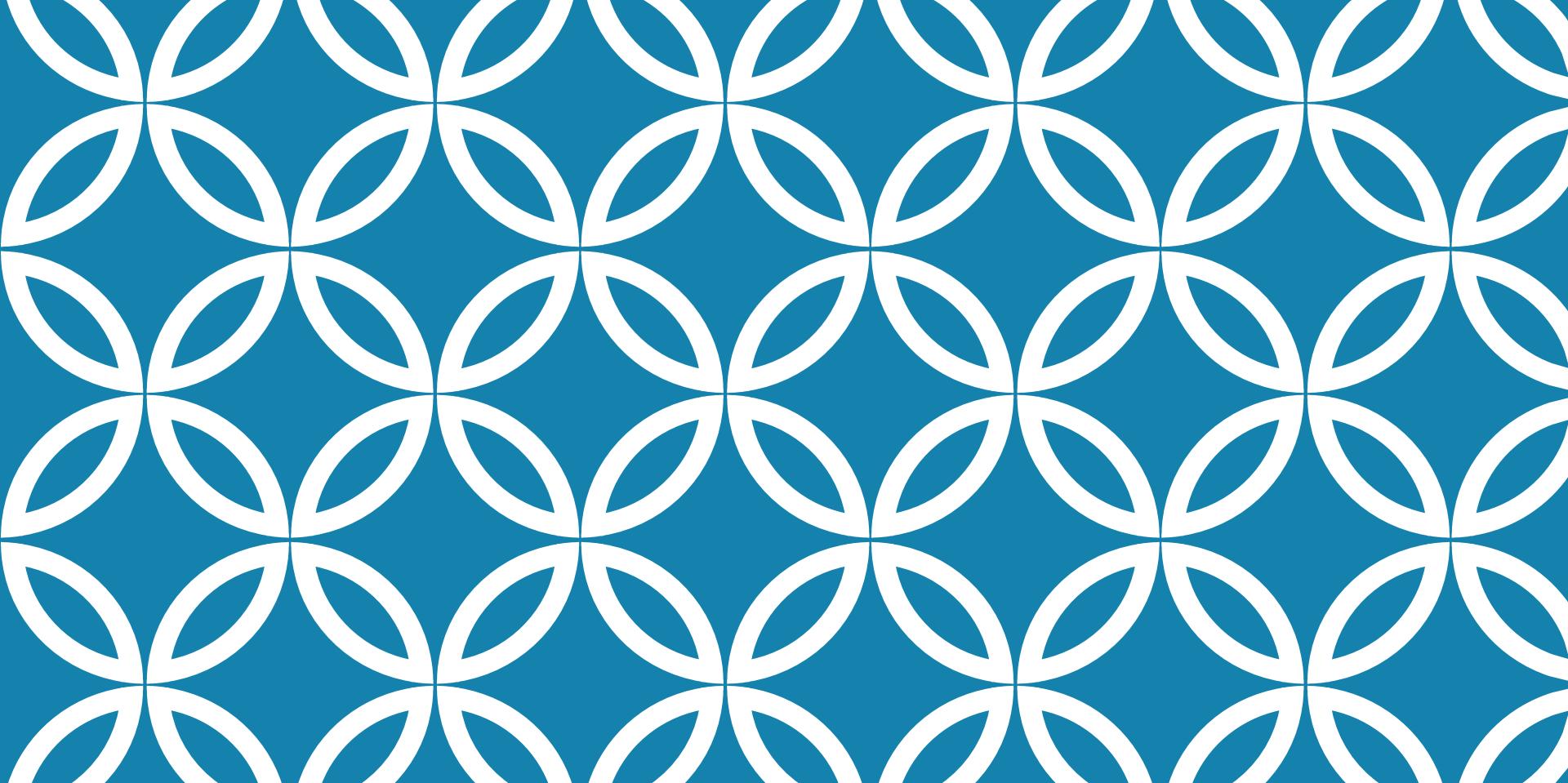


EUROPEAN 2016: TUTORIAL

Vašek Lorenc



2ND PART

GrayLog2 + ELK

CONTENTS

Open source solutions

- GrayLog2
- Elastic Stack

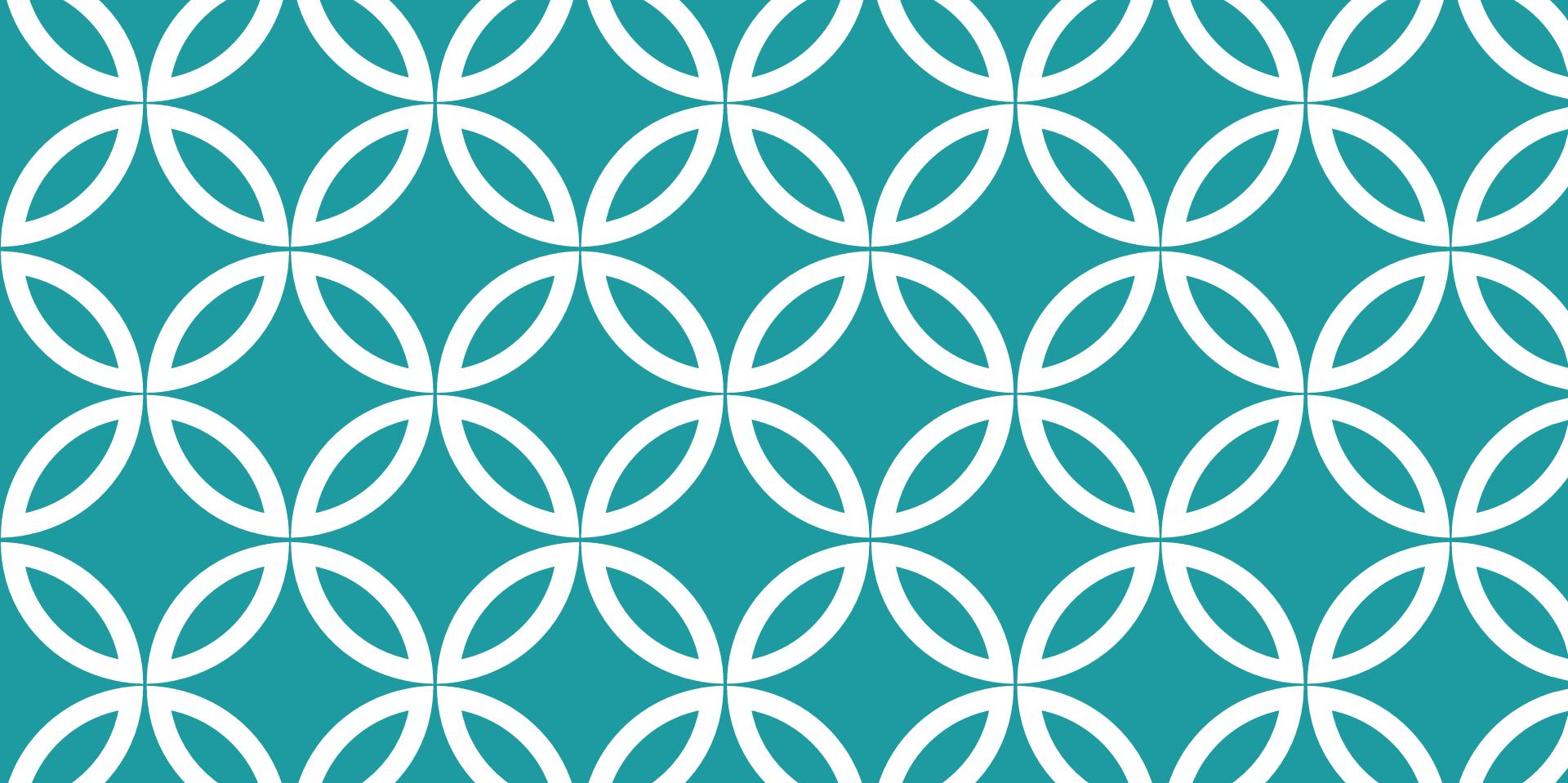
Hands-on Examples

- Cisco (ASA) logy
- Twitter API
- E-mails processing (IMAP)
- VirusShare/APTNotes

Further sources of (security) data

Related tools

Q&A



GRAYLOG2

Brief Introduction

GRAYLOG2

VM:

[graylog-box.ova](#)

Accesses:

admin/admin (web)

europen/europen (ssh)

Ports on your host:

8081/tcp (web)

12201/{udp,tcp}

5044/{udp,tcp}

2222/tcp (ssh)

Description

Inputs

Extractors

Dashboards

Streams

Alerts



Search in the last 8 hours

Saved searches

gl2_source_input:563bec4be4b09abf23ab88b0

Search result

Found 64 messages in 61 ms, searched in 1 index.

