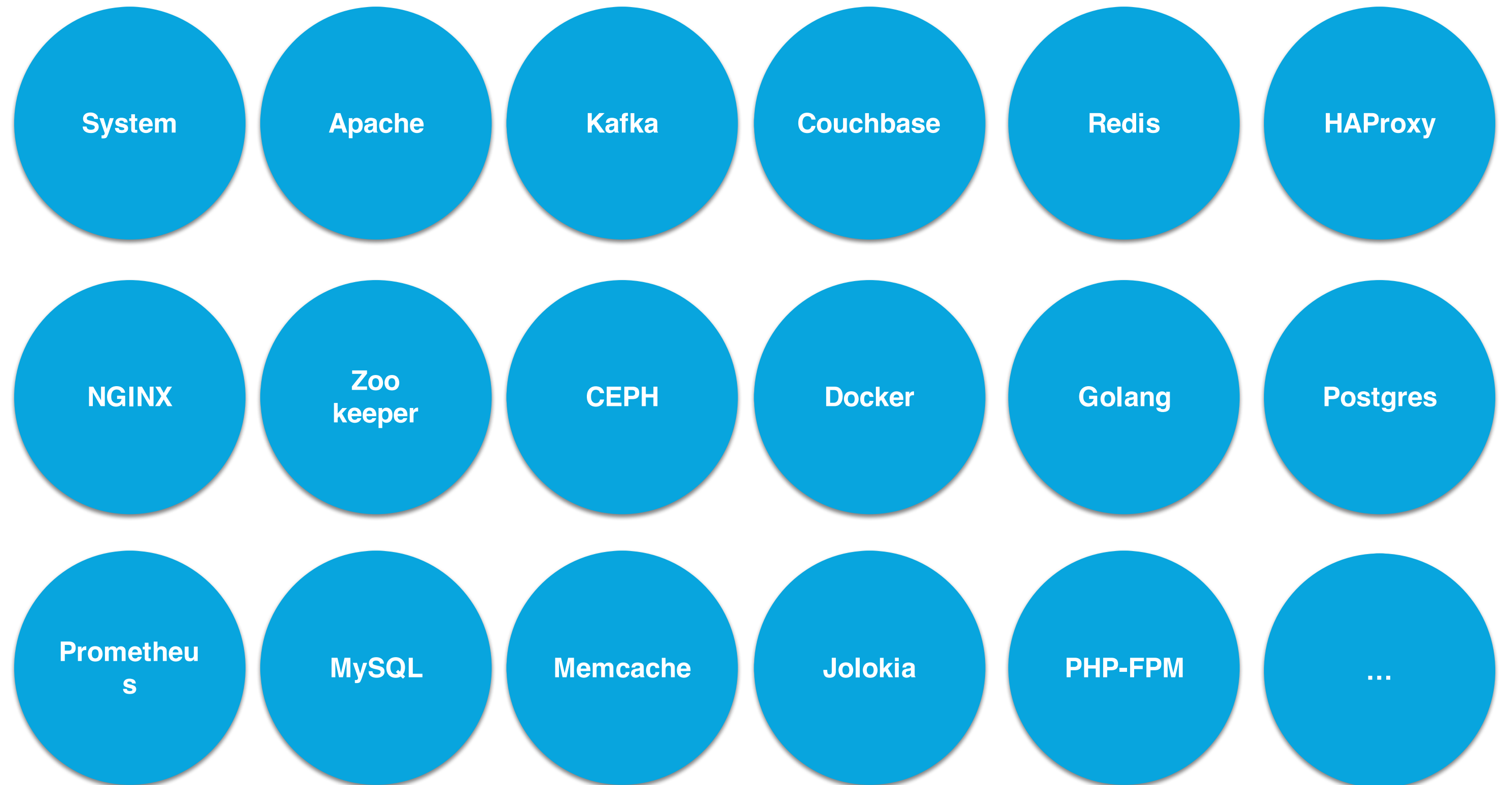
- Because simple things should be simple
- Prepackaged configs for common log formats
- Get to a dashboard in minutes
- First release includes Apache, Nginx, MySQL, system modules. More to come.



