Infrastruktúranaplózás ElasticSearch-csel

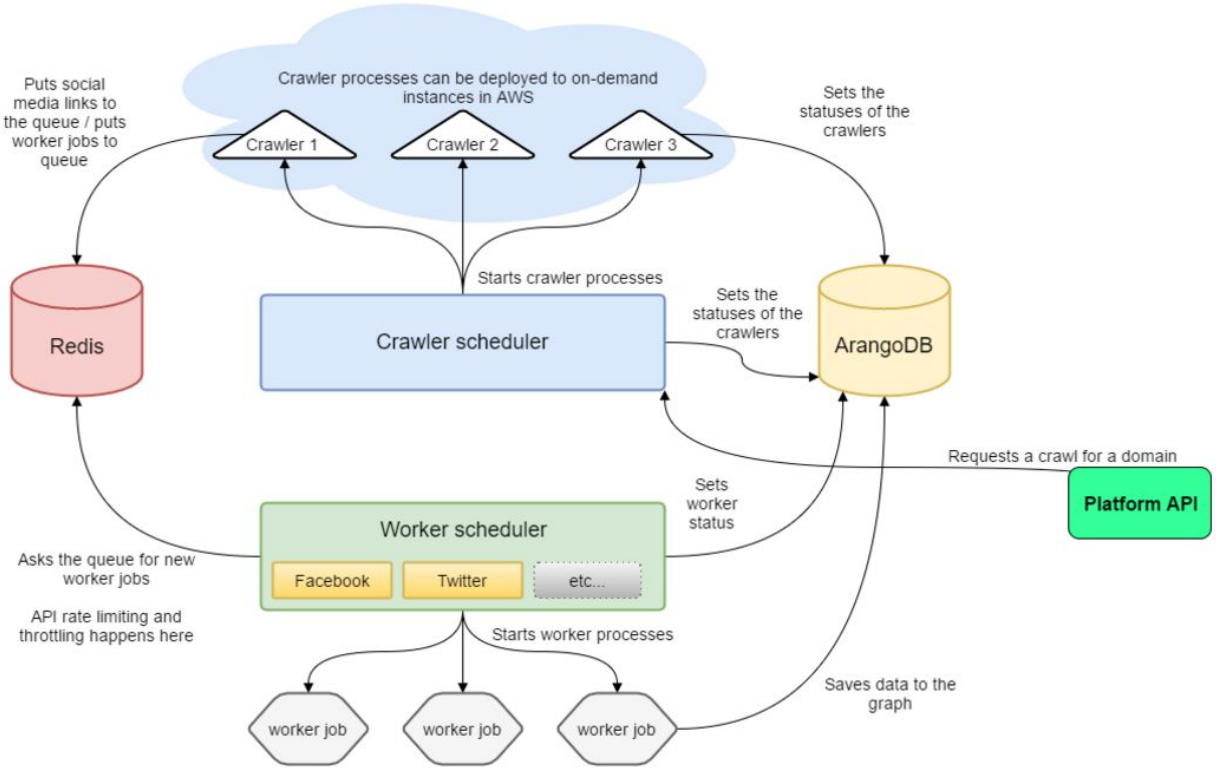
Bencs Balázs
Attrecto Zrt.

attrecto

Miért?

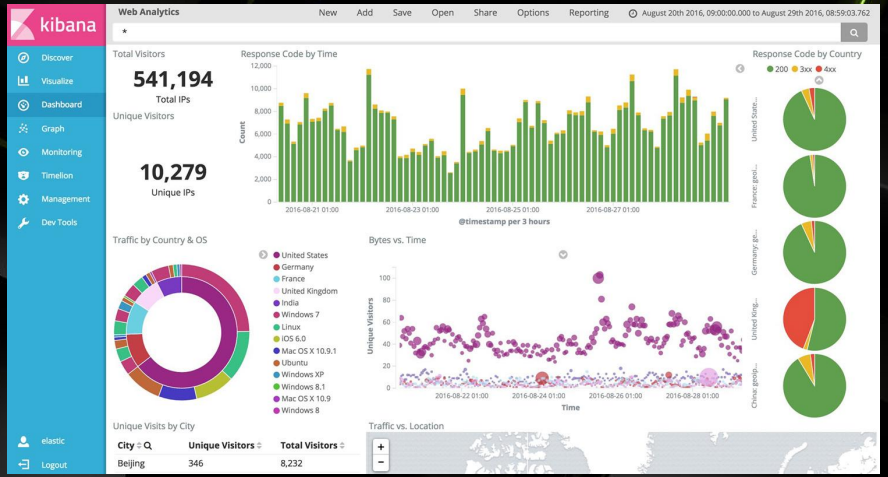
“Kellenének a production server
tegnapi logjaiból a hibák 10 és 1 óra között”

