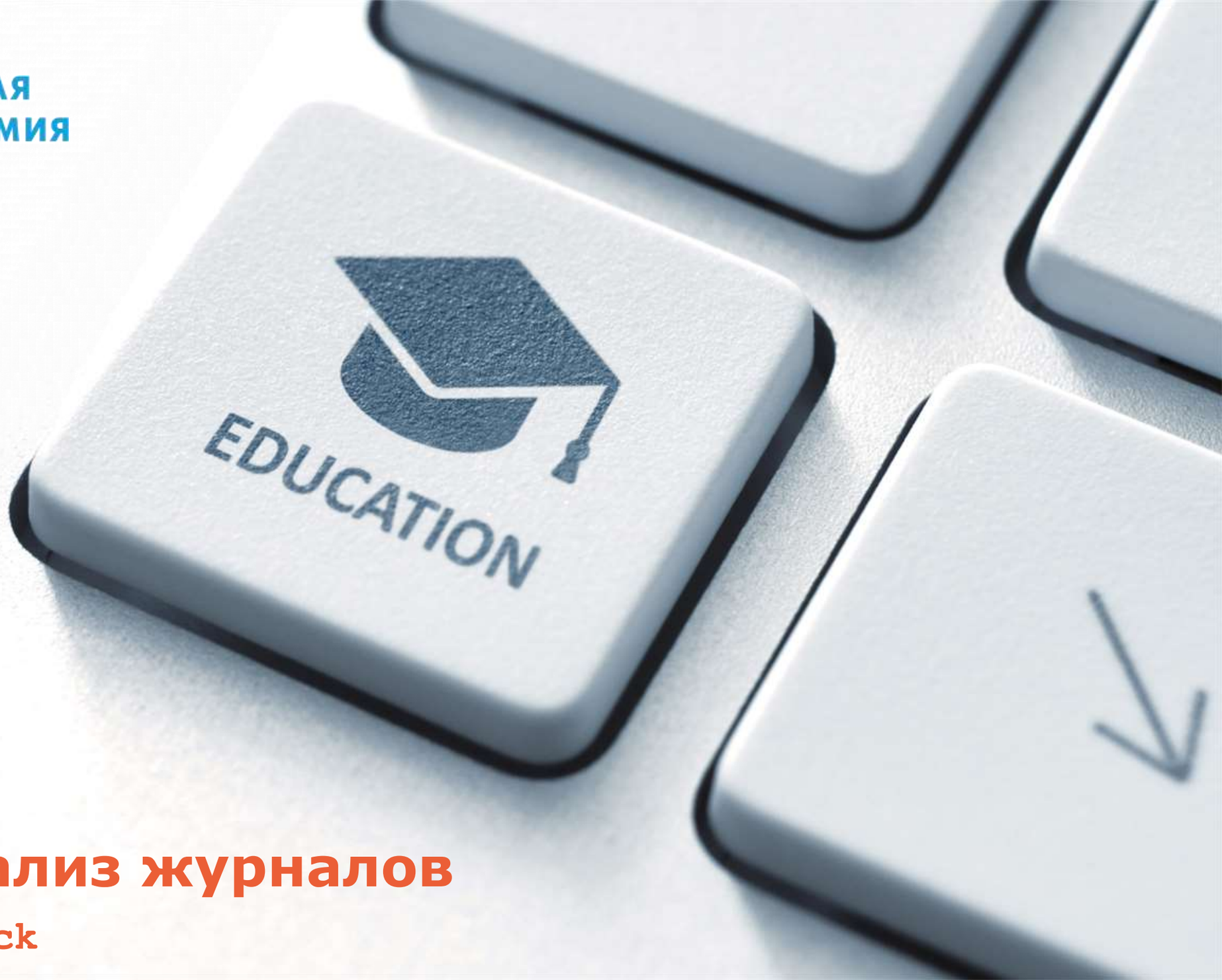
СЕТЕВАЯ
АКАДЕМИЯ
ЛАНИТ




# Сбор и анализ журналов

# Elastic stack

# ПЛАН НА СЕГОДНЯ

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

- О сборе журналов
- ELK stack: ElasticSearch, Logstash, Kibana

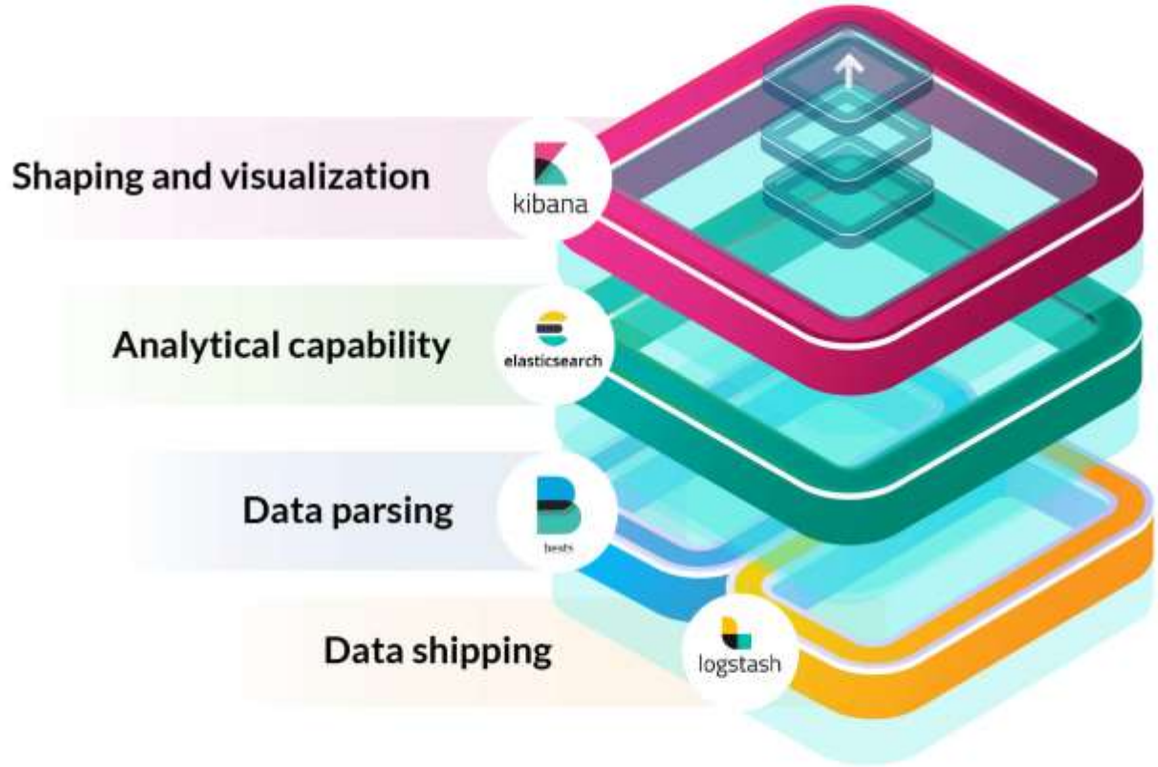