# More Modules in Metricbeat

## Modules are growing

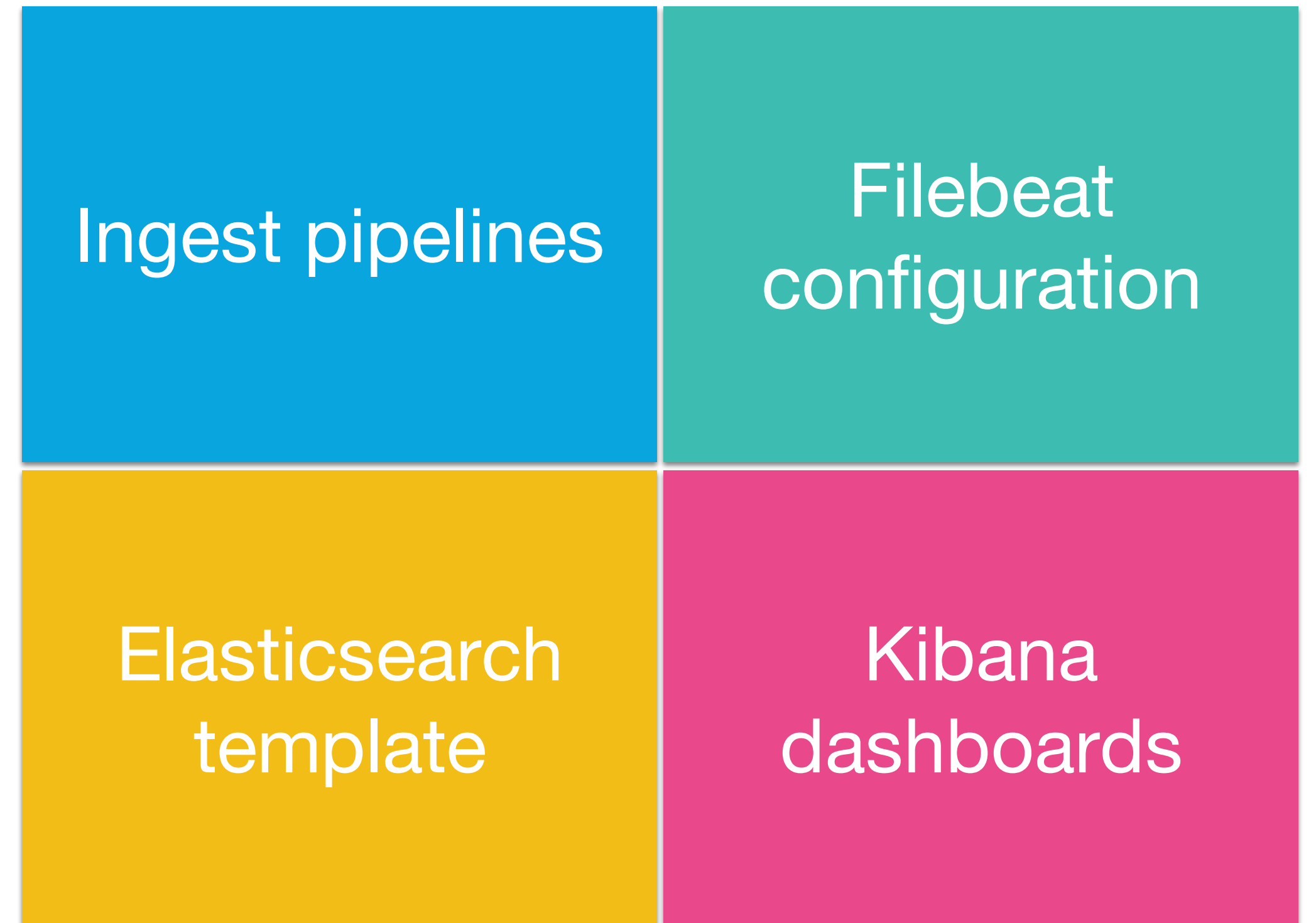
- Docker
- Kubernetes
- Kafka
- Prometheus
- Elasticsearch
- Kibana
- vSphere
- RabbitMQ
- ... ..