Kell egy központi tároló


Crawler processes can be deployed to on-demand instances in AWS
Crawler 1
Crawler 2
Crawler 3
Puts social media links to the queue / puts worker jobs to queue
Sets the statuses of the crawlers
Starts crawler processes
Redis
Crawler scheduler
Sets the statuses of the crawlers
ArangoDB
Requests a crawl for a domain
Platform API
Sets worker status
Worker scheduler
Facebook
Twitter
etc...
Asks the queue for new worker jobs
API rate limiting and throttling happens here
Starts worker processes
worker job
worker job
worker job
Saves data to the graph

nginx access log

```
192.168.1.16 - - [10/Jan/2017:18:06:59 +0000] "POST
/kibana/elasticsearch/_msearch?timeout=30000&ignore
_unavailable=true&preference=1447070343481
HTTP/1.1" 200 8352
"https://elastic/kibana/index.html" "Mozilla/5.0
(X11; Linux armv7l) AppleWebKit/537.36 (KHTML, like
Gecko) Ubuntu Chromium/45.0.2454.101
Chrome/45.0.2454.101 Safari/537.36" 0.465 0.454
```



ElasticSearch

Java alapú

Fulltext search engine és document storage

Keresés és indexelés

Elosztott - Sharding és replication támogatott

Clustering

Egyszerű REST API

Logstash

Real-time naplófájl feldolgozás

Többféle input, többféle output

Központosítja a logokat

Begyűjti az adatokat

Lehetnek szöveges logfájlok, syslog, redis, filebeat, vagy tetszőleges..
(a filebeat-ről később)

