Fundamentos da Arquitetura Distribuída do Elasticsearch

Clusters, Nós, Índices, Shards e Réplicas

05 de Setembro de 2025

Agenda da Apresentação

01

O que é Elasticsearch?

E por que ele é distribuído?

02

Conceitos Fundamentais

Cluster, Nó, Índice, Shard e Réplica

03

Diagrama da Arquitetura

Visualizando a estrutura distribuída

04

Próximos Passos

Recursos e considerações finais

O que é Elasticsearch e Por Que é Distribuído?

Definição

Um motor de busca e análise de dados, de código aberto, construído sobre a biblioteca Apache Lucene, com alta capacidade de indexação e busca em tempo real.

Funcionalidades

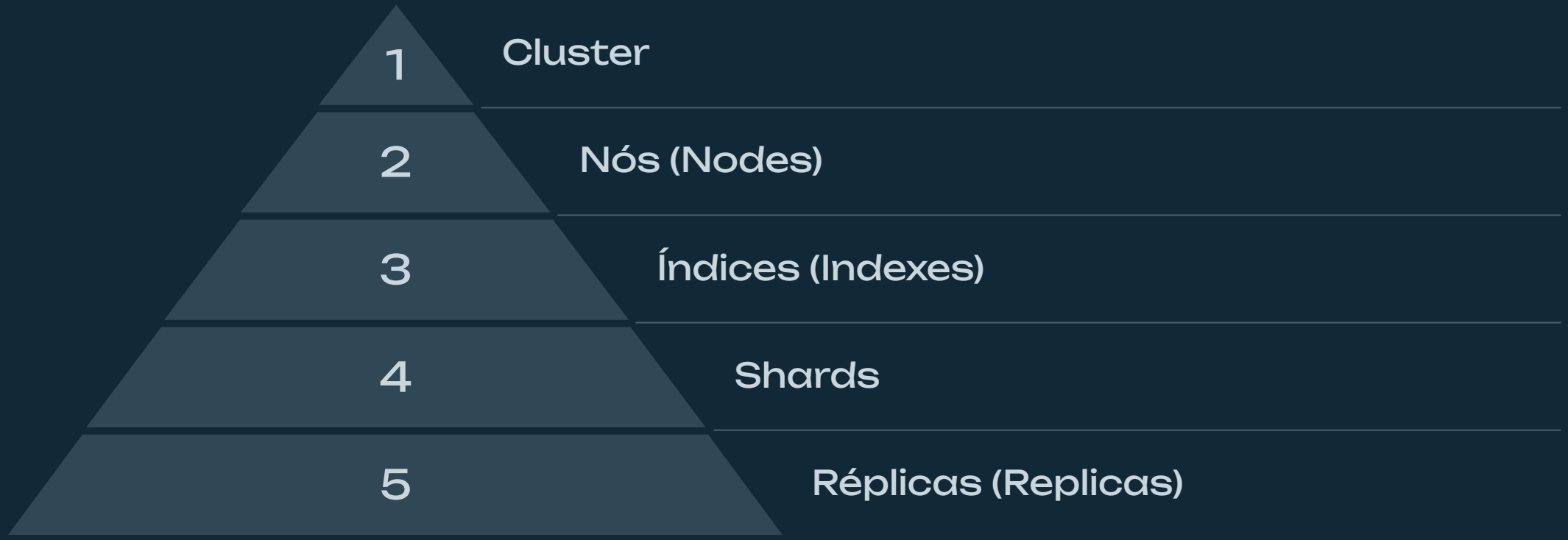
Busca de texto completo, análise de logs, monitoramento de métricas, inteligência de segurança (SIEM) e muito mais.

Por que Distribuído?

- **Escalabilidade Horizontal:** Adicione mais máquinas (nós) para aumentar a capacidade de forma linear.
- **Alta Disponibilidade e Resiliência:** O sistema continua funcionando mesmo se houver falhas em um ou mais nós, devido à replicação de dados.

A Hierarquia dos Componentes do Elasticsearch

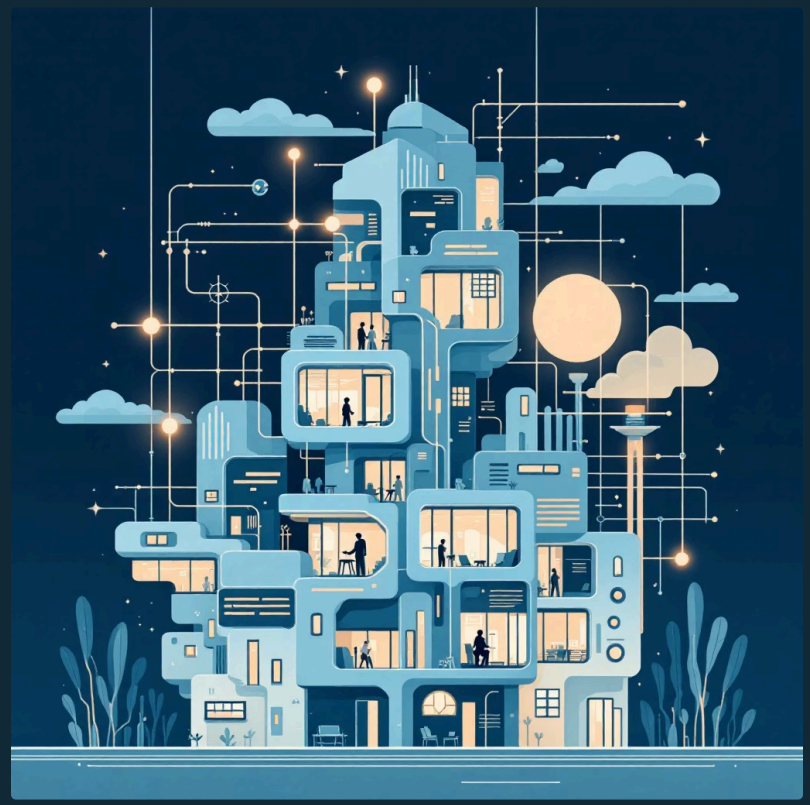
Vamos explorar a arquitetura do Elasticsearch de uma perspectiva "macro" para "micro", seguindo a visão do "Elasticsearch from the Top Down".



