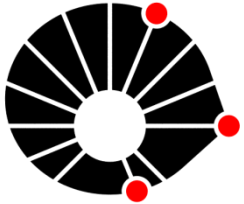
Elasticsearch, Logstash e Kibana (Pilha ELK)

Transformando dados brutos em informações poderosas de forma rápida e eficiente

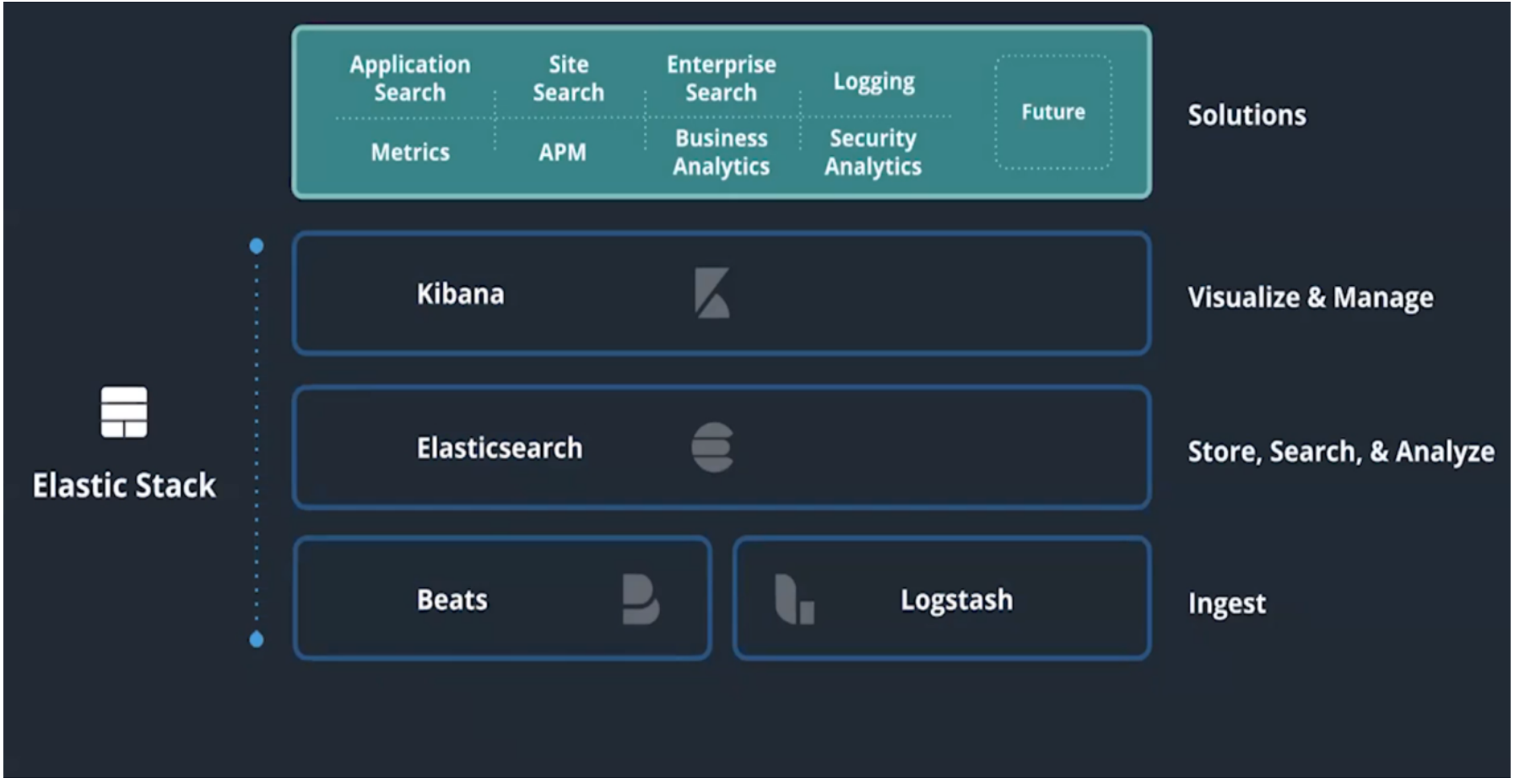


Equipes Conectividade e Data Center e Desenvolvimento de Sistemas

Agenda

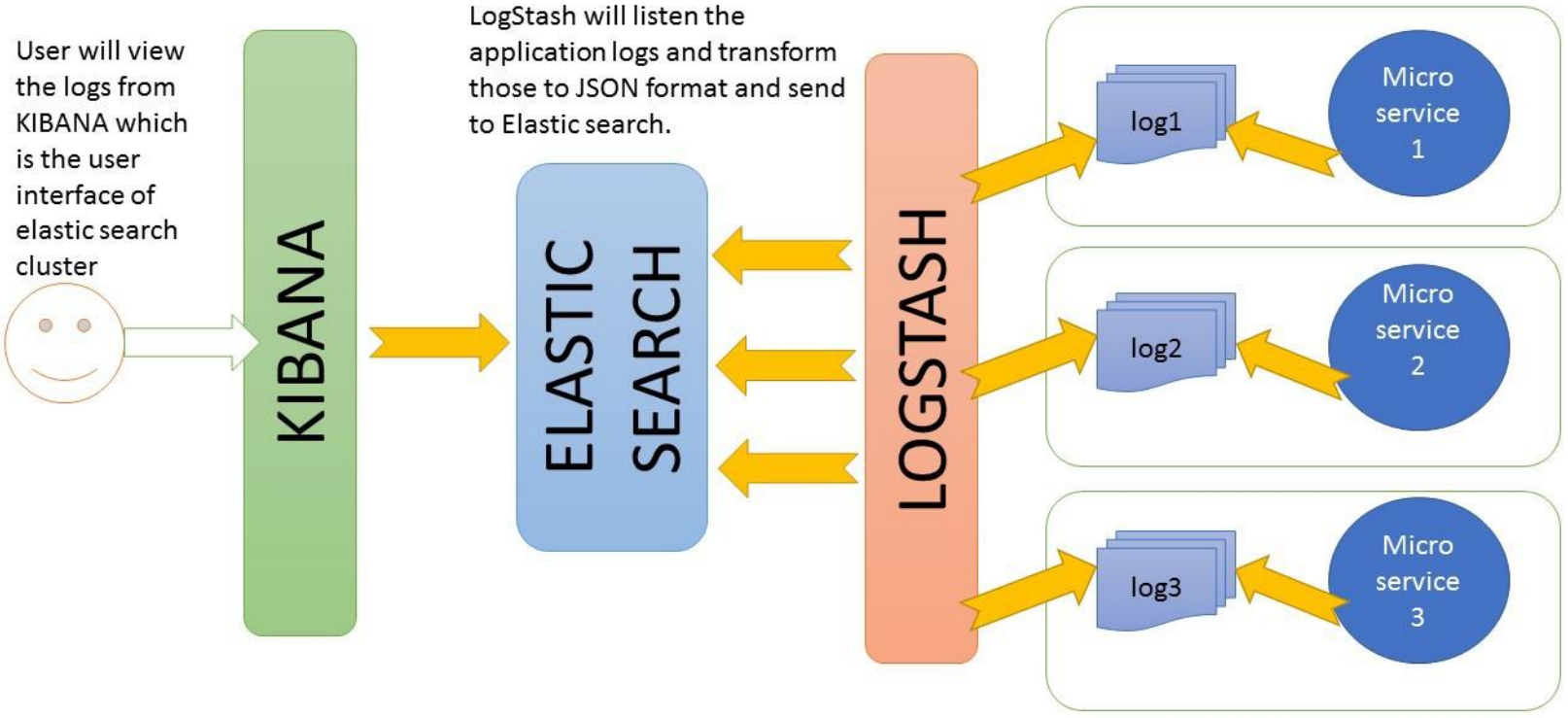
- Ferramentas que compõem a pilha ELK
- Cenário pré e pós implantação
- Desafios encontrados
- Demonstração prática no Kibana
- Recados
- Discussão geral

ELK – Visão geral



Fonte: https://www.elastic.co/

ELK – Visão geral



ELK stack interaction with different applications based on Log file

Fonte: https://howtodoinjava.com/microservices/elk-stack-tutorial-example/

Elasticsearch - características

- Servidor de buscas distribuído baseado no Apache Lucene.
- Armazena, centraliza e processa os dados
- Linguagem Java
- Open Source (licença apache 2.0)
- Desenvolvido em 2010 por Shay Bannon
- Escalável, suporta imensa quantidade de dados

Elasticsearch - características

- Consegue processar grandes quantidades de dados em tempo real, permitindo buscas instantâneas.
- Utilizado por empresas como Google, GitHub, Twitter entre outras.
- Possui outros recursos como geolocalização, analytics, plugins para machine learning etc.
- Versões p/ Linux, Windows e Mac

Elasticsearch - armazenamento

- Estrutura de armazenamento: JSON

Objeto JavaScript . Formato compacto, de padrão aberto, permite troca de dados de forma rápida e simples entre sistemas.

