Elastic stack

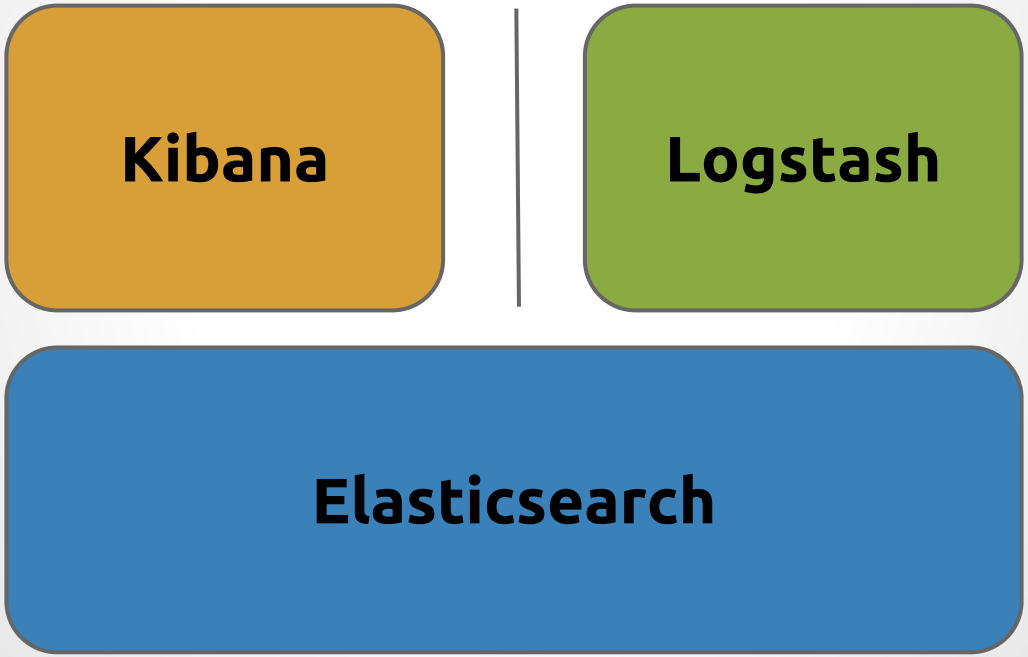
Elasticsearch Logstash Kibana

Marius Bieliauskas

VilniusPHP 0x20

estina

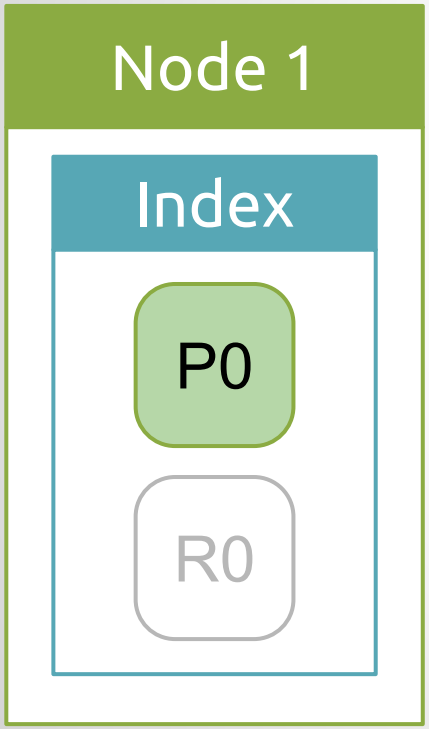
What is Elastic stack (ELK)?



What is Elasticsearch?