# JMX Monitoring

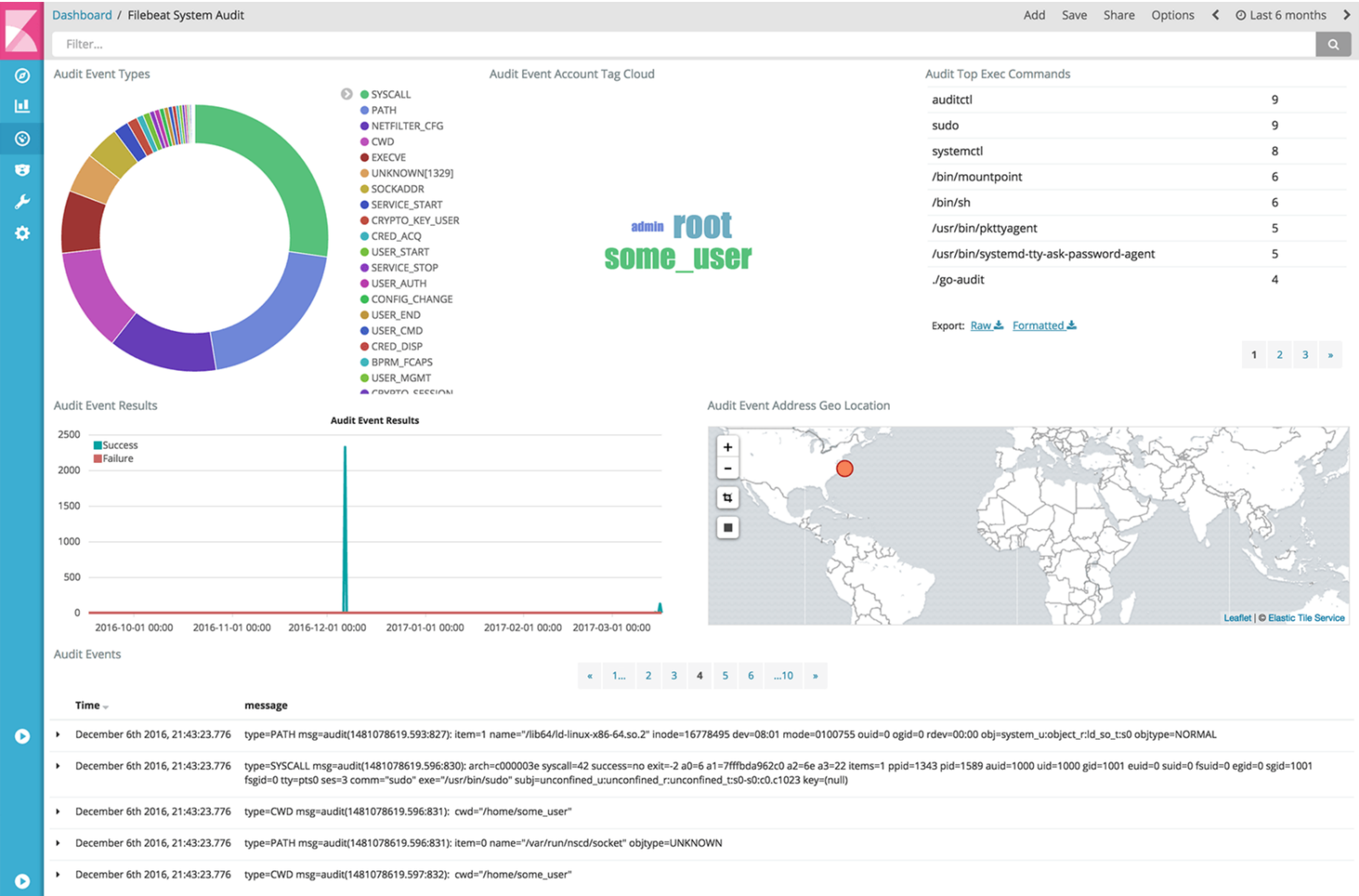
Modules make the common tasks easy.

- Metricbeat 5.4 Jolokia Module
- Jolokia provides REST-like access to JMX with JSON over HTTP.

