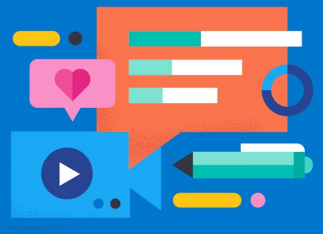
Tratamento de dados usando Logstash com filter Ruby

13/09/2022

Weberth Oliveira

Analista Elasticsearch | Desenvolvedor full stack

Índice



- Contexto
- Arquitetura
- Implementação
- Aprendizado
- Perguntas
- Contatos
- Referências

elastic

Contexto

Case profissional:

- Case profissional que atualmente está em homologação pelo cliente;
- Iniciou com a necessidade de criação de API usando o ElasticSearch;
- Dados proveniente de um índice Elasticsearch;
- Necessidade de anonimização de campos e em alguns casos aplicar certas condições para apresentação.

Dificuldades:

- **Origem continha informações inseridas em objetos e arrays aninhados. Isso dificulta o processamento dos dados através de plugins mais simples.**
- **Nível de segurança do ambiente muito alto, dificultando então implementação/uso de scripts, bibliotecas e outras ferramentas**
- **Falta de padronização nos dados. Origem formada pela união de diversas fontes.**

Ação Tomada:

Definido a utilização da linguagem de programação Ruby, pois é a linguagem aceita no filtro do Logstash.

Justificativas:

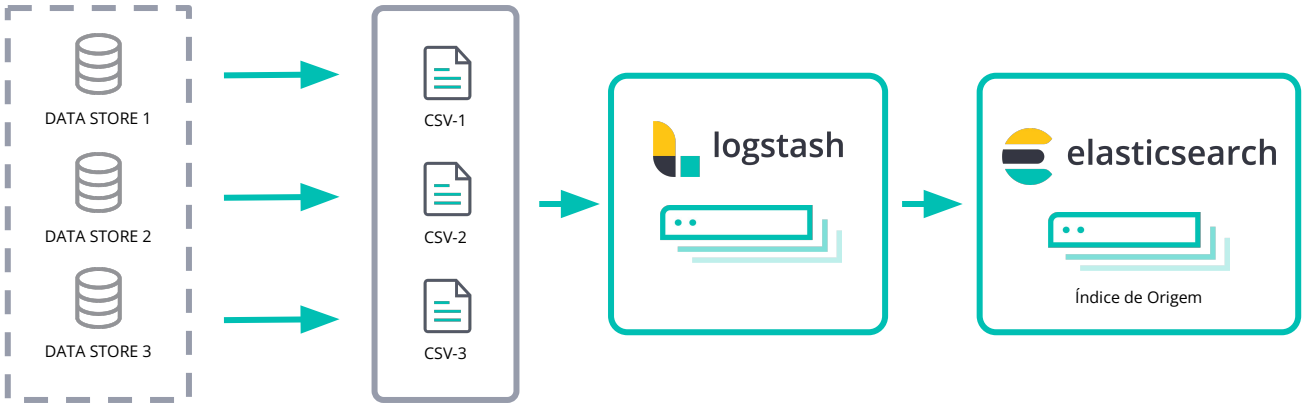
- **Necessário aplicar lógica de programação**
- **Inviabilidade de usar Javascript :(**
  - Restrição do ambiente
  - Serviços adicionais para monitorar
- **Tempo de entrega**

Não entendi! Pode desenhar?