Documento

```
{
    "_index": "tickets",
    "_type": "doc",
    "_id": "14",
    "_score": 1,
    "_source": {
        "type": "ticket",
        "ticketKey": 14,
        "titulo": "Titulo 14",
        "dataDeCriacao": "2018-06-07 22:45:51"
    }
}
```

Elasticsearch – armazenamento

Forma como o registro é armazenado e disponibilizado no Elasticsearch

Table | JSON


```
{
  "_index": "logs",
  "_type": "doc",
  "_id": "MR6znGoBH6EnMjmD1d5Z",
  "_version": 1,
  "_score": null,
  "_source": {
    "message": "Interface ethernet 1/1/19, state up ",
    "@version": "1",
    "tag": "2e",
    "path": "/var/log/all_syslog_in_json.log",
    "type": "syslog",
    "@timestamp": "2019-05-09T13:06:31.000Z",
    "facility": "syslog",
    "level": "info",
    "host": "192.168.210.160",
    "program": "System"
  },
  "fields": {
    "@timestamp": [
      "2019-05-09T13:06:31.000Z"
    ]
  },
  "sort": [
    1557407191000
  ]
}
```


Table | JSON

@timestamp	09/05/2019, 10:06:31
t @version	1
t _id	MR6znGoBH6EnMjmD1d5Z
t _index	logs
# _score	-
t _type	doc
t facility	syslog
t host	192.168.210.160
t level	info
t message	Interface ethernet 1/1/19, state up
t path	/var/log/all_syslog_in_json.log
t program	System
t tag	2e
t type	syslog

Elasticsearch – busca

- Ao importar um documento, as frases são identificadas por um ID. As palavras então são separadas e alocadas em uma tabela, que contém os IDs associados e a frequência de repetição.
- Para cada documento JSON é gerado um hash (index invertido), baseado nessa tabela.
- Ao fazer uma busca, o processamento é muito rápido, pois o hash já contém a relação dos documentos que contemplam a pesquisa feita.

ID	Text
1	Baseball is played during summer months.
2	Summer is the time for picnics here.
3	Months later we found out why.
4	Why is summer so hot here

↑ Sample document data

Dictionary and posting lists →

Term	Freq	Document ids
baseball	1	[1]
during	1	[1]
found	1	[3]
here	2	[2], [4]
hot	1	[4]
is	3	[1], [2], [4]
months	2	[1], [3]
summer	3	[1], [2], [4]
the	1	[2]
why	2	[3], [4]

Fonte: https://www.mundipagg.com/blog/mandamentos-do-elasticsearch-em-producao/

Kibana

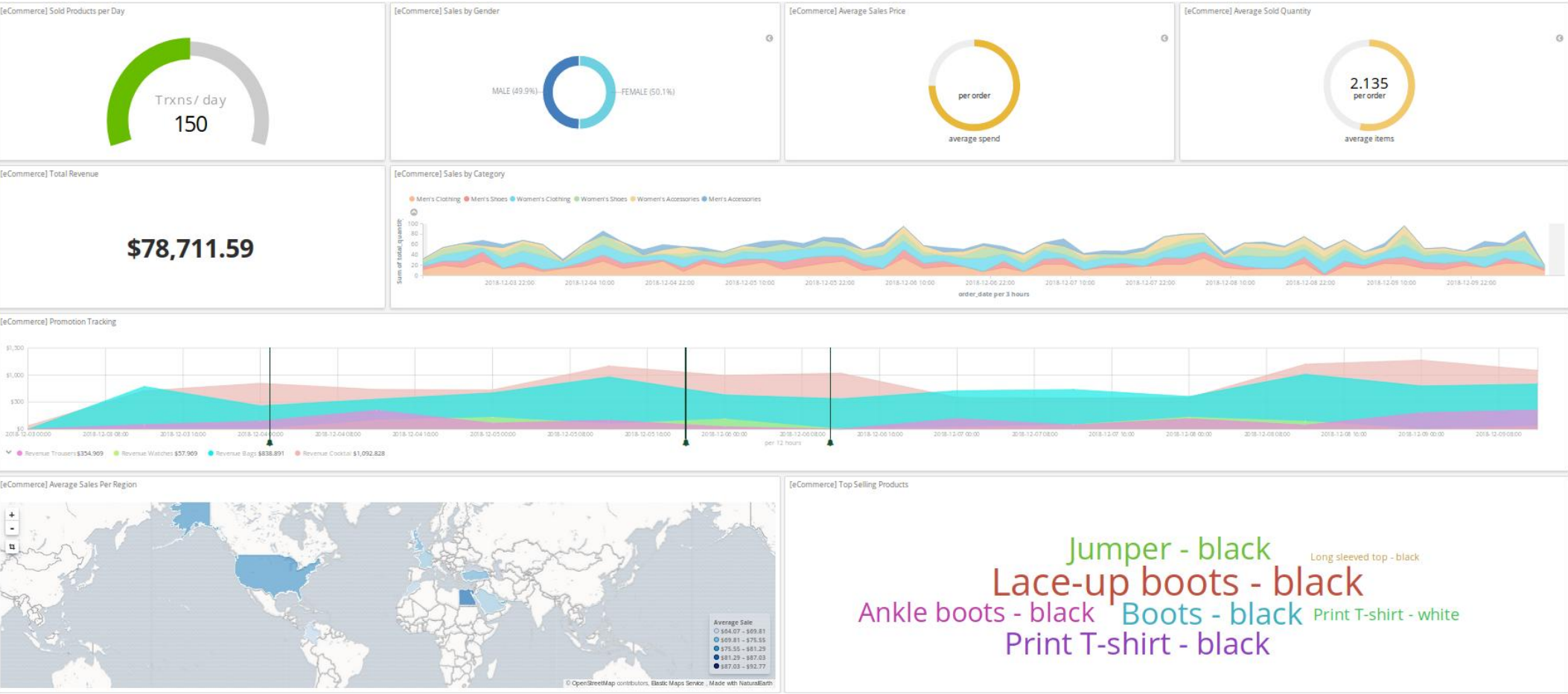
- Plataforma de análise e visualização dos dados armazenados no Elasticsearch
- Permite criar e compartilhar gráficos, dashboards, apresentações e relatórios em tempo real
- Opensource
- Excelente usabilidade e desempenho
- Versões p/ Windows, Linux e Mac

