Agenda

- Logging
- Splunk vs Open Source (ELK stack)
- Logstash
- Elasticsearch
- Kibana
- Getting started
- Most asked/ upvoted questions and answers on Quora
- DEMO
- Q/A

Logging

- Logging
  - Log (file) created by server/ app
    - Information about the requests, date, bytes served, user agent, etc. It’s variable.

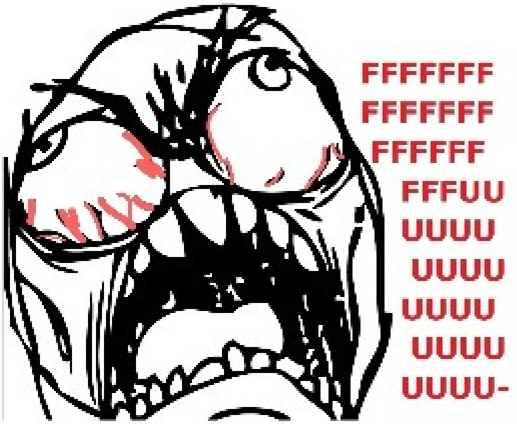
- Application runs
  - Produces errors, warnings, debug, telemetry, analytics events, and other information
  - How to make sense of it?

$plunk

Business as usual, until…

#Outage @03:00 AM

Massive RAGE



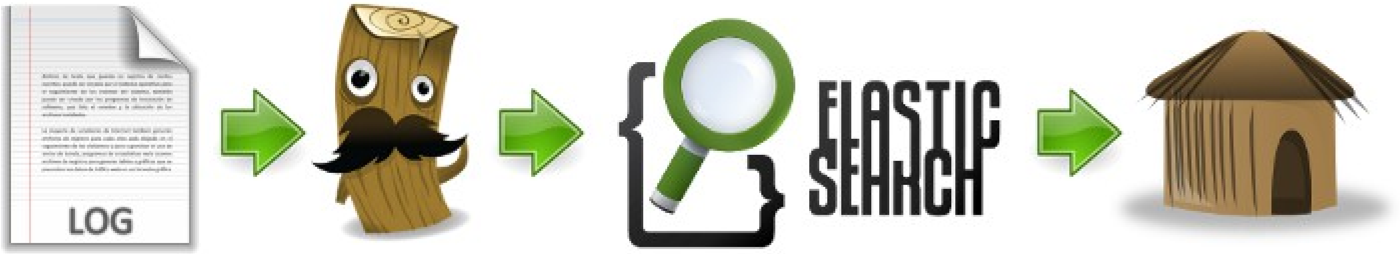
Or the old school style: Cat, grep, awk, cut via the terminal …

Good luck with that on 200 GB of unstructured logs. Think lots of coffee breaks.

The fix: ELK stack (it is Open Source)

Splunk vs. Open Source (ELK)

splunk>™ ☒



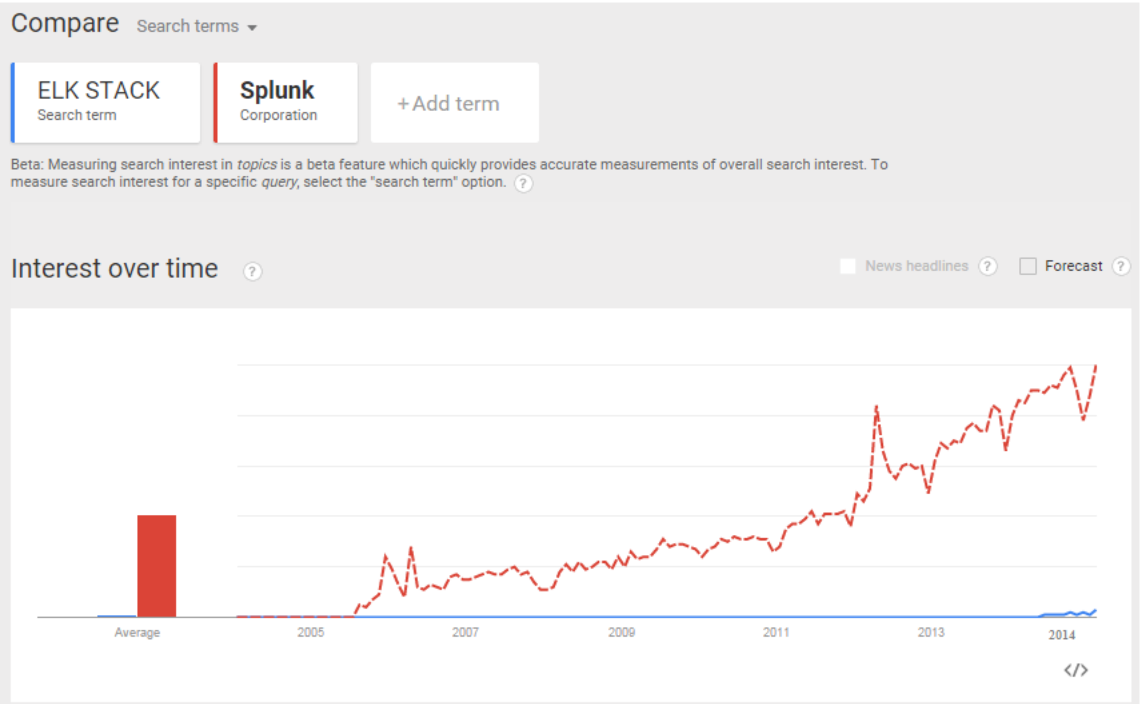
☑

accenture

Why should I use Splunk when I can use Open Source?