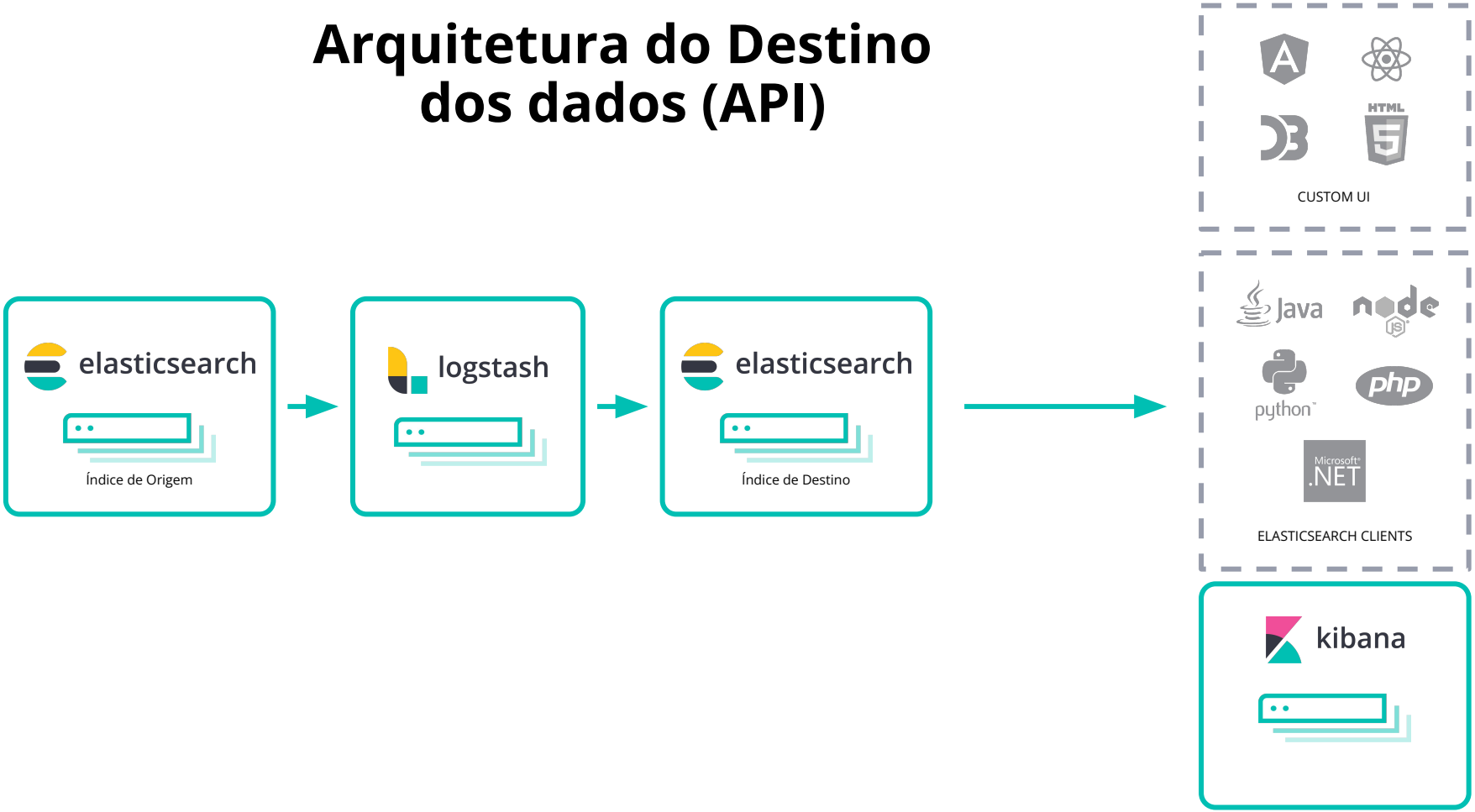
elastic

Arquitetura

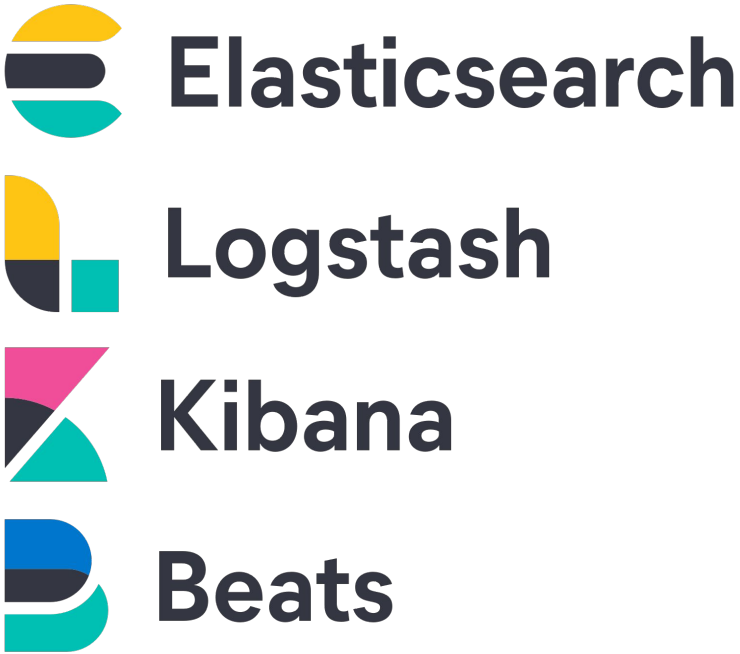
Arquitetura da Origem dos dados



Arquitetura do Destino dos dados (API)

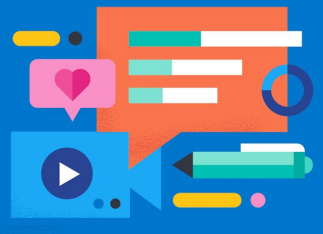


Ferramentas Utilizadas



elastic

Elasticsearch (Indexação e API REST)



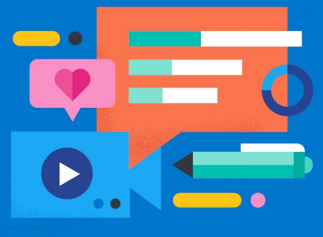
Elasticsearch

- Mecanismo de busca open source;
- Escalável e Flexível;
- **Fornece um conjunto completo e poderoso de APIs REST.**

Create **R**ead **U**pdate **D**elete

elastic

Logstash (Processar dados)

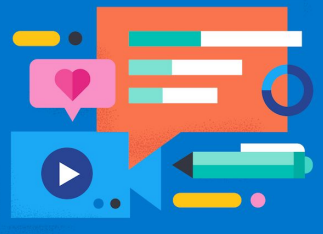


Logstash

- Logstash é um pipeline de processamento de dados open source do lado do servidor que permite fazer a ingestão de dados de várias fontes simultaneamente e enriquecer e transformar esses dados antes de serem indexados no Elasticsearch.