Fields

- facility
- level
- message
- source

[List fields of current page or all fields.](#)

Histogram

Year, Quarter, Month, Week, Day, Hour, Minute

12:00 15:00 18:00

Messages

Previous Next

Timestamp	source
2015-11-06 18:35:33.000	graylog graylog ntpd_intres[891]: host name not found: 3.pool.ntp.org
2015-11-06 18:35:33.000	graylog graylog ntpd_intres[891]: host name not found: 1.pool.ntp.org
2015-11-06 18:35:33.000	graylog graylog ntpd_intres[891]: host name not found: 2.pool.ntp.org
2015-11-06 18:35:33.000	graylog

BASIC USER INTERFACE

All-In-One

GRAYLOG2: BRIEF INTRODUCTION

Easy-to-use Solution (just like Splunk)

- Powerful search capabilities + syntax completion
- Dashboards
- Authentication, roles
- Alerting, notifications(Slack, JIRA, PagerDuty, ...)
- Company-wide deployment scenarios, documented
- Scalable
- Plenty of various input modules, easy setup
- API

Drawbacks

- Aggregation functions missing

EXAMPLE: CISCO ASA

Input configuration

- Syslog TCP, port 5044

Extractors

Data processing configuration

- (See the next picture)

Searching

- Do your own searches!

Dashboard

- First, create an empty dashboard, second add widgets from searches

Streams + Alerting

Example message

```
Sep 23 00:09:00 europen-fw001.brno.company.com %ASA-6-106102: access-list brno_olomouc_vpn_filter permitted udp for user '<unknown>' outside/147.1.55.233(30467) -> core_conduit/147.16.55.54(161) hit-cnt 1 first hit [0x9c5cadac, 0x0]
```

Wrong example? You can [load another message](#).

Extractor configuration

Extractor type Grok pattern

Source field message

Named captures only

Only put the explicitly named captures into the message.

Grok pattern

```
%{SYSLOGTIMESTAMP:timestamp} (%:{SYSLOGFACILITY} )%{SYSLOGHOST:logsource} %%{SYSLC
```

Try

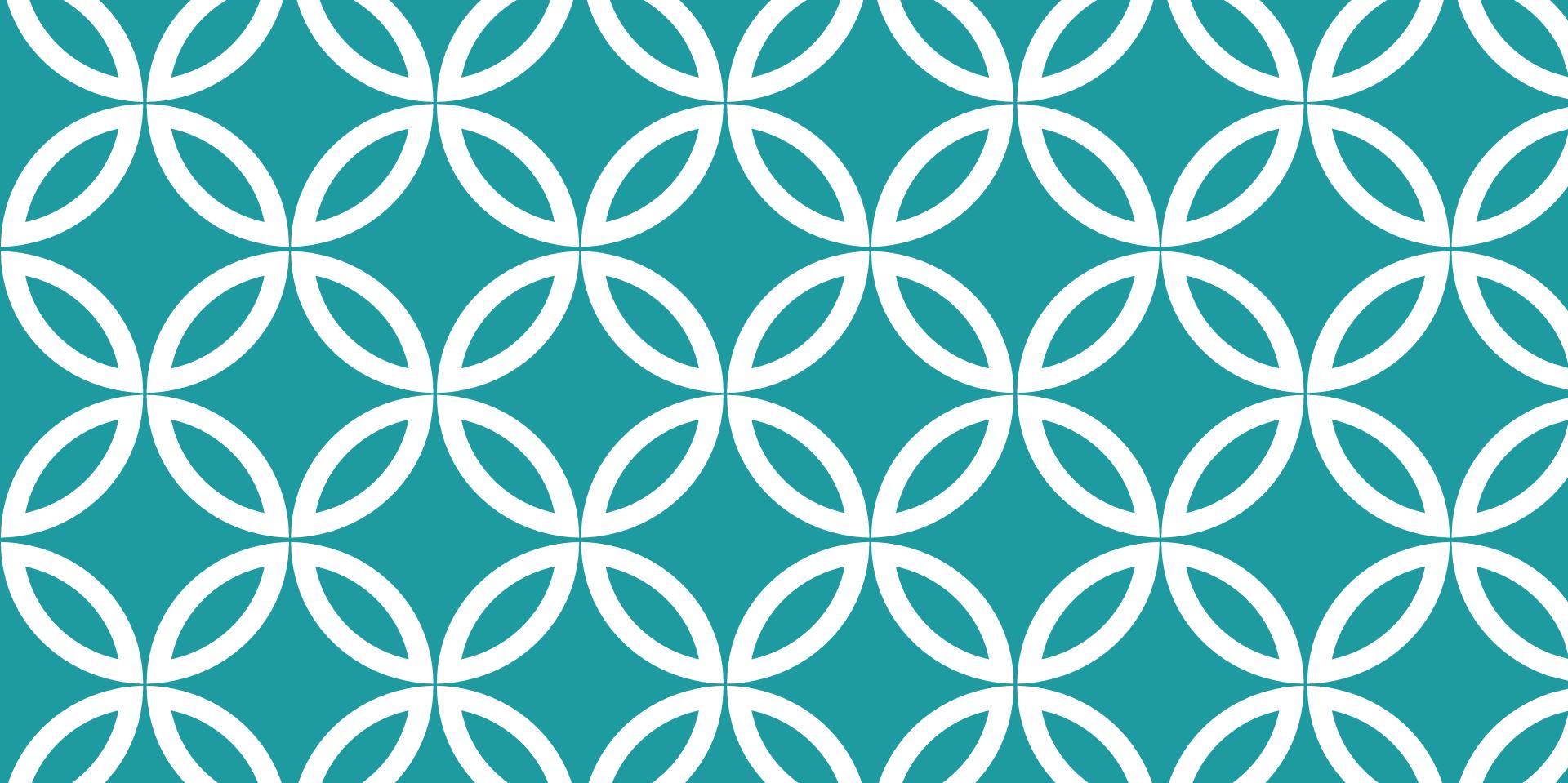
Matches the field against the current Grok pattern list, use **%{PATTERN-NAME}** to refer to a [stored pattern](#).

Extractor preview

```
aclname  
brno_olomouc_vpn_filter  
ciscotag  
ASA-6-106102  
logsource  
europen-fw001.brno.company.com  
program  
ASA-6-106102  
timestamp  
Sep 23 00:09:00
```

CREATING AN EXTRACTOR

... in Graylog



ELASTIC STACK

Not-so-brief Introduction

ELASTIC STACK

VM:

[vagrant-elk-box.ova](#)

Accesses:

european/european (ssh)

Ports on your host:

5601/tcp (kibana)

9200/tcp (elasticsearch)

2222/tcp (ssh)

Versions installed:

Elasticsearch 2.3.5

Kibana 4.5.0

Logstash 2.4

Elasticsearch

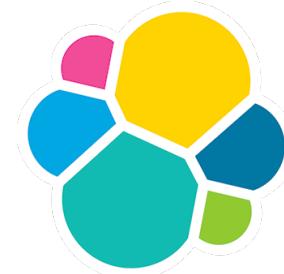
- Capabilities, options, indexes
- Data types
- Data mapping, templates

Logstash

- Samples of various inputs and configs

Kibana

- Visualizations (simple or powerful)
- Additional applications



elastic

{ELASTIC STACK, ELK, ELK STACK}

Elasticsearch + ...

Elastisearch

- Data processing and indexing
- Apache Lucene
- Full-text
- Document database
- JSON

Beats

- Input plugins (Filebeat, ...)