Feldolgozza az adatokat beolvasás közben

Rengeteg kész pattern különböző log típusokhoz

Grok pattern engine

Sok plugin gyárilag, de tovább bővíthető

GeoIP feloldás

Továbbítja és tárolja Elastic-ban

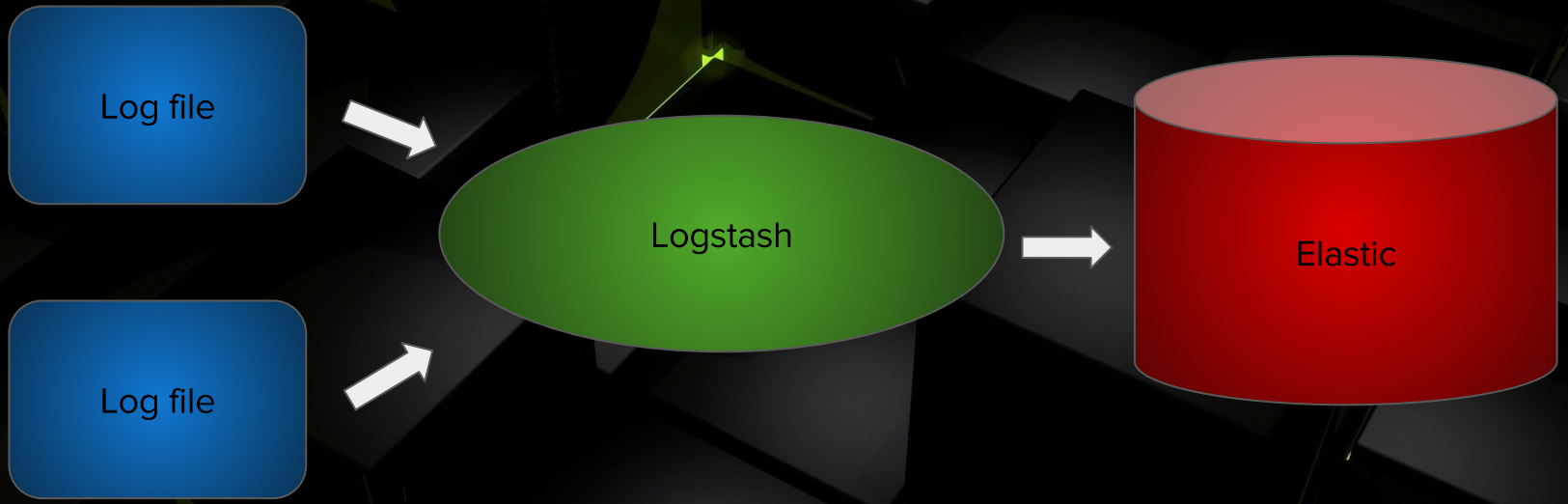
Log feldolgozás

Logstash

Grok pattern

```
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}
```

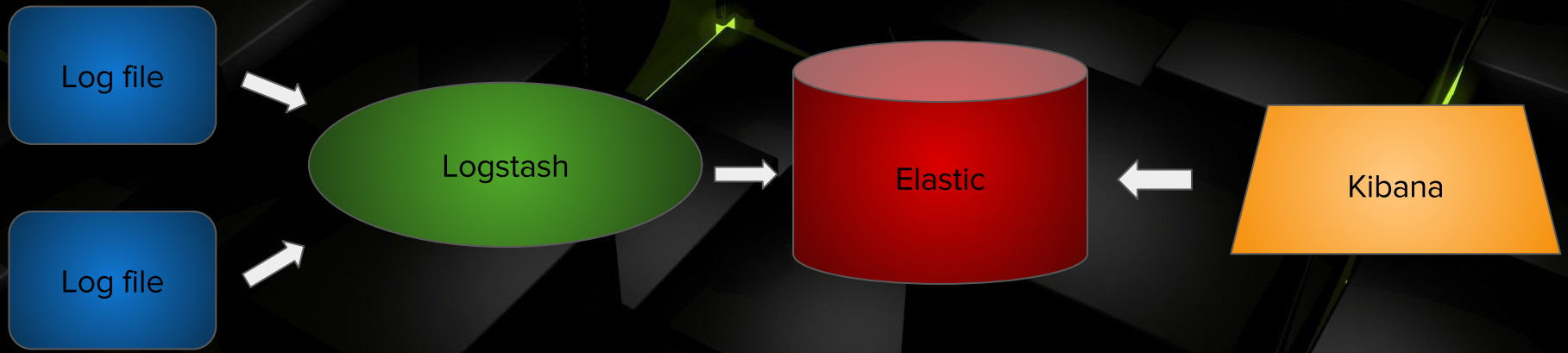
Logstash



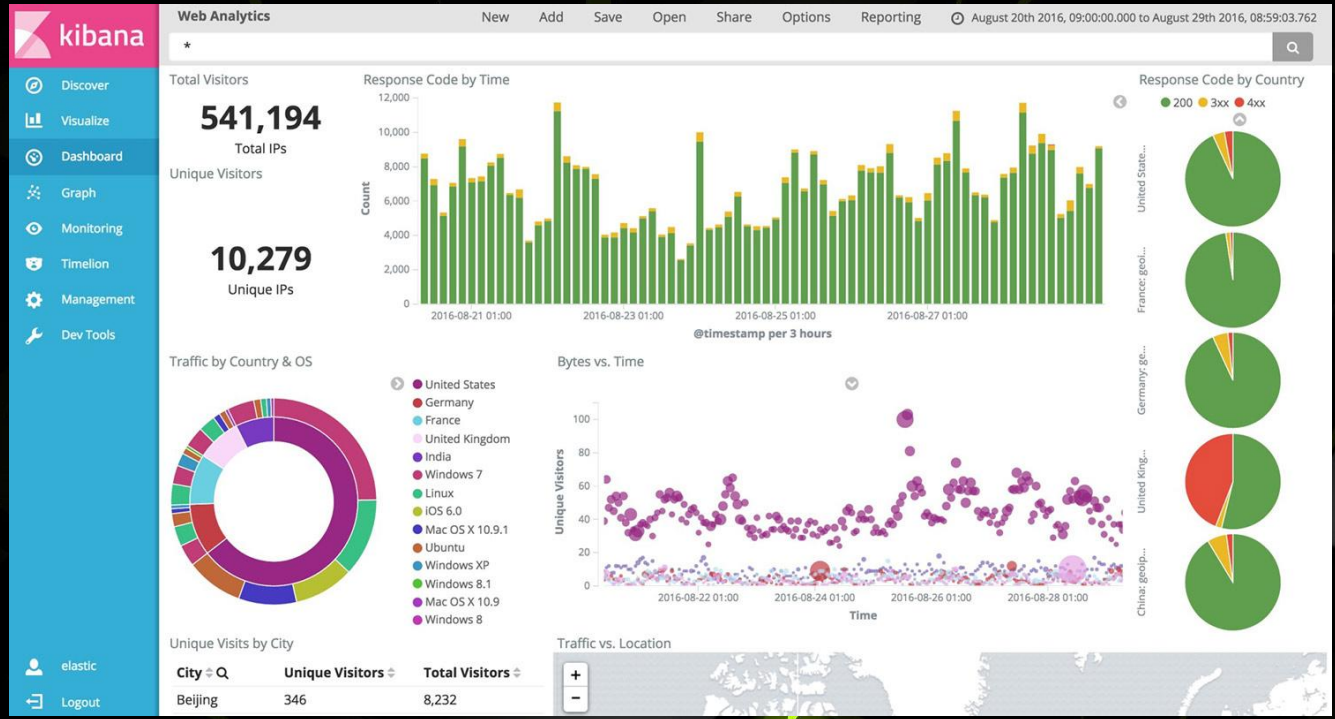
Logstash filter

```
grok {
  match => [ "message" , "%{NGINXACCESS}"]
  overwrite => [ "message" ]
}

mutate {
  convert => ["response", "integer"]
  convert => ["bytes", "integer"]
  convert => ["responsetime", "float"]
}

geoip {
  source => "clientip"
  target => "geoip"
  add_tag => [ "nginx-geoip" ]
}

date {
  match => [ "timestamp" , "dd/MMM/YYYY:HH:mm:ss Z" ]
  remove_field => [ "timestamp" ]
}
```