**Журналирование / Логирование:**

Процесс ведения файлов, которые содержат системную информацию о работе сервера или сервиса и события, такие как действия пользователя или приложения.

# ELK в сборе



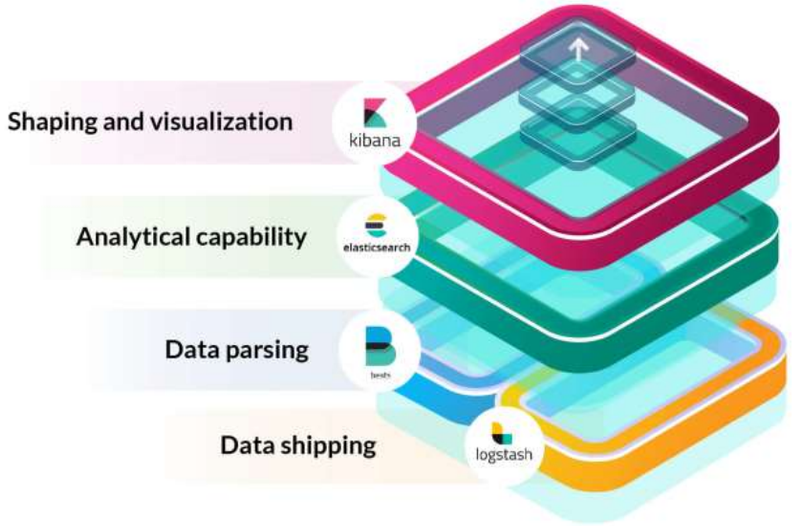
- Сбор и хранение журналов
- Сбор метрик
- Индексация
- Анализ
- Визуализация