Kibana

- Dashboard baseado em dados de voos comerciais

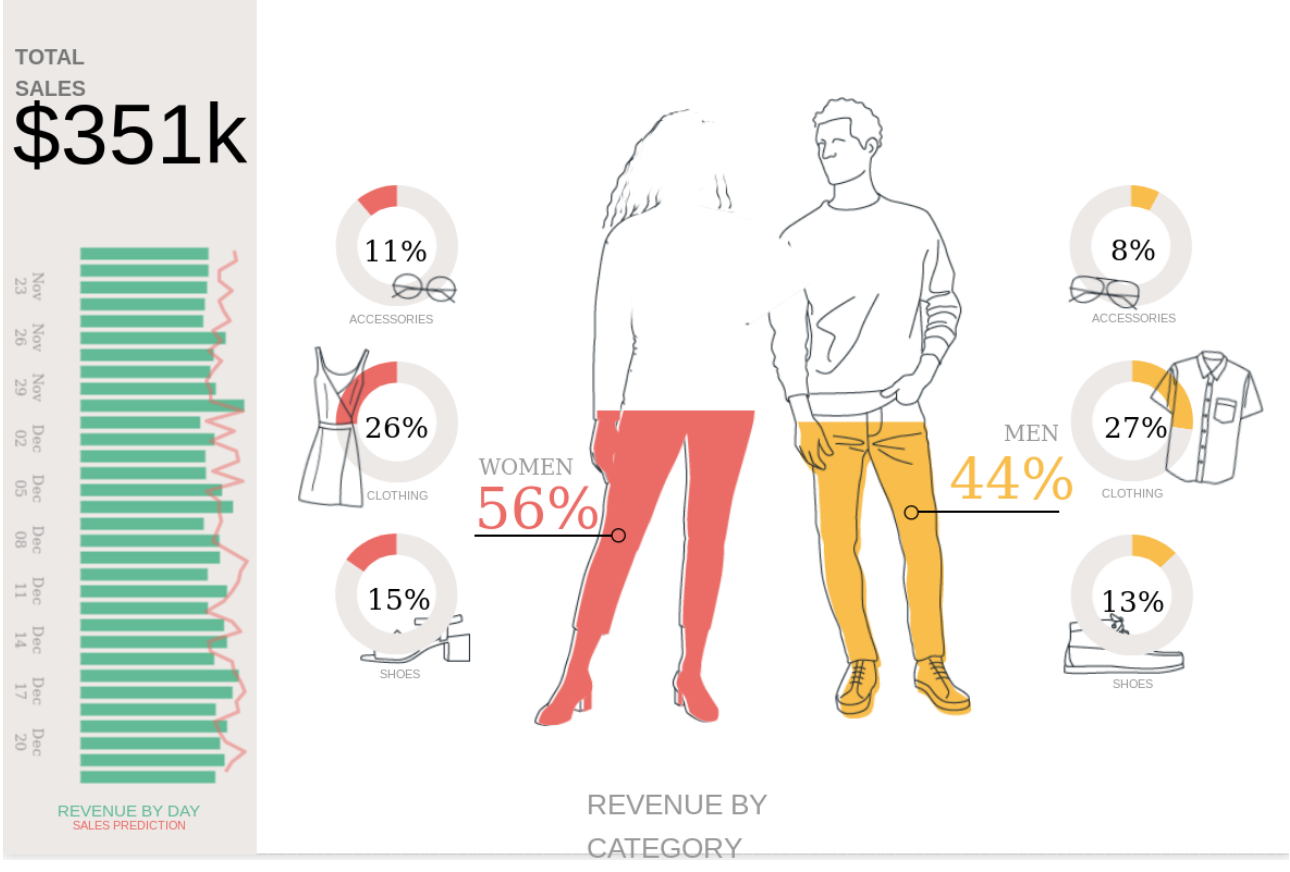
Kibana

- Dashboard baseado em dados de um e-Commerce



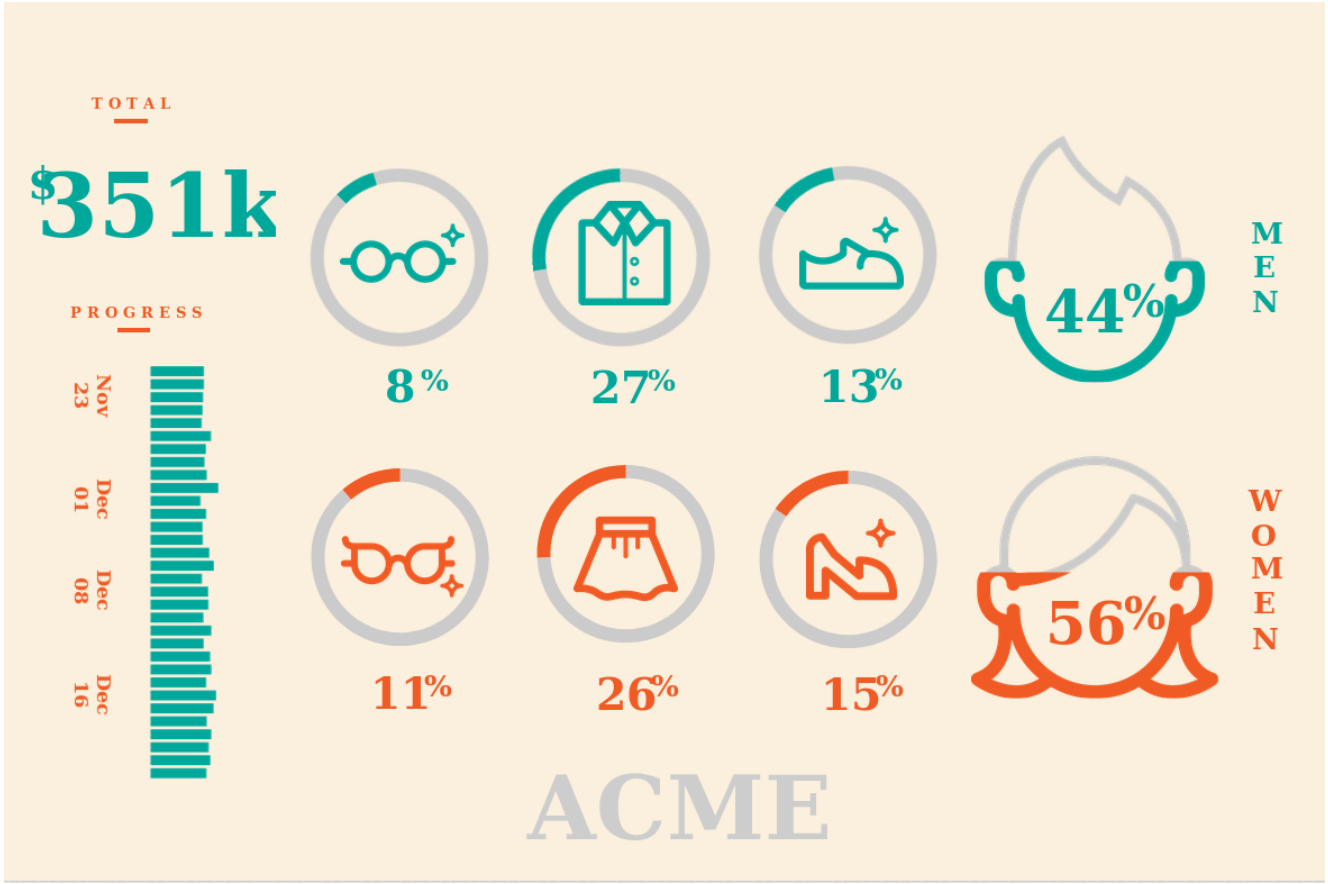
Kibana

- Modo Canvas (apresentação) alimentado com dados de um e-Commerce



Kibana

- Modo Canvas (apresentação) alimentado com dados de um e-Commerce



Logstash

- Poderosa ferramenta utilizada para coletar, modificar e enviar dados ao Elasticsearch
- Opensource
- Suporte inúmeros plugins
- Versões p/ Windows, Linux e Mac
- Pipeline: input => filter => output

Logstash - pipeline

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

Logstash – pipeline input

Entradas geram eventos.

- file
- syslog
- http
- beats
- github

+ infos em https://www.elastic.co/guide/en/logstash/6.7/input-plugins.html

Logstash – pipeline input

Exemplo:

```
input {
  file {
   path => "/var/log/all_syslog_in_json.log"
    start_position => "beginning"
    codec => "json"
   sincedb_path => "/var/log/db_for_watched_files.db"
    type => "syslog"
  }
}
```

Logstash – pipeline filter

Utilizados para modificar os dados.

- grok
- mutate
- drop
- clone
- geoip

+ infos em
https://www.elastic.co/guide/en/logstash/6.7/filter-plugins.html

Logstash – pipeline Filter

Os dados de entrada precisam estar no formato JSON? Não

Plugin Filter Grok: %{SYNTAX:SEMANTIC}

Exemplo de log: 55.3.244.1 GET /index.html 15824 0.043

```
input {
 file {
        path => "/var/log/http.log“
       }
}
filter {
     grok {
          match => { "message" => "%{IP:client} %{WORD:method} %{URIPATHPARAM:request} %{NUMBER:bytes} %{NUMBER:duration}" }
    }
}
```

Resultado: {“client”: “55.3.244.1”, “method”: “GET”, “ request”: “/index.html”, “bytes”: “15824”, “duration”: “0.043”}

Logstash – pipeline output

Encaminham os dados para o local definido

- elasticsearch
- S3 (Amazon)
- zabbix
- mail
- mongodb

+ infos em
https://www.elastic.co/guide/en/logstash/6.7/output-plugins.html

Logstash – pipeline output

Exemplo:

```
output {
if [type] == "syslog" {
      elasticsearch {
      hosts => [ "143.106.221.91:9200" ]
      index => "logs"
        }
    }
    #stdout {
       # codec => rubydebug
     #}
}
```