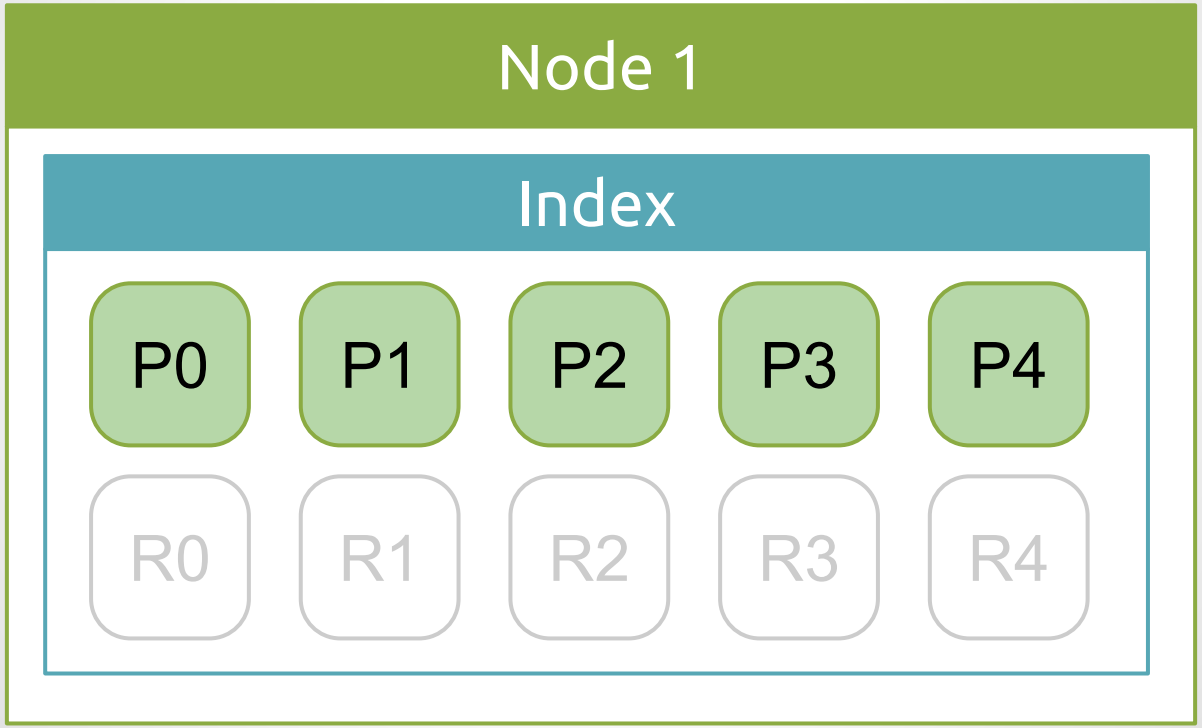
Scalable data **storage**
with search and analytics engine

Elasticsearch basic setup (1 P, 1 R)

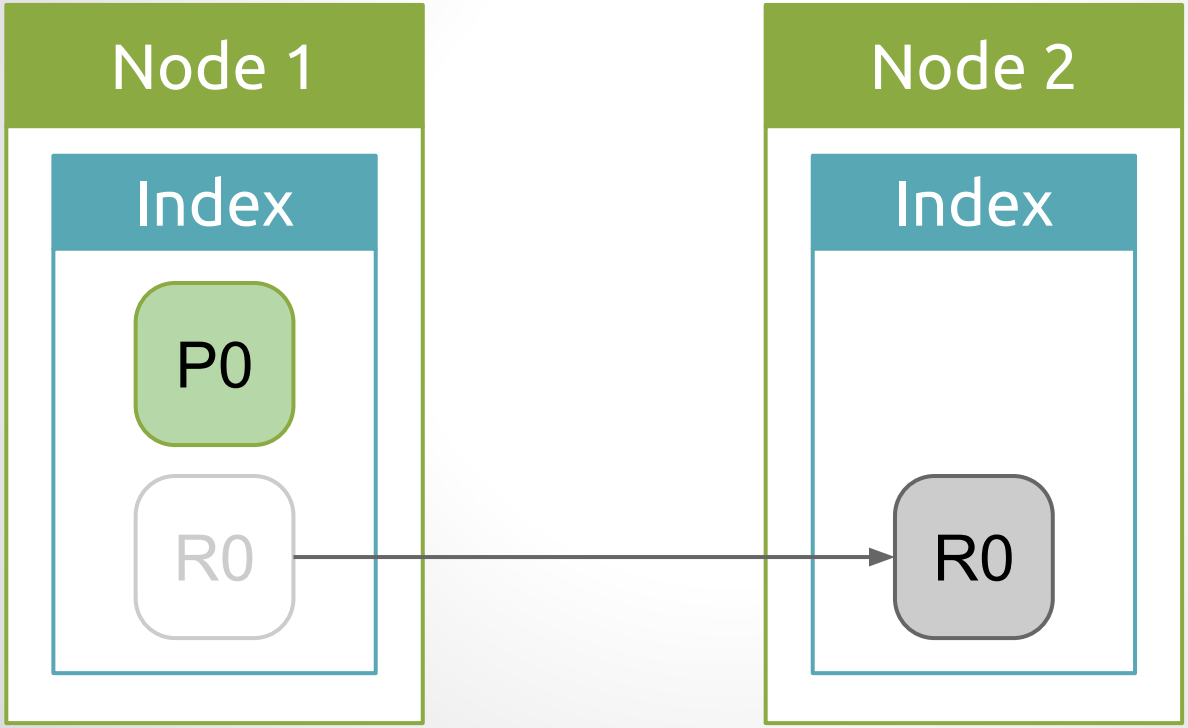


- Node - Elasticsearch instance;
- Index - Logical namespace - “database”;
- Type - Field mappings - “table”;
- Shard - Single Lucene instance (Primary/Replica);