elastic

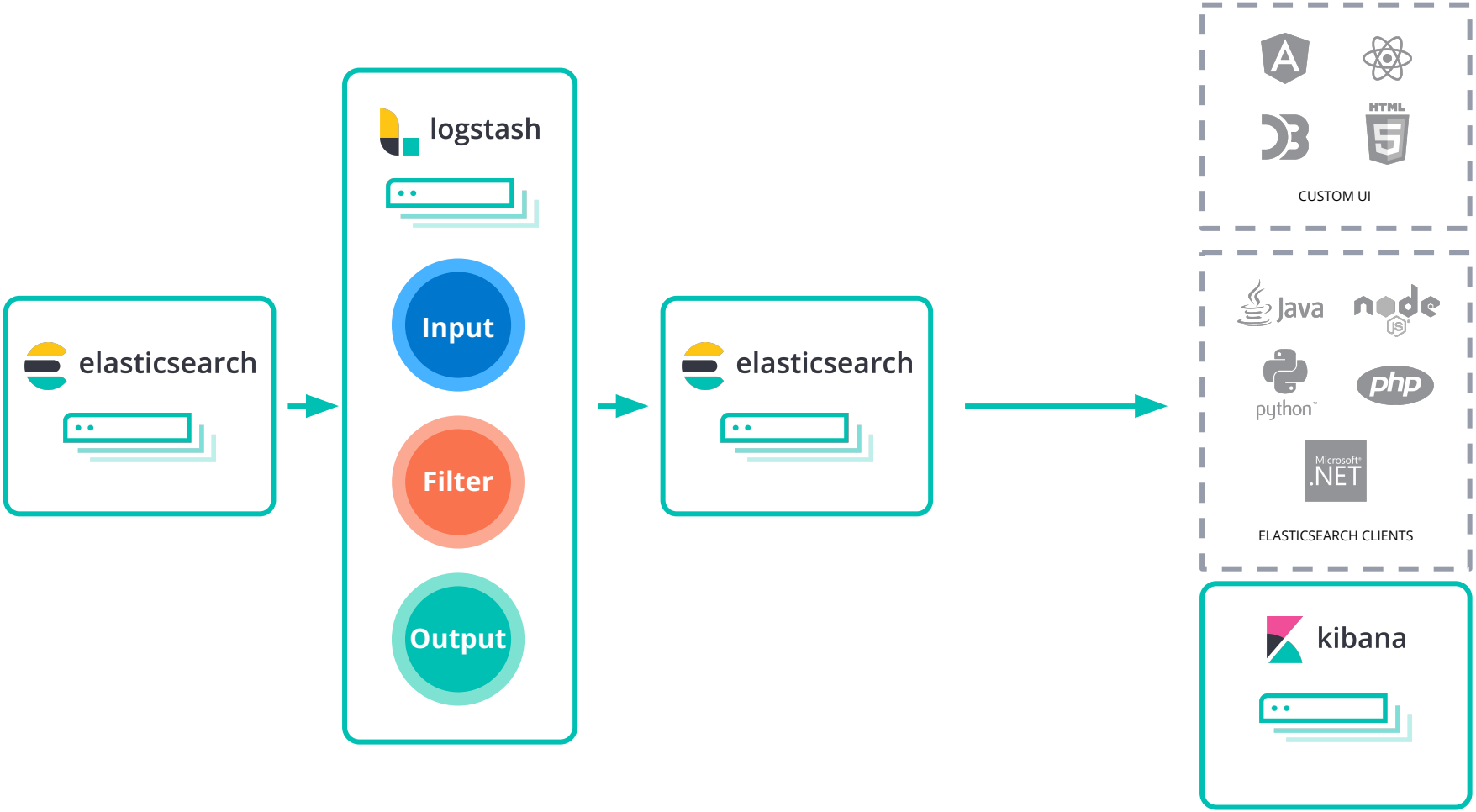
Acessando os dados com o Kibana



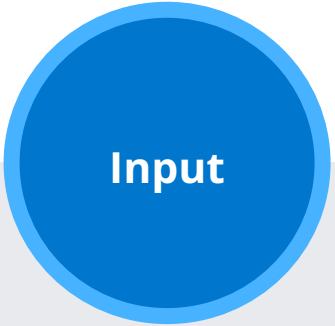
Kibana

- O Kibana é uma ferramenta de visualização e gerenciamento de dados para o Elasticsearch.
- Ferramenta de análise e gerenciamento de dados.

Implementação

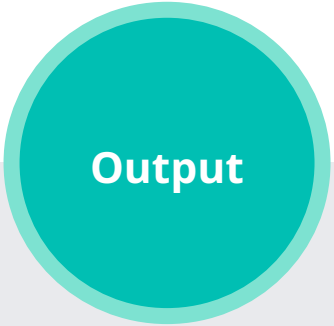

elasticsearch
logstash
Input
Filter
Output
elasticsearch
CUSTOM UI
Java
node
python
php
Microsoft .NET
ELASTICSEARCH CLIENTS
kibana

Estágios de uma Pipeline Logstash



- Entrada de Dados
- Exemplos de Plugins:
  - Jdbc
  - File
  - exec
  - s3
  - sqlite
  - elasticsearch

- Processamento de dados
- Exemplos de Plugins:
  - xml
  - csv
  - split
  - json
  - mutate
  - ruby



- Saída de dados
- Exemplos de Plugins:
  - csv
  - email
  - file
  - mongodb
  - http
  - elasticsearch

SHOW ME THE CODE

elastic

Aprendizado

1º

Realmente não existe a tal bala de prata.

2º

Avalie a situação e use novas ferramentas se necessário.

3º

Novos conhecimentos aguçam o poder de decisão.

Perguntas?

Contatos

Contatos

Linkedin

@weberthmo

Github

@weberthmo

Obrigado!

Referências

- https://www.elastic.co/guide/en/logstash/8.4/index.html
- https://www.elastic.co/guide/en/elasticsearch/reference/8.4/index.html
- https://www.elastic.co/pt/what-is/elk-stack