- Splunk
  - Widely used
  - Easy to use
  - Cross platform
  - Expensive
  - Complex set up process
- ELK stack
  - Easy installation
  - Open Source
  - Extend functionality via plugins
  - Simple web interface
  - Prod, dev support and trainings paid

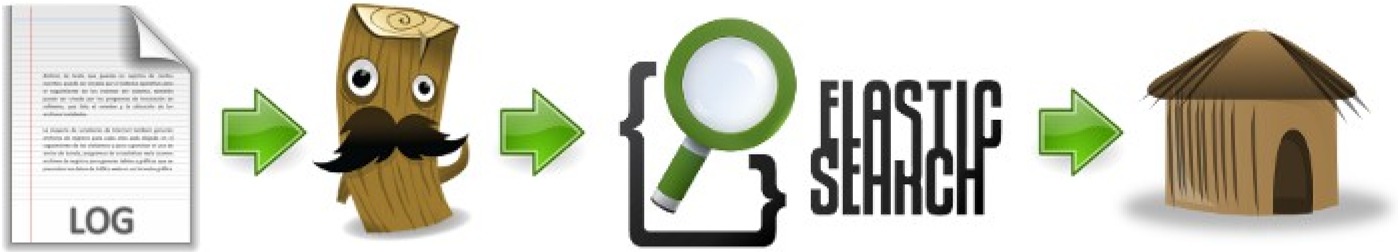
ELK is of pretty new about google trends (since 2013 is used)



Beta: Measuring search interest in *topics* is a beta feature which quickly provides accurate measurements of overall search interest. To measure search interest for a specific *query*, select the "search term" option.

ELK Stack?

- Elasticsearch
- Logstash
- Kibana



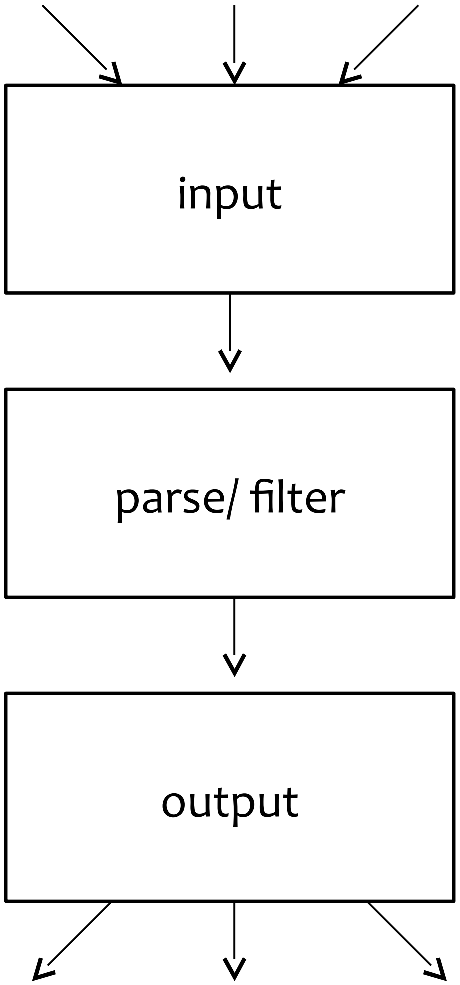
logstash

1. collect data
2. parse/ filter
3. send data

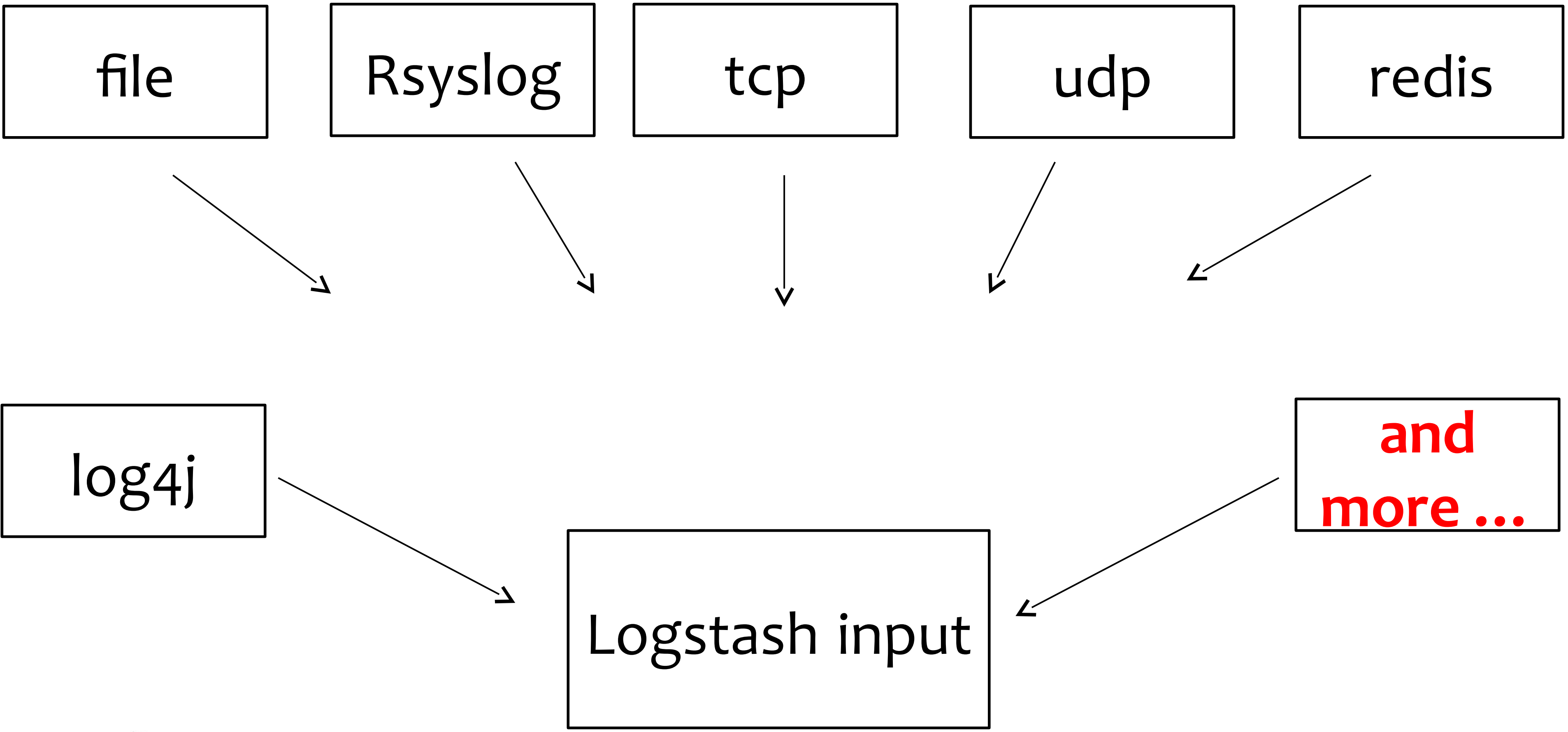
Logstash is part of the family of elasticsearch.