... Logstash + Kibana

Logstash

- Powerful tool for data transformations
- Filtering, input and output plugins

Kibana

- Visualization framework
- Plugins available
- Whole app platform (Kibana apps)

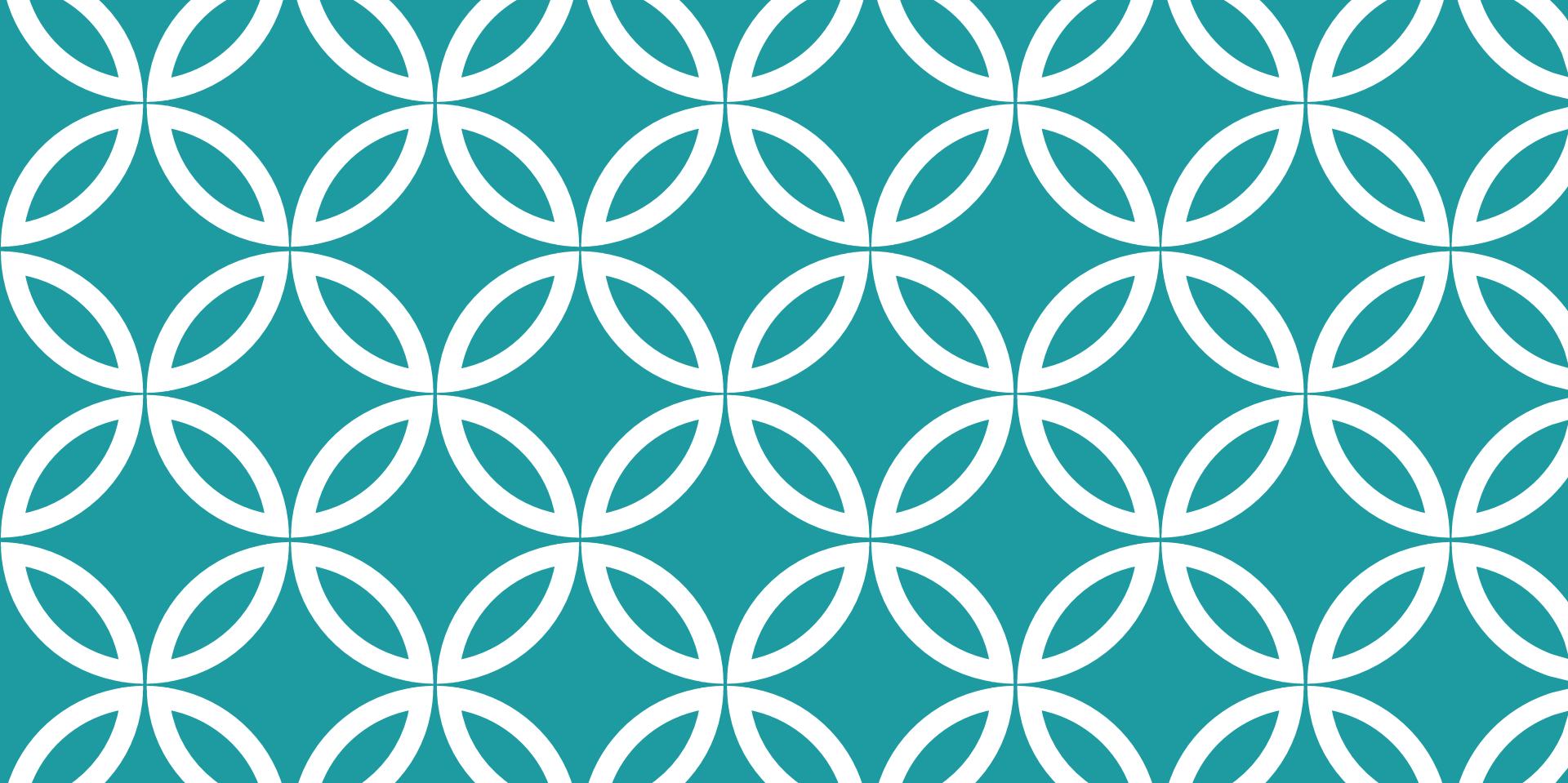
ELK VS. {GRAYLOG2, SPLUNK}

What may bother?

- No auth and roles by default
- No alerting
- No cross-joins and lookups
- Complicated to set up
- Rapid development

What's nice?

- Price
- Flexibility
- Logstash included
- Detailed data management
- Rapid development
- Community
- Similarity search(!)
- REST API



LOGSTASH

Transforming Data

LOGSTASH: BASIC WORKFLOW

1. INPUT

Variety of input plugins and codecs

2. FILTER

Programming language, in fact (Ruby)

With all the benefits and drawbacks

“grok is the new grep”

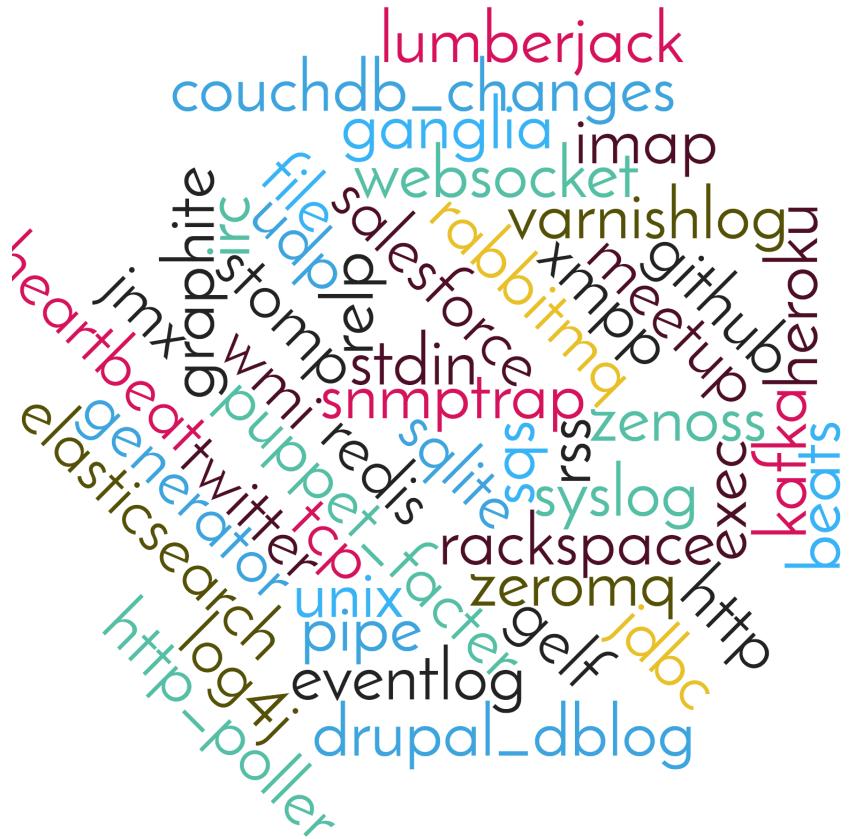
3. OUTPUT

Variety of output plugins again

We need just one for Elastic output though...

LOGSTASH: INPUT

imap
file
stdin
tcp
twitter
...



A word cloud visualization showing various Logstash input plugin names. The words are arranged in a cluster, with larger and more prominent words indicating higher frequency or importance. The colors of the words vary randomly.

lumberjack
couchdb_changes
ganglia
imap
websocket
varnishlog
github
xmpp
meetup
salesforce
rabbitmq
kafka
heroku
beats
graphite
file
udp
stomp
stdin
jmx
wmi
redis
snmptrap
salite
bs
rss
zenoss
syslog
exec
stomp_src
tcp
puppet
tcp
fact
rackspace
zeromq
http
jdbc
generator
heartbeat
twitter
unix
pipe
gelf
eventlog
elasticsearch
log4j
drupal_dblog
http_poller

LOGSTASH: FILTER

grok

mutate

cidr

geoip

csv

...

A word cloud diagram centered around the word "environment". The words are colored in various shades of blue, green, yellow, red, and purple. The size of each word represents its frequency or importance within the Logstash filter ecosystem. The words include: grok, mutate, cidr, geoip, csv, ..., environment, extractnumbers, elapsed, anonymize, urldecode, cipher, geoip, environment, aggregate, checksum, throttle, json_encode, json_decode, split, de_dot, dns, ilog, log, sleep, prune, multiline, range, event, timestamp, useragent, cidr, grok, alter, metrics, drop, clone, date, xml, translate, fingerprint, collate, elasticsearch.