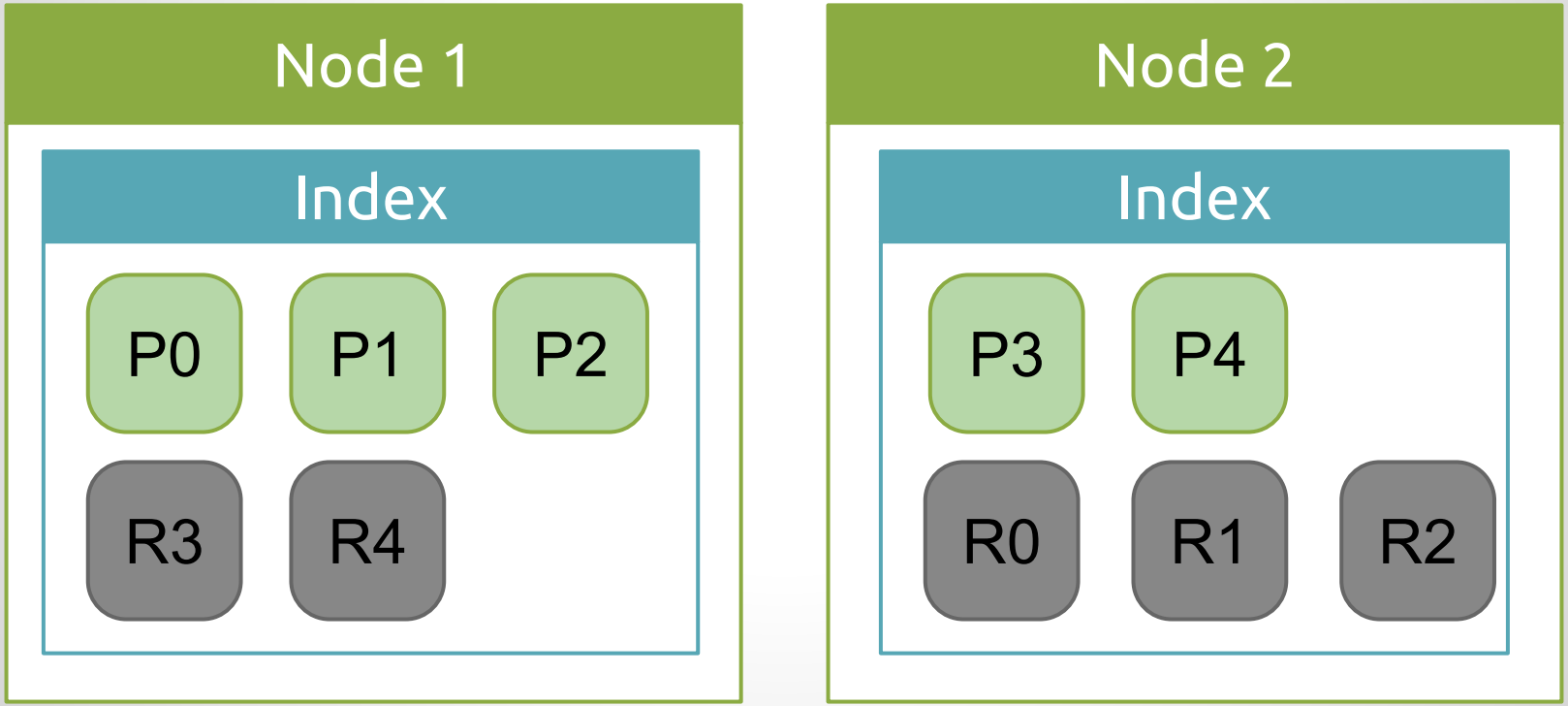
Elasticsearch default setup (5 P, 5 R)