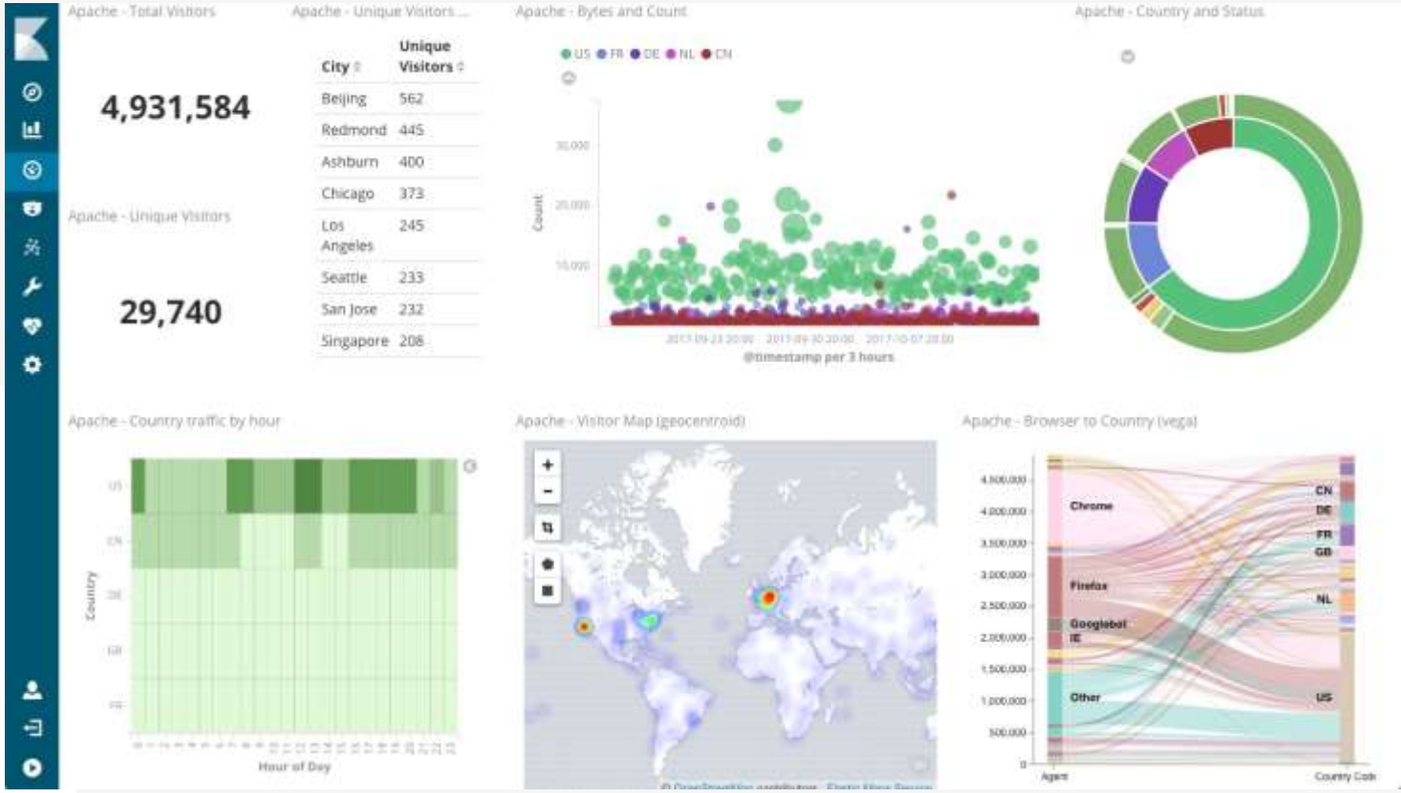


# ELK в сборе

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

- Сбор и хранение журналов
- Сбор метрик
- Индексация
- Анализ
- Визуализация




- Единый интерфейс для работы с журналами
- Анализ и визуализация его результатов
- Повышение эффективности работы специалистов
- Сокращение времени разрешения инцидентов

# Варианты использования Elastic Stack

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

- Управление и анализ журналов
- Сервис полнотекстового поиска
- Мониторинг производительности приложений
- SIEM – Управление информацией в целях безопасности
- Безопасность точек входа
- Машинное обучение

- Используется всеми публичными облачными провайдерами

# Shay Banon


2010 – **Elasticsearch 0.4**

2012 – компания **Elasticsearch BV**, начало платной поддержки

2014 – привлечено $104 млн **инвестиций**

2019 – изменение **лицензии (Apache 2.0)**

2021 – снова изменение **лицензии (ELv2)**

Актуальная версия: **8.13**

# Основные особенности ElasticSearch

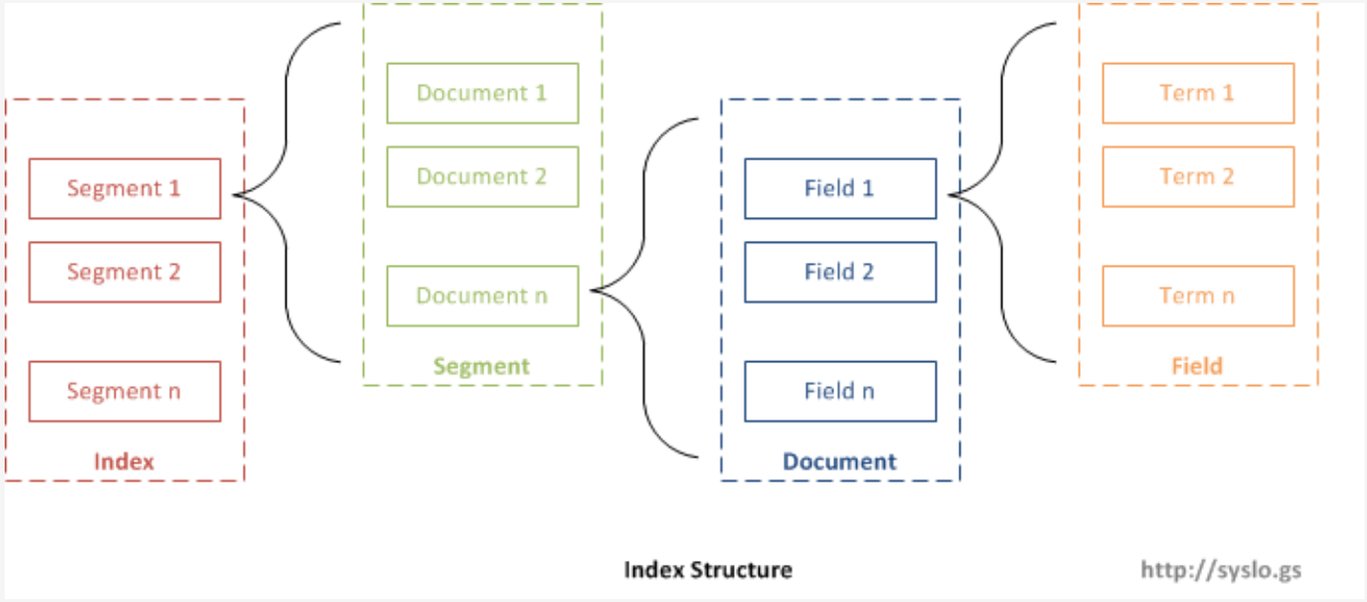
СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