accenture

logstash architecture



1. collect data



Sample conf

When 1 input

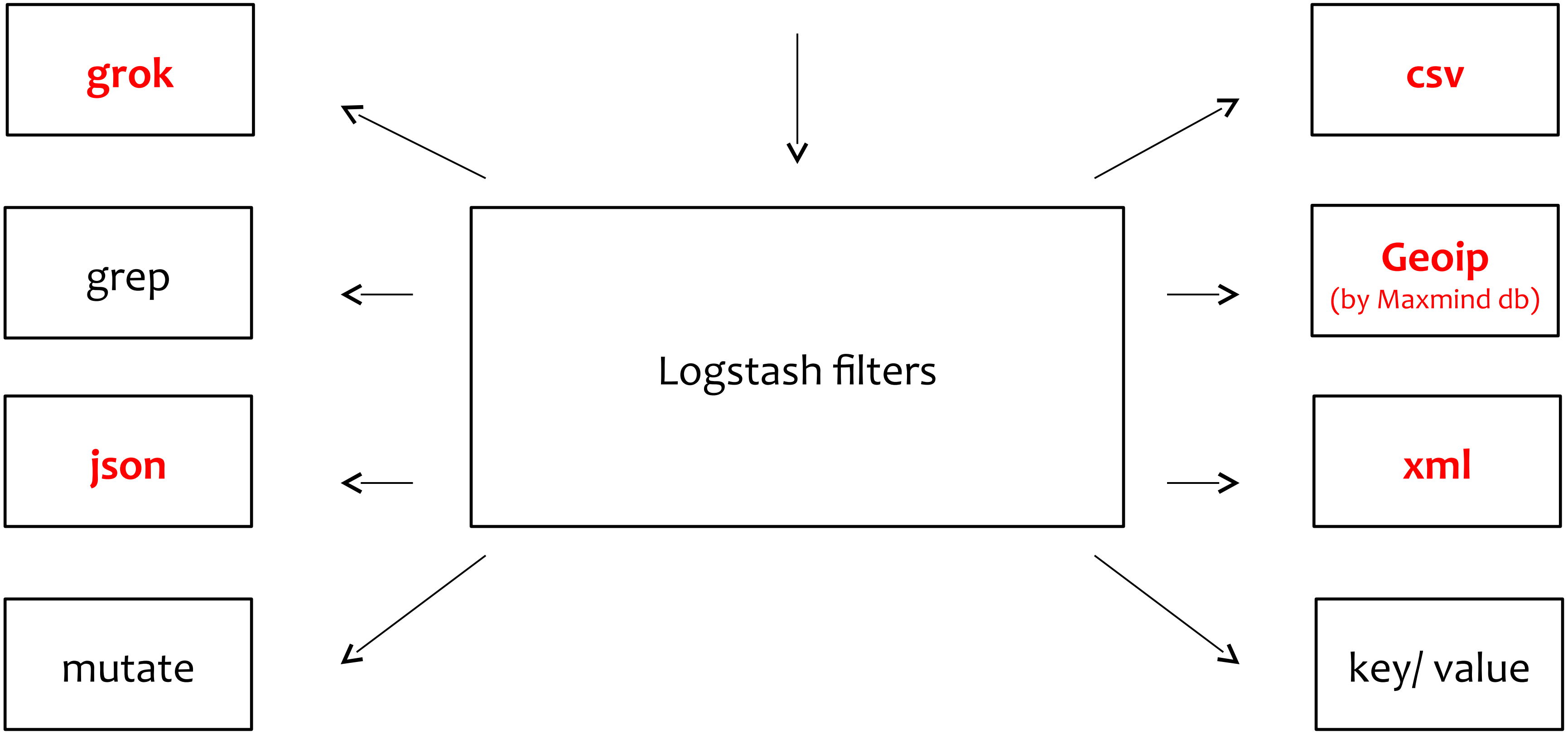
```
input{
  tcp{
    type=> “server1”
    host=> “192.168.1.1”
    port=> “5555”
     }
   }
```

When multiple inputs

```
 input{
  tcp{
    type=> “server1”
    host=> “192.168.1.1”
      port=> “5555”
      }
   }

  file{
type => "my-log"
path => [ "C:/dev/Log/*.log*" ]
   }
    .
    .
    .
```