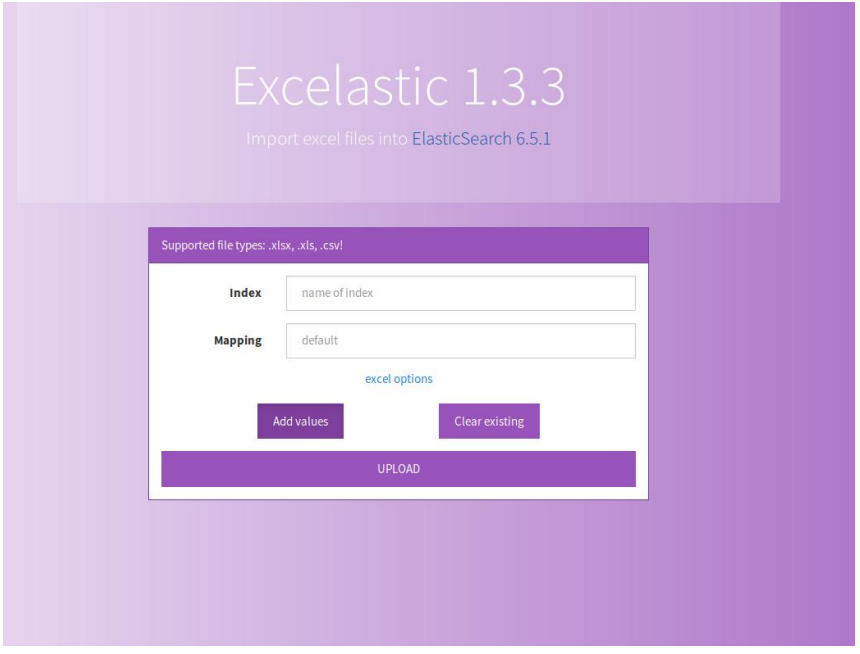
Importação de tabelas com Excelastic

Plugin Excelastic
https://github.com/codingchili/excelastic/releases

Permite a importação de arquivos com extensões .xls, .xlsx e .csv

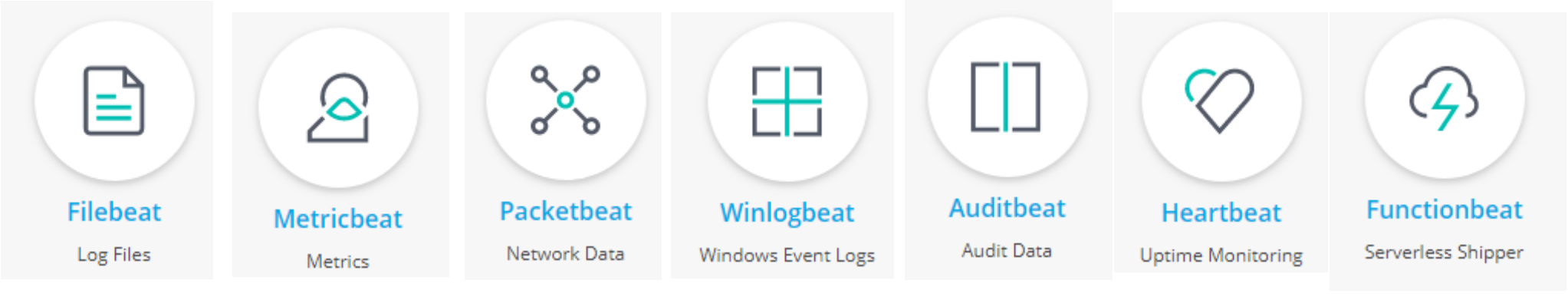
Para iniciar: java -jar excelastic-1.3.4.jar

Serviço roda na porta 9999



Beats

- Agentes que coletam métricas, como tráfego de rede, logs de Windows ou Linux etc.
- Encaminham os dados para o Logstash ou Elasticsearch
- Permitem coletar informações de vários serviços, como Apache, DNS, Bancos de dados, NFS etc.



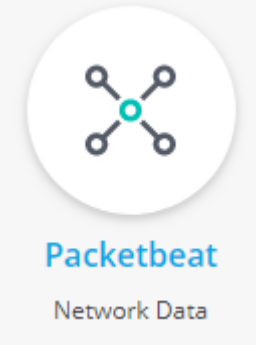
Beats - Configuração

```
packetbeat.protocols:

- type: dhcpv4
  ports: [67, 68]

- type: dns
  ports: [53]

- type: http
  ports: [80, 8080, 8081, 5000, 8002]

- type: memcache
  ports: [11211]

- type: mysql
  ports: [3306,3307]

- type: pgsql
  ports: [5432]

- type: redis
  ports: [6379]

- type: thrift
  ports: [9090]

- type: mongodb
  ports: [27017]

- type: cassandra
  ports: [9042]

- type: tls
  ports: [443, 993, 995, 5223, 8443, 8883, 9243]
```

```
output.elasticsearch:
  hosts: ["myEShost:9200"]
```

```
setup.kibana:
  host: "mykibanahost:5601"
```

https://www.elastic.co/guide/en/beats/packetbeat/current/packetbeat-configuration.html



Beats – Dashboard Kibana

Comparação das licenças

+ informações em
https://www.elastic.co/pt/subscriptions

	FREE		GOLD	PLATINUM
	OPEN SOURCE	BASIC		
	Download		Request Info	Request Info
ELASTIC STACK				
Elasticsearch				
Scalability & Resiliency	✓	✓	✓	✓
Query & Analytics	✓	✓	✓	✓
Data Enrichment	✓	✓	✓	✓
Management & Tooling	✓	✓	✓	✓
Security			✓	✓
Alerting			✓	✓
Machine Learning				✓
SUPPORT				
Support coverage			Business hours	24/7/365
Response times			Critical: 4 hrs L2: 1 day L3: 2 days	Critical: 1 hr L2: 4 hrs L3: 1 day
Unlimited # of incidents			✓	✓
# of support contacts			6	8
Web and phone support			✓	✓
Emergency patches				✓