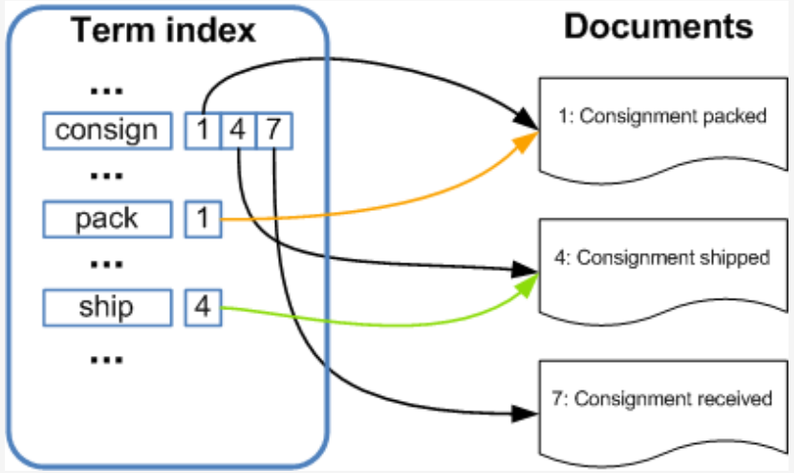
- Open Source*
- Документоориентированность
- Полнотекстовый поиск
- RESTful API
- JSON over HTTP
- Ответ в режиме, близком к реальному времени
- Распределенность
- Масштабируемость
- Высокая доступность

**Из коробки**

- Быстрый поиск документов
- Нечеткий поиск: suggesters
- Более 50 видов агрегации
- Гео-поиск: shape, bounding, distance, polygon
- и много чего другого.

# Основные понятия: обратный индекс




[С чего начинается Elasticsearch / Хабр (habr.com)]

# Основные понятия

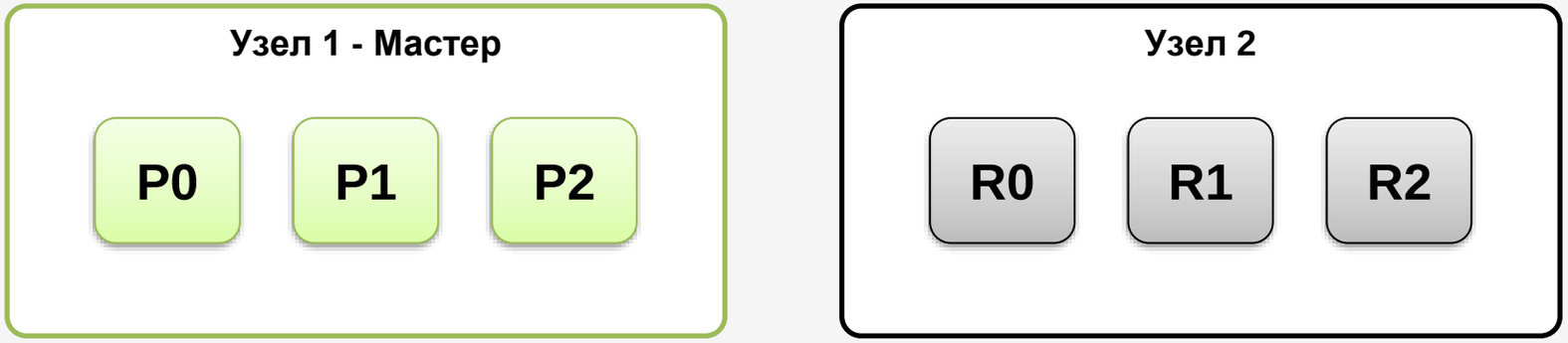


| В реляционных СУБД | ElasticSearch |
|---|---|
| База данных (Database) | Индекс (Index) |
| Таблица (Table) | Тип (Type) |
| Запись (Row) | Документ (Document) |
| Поле (Column) | Поле (Field) |
| Схема (Schema) | Отображение (Mapping) |
| SQL | DSL Query |

| В реляционных СУБД | ElasticSearch - метод | Действие |
|---|---|---|
| Select | GET | Чтение |
| Update | PUT | Обновление |
| Insert | POST | Вставка и обновление |
| Delete | DELETE | Удаление |