2. parse/ filter



Logstash filters

grok

grep

json

mutate

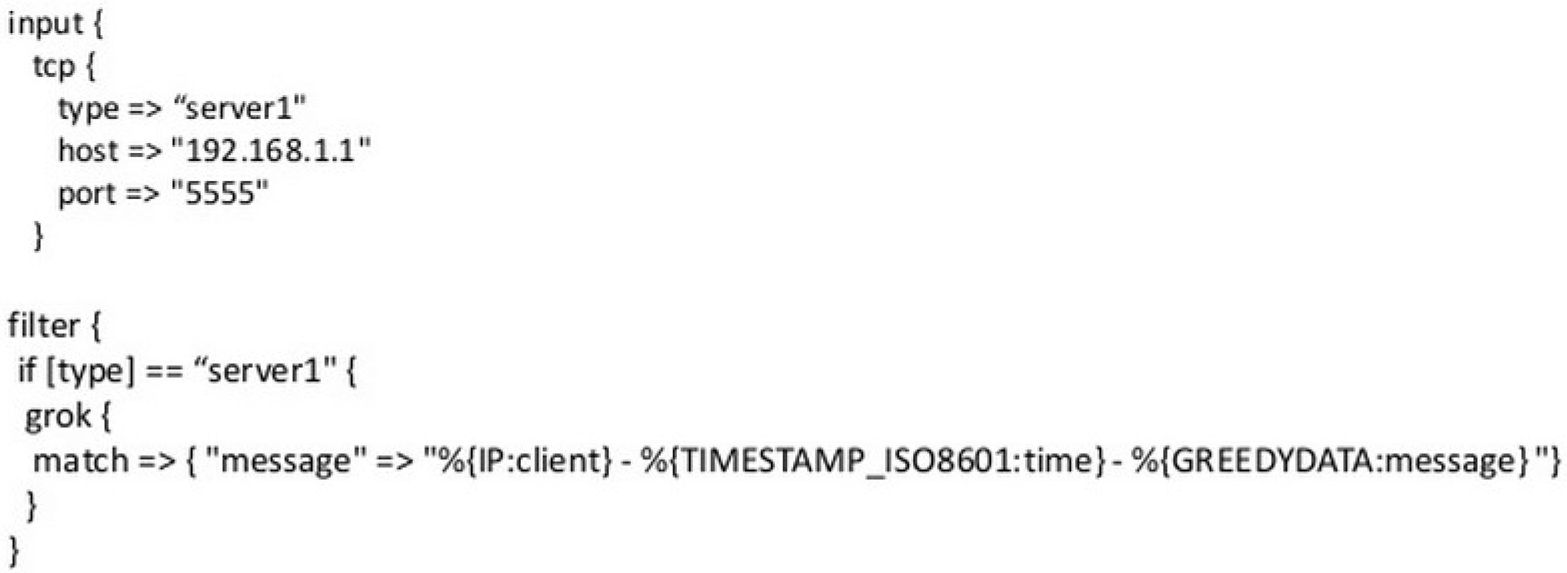
csv

Geoip
(by Maxmind db)

xml

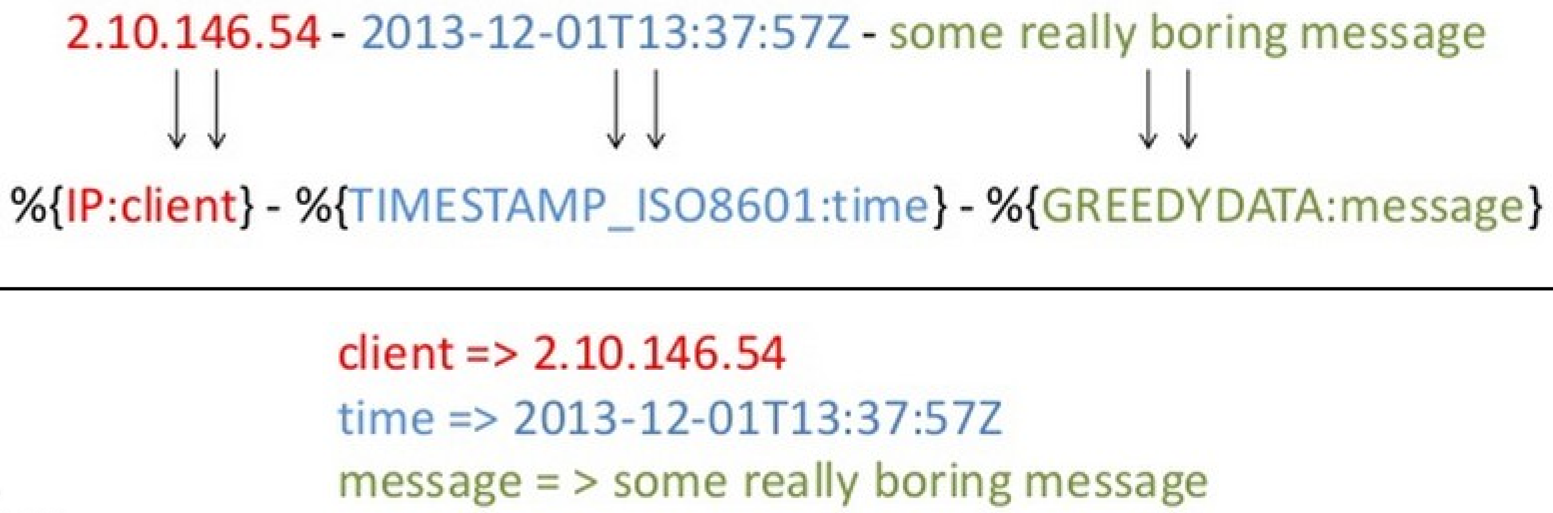
key/ value

Grok filter (example)



```
input {
  tcp {
    type => “server1"
    host => "192.168.1.1"
    port => "5555"
  }

filter {
 if [type] == “server1" {
  grok {
   match => { "message" => "%{IP:client} - %{TIMESTAMP_ISO8601:time} - %{GREEDYDATA:message} "}
  }
}
```