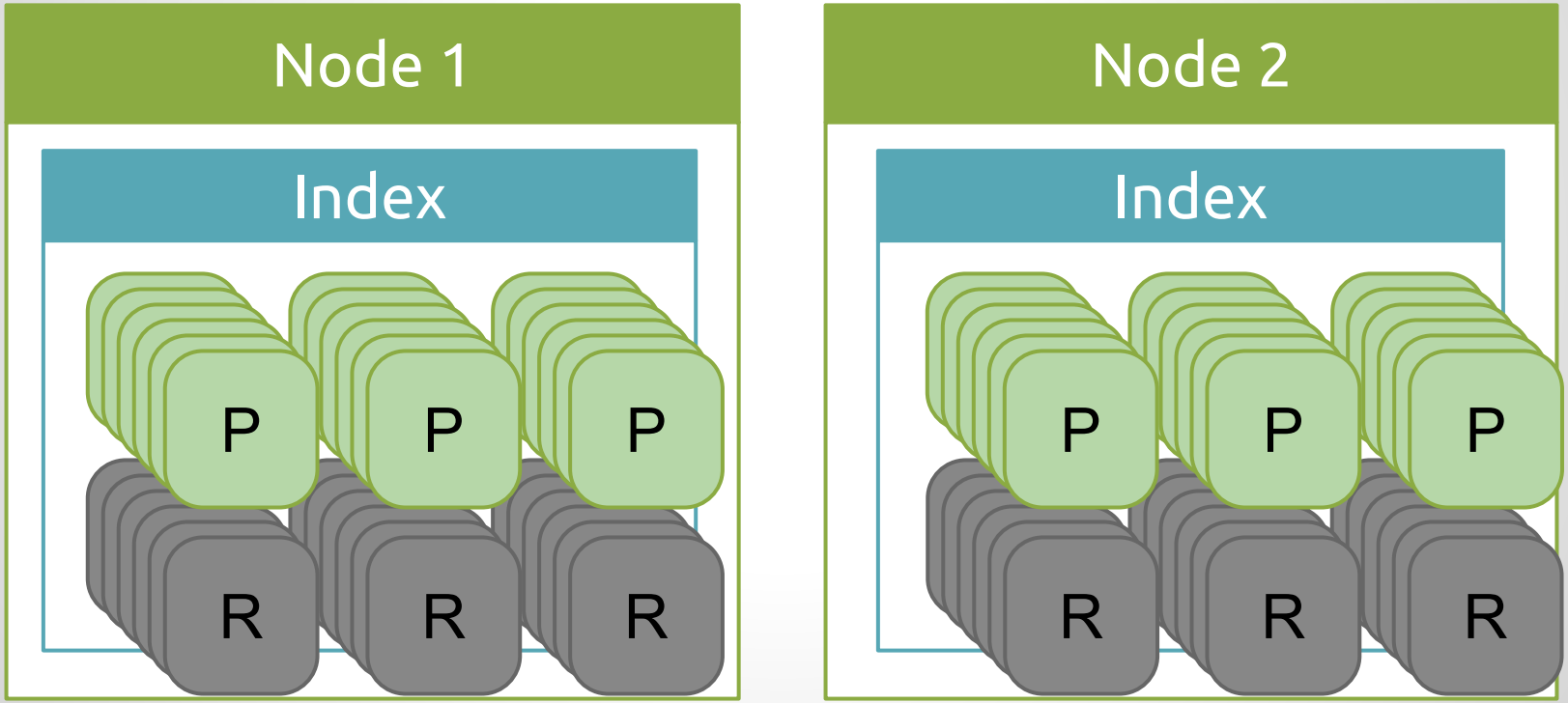
Elasticsearch basic scaling



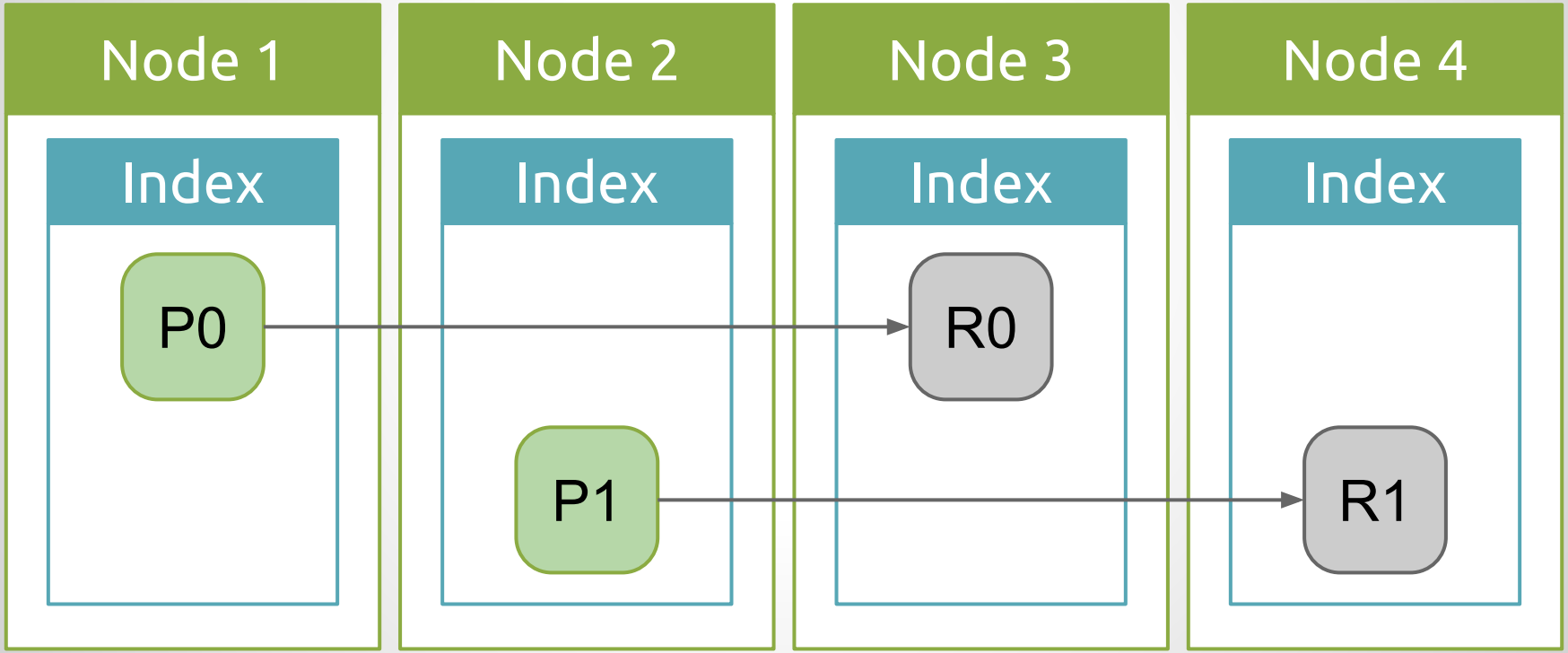
Elasticsearch default scaling



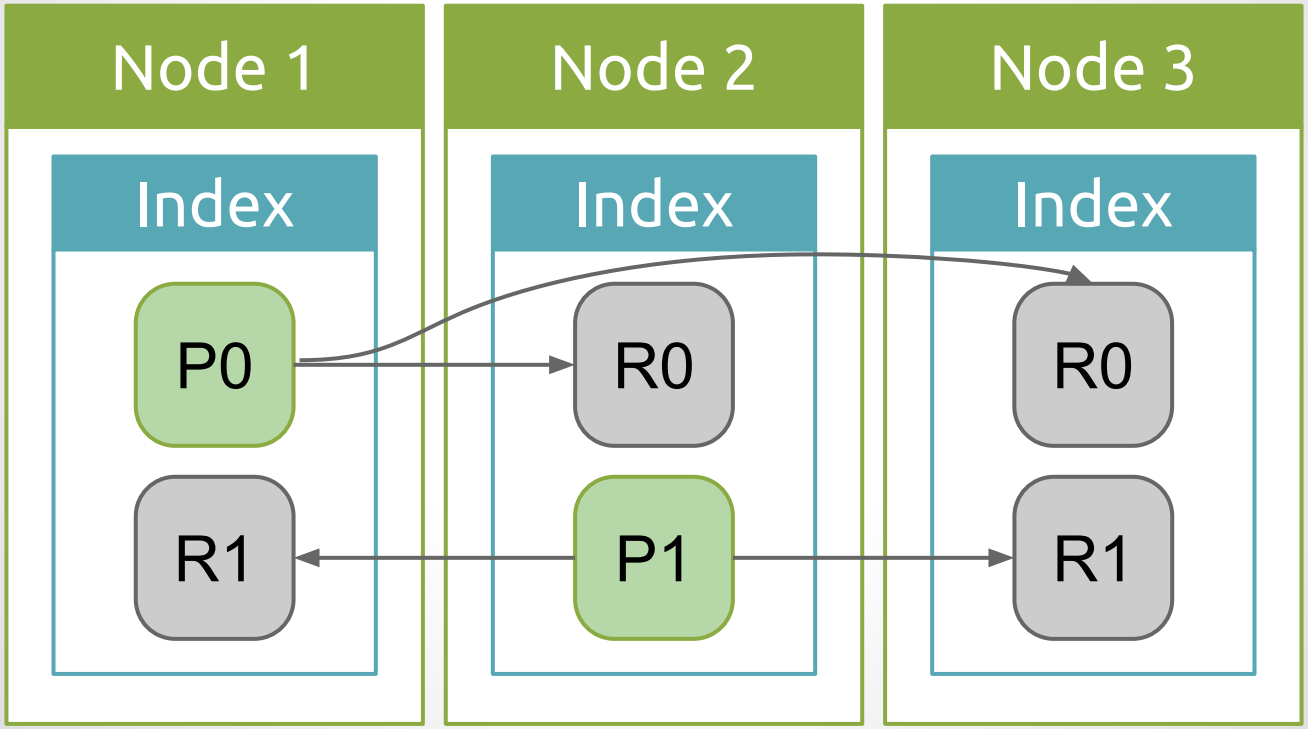
Kagillion shards?



Elasticsearch replication



Elasticsearch replication



Elasticsearch vs RDBMS vs JSON store

No transactions;

Designed for write once read many;

Can be inconsistent;

Elasticsearch eats memory!

< 32GB and < 50% total memory

Elasticsearch is built for speed, with the assumption that memory is abundant.

Logstash

Data collection and transformation agent

Logstash: Plugins

Input - file, jdbc, heartbeat, elasticsearch, rabbitmq, stdin, irc, exec etc.

Filter - grok, csv, date, anonymize, json, mutate, geoip...

Output - file, csv, email, stdout, elasticsearch, rabbitmq, xmpp, redis...