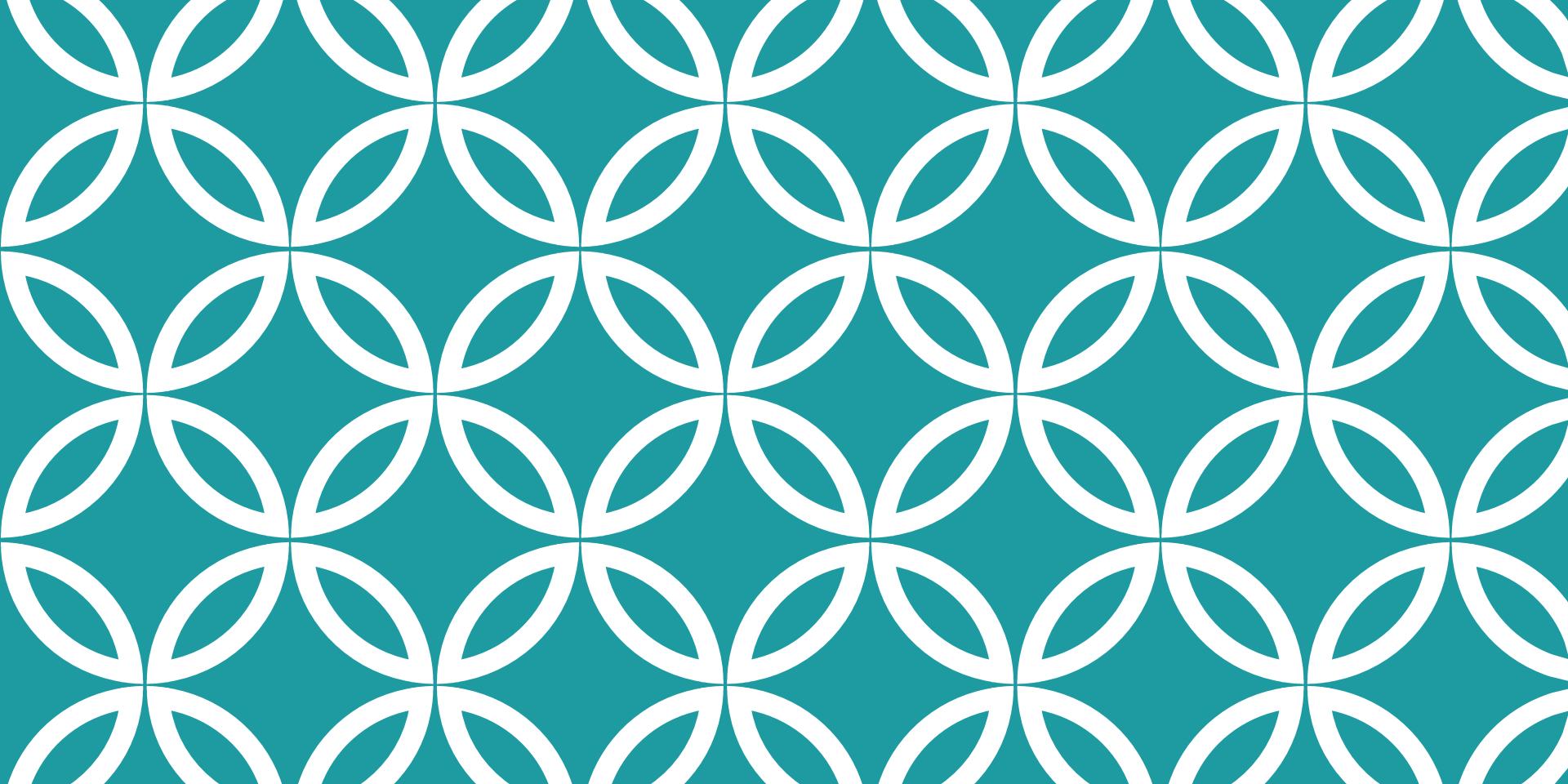
JSON

A scene from Toy Story featuring Woody and Buzz Lightyear. Woody, on the left, has a concerned expression and is looking towards the right. Buzz, on the right, is in his signature green space ranger suit with purple highlights, including a purple headband and purple nail polish. He is waving his right hand, which features a purple ring on the middle finger. The background shows a room with a chalkboard and some toys.

JSON EVERYWHERE

JSON EVERYWHERE

“Deal with it”



KIBANA

Showing Data
Versions 3, v4, v5, ...

KIBANA: VISUALIZATION FRAMEWORK

UI for queries over the data

Visualization

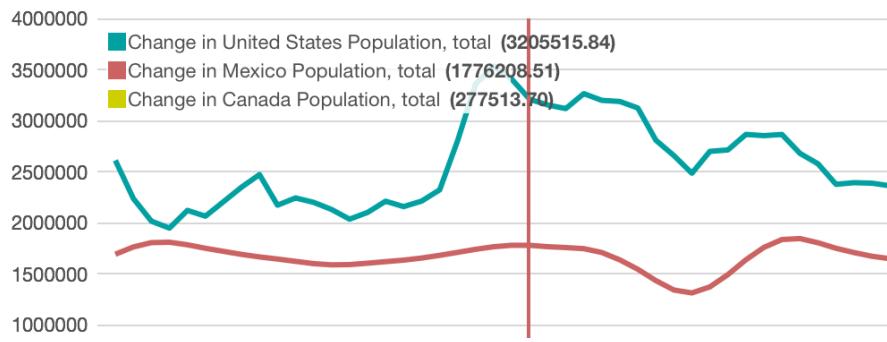
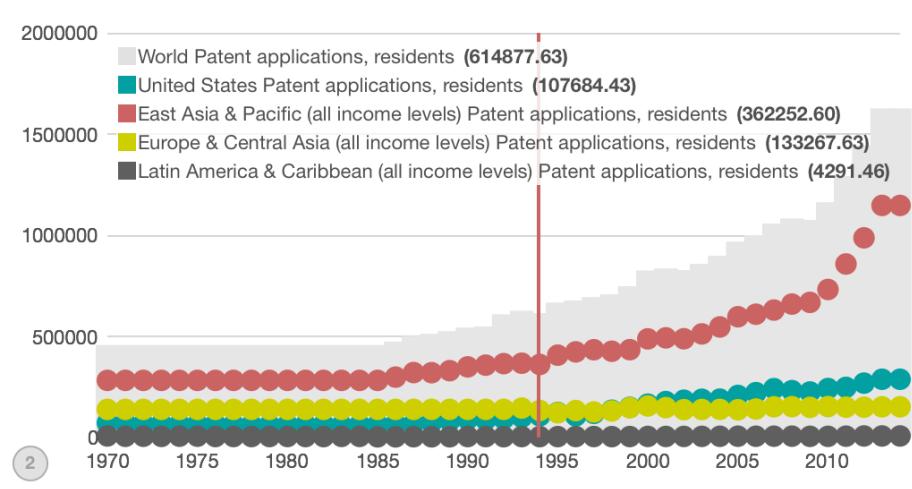
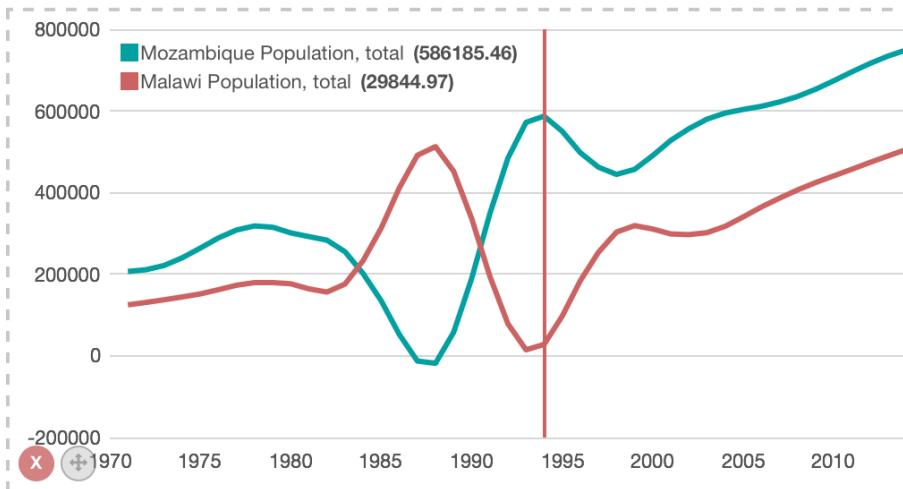
- Variety of visualization plugins
- Some of the non-standard were added to the VM prepared for this tutorial