Cenário IB – pré implementação

Servidor Syslog

```
ARUBA  BROCADE  iDRAC  SERVICOS  SRV-LINUX  SRV-WIN  SWITCHES  WIN-ADs  XENSERVER
```

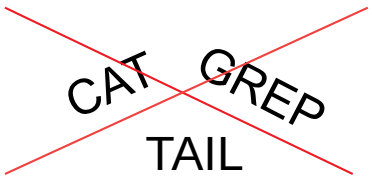
Cenário IB – pré implementação

```
root@srv-syslog38:/var/log/000-REMOTE_LOGS/SRV-LINUX# ls
143.106.179.6.log                192.168.240.26.log-20180103.gz  192.168.240.58.log-20180331.gz  192.168.240.70.log-20180303.gz          intranet.ib.unicamp.br.log-20180908.gz      srv-duplicati.ib.unicamp.br.log-20180821.gz
143.106.179.6.log-20181116.gz    192.168.240.26.log-20180104.gz  192.168.240.58.log-20180401.gz  192.168.240.70.log-20180304.gz          intranet.ib.unicamp.br.log-20180909.gz      srv-duplicati.ib.unicamp.br.log-20180822.gz
143.106.221.68.log               192.168.240.26.log-20180105.gz  192.168.240.58.log-20180402.gz  192.168.240.70.log-20180305.gz          intranet.ib.unicamp.br.log-20180910.gz      srv-duplicati.ib.unicamp.br.log-20180823.gz
143.106.221.68.log-20180720.gz   192.168.240.26.log-20180106.gz  192.168.240.58.log-20180403.gz  192.168.240.70.log-20180306.gz          intranet.ib.unicamp.br.log-20180911.gz      srv-duplicati.ib.unicamp.br.log-20180824.gz
143.106.221.69.log               192.168.240.26.log-20180107.gz  192.168.240.58.log-20180404.gz  192.168.240.70.log-20180307.gz          intranet.ib.unicamp.br.log-20180912.gz      srv-duplicati.ib.unicamp.br.log-20180825.gz
143.106.221.69.log-20180720.gz   192.168.240.26.log-20180108.gz  192.168.240.58.log-20180405.gz  192.168.240.70.log-20180308.gz          intranet.ib.unicamp.br.log-20180913.gz      srv-duplicati.ib.unicamp.br.log-20180826.gz
143.106.221.70.log               192.168.240.26.log-20180109.gz  192.168.240.58.log-20180406.gz  192.168.240.70.log-20180309.gz          intranet.ib.unicamp.br.log-20180914.gz      srv-duplicati.ib.unicamp.br.log-20180827.gz
143.106.221.70.log-20180720.gz   192.168.240.26.log-20180110.gz  192.168.240.58.log-20180407.gz  192.168.240.70.log-20180310.gz          intranet.ib.unicamp.br.log-20180915.gz      srv-duplicati.ib.unicamp.br.log-20180828.gz
143.106.221.71.log               192.168.240.26.log-20180111.gz  192.168.240.58.log-20180408.gz  192.168.240.70.log-20180311.gz          intranet.ib.unicamp.br.log-20180916.gz      srv-duplicati.ib.unicamp.br.log-20180829.gz
143.106.221.71.log-20180720.gz   192.168.240.26.log-20180112.gz  192.168.240.58.log-20180409.gz  192.168.240.70.log-20180312.gz          intranet.ib.unicamp.br.log-20180917.gz      srv-duplicati.ib.unicamp.br.log-20180830.gz
143.106.221.80.log               192.168.240.26.log-20180113.gz  192.168.240.58.log-20180410.gz  192.168.240.70.log-20180313.gz          intranet.ib.unicamp.br.log-20180918.gz      srv-duplicati.ib.unicamp.br.log-20180831.gz
143.106.221.80.log-20180720.gz   192.168.240.26.log-20180114.gz  192.168.240.58.log-20180411.gz  192.168.240.70.log-20180314.gz          intranet.ib.unicamp.br.log-20180919.gz      srv-duplicati.ib.unicamp.br.log-20180901.gz
143.106.62.12.log                192.168.240.26.log-20180115.gz  192.168.240.58.log-20180412.gz  192.168.240.70.log-20180315.gz          intranet.ib.unicamp.br.log-20180920.gz      srv-duplicati.ib.unicamp.br.log-20180902.gz
143.106.62.12.log-20180720.gz    192.168.240.26.log-20180116.gz  192.168.240.58.log-20180413.gz  192.168.240.70.log-20180316.gz          intranet.ib.unicamp.br.log-20180921.gz      srv-duplicati.ib.unicamp.br.log-20180903.gz
143.106.62.1.log                 192.168.240.26.log-20180117.gz  192.168.240.58.log-20180414.gz  192.168.240.70.log-20180317.gz          intranet.ib.unicamp.br.log-20180922.gz      srv-duplicati.ib.unicamp.br.log-20180904.gz
143.106.62.1.log-20180720.gz     192.168.240.26.log-20180118.gz  192.168.240.58.log-20180415.gz  192.168.240.70.log-20180318.gz          intranet.ib.unicamp.br.log-20180923.gz      srv-duplicati.ib.unicamp.br.log-20180905.gz
143.106.62.1.log-20180721.gz     192.168.240.26.log-20180119.gz  192.168.240.58.log-20180416.gz  192.168.240.70.log-20180319.gz          intranet.ib.unicamp.br.log-20180924.gz      srv-duplicati.ib.unicamp.br.log-20180906.gz
143.106.62.1.log-20180722.gz     192.168.240.26.log-20180120.gz  192.168.240.58.log-20180417.gz  192.168.240.70.log-20180320.gz          intranet.ib.unicamp.br.log-20180925.gz      srv-duplicati.ib.unicamp.br.log-20180907.gz
143.106.62.1.log-20180723.gz     192.168.240.26.log-20180121.gz  192.168.240.58.log-20180418.gz  192.168.240.70.log-20180321.gz          intranet.ib.unicamp.br.log-20180926.gz      srv-duplicati.ib.unicamp.br.log-20180908.gz
143.106.62.1.log-20180724.gz     192.168.240.26.log-20180122.gz  192.168.240.58.log-20180419.gz  192.168.240.70.log-20180322.gz          intranet.ib.unicamp.br.log-20180927.gz      srv-duplicati.ib.unicamp.br.log-20180909.gz
143.106.62.1.log-20180725.gz     192.168.240.26.log-20180123.gz  192.168.240.58.log-20180420.gz  192.168.240.70.log-20180323.gz          intranet.ib.unicamp.br.log-20180928.gz      srv-duplicati.ib.unicamp.br.log-20180910.gz
143.106.62.1.log-20180726.gz     192.168.240.26.log-20180124.gz  192.168.240.58.log-20180421.gz  192.168.240.70.log-20180324.gz          intranet.ib.unicamp.br.log-20180929.gz      srv-duplicati.ib.unicamp.br.log-20180911.gz
143.106.62.1.log-20180727.gz     192.168.240.26.log-20180125.gz  192.168.240.58.log-20180422.gz  192.168.240.70.log-20180325.gz          intranet.ib.unicamp.br.log-20180930.gz      srv-duplicati.ib.unicamp.br.log-20180912.gz
143.106.62.1.log-20180728.gz     192.168.240.26.log-20180126.gz  192.168.240.58.log-20180423.gz  192.168.240.70.log-20180326.gz          intranet.ib.unicamp.br.log-20181001.gz      srv-duplicati.ib.unicamp.br.log-20180913.gz
143.106.62.1.log-20180729.gz     192.168.240.26.log-20180127.gz  192.168.240.58.log-20180424.gz  192.168.240.70.log-20180327.gz          intranet.ib.unicamp.br.log-20181002.gz      srv-duplicati.ib.unicamp.br.log-20180914.gz
143.106.62.1.log-20180730.gz     192.168.240.26.log-20180128.gz  192.168.240.58.log-20180425.gz  192.168.240.70.log-20180328.gz          intranet.ib.unicamp.br.log-20181003.gz      srv-duplicati.ib.unicamp.br.log-20180915.gz
143.106.62.1.log-20180731.gz     192.168.240.26.log-20180129.gz  192.168.240.58.log-20180426.gz  192.168.240.70.log-20180329.gz          intranet.ib.unicamp.br.log-20181004.gz      srv-duplicati.ib.unicamp.br.log-20180916.gz
143.106.62.1.log-20180801.gz     192.168.240.26.log-20180130.gz  192.168.240.58.log-20180427.gz  192.168.240.70.log-20180330.gz          intranet.ib.unicamp.br.log-20181005.gz      srv-duplicati.ib.unicamp.br.log-20180917.gz
143.106.62.1.log-20180802.gz     192.168.240.26.log-20180131.gz  192.168.240.58.log-20180428.gz  192.168.240.70.log-20180331.gz          intranet.ib.unicamp.br.log-20181006.gz      srv-duplicati.ib.unicamp.br.log-20180918.gz
143.106.62.1.log-20180803.gz     192.168.240.26.log-20180201.gz  192.168.240.58.log-20180429.gz  192.168.240.70.log-20180401.gz          intranet.ib.unicamp.br.log-20181007.gz      srv-duplicati.ib.unicamp.br.log-20180919.gz
143.106.62.1.log-20180804.gz     192.168.240.26.log-20180202.gz  192.168.240.58.log-20180430.gz  192.168.240.70.log-20180402.gz          intranet.ib.unicamp.br.log-20181008.gz      srv-duplicati.ib.unicamp.br.log-20180920.gz
143.106.62.1.log-20180805.gz     192.168.240.26.log-20180203.gz  192.168.240.58.log-20180501.gz  192.168.240.70.log-20180403.gz          intranet.ib.unicamp.br.log-20181009.gz      srv-duplicati.ib.unicamp.br.log-20180921.gz
143.106.62.1.log-20180806.gz     192.168.240.26.log-20180204.gz  192.168.240.58.log-20180502.gz  192.168.240.70.log-20180404.gz          intranet.ib.unicamp.br.log-20181010.gz      srv-duplicati.ib.unicamp.br.log-20180922.gz
143.106.62.1.log-20180807.gz     192.168.240.26.log-20180205.gz  192.168.240.58.log-20180503.gz  192.168.240.70.log-20180405.gz          intranet.ib.unicamp.br.log-20181011.gz      srv-duplicati.ib.unicamp.br.log-20180923.gz
143.106.62.1.log-20180808.gz     192.168.240.26.log-20180206.gz  192.168.240.58.log-20180504.gz  192.168.240.70.log-20180406.gz          intranet.ib.unicamp.br.log-20181012.gz      srv-duplicati.ib.unicamp.br.log-20180924.gz
143.106.62.1.log-20180809.gz     192.168.240.26.log-20180207.gz  192.168.240.58.log-20180505.gz  192.168.240.70.log-20180407.gz          intranet.ib.unicamp.br.log-20181013.gz      srv-duplicati.ib.unicamp.br.log-20180925.gz
143.106.62.1.log-20180810.gz     192.168.240.26.log-20180208.gz  192.168.240.58.log-20180506.gz  192.168.240.70.log-20180408.gz          intranet.ib.unicamp.br.log-20181014.gz      srv-duplicati.ib.unicamp.br.log-20180926.gz
143.106.62.1.log-20180811.gz     192.168.240.26.log-20180209.gz  192.168.240.58.log-20180507.gz  192.168.240.70.log-20180409.gz          intranet.ib.unicamp.br.log-20181015.gz      srv-duplicati.ib.unicamp.br.log-20180927.gz
143.106.62.1.log-20180812.gz     192.168.240.26.log-20180210.gz  192.168.240.58.log-20180508.gz  192.168.240.70.log-20180410.gz          intranet.ib.unicamp.br.log-20181016.gz      srv-duplicati.ib.unicamp.br.log-20180928.gz
143.106.62.1.log-20180813.gz     192.168.240.26.log-20180211.gz  192.168.240.58.log-20180509.gz  192.168.240.70.log-20180411.gz          intranet.ib.unicamp.br.log-20181017.gz      srv-duplicati.ib.unicamp.br.log-20180929.gz
143.106.62.1.log-20180814.gz     192.168.240.26.log-20180212.gz  192.168.240.58.log-20180510.gz  192.168.240.70.log-20180412.gz          intranet.ib.unicamp.br.log-20181018.gz      srv-duplicati.ib.unicamp.br.log-20180930.gz
143.106.62.1.log-20180815.gz     192.168.240.26.log-20180213.gz  192.168.240.58.log-20180511.gz  192.168.240.70.log-20180413.gz          intranet.ib.unicamp.br.log-20181019.gz      srv-duplicati.ib.unicamp.br.log-20181001.gz
143.106.62.1.log-20180816.gz     192.168.240.26.log-20180214.gz  192.168.240.58.log-20180512.gz  192.168.240.70.log-20180414.gz          intranet.ib.unicamp.br.log-20181020.gz      srv-duplicati.ib.unicamp.br.log-20181002.gz
143.106.62.1.log-20180817.gz     192.168.240.26.log-20180215.gz  192.168.240.58.log-20180513.gz  192.168.240.70.log-20180415.gz          intranet.ib.unicamp.br.log-20181021.gz      srv-duplicati.ib.unicamp.br.log-20181003.gz
143.106.62.1.log-20180818.gz     192.168.240.26.log-20180216.gz  192.168.240.58.log-20180514.gz  192.168.240.70.log-20180416.gz          intranet.ib.unicamp.br.log-20181022.gz      srv-duplicati.ib.unicamp.br.log-20181004.gz
143.106.62.1.log-20180819.gz     192.168.240.26.log-20180217.gz  192.168.240.58.log-20180515.gz  192.168.240.70.log-20180417.gz          intranet.ib.unicamp.br.log-20181023.gz      srv-duplicati.ib.unicamp.br.log-20181005.gz
143.106.62.1.log-20180820.gz     192.168.240.26.log-20180218.gz  192.168.240.58.log-20180516.gz  192.168.240.70.log-20180418.gz          intranet.ib.unicamp.br.log-20181024.gz      srv-duplicati.ib.unicamp.br.log-20181006.gz
143.106.62.1.log-20180821.gz     192.168.240.26.log-20180219.gz  192.168.240.58.log-20180517.gz  192.168.240.70.log-20180419.gz          intranet.ib.unicamp.br.log-20181025.gz      srv-duplicati.ib.unicamp.br.log-20181007.gz
143.106.62.1.log-20180822.gz     192.168.240.26.log-20180220.gz  192.168.240.58.log-20180518.gz  192.168.240.70.log-20180420.gz          intranet.ib.unicamp.br.log-20181026.gz      srv-duplicati.ib.unicamp.br.log-20181008.gz
143.106.62.1.log-20180823.gz     192.168.240.26.log-20180221.gz  192.168.240.58.log-20180519.gz  192.168.240.70.log-20180421.gz          intranet.ib.unicamp.br.log-20181027.gz      srv-duplicati.ib.unicamp.br.log-20181009.gz
143.106.62.1.log-20180824.gz     192.168.240.26.log-20180222.gz  192.168.240.58.log-20180520.gz  192.168.240.70.log-20180422.gz          intranet.ib.unicamp.br.log-20181028.gz      srv-duplicati.ib.unicamp.br.log-20181010.gz
143.106.62.1.log-20180825.gz     192.168.240.26.log-20180223.gz  192.168.240.58.log-20180521.gz  192.168.240.70.log-20180423.gz          intranet.ib.unicamp.br.log-20181029.gz      srv-duplicati.ib.unicamp.br.log-20181011.gz
```