2.10.146.54 - 2013-12-01T13:37:57Z - some really boring message

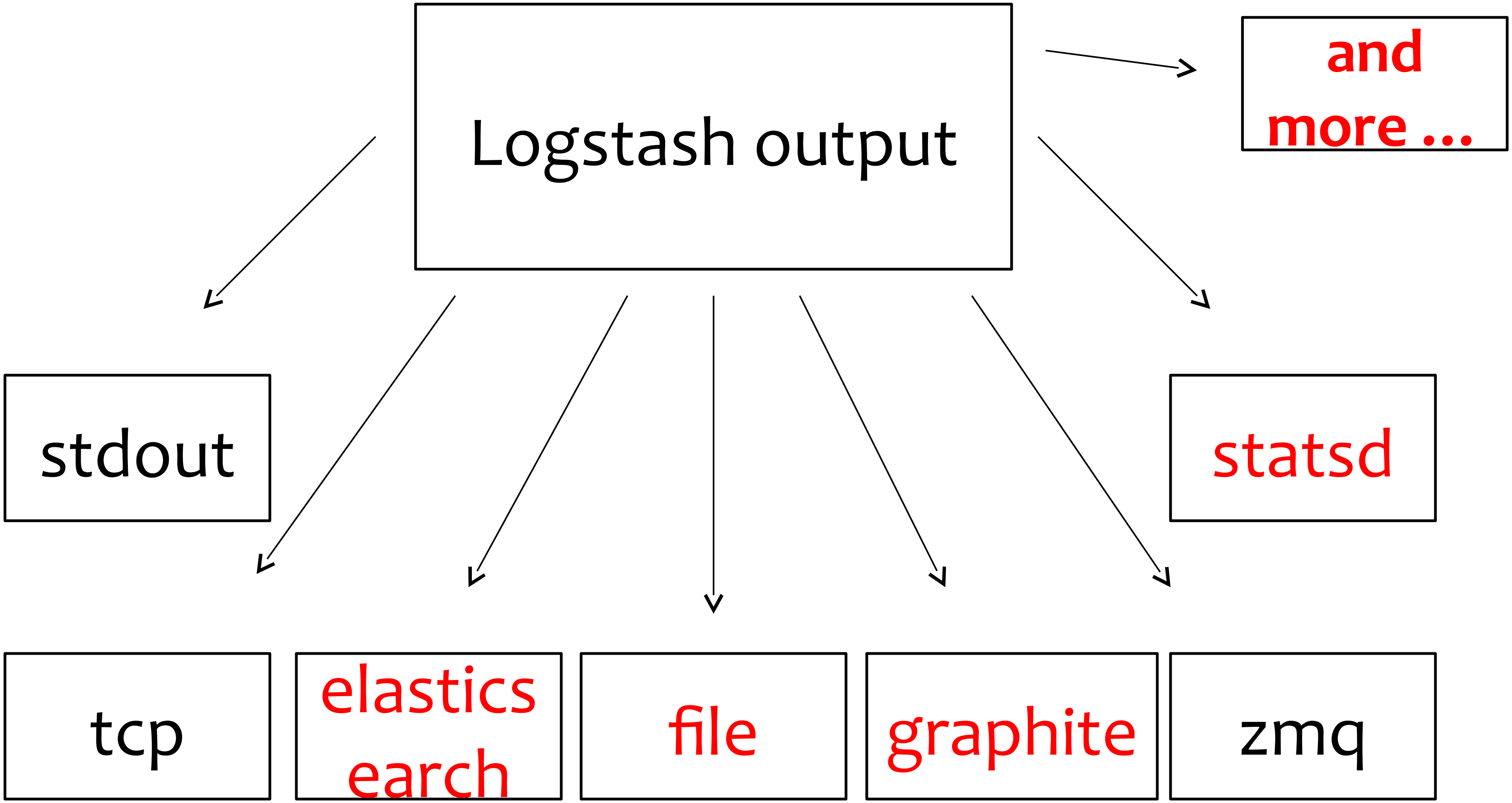
%{IP:client} - %{TIMESTAMP_ISO8601:time} - %{GREEDYDATA:message}