# Работа кластера

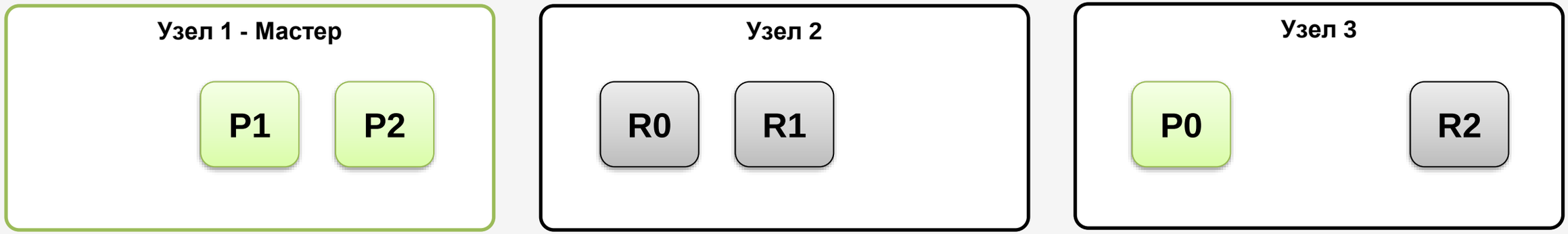
"index": { "number_of_replicas": 2 }



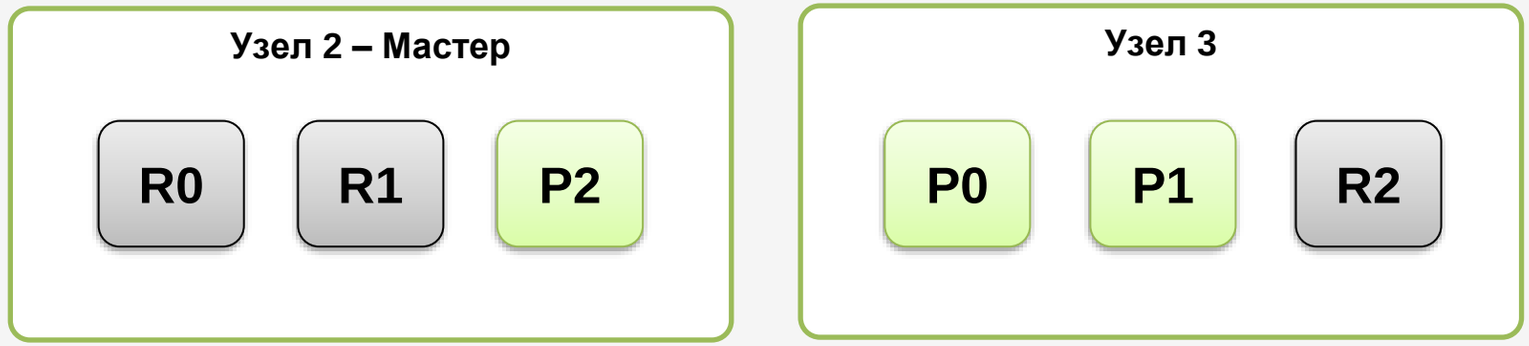

node.master: **true**
node.data: **true**

node.master: **false**
node.data: **true**

# Работа кластера



# Работа кластера



# Как выглядит поисковый запрос

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

```
curl -X GET "localhost:9200/my-index-000001/_search?pretty
```

```
GET /my-index-000001/_search?from=40&size=20
{
  "query": {
    "term": {
      "user.id": "kimchy"
    }
  }
}
```

```
POST my-index-000001/_search
{
  "query" : {
    "match": {
      "message": "tring out Elasticsearch"
    }
  },
  "suggest" : {
    "my-suggestion" : {
      "text" : "tring out Elasticsearch",
      "term" : {
        "field" : "message"
      }
    }
  }
}
```