Cenário IB – pré implementação

PROBLEMAS

- Como buscar as informações?
- Como monitorar os dados em tempo real?
- Como gerar relatórios?
- Como agregar diferentes tipos de dados?



Solução

Cenário IB – pós implementação

Equip. de rede enviam logs p/ Srv-Syslog

Srv-syslog grava os dados recebidos em um único arquivo, já no formato JSON

logstash

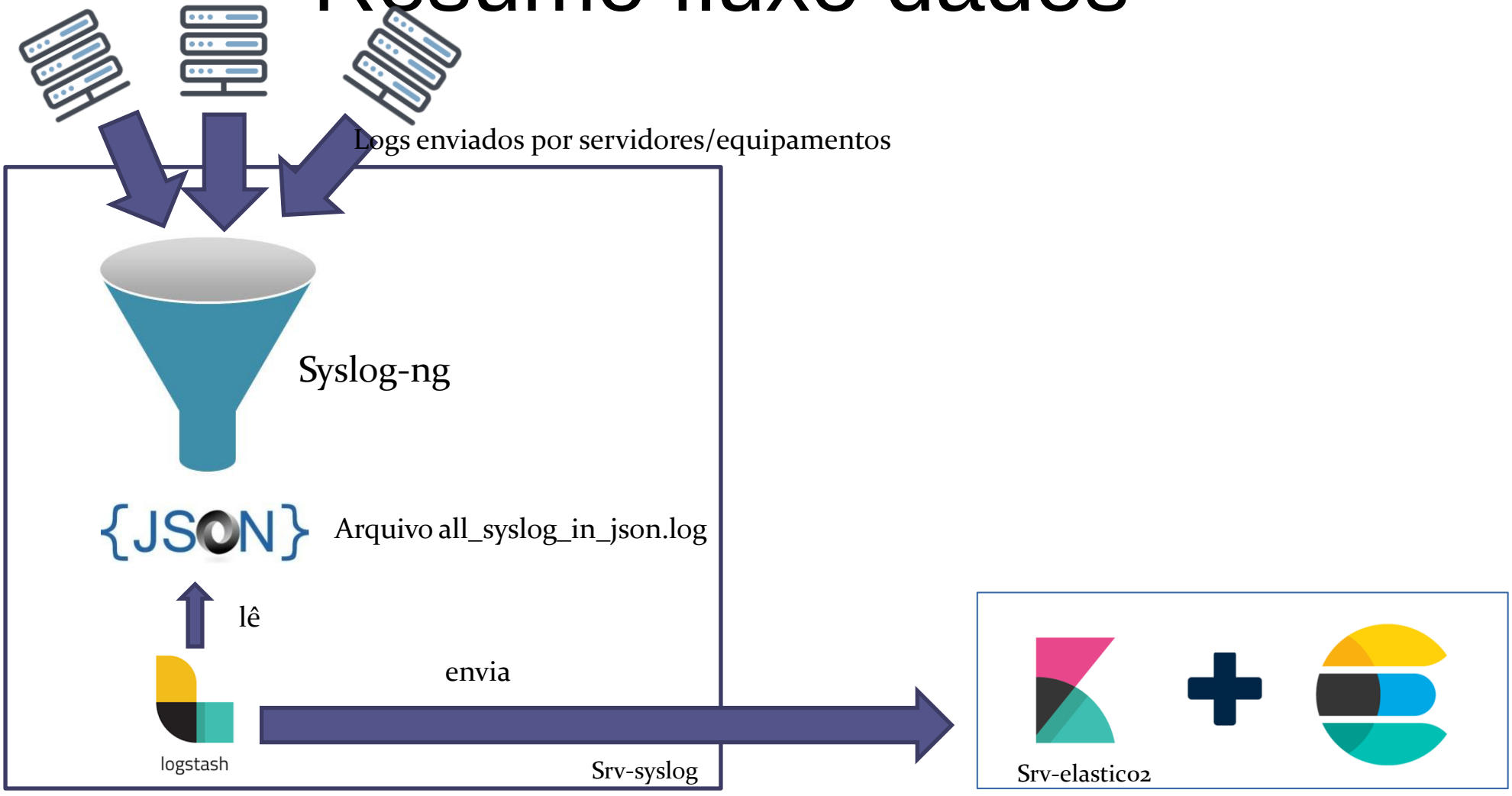
Logstash lê o arquivo e encaminha os dados p/ o servidor Elasticsearch