Grok Debugger | Debugger | Discover | Patterns

127.0.0.1 - - [30/Aug/2015:00:56:43 +0300] "GET /api/v2/users/64/assigned_bookings?timestamp=1440885382&lang=lt HTTP/1.0" 200 169 "-" "-"

%{NGINXACCESS}

☑ Add custom patterns ☐ Keep Empty Captures ☑ Named Captures Only ☐ Singles ☐ Autocomplete

One per line, the syntax for a grok pattern is `%{SYNTAX:SEMANTIC}`

```
NGUSERNAME [a-zA-Z\.\@\-\+_%]+
NGUSER %{NGUSERNAME}
NGINXACCESS %{IPORHOST:clientip} %{NGUSER:ident} %{NGUSER:auth} \[%{HTTPDATE:timestamp}\] "%{WORD:verb} %{URIPATHPARAM:request} HTTP/%{NUMBER:httpversion}" %{NUMBER:response} (?:%{NUMBER:bytes}|-) (?:"(?:%{URI:referrer}|-)"|%{QS:referrer}) %{QS:agent}
```


```
{
  "clientip": [
    [
      "127.0.0.1"
    ]
  ],
  "ident": [
    [
```

Logstash: File input

```
input {
  file {
    path => "/opt/logstash/configs/*.csv"
    start_position => "beginning"
    type => "call"
  }
}
```

Logstash: CSV filter

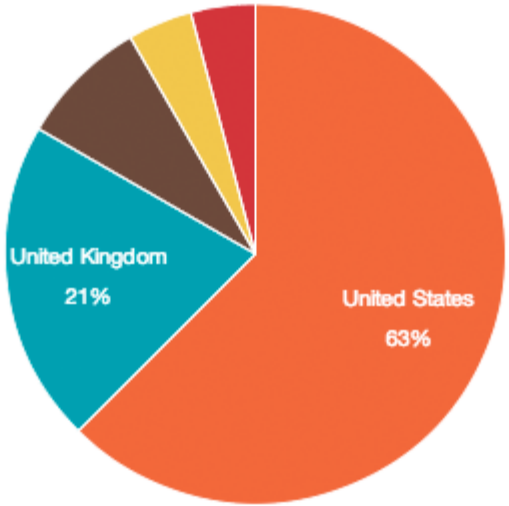
```
filter {
  csv {
    columns => ["Id","LineId","CalledId","CallerId","
DateOffering","CallWait","CallDuration","
CallTotalDuration"]
  }

  date {
    match => [ "DateOffering" , "UNIX" ]
  }
}
```