# Обратить внимание

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

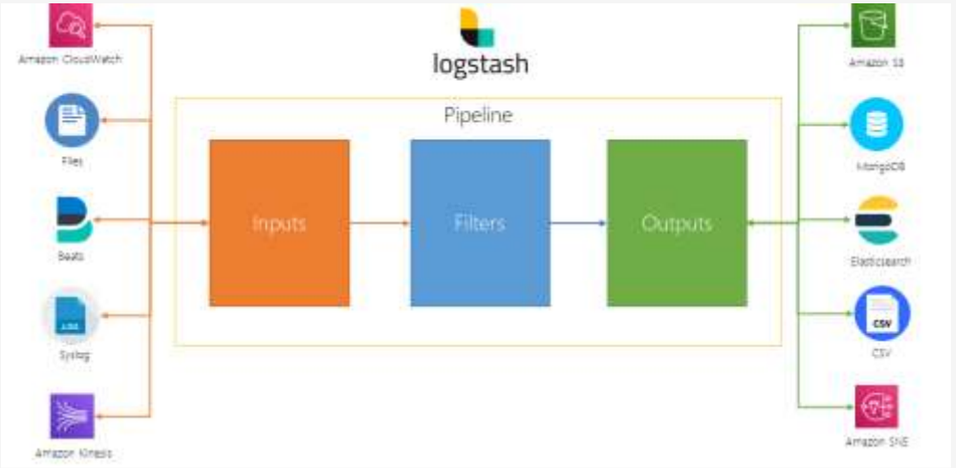
- Документации может быть недостаточно
- Документацию читать внимательно
- Маппинги нужны и как можно точнее, но в управляемом количестве
- Релевантность поиска зависит от количества записей на шарде
- Планирование нагрузки, железо
- Сюрпризы лицензии

Quickstart — Rally 2.4.0 documentation (esrally.readthedocs.io)

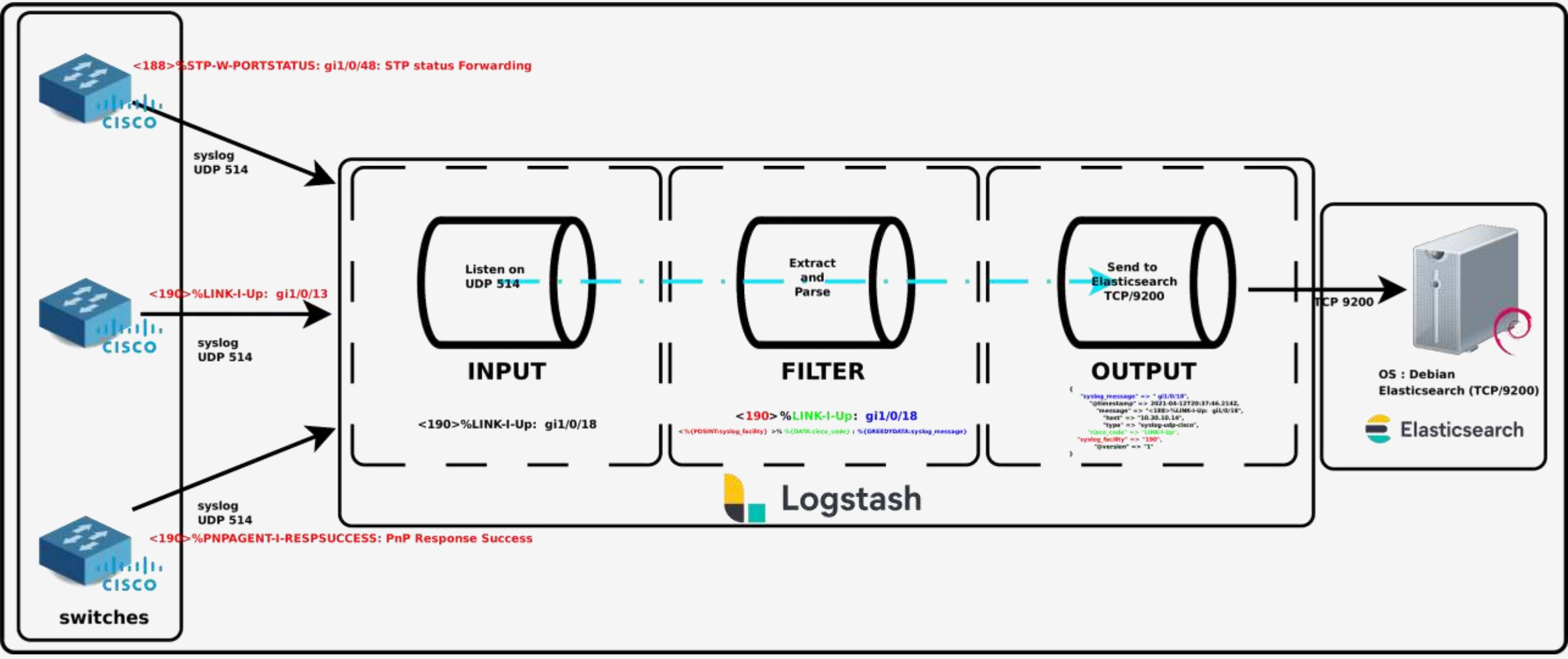
# Logstash



- Легковесный конвейер обработки данных
- Server-side
- Собирает данные из различных источников
- Преобразовывает на лету
- Отправляет данные на указанный вывод
- Огромные возможности по обработке журналов
- Помогает индексировать данные
- Open Source



# Logstash: архитектура обработки



# Logstash: структура файла конфигурации

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

```
input {
  ..
}
filter {
  ..
}
output {
  ..
}
```

Input plugins | Logstash Reference [8.2] | Elastic

Filter plugins | Logstash Reference [8.2] | Elastic

Output plugins | Logstash Reference [8.2] | Elastic

```
$ /usr/share/logstash/bin/logstash-plugin [list | update | install]
```

# Logstash: структура файла конфигурации

## Input

```
input {
  ..
}
filter {
  ..
}
output {
  ..
}
```

```
input {
  stdin {
    type => "rails"
  }
}
```

```
input {
  file {
    path => "/var/log/http.log"
  }
}
```

```
input {
  beats {
    port => 5044
    tags => "beats"
  }
}
```

```
input {
  generator {
    type => "generated"
  }
}
```

# The Beats family

All kinds of shippers for all kinds of data.