Application Platform

- Running on node.js, you can write your own plugins and apps

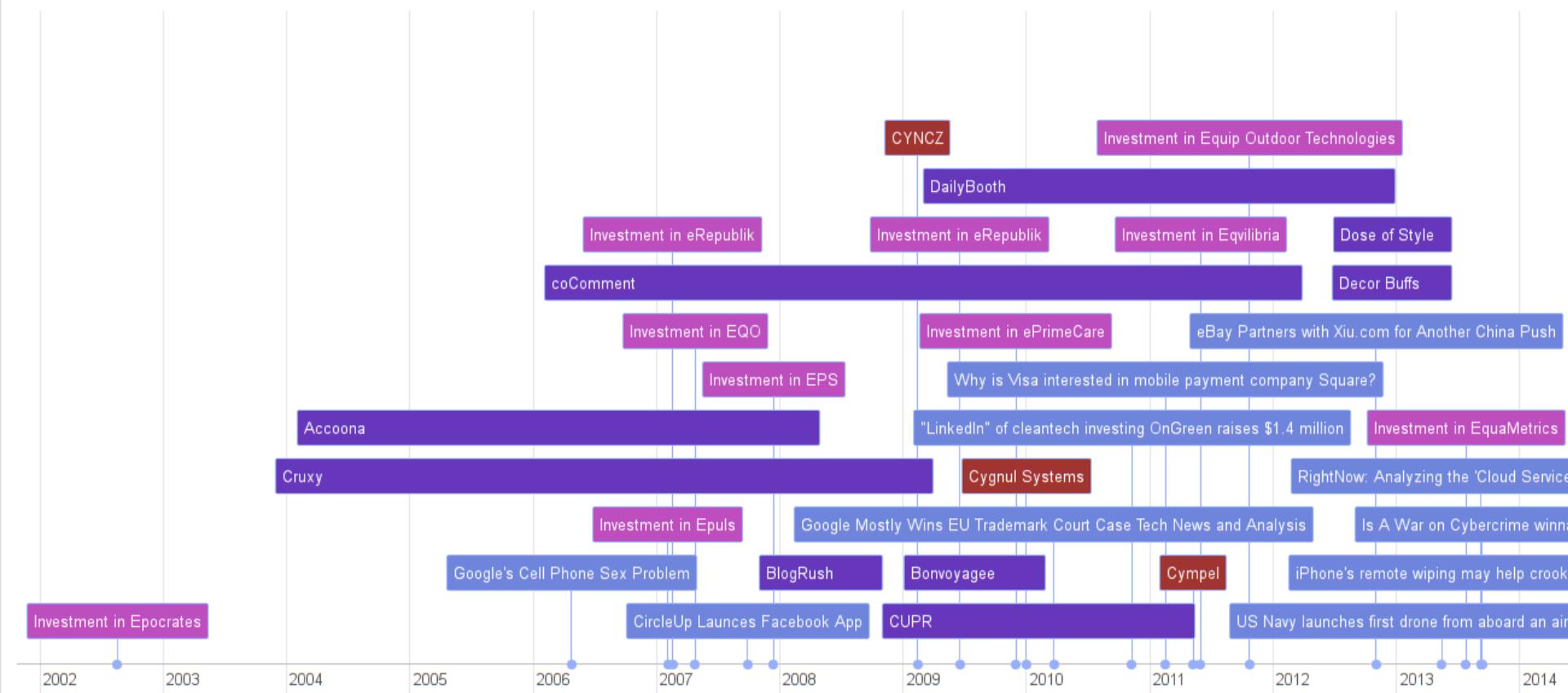
Pre-installed Apps

- Kibana
- Timelion
- Sense
- KAAE (Kibana Alert & Report App for Elastic)



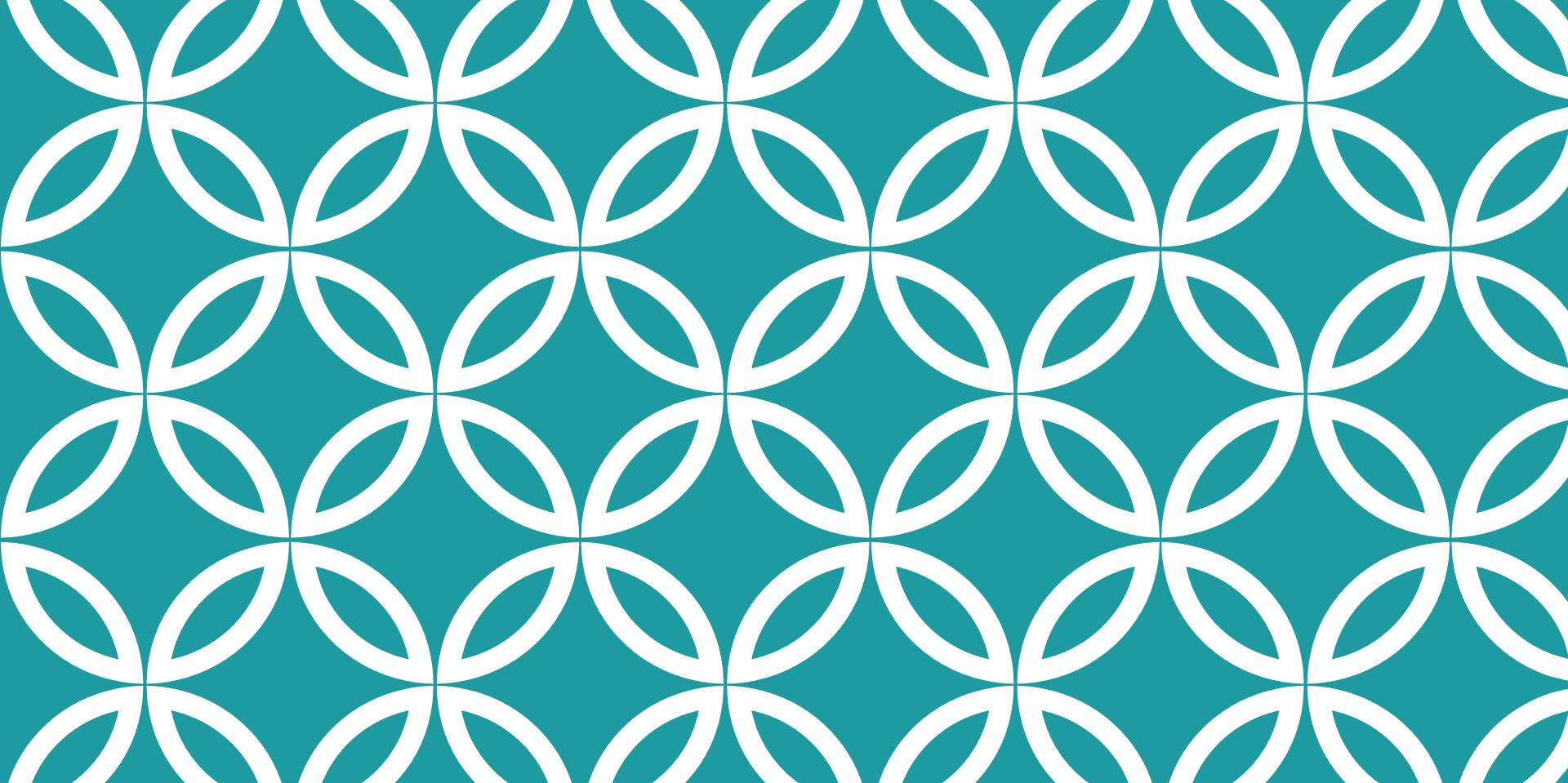
TIMELION

Complex Visualization
Example



KIBI TIMELINE

Complex Visualization
Example



HANDS-ON EXAMPLES!

Elastic Stack

WORKFLOW FOR EACH EXAMPLE

1. Let's check the raw input data
2. Find a name for the Elasticsearch index
3. Try to configure Logstash
4. Configure Kibana for our index
5. Searching, visualization
6. (*Troubleshooting*)
7. Fixing all issues found, goto 2
8. Another example is waiting!

SYSLOG: CISCO LOGS IN ELASTIC STACK

The data processed in Graylog, now on ELK

Leveraging Logstash

- Configuration of input, filter & output sections
- Grok filter & debugger

Visualization

Tweaking of Logstash configuration files

- Same config can be used for output to Graylog and ELK!