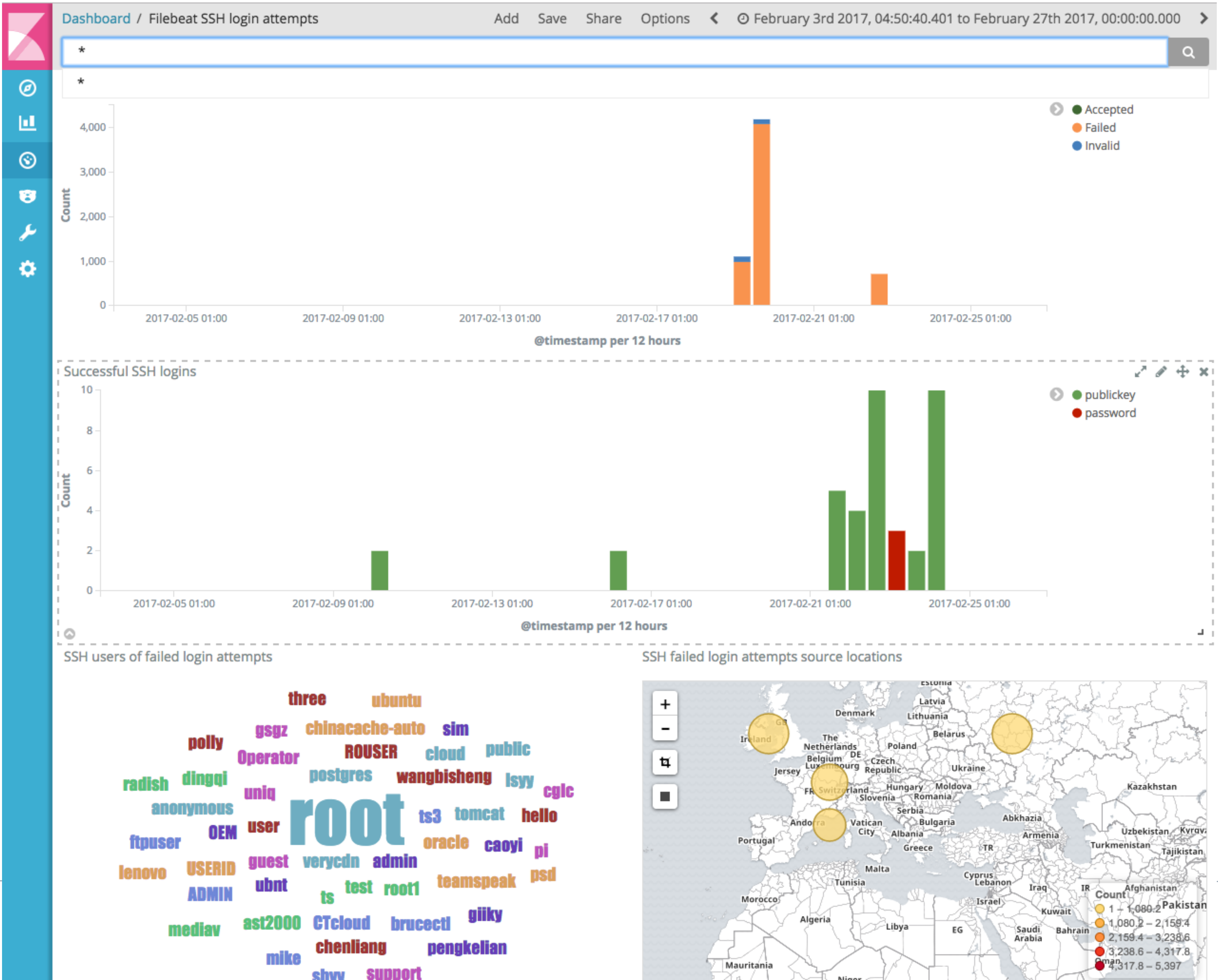




# Linux auditd Filebeat Module



# Linux system authentication logs





# Elastic & Community

- 中文权威指南已上线！
- References 翻译已启动！
- 欢迎加入！
- QQ群：109764489



Thanks !