## Filebeat

Lightweight shipper for logs and other data

## Metricbeat

Lightweight shipper for metric data

## Packetbeat

Lightweight shipper for network data

## Winlogbeat

Lightweight shipper for Windows event logs

## Auditbeat

Lightweight shipper for audit data

## Heartbeat

Lightweight shipper for uptime monitoring

## Functionbeat

Serverless shipper for cloud data

# Logstash: структура файла конфигурации
## Filter

```
input {
  ..
}
filter {
  ..
}
output {
  ..
}
```

```
filter {
  csv {
      separator => ","
      columns => ["CustID", “CustName", “CustEmail", “CustCity"]
  }
  mutate{ convert => ["CustID" => "Integer"] }
}
```

# Logstash: структура файла конфигурации
## Filter

```
input {
  ..
}
filter {
  ..
}
output {
  ..
}
```

```
filter {
  multiline {
      pattern => "(?m)Started"
      negate => true
      what => "previous"
  }
  grok {
    match => [
      "message", "%{RAILS}"
    ]
  }
  if "_grokparsefailure" in [tags] {
    drop { }
  }
}
```

grok-patterns (github.com)

# Logstash: структура файла конфигурации

## Output

```
input {
  ..
}
filter {
  ..
}
output {
  ..
}
```

```
output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index => "index-%{+YYYY.MM.dd}"]
  }
  stdout {
    codec => rubydebug
  }
  file {
    path => "/tmp/output.log"
  }
}
```

# Logstash: пример файла конфигурации

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

```
input {
  stdin{}
}
filter {
  grok {
    match => { "message" => "%{COMBINEDAPACHELOG}" }
  }
  date {
    match => { "timestamp", "dd/MMM/yyyy:HH:mm:ss Z" }
}
output {
  elasticsearch {hosts => ["http://localhost:9200"]}
  stdout {codec => rubydebug}
}
```