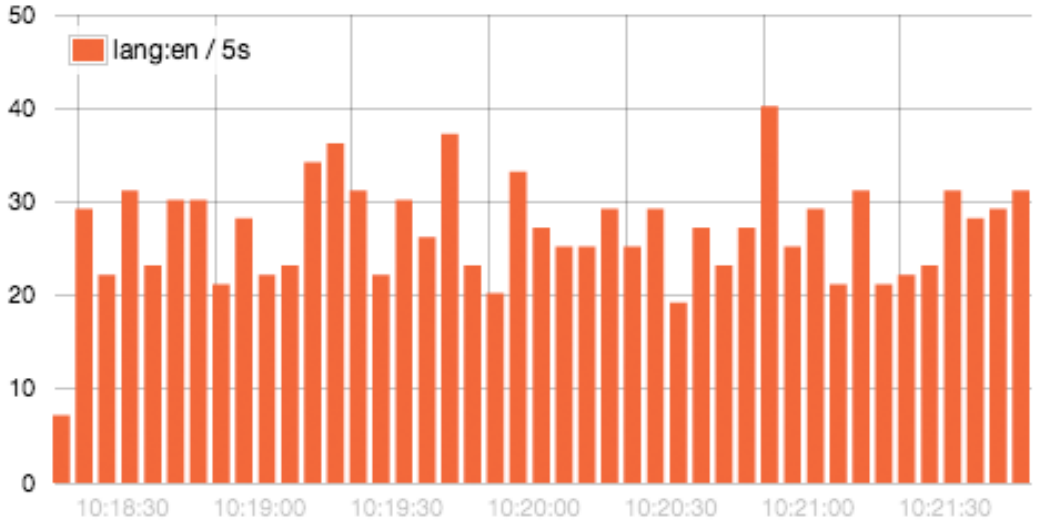
Logstash: Elasticsearch output

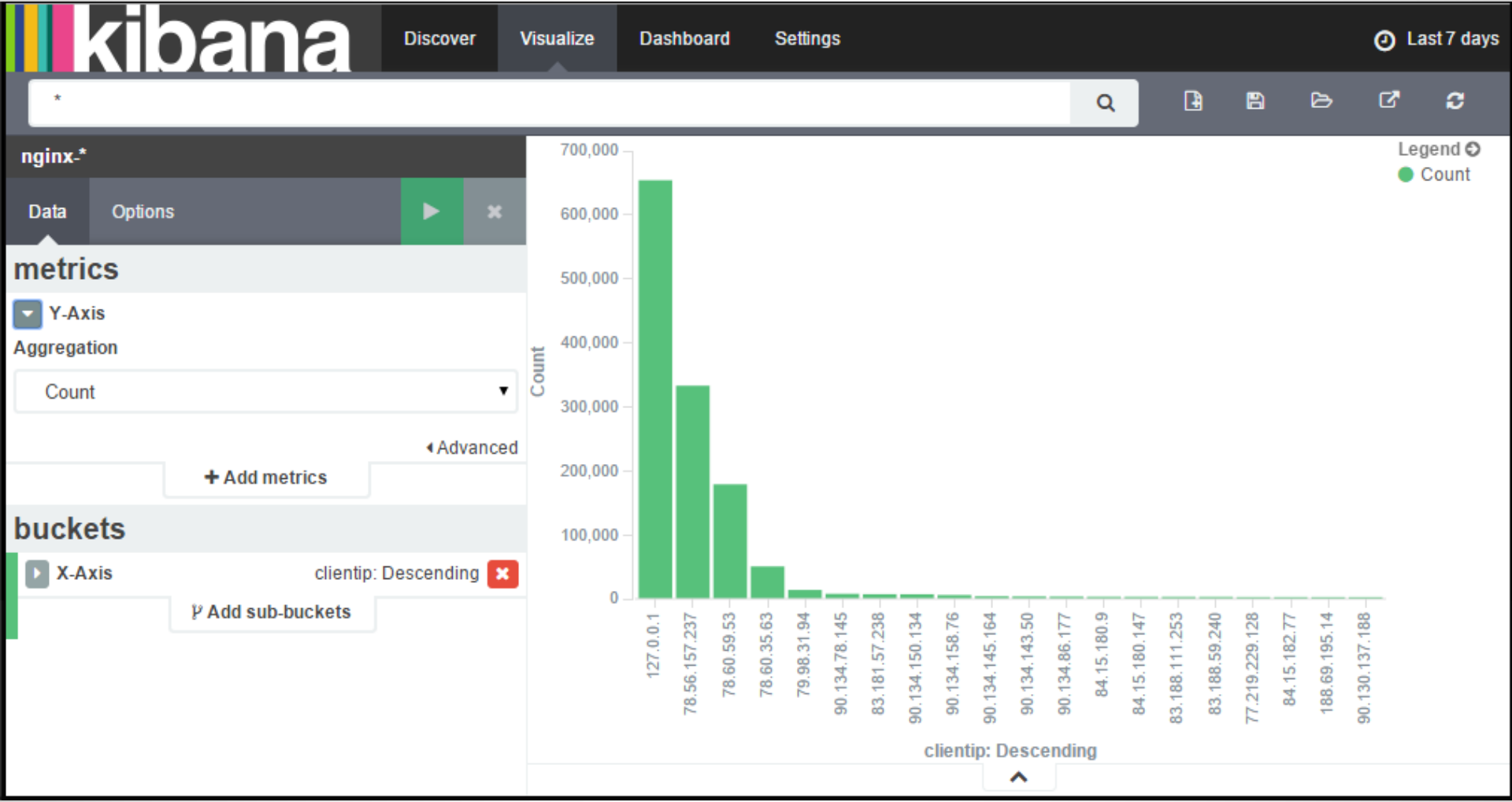
```
output {
  elasticsearch {
    host => "localhost"
    protocol => "http"
    index => "calls-%{+YYYY.MM.DD}"
  }
  #stdout { codec => rubydebug }
}
```