client => 2.10.146.54

time => 2013-12-01T13:37:57Z

message = > some really boring message

3. send data



Logstash output

and more …

stdout

statsd

tcp

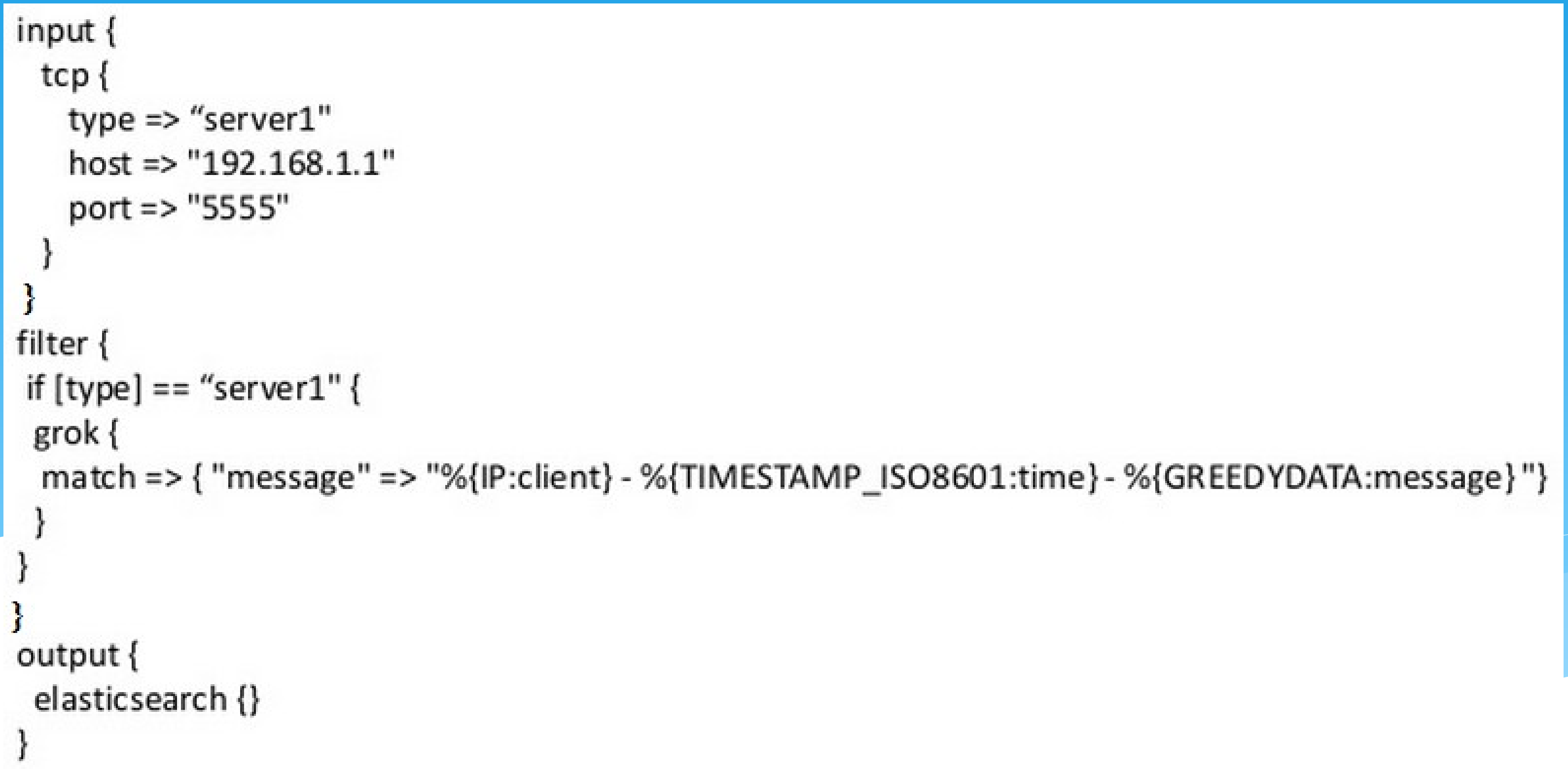
elastics
earch

file

graphite

zmq

logstash => elasticsearch sample



```
input {
  tcp {
    type => “server1"
    host => "192.168.1.1"
    port => "5555"
  }
}
filter {
 if [type] == “server1" {
  grok {
   match => { "message" => "%{IP:client} - %{TIMESTAMP_ISO8601:time} - %{GREEDYDATA:message} "}
  }
 }
}
output {
  elasticsearch {}
}
```

elasticsearch

Distributed RESTful
search server

1. JSON based REST API
2. Schema-less database
3. Indexes every single field
4. Full text search
5. Relational DB/ JSON document ("NoSQL" world)