LOGSTASH-CISCO-ELK.CONF (-GELF.CONF)

```
input {
    tcp {
        port => 5602
        type => europeninput
    }
}

filter {

}

output {
    elasticsearch { }
}
```

```
Oct 07 00:00:23 %ASA-6-106102: access-list brno.olomouc.vpn.filter permitted tcp for user '<unknown>' outside/12.130.60.5(35335) -> inside/10.174.96.92(5666) hit-cnt 1 first hit [0x9c5cadac, 0x0]
```

```
%{CISCOTIMESTAMP} %{CISCOTAG}: access-list %{WORD:policy_id} %{CISCO_ACTION:action} %{WORD:protocol} for user \%
{DATA:src_fwuser}\' %{DATA:src_interface}/%{IP:src_ip}\(%{INT:src_port}\)(\(%{DATA:src_fwuser}\'))? -> %{DATA:dst_interface}/%
{IP:dst_ip}\(%{INT:dst_port}\) hit-cnt %{INT:hit_count} %{CISCO_INTERVAL:interval} \[%{DATA:hashcode1}, %{DATA:hashcode2}\]
```

Add custom patterns Keep Empty Captures Named Captures Only Singles

Autocomplete

Go

One per line, the syntax for a grok pattern is [%{SYNTAX:SEMANTIC}](#)

Custom patterns

GROK DEBUGGER

Making your life easier a bit

GROK CONFIGURATION

```
filter {
    grok {
        match => ["message", "%{CISCOTIMESTAMP:timestamp}
%{SYSLOGHOST:sysloghost} %{CISCOTAG:ciscotag}:
%{GREEDYDATA: cisco_message}"]
    }

    grok {
        match => [
            "cisco_message", "%{CISCOFW106100_2_3}",
            "cisco_message", "%{CISCOFW106001}", ]
    }
}
```

API: TWITTER FEEDS PROCESSING

Working with live data

Logstash input module for Twitter

- If you have your own Twitter, you can use it
- Processing tweets with europeen, europeen2016 hashtags

Searching, Visualization

Sense Application

- Direct access to Elasticsearch REST API

Detailed insight into ELK

- Data types of Elaticsearch
- Mapping, templates

```
1 # Delete all data in the `website` index
2 DELETE /website
3
4 # Create a document with ID 123
5 PUT /website/blog/123
6 {
7   "title": "My first blog entry",
8   "text": "Just trying this out...",
9   "date": "2014/01/01"
10 }
11
12 # Search!
13 GET website/_search
14 {
15   "query": {
16     "match": {
17       "title": "blog"
18     }
19   }
20 }
21
22
23
24 # Delete all data in the `website` index
25 DELETE /website
26
```

1

SENSE APPLICATION UI

REST API exposed

ELASTICSEARCH: {DATA TYPES, MAPPING}

1. Basic data types

1. string
2. long, integer, short, byte, double, float
3. {date, boolean, binary}

2. Complex types

1. object, nested
2. *No arrays? – each field can contain 0, 1 or more items by default.*

3. And there's more!

1. {geo_point, geo_shape, ip}
2. Support for IPv4 now, IPv6 as strings only (will be implemented in 5.x)
3. multifields (analyzed, not_analyzed indexes)

ELASTICSEARCH: {DATA TYPES, MAPPING}

1. Dynamic mapping

1. Automatic data types guessed by Elasticsearch

2. Static mapping

1. That's what we are up to now!
2. Fixing Elasticsearch's wrong guesses
3. Or if we need more complex data types ("nested")

3. Both can be combined

1. .. So that you can just tweak indexes for selected fields
2. You can adjust for languages (English, Czech, ...)

It's always better to start with the default mapping via Sense app.

API: IMAP MAILS PROCESSING

Working with live data

Leveraging Logstash input module for IMAP

- You can send e-mails to europen2016.elk@seznam.cz

Searching, Visualizations

Further improvements

- Detailed and more reliable mail parsing (py-cgmail project, e.g.)
- E-mails sandbox logs processing
- ...

JSON: APTNOTES DATA PROCESSED

Working with JSON data

- PDF parsed by “jager” project
- <https://github.com/sroberts/jager>

Logstash input module for JSON

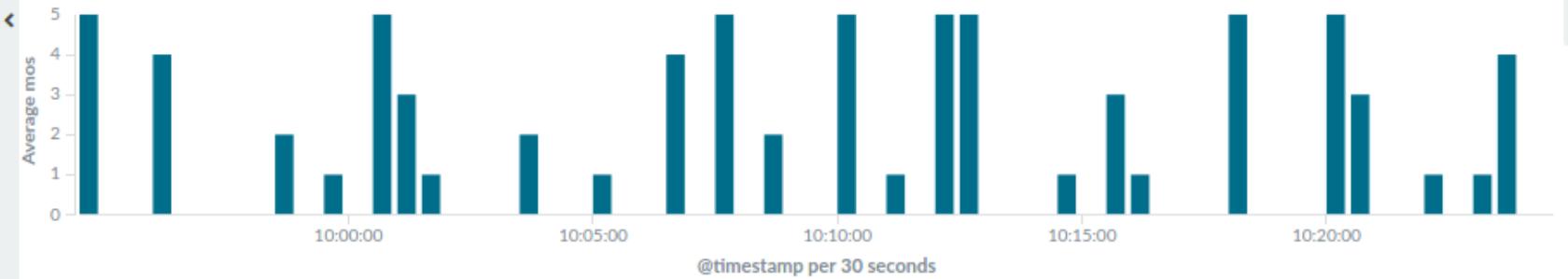
Searching, Visualizations

Better insight into ELK functionality

- Manual data mapping and templates

Another round with data from APTNotes

Average mos



Table

Request

Response

Statistics

Set Watcher

Save Watcher

Save + Edit Watcher in Kaae

Configure the new Watcher preferences before sending to Kaae for processing and scheduling.

Watcher ID:

new_saved

Watcher Repeat Interval:

1m

Watcher Query Range:

5m

Watcher Indices:

<mos-[now/d],<mos-[now/d-1d]>

Watcher Script:

payload.hits.total > 100

Watcher Script - Available Keys:**Email To:**

root@localhost

Email Subject:

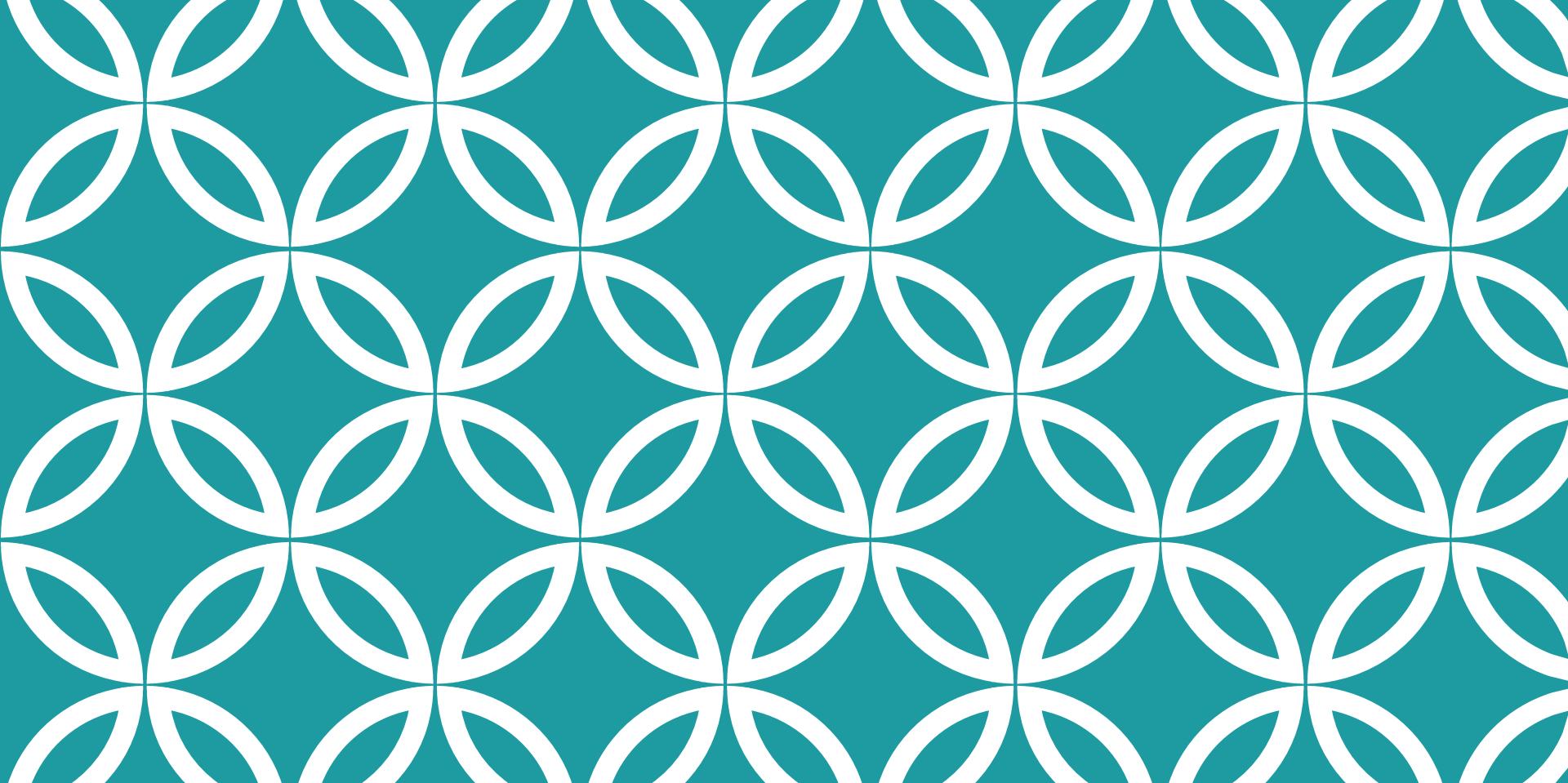
KAAE ALARM {{ payload._id }}

Email Body:

Series Alarm {{ payload._id }}: {{ payload.hits.total }}

KAAE: KIBANA ALERT APP

A new alerting app for
Kibana



KIBANA 5

Sneak Peek

*



Discover



Visualize



Dashboard



Timelion



Graph

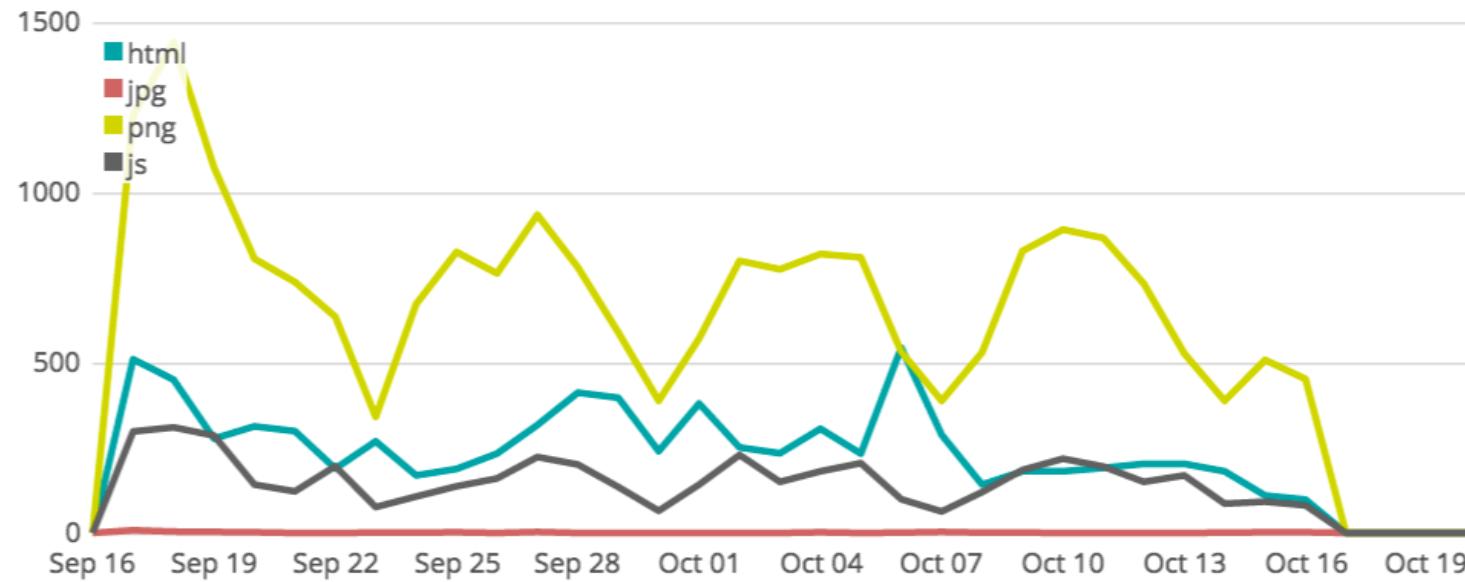


Monitoring



Settings

Content Types



Country by OS



BRAND NEW UI

Sense & Timelion
Integrated



1. Select

2. Review

3. Upload

Pick a CSV file to get started. Please follow the instructions below.

Drop your file here

or

Select File

Maximum upload file size: 1 GB

Reset

Next

EASY CSV PROCESSING

Finally!



Review the index pattern. Here we'll define how and where to store your parsed events. We've made some intelligent guesses for you, but most fields can be changed if we got it wrong!

Index name

The name of the Elasticsearch index you want to create for your data.

Name	Type	Example
Transaction_date	string	1/2/09 6:17
Product	string	Product1
Price	number	1200
Payment_Type	string	Mastercard
Name	string	carolina
City	string	Basildon
State	string	England
Country	string	United Kingdom
Account_Created	string	1/2/09 6:00
Last_Login	string	1/2/09 6:08

1 2 »

Page Size 10

TYPE MAPPING IN UI

Data types made easy

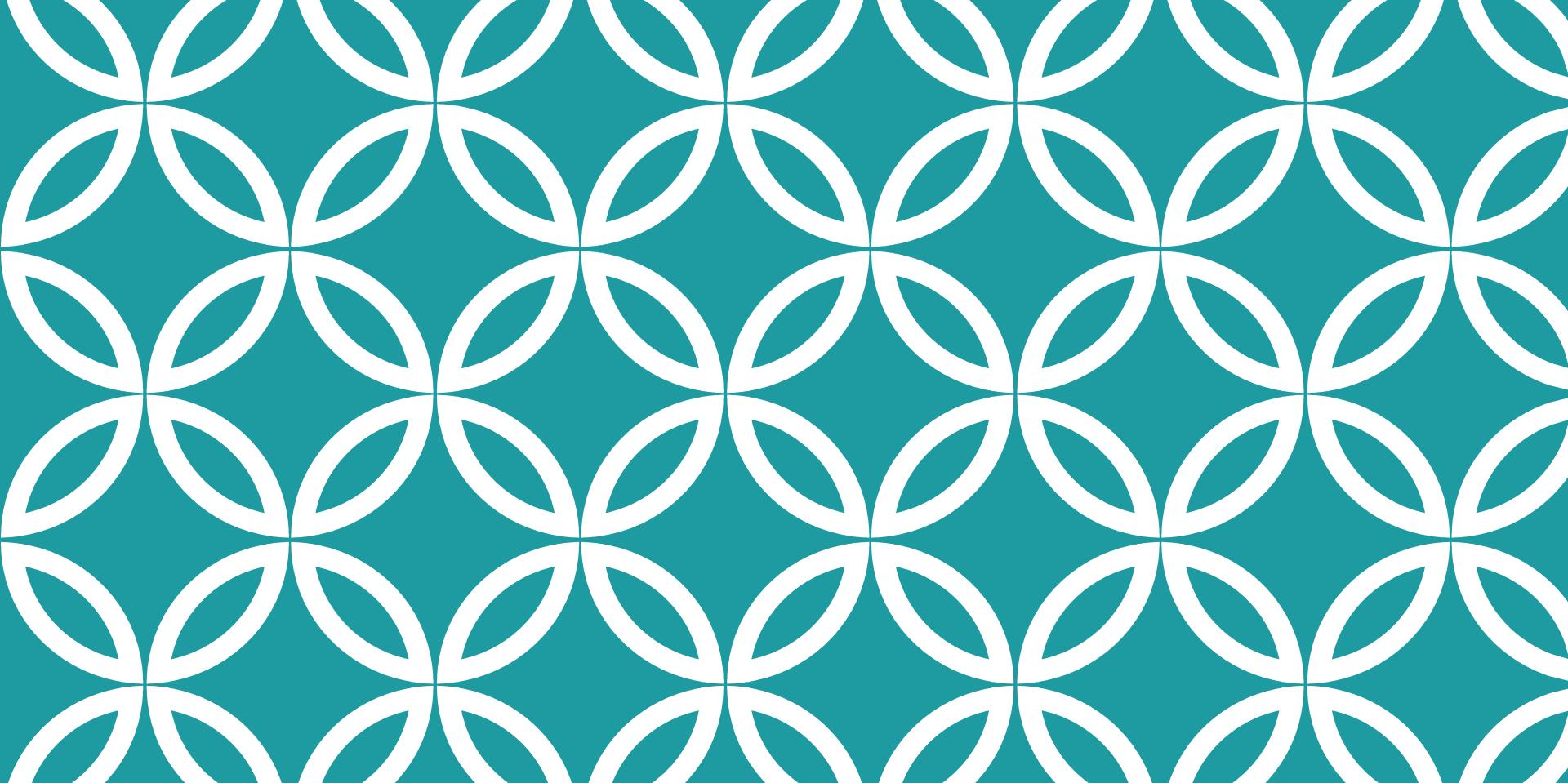


```
1 GET _search
2 {
3   "query": {
4     "match_all": {}
5   }
6 }
```

```
1 {
2   "took": 15,
3   "timed_out": false,
4   "_shards": {
5     "total": 21,
6     "successful": 21,
7     "failed": 0
8   },
9   "hits": {
10    "total": 38638,
11    "max_score": 1,
12    "hits": [
13      {
14        "_index": ".kibana",
15        "_type": "visualization",
16        "_id": "test1",
17        "_score": 1,
18        "_source": {
19          "title": "test1",
20          "visState": "{\"title\":\"test1\",\"type\":\"histogram\",\"params\":{\"shareYAxis\":true,\"addTooltip\":true,\"addLegend\":true,\"scale\":\"linear\",\"mode\":\"stacked\",\"times\":[],\"addTimeMarker\":false,\"defaultYExtents\":false,\"setYExtents\":false,\"yAxis\":[]},\"aggs\":[{\"id\":\"1X\",\"enabled\":true,\"type\":\"count\",\"schema\":\"metric\",\"params\":{}},{\"id\":\"2X\",\"enabled\":true,\"type\":\"terms\",\"schema\":\"segment\",\"params\":{\"field\":\"policyID\",\"size\":5,\"order\":[\"desc\"],\"orderBy\":1}],{\"id\":\"3X\",\"enabled\":true,\"type\":\"cardinality\",\"schema\":\"metric\",\"params\":[\"field\":\"fr_site_limit\"]}],\"listeners\":[]}",
21          "uiStateJSON": "{\"vis\":{\"legendOpen\":true}}",
22          "description": "",
23          "version": 1,
24        },
25        "kibanaSavedObjectMeta": {
26          "searchSourceJSON": "{\"index\":\"insurance\",\"query\":{\"query_string\":{\"query\":\"*\"},\"analyze_wildcard\":true}},\"filter\":[]}"
27        }
28      },
29    ],
30  }
```