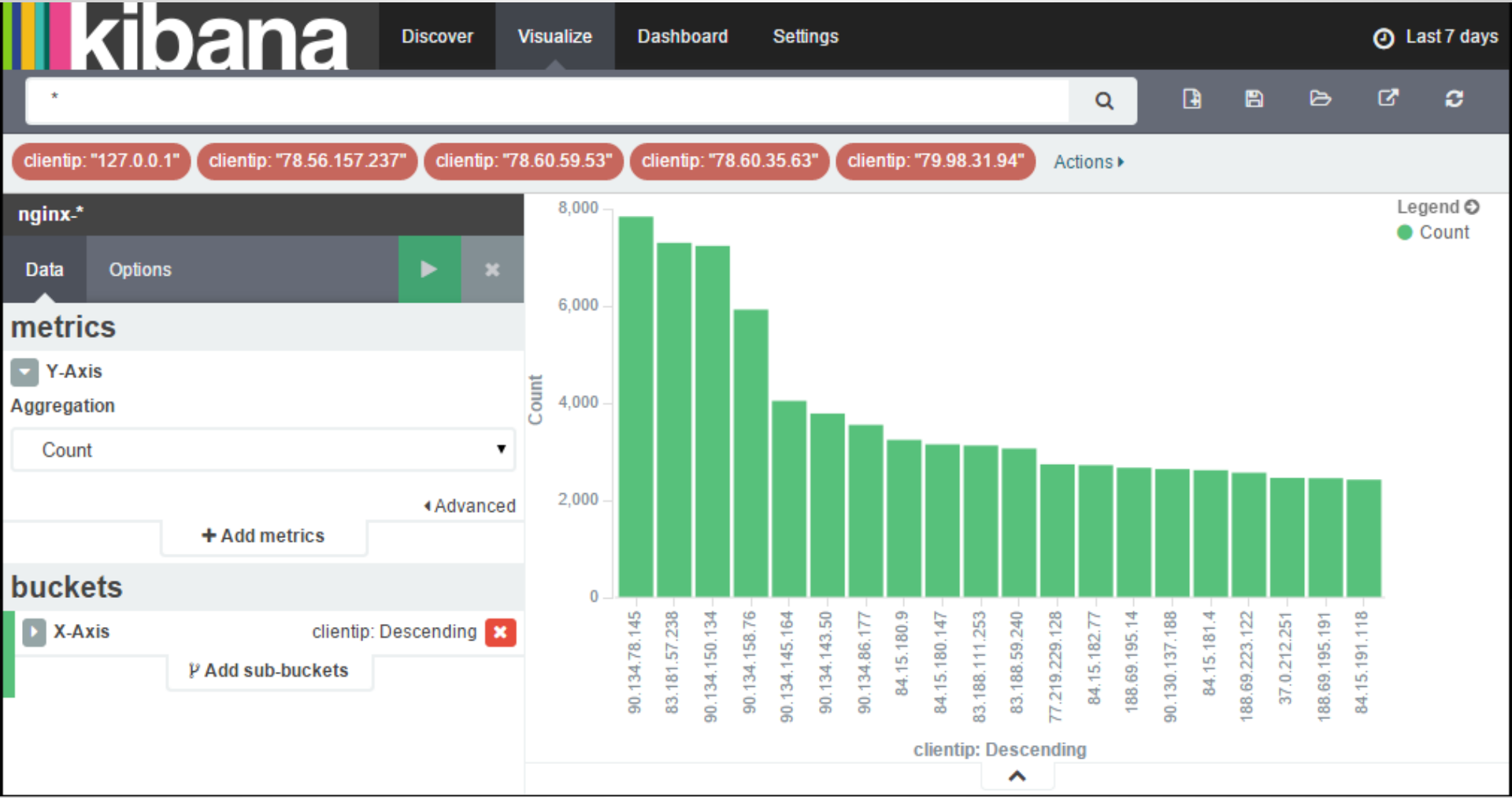
Elastic stack: Kibana

Powerful and beautiful data visualizations






kibana
Discover
Visualize
Dashboard
Settings
Last 7 days
*
nginx-*
Data
Options
metrics
Y-Axis
Aggregation
Count
Advanced
+ Add metrics
buckets
X-Axis
clientip: Descending
Add sub-buckets
Legend
Count
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0
Count
127.0.0.1
78.56.157.237
78.60.59.53
78.60.35.63
79.98.31.94
90.134.78.145
83.181.57.238
90.134.150.134
90.134.158.76
90.134.145.164
90.134.143.50
90.134.86.177
84.15.180.9
84.15.180.147
83.188.111.253
83.188.59.240
77.219.229.128
84.15.182.77
188.69.195.14
90.130.137.188
clientip: Descending


kibana
Discover
Visualize
Dashboard
Settings
Last 7 days
*
clientip: "127.0.0.1"
clientip: "78.56.157.237"
clientip: "78.60.59.53"
clientip: "78.60.35.63"
clientip: "79.98.31.94"
Actions
nginx-*
Data
Options
metrics
Y-Axis
Aggregation
Count
Advanced
+ Add metrics
buckets
X-Axis
clientip: Descending
Add sub-buckets
Legend
Count
8,000
6,000
4,000
2,000
0
Count
90.134.78.145
83.181.57.238
90.134.150.134
90.134.158.76
90.134.145.164
90.134.143.50
90.134.86.177
84.15.180.9
84.15.180.147
83.188.111.253
83.188.59.240
77.219.229.128
84.15.182.77
188.69.195.14
90.130.137.188
84.15.181.4
188.69.223.122
37.0.212.251
188.69.195.191
84.15.191.118
clientip: Descending

Links

- https://vimeo.com/44716955
- https://www.korekontrol.eu/blog/tips-for-centralized-logging-infrastructure-with-logstash
- http://logz.io/blog/5-great-reasons-to-upgrade-to-kibana-4/
- https://github.com/vilniusphp/vilniusphp-meetups/blob/master/2014-02-06/Log.pdf
- http://www.cubrid.org/blog/dev-platform/our-experience-creating-large-scale-log-search-system-using-elasticsearch/
- https://blog.codecentric.de/en/2014/05/elasticsearch-indexing-performance-cheatsheet/
- http://grokdebug.herokuapp.com/