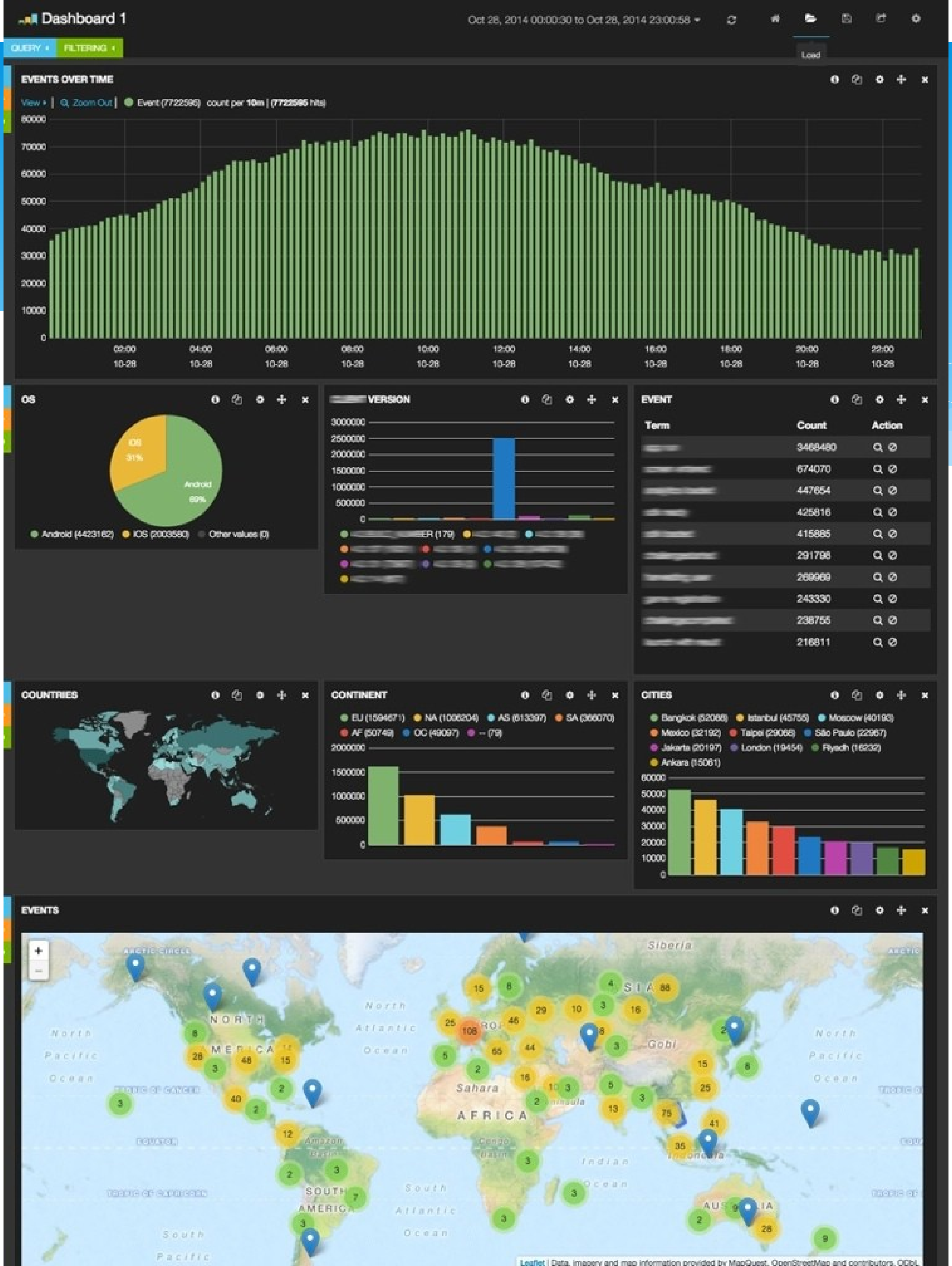
Kibana

Web UI for the logs

1. Clean and simple UI
2. Data discovery
2. Fully customizable
3. Boostrap based

Kibana is part of the family of elasticsearch.

Kibana when it is heavily set up

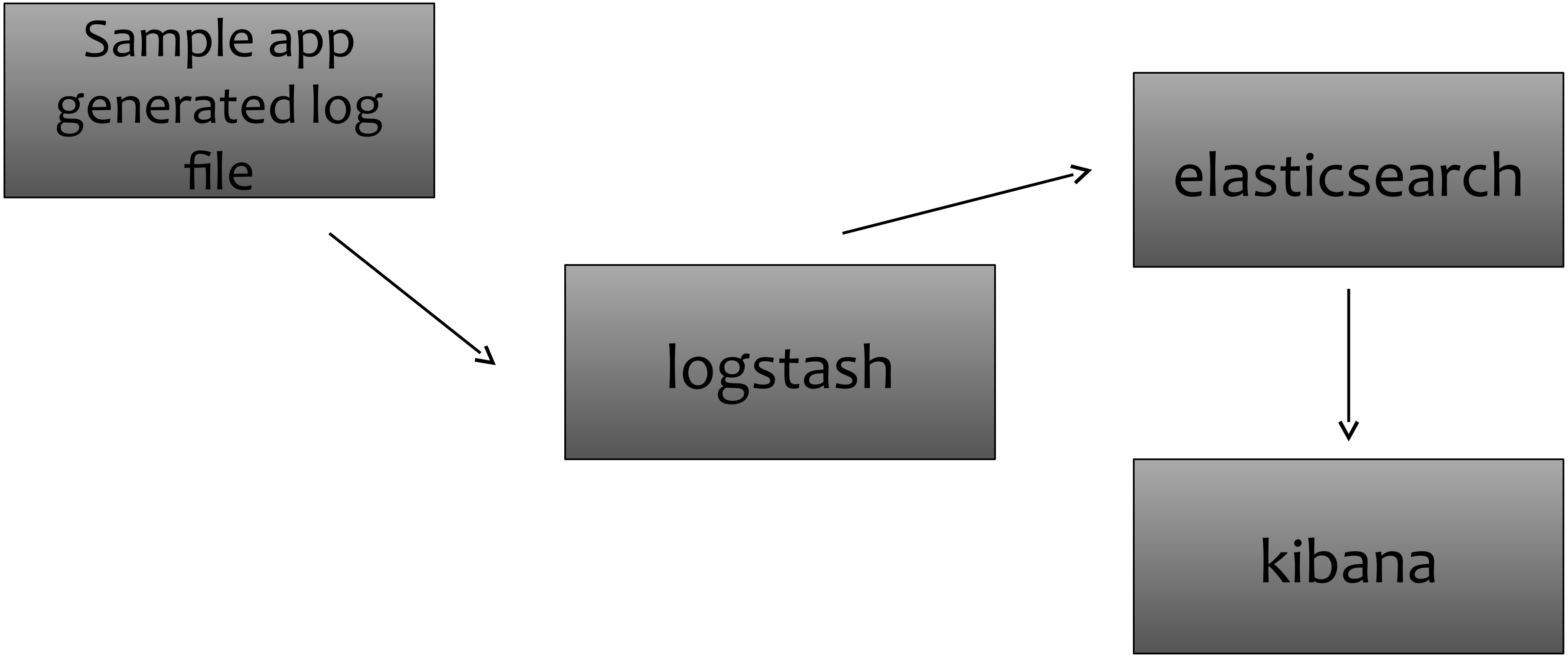


Term	Count	Action
[redacted]	3468480	🔍 ⊘
[redacted]	674070	🔍 ⊘
[redacted]	447654	🔍 ⊘
[redacted]	425816	🔍 ⊘
[redacted]	415885	🔍 ⊘
[redacted]	291798	🔍 ⊘
[redacted]	269969	🔍 ⊘
[redacted]	243330	🔍 ⊘
[redacted]	238755	🔍 ⊘
[redacted]	216811	🔍 ⊘

Getting started

1) Download Elasticsearch (ES)/ Logstash/ Kibana to your computer. The download links on the "Sources" slide of the presentation*.
2) Simply run ES as is, worry about config later.
3) Follow logstash cookbook to get started.
4) Setup some inputs.
5) Install kibana plugin in ES.
6) Open your browser and type "host:port" where kibana is running and try out the fresh log server.