CONSOLE INTEGRATED

Batteries included!



FURTHER DATA SOURCES?

Sky is the limit

WHAT COULD YOU POSSIBLY PROCESS?

Antivirus | [HIDS/HIPS logs](#)

Endpoint Security Protection | [common support for {Splunk, ElasticSearch}](#)

DNS logs | [PassiveDNS in some form](#)

Proxy logs | [+ pertinent stuff \(WAF, Apache, IIS, haproxy\)](#)

Firewalls | [Cisco ASA and similar](#)

IDS | [Suricata has direct JSON support](#)

Honeypots | [log at least the metadata](#)

Windows logy | [events, RDP, AD, ...](#)

WHAT COULD YOU POSSIBLY PROCESS?

VPN | Logons, certificate details, IPs, ...

Forensic artifacts, IOCs | From the previous incidents

Sandbox data | Cuckoo, Lastline, ...

Ticketing Tools | Metrics, Dashboards, correlations with other tools

Backup info | Ransomware infections can be defeated by good backups

Threat Intel | Correlation source; you can measure their quality too!

CIRT emails | great for phishing campaigns analysis

Anything you could use in your company!



FURTHER PROJECTS

Strong Community Support

ELASTICSEARCH BASED

UI over the data stored

ElasticUI

SearchKit

Metrics

Grafana

Alerting

ElastAlert

411

PCAP collector

Moloch

Incident Response

nightHawkResponse

Kibana plugins

Kibana: Known Plugins

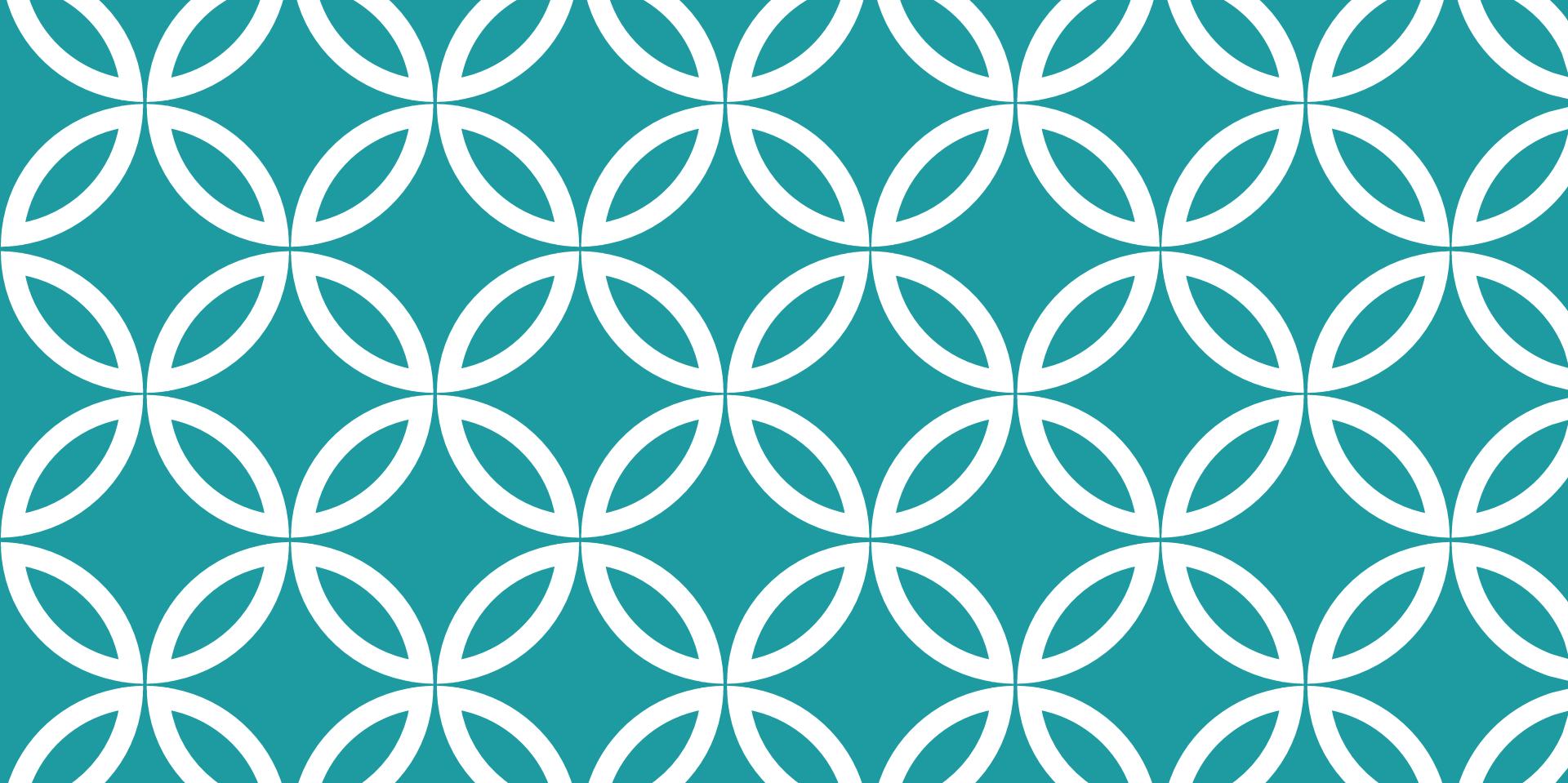
ELASTIC AS A SERVICE

logz.io

- <http://site2.logz.io/pricing/>
- Free: 1 GB/day | 3 days retention | alerting

LogSene

- <http://sematext.com/logsene/>
- Free: 512 MB/day | 7 days data retention | role-based access



QUESTIONS & ANSWERS

Questions
Answers
Comments