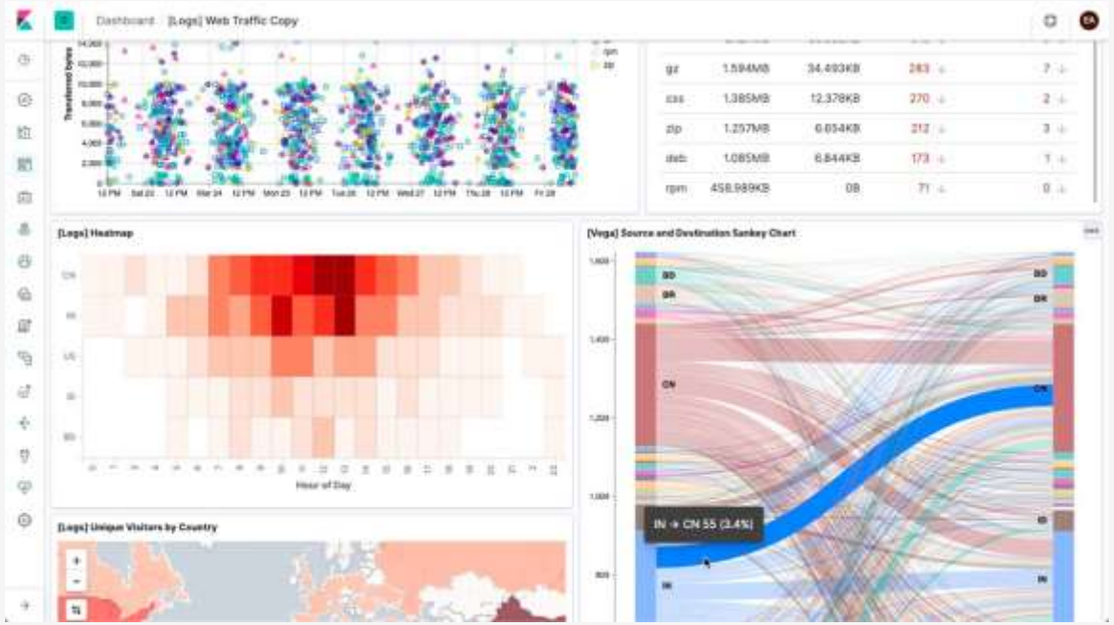
# Kibana


kibana

# Kibana

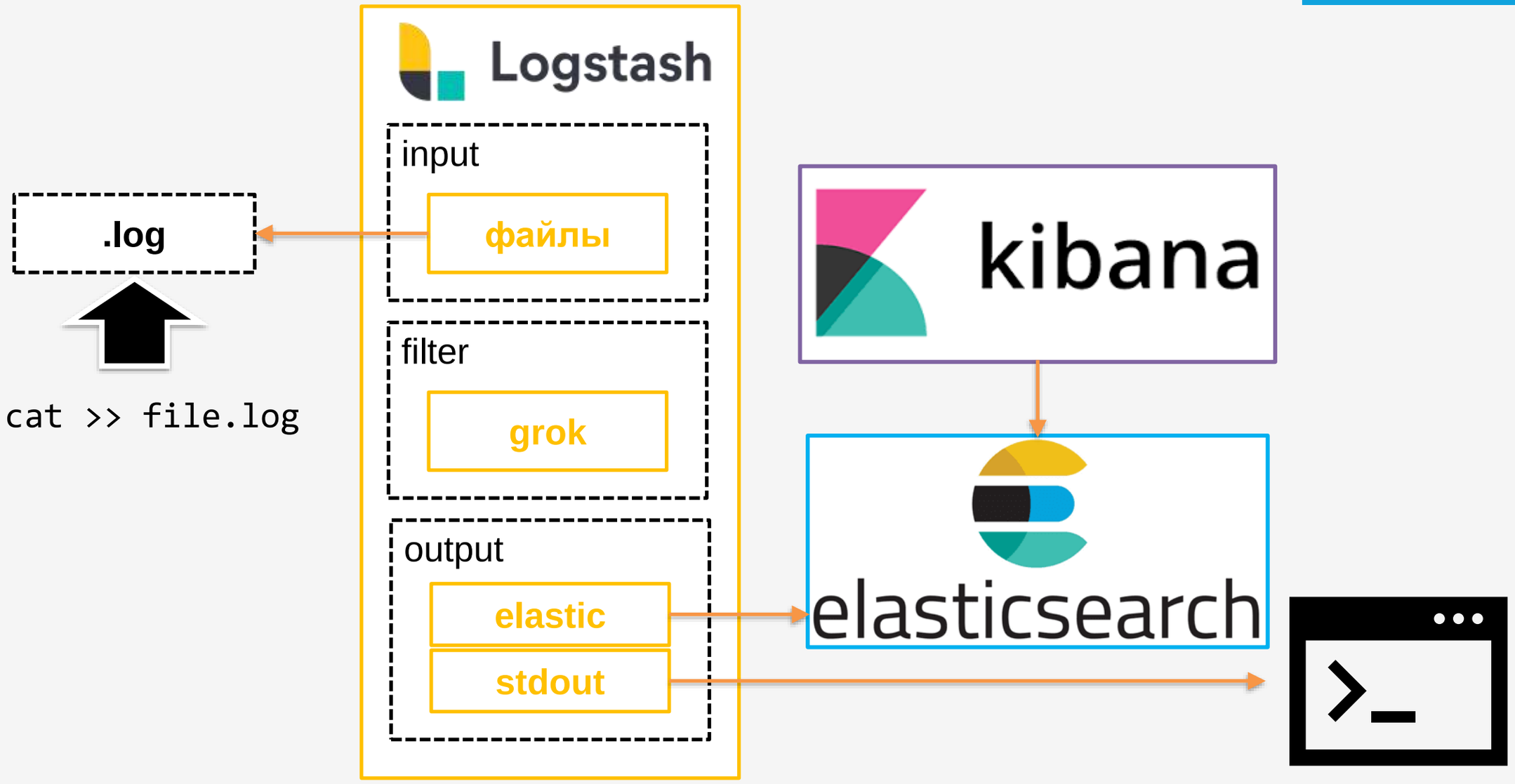
СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

- Инструмент визуализации и исследования собранных данных
- Мониторинг приложений
- Для анализа журналов или временных рядов
- Operational Intelligence
- Простые и мощные функции
- Множество вариантов визуализации
- Open Source



# Демо: схема




Logstash
input
файлы
.log
filter
grok
output
elastic
stdout
kibana
elasticsearch
cat >> file.log

# Ресурсы

СЕТЕВАЯ АКАДЕМИЯ ЛАНИТ

| | |
|---|---|
| ELK stack guide | https://www.elastic.co/guide |
| Репозиторий кода Elastic | https://github.com/elastic |
| Elastic CRUD | https://developer.ibm.com/tutorials/perform-crud-operations-with-databases-for-elasticsearch/ |
| Подробные инструкции по установке на Хабре | https://habr.com/ru/post/538840/ |
| Советы по преобразованию данных из логов в ELK Stack (grok) | https://habr.com/ru/post/509632/ |
| Open Distro (Amazon) | https://opendistro.github.io/for-elasticsearch/ |
| Grok Debugger | https://grokdebugger.com/ |
| Как настроить Elasticsearch, чтобы не было утечек | https://habr.com/ru/company/dataline/blog/487210/ |
| Elasticsearch Architecture Best Practices (Eric Westberg) | https://www.elastic.co/webinars/elasticsearch-architecture-best-practices |