Demo scenario



Most asked/ voted Q/ As on Quora*

Who are the biggest direct competitors to Splunk?

- “ELK is a free alternative to Splunk. Needless to say, the officiall support ELK (Elasticsearch, Logstash, Kibana) stack is an open-source alternative to Splunk's log-forwader/indexer/dashboard combo.”

Can Elasticsearch be used to replace your business's existing business intelligence system?

- “Works pretty well but it has a downside, the security shield is still very nascent but also the releases are coming quickly so it is improved over the time.”

What are the most latest recommended tools and technologies for real time analysis and visualization using Twitter data?

- “The ELK (ElasticSearch) stack is an open source option to do real time search on Twitter data. Logstash has a plugin for Twitter that can be used to collect, parse and store the data.”

***www.Quora.com** is a question-and-answer website where (**mostly IT**) questions are created, answered, edited and organized by its community of users. It had around **50 Million visitors in Jan 2015.

Q/ A

accenture

Sources

* ELK stack tools to download- http://www.elasticsearch.org/overview/elkdownloads/

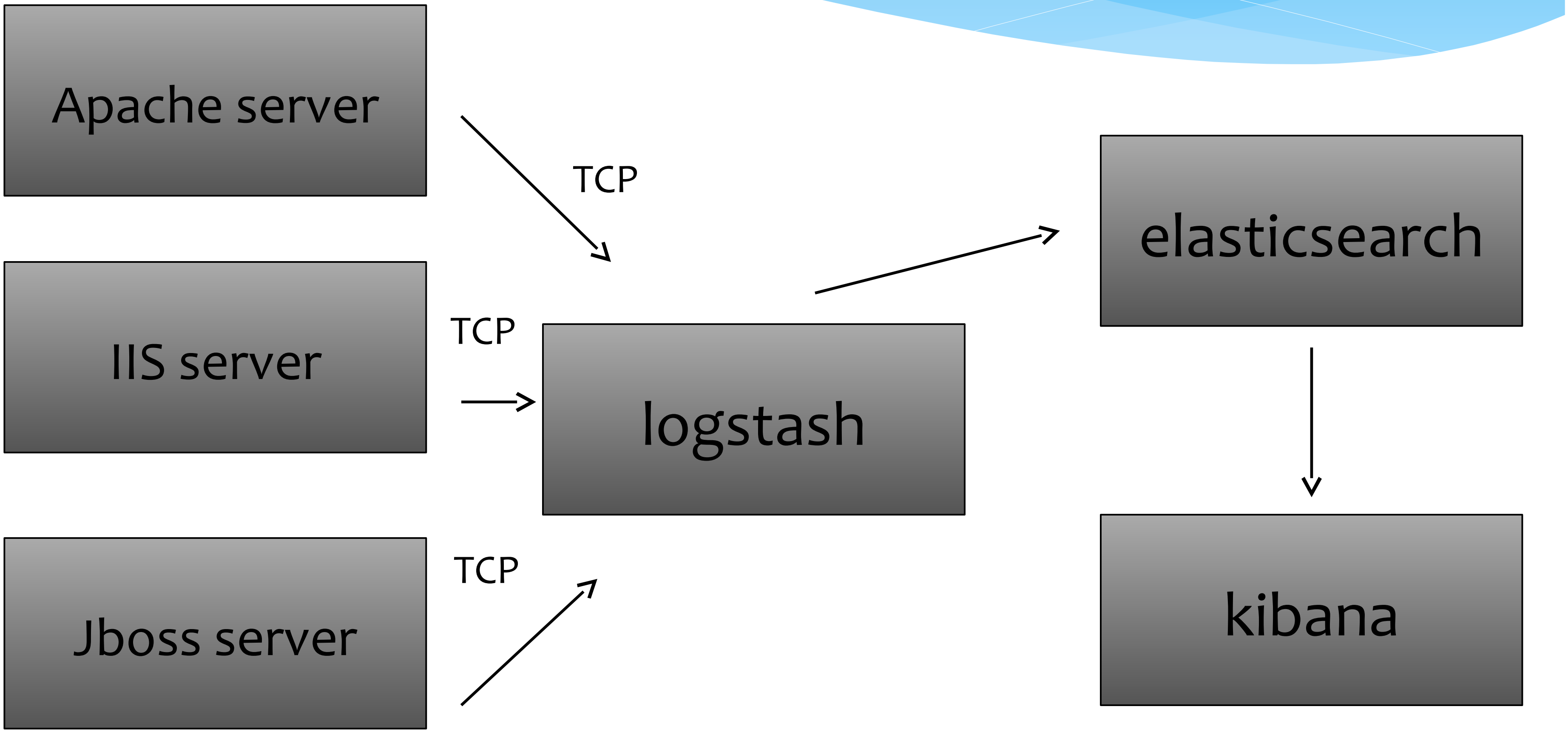
Installation guide for Windows - https://community.ulyaoth.net/threads/how-to-install-logstash-on-a-windows-server-with-kibana-in-iis.17/

Installation guide for Linux- http://everythingshouldbevirtual.com/highly-available-elk-elasticsearch-logstash-kibana-setup

Logstash documentation- http://logstash.net/docs/1.4.2/
Kibana documentation- http://www.elasticsearch.org/guide/en/kibana/current/index.html
Elasticsearch documentation- http://www.elasticsearch.org/guide/

*’’Quora number of visitors- http://www.similarweb.com/website/quora.com

Multiple schema example (no demo)


Apache server
TCP
IIS server
TCP
logstash
elasticsearch
Jboss server
TCP
kibana