Logstash



Kibana


kibana
Web Analytics
New
Add
Save
Open
Share
Options
Reporting
August 20th 2016, 09:00:00.000 to August 29th 2016, 08:59:03.762
*
Discover
Visualize
Dashboard
Graph
Monitoring
Timelion
Management
Dev Tools
elastic
Logout
Total Visitors
541,194
Total IPs
Unique Visitors
10,279
Unique IPs
Response Code by Time
Count
@timestamp per 3 hours
Response Code by Country
200
3xx
4xx
Traffic by Country & OS
United States
Germany
France
United Kingdom
India
Windows 7
Linux
iOS 6.0
Mac OS X 10.9.1
Ubuntu
Windows XP
Windows 8.1
Mac OS X 10.9
Windows 8
Bytes vs. Time
Unique Visitors
Time
Unique Visits by City
City
Unique Visitors
Total Visitors
Beijing
346
8,232
Traffic vs. Location

Logstash kliensek

In-App log

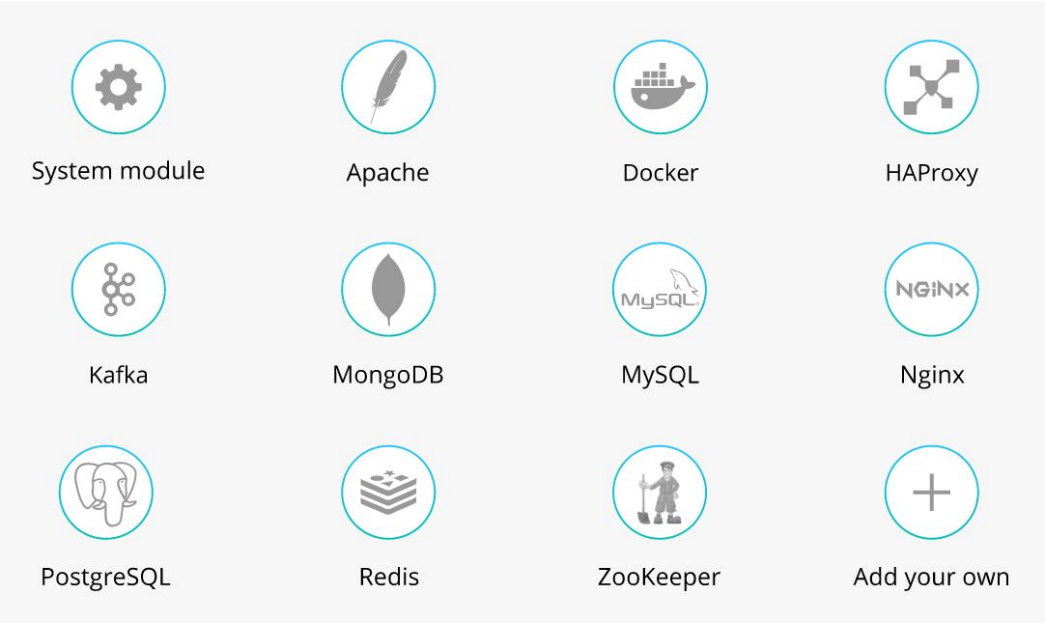
log4j.properties:

```
log4j.rootLogger=daily
log4j.appender.daily=org.apache.log4j.rolling.RollingFileAppender
log4j.appender.daily.RollingPolicy=org.apache.log4j.rolling.TimeBasedRollingPolicy
log4j.appender.daily.RollingPolicy.FileNamePattern=/var/log/application/app.%d.log
log4j.appender.daily.layout = org.apache.log4j.PatternLayout

log4j.appender.daily.layout.ConversionPattern=%d{YYYY-MM-dd HH:mm:ss,SSSZ} %p %c{1}:%L
- %m%n
```