srv-syslog

srv-elastic02

Kibana + Elasticsearch permitem a manipulação dos dados

Resumo fluxo dados

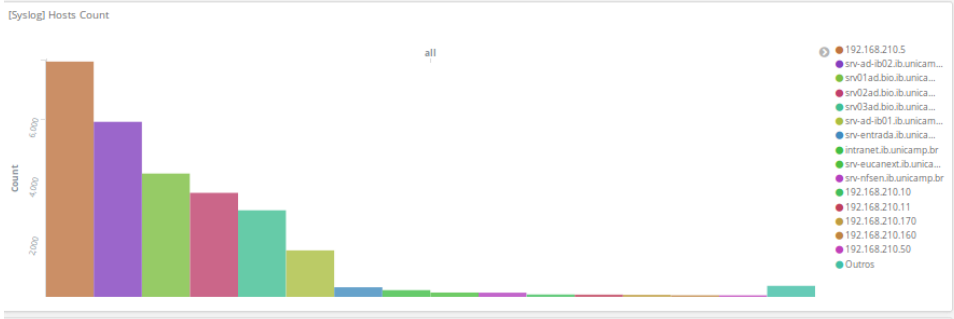


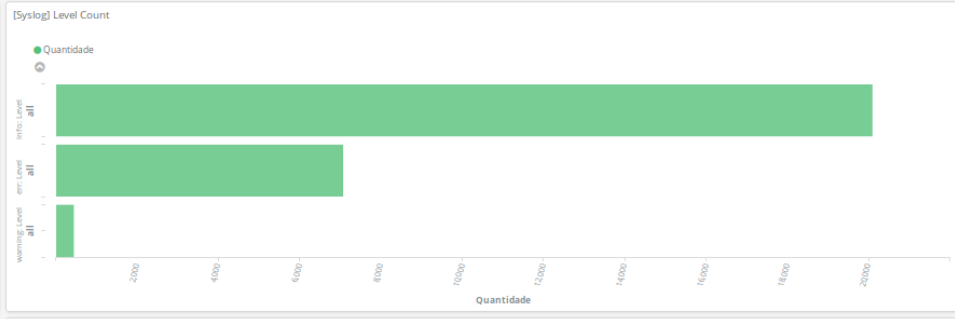
Ganhos

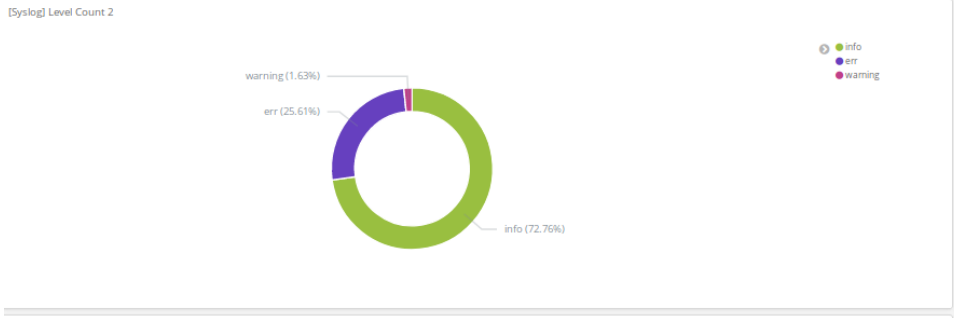
- [x] Centralização
- [x] Agilidade na análise da informação
- [x] Busca instantânea
- [x] Monitoramento em tempo real
- [x] Combinação de diferentes filtros

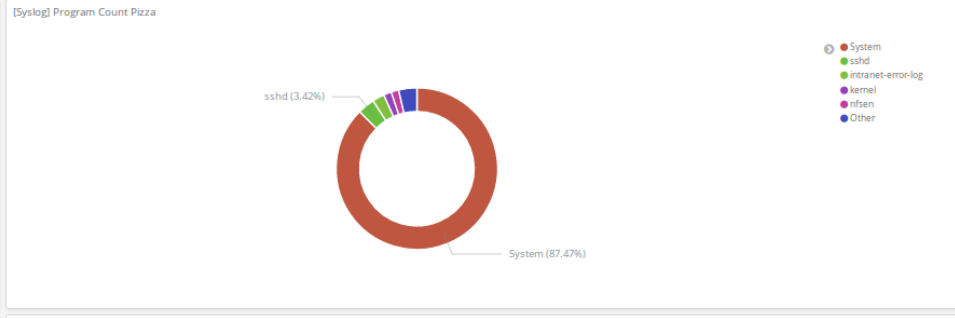
Kibana

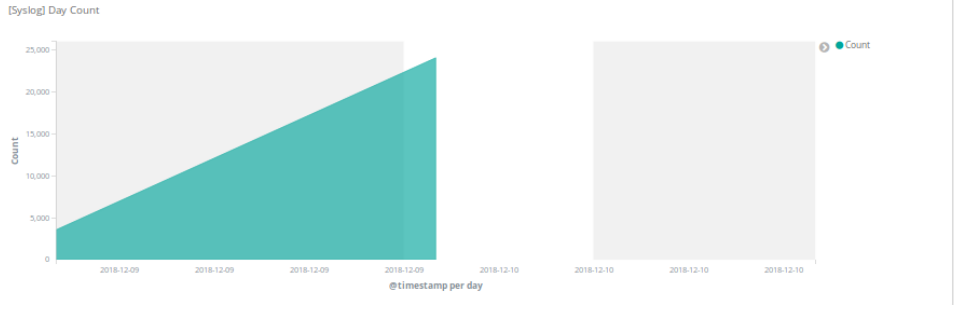
- Dashboard criado para visualizar dinamicamente os dados dos logs equip. de rede

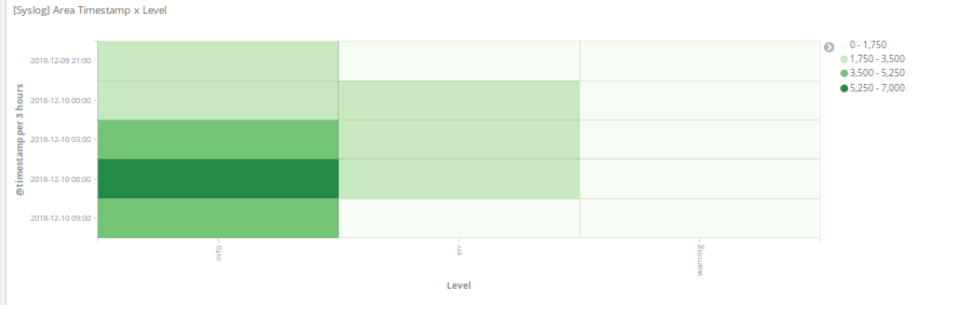












Principais desafios encontrados

1) Gravar os dados do syslog em um único arquivo já no formato JSON

SYSLOG-NG.conf

```
template JSON {
    template ("{\"@timestamp\": \"$ISODATE\", \"facility\": \"$FACILITY\", \"level\": \"$LEVEL\",
\"host\": \"$HOST\", \"tag\": \"$TAG\", \"program\": \"$PROGRAM\", \"message\": \"$MSG\"}\n");
};
```

Principais problemas encontrados

2) Deixar o JSON em um formato válido

SYSLOG-NG.conf

```
destination d_json { file("/var/log/all_syslog_in_json.log" perm(0660) template(JSON)); };

rewrite rewrite_remove_doublequotes{
    Subst(' " ', " ' ", value("MESSAGE"), flags("global"));
    Subst(' \ " ', " ' ", value("MESSAGE"), flags("global"));
    Subst(' \\ ', " / ", value("MESSAGE"),  flags("global"));
};


#grava arquivo log no formato JSON
log { source(s_net); rewrite(rewrite_remove_doublequotes); destination(d_json); };
```

Arquivo all_syslog_in.json

{"@timestamp": "2018-12-06T13:42:23-02:00", "facility": "local1", "level": "err", "host": "controlador2.ib.unicamp.br", "tag": "8b", "program": "nanny", "message": "<303085> <ERRS> <aruba-62.23 143.106.62.24> Process Manager (nanny) shutting down - Machine will reboot!"}

Arquivo all_syslog_in.json

{"@timestamp": "2018-12-06T15:16:29-02:00", "facility": "user", "level": "info", "host": "srv03ad.bio.ib.unicamp.br", "tag": "0e", "program": "1", "message": "2018-12-06T15:16:28.882936-02:00 srv03ad Microsoft-Windows-Security-Auditing 852 - [NXLOG@14506 Keywords='-9214364837600034816' EventType='AUDIT_SUCCESS' EventID='4634' ProviderGuid='{54849625-5478-4994-A5BA-3E3B0328C30D}' Version='0' Task='12545' OpcodeValue='0' RecordNumber='27703076' ThreadID='4968' Channel='Security' Category='Logoff' Opcode='Informações' TargetUserSid='S-1-5-21-2820607236-2464964702-1108132186-2374' TargetUserName='ra163532' TargetDomainName='BIO' TargetLogonId='0xe4b12f7' LogonType='3' EventReceivedTime='2018-12-06 15:16:29' SourceModuleName='in' SourceModuleType='im_msvistalog'] Foi efetuado o logoff de uma conta. Requerente: Identificação de segurança: S-1-5-21-2820607236-2464964702-1108132186-2374 Nome da conta: ra163532 Domínio da conta: BIO Identificação de logon: 0xE4B12F7 Tipo de logon: 3 Este evento é gerado quando uma sessão de logon é destruída. Ele pode ser positivamente correlacionado com um evento de logon, utilizando o valor Identificação de logon. As identificações de logon são exclusivas apenas entre as reinicializações do mesmo computador."}

Logstash - logstash-syslog.conf

```
input {
  file {
   path => "/var/log/all_syslog_in_json.log"
    start_position => "beginning"
    codec => "json"
    sincedb_path => "/var/log/db_for_watched_files.db"
    type => "syslog"
  }
}
filter {
      if [host] == "srv01ad.bio.ib.unicamp.br" or [host] == "srv02ad.bio.ib.unicamp.br" or [host] == "srv03ad.bio.ib.unicamp.br" or [host] == "srv-ad-ib02.ib.unicamp.br" or [host] == "srv-winadk.bio.ib.unicamp.br" or [host] =="srv-ad-ib01.ib.unicamp.br" {
            mutate { remove_field => "program" }
     }
     }
```

Logstash – logstash-syslog.conf

```
output {
if [type] == "syslog" {
      elasticsearch {
      hosts => [ "143.106.221.91:9200" ]
      index => "logs"
        }
    }
    #stdout {
      # codec => rubydebug
     #}
}
```

Demonstração prática

Resultados

- Transformação de dados brutos em informações que auxiliam os gestores na tomada de decisões estratégicas;
- Agilidade na análise dos dados;
- Acompanhamento em tempo real dos alertas e erros dos equipamentos de infraestrutura, o que leva a uma rápida ação para a mitigação do problema;
- Otimização no tempo de trabalho e aumento da produtividade dos colaboradores;
- Diminuição no número das ordens de serviços direcionadas à equipe de TI, devido a independência que os usuários passaram a ter para a geração de relatórios que antes eram solicitados e criados pela equipe de desenvolvimento.