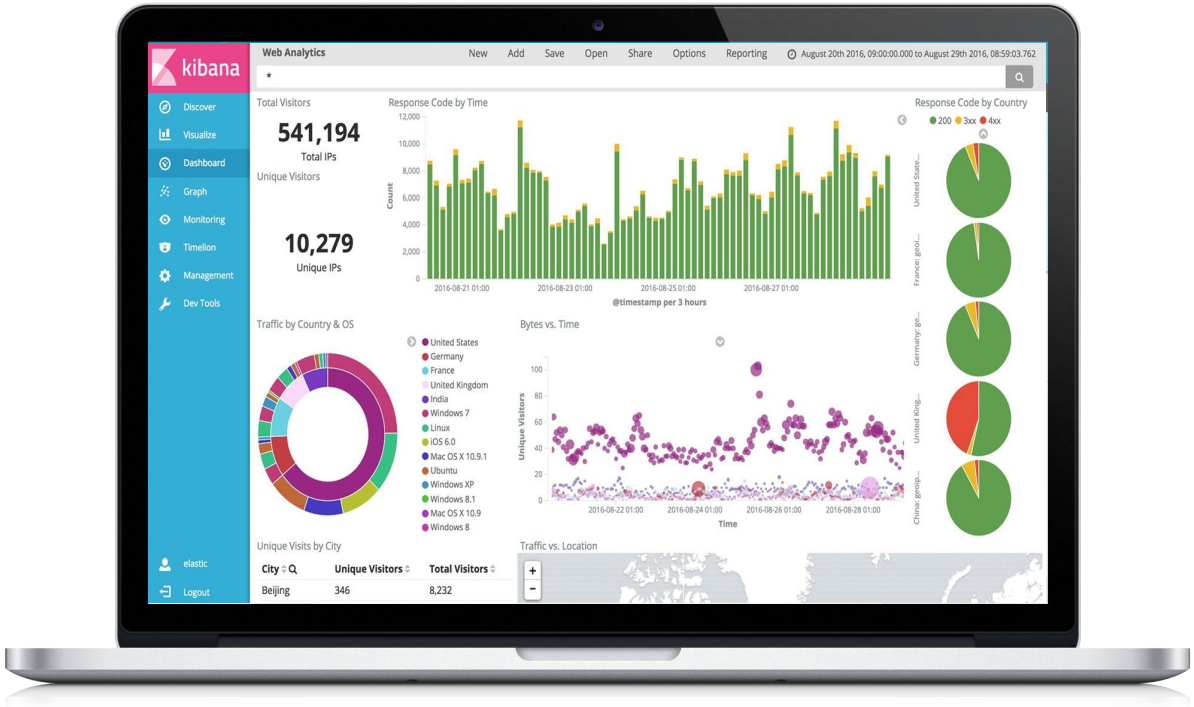
Logstash filter config:

```
input {
  log4j {
    mode => server
    host => "0.0.0.0"
    port => [LOGSTASH_PORT]
    type => "log4j"
  }
}
output {
  elasticsearch {
    protocol => "http"
    host => "[IP_ADDRESS]"
    port => "[PORT]"
  }
}
```

MetricsBeat



DEMO


kibana
Web Analytics
New
Add
Save
Open
Share
Options
Reporting
August 20th 2016, 09:00:00.000 to August 29th 2016, 08:59:03.762
Discover
Visualize
Dashboard
Graph
Monitoring
Timelion
Management
Dev Tools
elastic
Logout
Total Visitors
541,194
Total IPs
Unique Visitors
10,279
Unique IPs
Response Code by Time
Count
@timestamp per 3 hours
Response Code by Country
200
3xx
4xx
Traffic by Country & OS
United States
Germany
France
United Kingdom
India
Windows 7
Linux
iOS 6.0
Mac OS X 10.9.1
Ubuntu
Windows XP
Windows 8.1
Mac OS X 10.9
Windows 8
Bytes vs. Time
Unique Visitors
Time
Unique Visits by City
City
Unique Visitors
Total Visitors
Beijing
346
8,232
Traffic vs. Location


Köszönöm a figyelmet!
attrecto