Banner Cinfotec

8º Cinfotec UNICAMP
26 e 27 de junho de 2019

Elasticsearch, Kibana e Logstash (Pilha ELK): transformando dados brutos em informações poderosas de forma rápida e eficiente

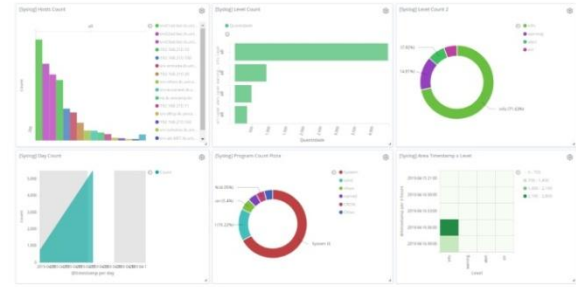
Autores: Allan Michel de Souza, Bruno Ximenes, Gustavo Bueno Bellini, Marcos Akira e Valcir Cabral

CTIC

UNICAMP

Introdução: Elasticsearch, Kibana e Logstash são softwares Open Source que compõem a pilha ELK. Com esse conjunto de ferramentas é possível centralizar e armazenar bilhões de registros, efetuar buscas instantâneas, importar dados de diferentes formatos, coletar métricas de desempenho, gerar relatórios combinando vários filtros e criar dashboards dinâmicos, que transformam dados em informação com apenas alguns cliques. São ferramentas essenciais que facilitam e auxiliam a tomada de decisão e que podem ser utilizadas em várias áreas, como a da tecnologia da informação, administração, finanças, saúde etc. Podemos dizer que trata-se de um Big Data muito fácil de utilizar, com excelente usabilidade e eficiência. No Instituto de Biologia da Unicamp a Pilha ELK foi integrada ao serviço de monitoramento de logs, Syslog, e também a 7 módulos do sistema Intranet: Almoxarifado, Compras, Financeiro, Informática, Manutenção, Nitrogênio e Transportes.

Objetivo: Transformar dados brutos em informações relevantes que possam ser utilizadas como base para a tomada de decisão, otimização e melhorias em processos de diferentes áreas.



Dashboard com os dados importados do Syslog, onde é possível monitorar em tempo real os alertas e erros dos equipamentos que compõem a infraestrutura de rede do Instituto de Biologia.

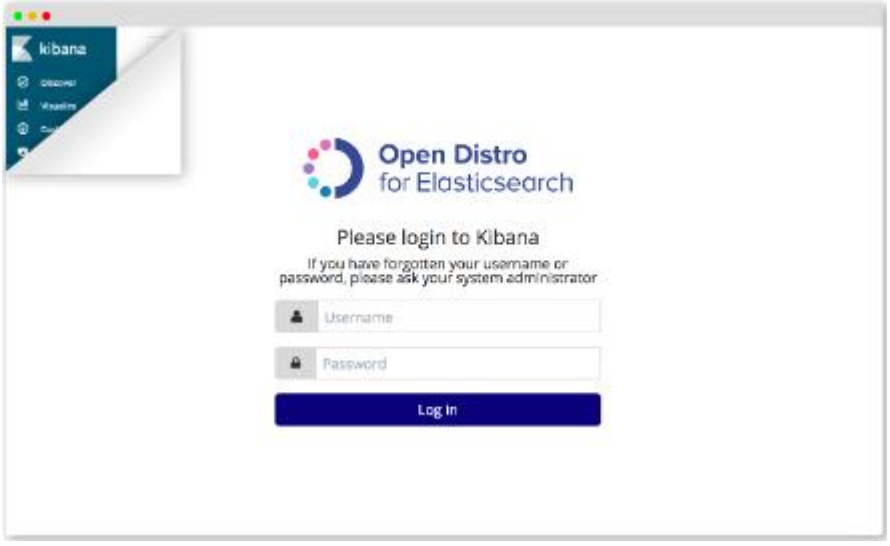
Resultados:

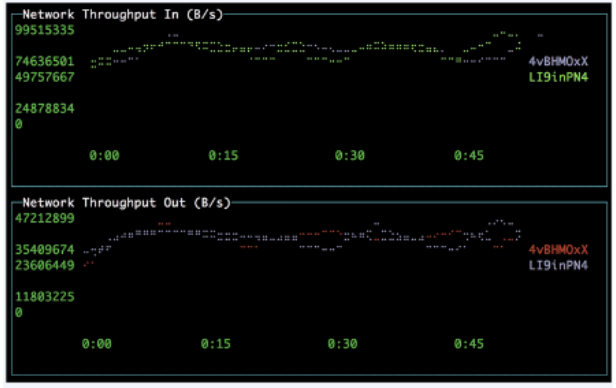
- Transformação de dados brutos em informações que auxiliam os gestores na tomada de decisões estratégicas;
- Agilidade na análise dos dados;
- Acompanhamento em tempo real dos alertas e erros dos equipamentos de infraestrutura, o que leva a uma rápida ação para a mitigação do problema;
- Otimização no tempo de trabalho e aumento da produtividade dos colaboradores;
- Diminuição no número das ordens de serviços direcionadas à equipe de TI, devido a independência que os usuários passaram a ter para a geração de relatórios que antes eram solicitados e criados pela equipe de desenvolvimento.

Conclusão: a pilha ELK facilitou a análise dos dados e permitiu, de forma rápida e eficiente, uma visão profunda e abrangente de informações até então desconhecidas.

Open Distro Elasticsearch - Amazon

- “Fork“ criado pela Amazon
- Acrescenta as seguintes funcionalidades: Security, Notificações e Análise de performance
- Amazon alegou que criou essa versão para sempre manter o ELK Open Source
- Link para download: https://opendistro.github.io/for-elasticsearch/





Open Distro
for Elasticsearch

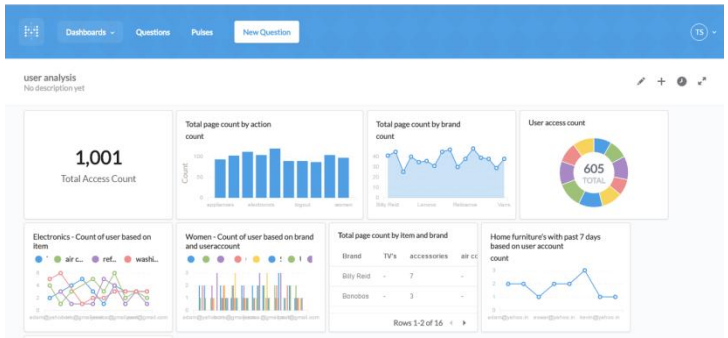
Concorrentes

Power BI

Microsoft

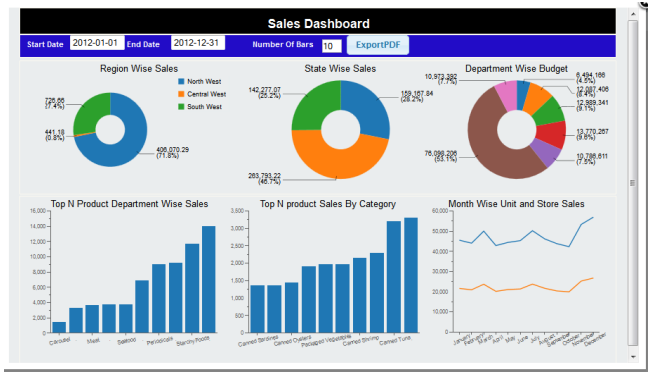
Plano premium: $5.000/mês

x



Metabase: Open Source

x

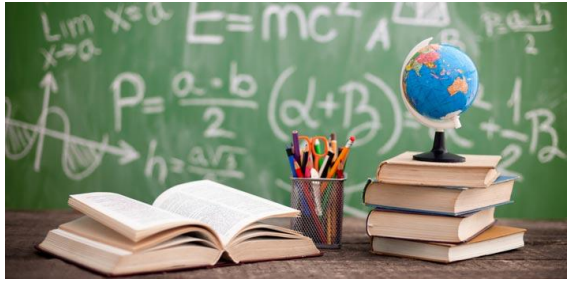


Pentaho: Open Source

Teste prático da ferramenta

- http://kibana-nuvem.ib.unicamp.br
- https://demo.elastic.co/app/kibana

Cenários possíveis de uso da ferramenta



Cenários possíveis de uso da ferramenta

Discussão geral

Dúvidas:
gbellini@unicamp.br