Cada nível é construído sobre o anterior, permitindo flexibilidade e robustez.

Cluster: O Ecossistema Completo

Definição: Um cluster é uma coleção de um ou mais servidores (nós) que, juntos, mantêm todos os seus dados e fornecem recursos de indexação e busca de forma federada. É o seu sistema Elasticsearch completo.



Analogia: Um Prédio de Escritórios Inteiro

Pense no cluster como um prédio de escritórios inteiro, representando toda a sua infraestrutura de dados.

Cada cluster elege automaticamente um **nó mestre (master)**, responsável por gerenciar o estado geral do cluster (criação/deleção de índices, alocação de shards).

Nó (Node): Uma Unidade de Trabalho

Definição: Um nó é uma única instância de servidor que faz parte de um cluster. Ele executa o processo do Elasticsearch e desempenha funções específicas.

Analogia: Um Andar do Prédio

Se o cluster é o prédio, um nó é como um andar desse prédio, com suas próprias salas e recursos.

Funções de um Nó:

- Armazena dados.
- Participa das capacidades de indexação e busca do cluster.

Nós se comunicam constantemente para manter o cluster coeso e os dados sincronizados.

Índice (Index): Organizando seus Dados

Definição: Um índice é uma coleção de documentos que possuem características semelhantes. É a unidade lógica de mais alto nível para organizar os dados no Elasticsearch.

Analogia: Um Grande Arquivo de Fichas

Um índice é como um grande arquivo de fichas ou uma tabela em um banco de dados relacional. Por exemplo, você pode ter um índice para `logs_aplicacao` ou outro para `produtos_ecommerce`.

Cada documento JSON que você insere no Elasticsearch pertence a um índice específico, fornecendo uma maneira estruturada de agrupar e gerenciar informações.

Shard: Fragmentando para Escalar

Definição: Como um índice pode crescer e se tornar muito grande para caber em um único nó, o Elasticsearch permite dividi-lo em múltiplos pedaços menores chamados **shards**.

Função Principal

Escalabilidade Horizontal: Os shards permitem que os dados de um índice sejam distribuídos por vários nós, otimizando o uso de recursos e o desempenho.

Analogia: Uma Gaveta no Arquivo de Fichas

Um shard é como uma gaveta específica dentro do seu grande arquivo de fichas (Índice).

Quando você cria um índice, você define o número de **shards primários**. Este número não pode ser alterado posteriormente, pois cada shard primário é uma instância independente e funcional do motor de busca Lucene.

Réplica (Replica Shard): Redundância e Desempenho

Definição: Uma réplica é uma cópia exata de um shard primário, garantindo segurança e eficiência.

Funções Principais

- **Alta Disponibilidade (Failover):** Se o nó que contém um shard primário falhar, sua réplica em outro nó pode ser promovida a primário, evitando perda de dados e garantindo a continuidade do serviço.
- **Escalabilidade de Leitura:** Requisições de busca podem ser atendidas tanto pelo shard primário quanto por suas réplicas, aumentando o throughput de leitura do cluster.

⊗ **Regra de Ouro:** Uma réplica **NUNCA** é alocada no mesmo nó que seu shard primário. Isso é crucial para a resiliência.

Hierarquia Lógica vs. Física do Elasticsearch

1. O Cluster (O Todo):

É o nível mais alto. Pense nele como seu serviço Elasticsearch completo. Ele é composto por um ou mais Nós.

2. O Nó (A Máquina):

É a instância física (ou virtual) do Elasticsearch. É um membro do Cluster que armazena dados e processa requisições.

3. O Índice (A Organização Lógica):

É uma abstração. Um índice não "mora" em um único nó; seus dados são distribuídos através do cluster na forma de shards.

4. O Shard (A Unidade Física de Dados):

Esta é a peça fundamental da distribuição. Cada Shard Primário é um "pedaço" do seu índice. O Elasticsearch distribui automaticamente os shards de um índice entre os diferentes Nós disponíveis.

5. A Réplica (A Cópia de Segurança):

É a chave para a resiliência. Uma réplica é uma cópia de um shard primário. Uma réplica nunca é alocada no mesmo nó que seu primário.

Exemplo Prático de Fluxo

Imagine um Cluster com 3 nós (Nó 1, Nó 2, Nó 3) e um Índice chamado pedidos. Ao criar o índice pedidos, você define que ele terá 2 shards primários (P0, P1) e 1 réplica.

Distribuição (Escalabilidade): O Elasticsearch aloca os primários. P0 vai para o Nó 1 e P1 vai para o Nó 2.

Resiliência (Alta Disponibilidade): O Elasticsearch aloca as réplicas em nós diferentes. R0 (cópia de P0) vai para o Nó 3 e R1 (cópia de P1) vai para o Nó 1.

Resultado: Se o Nó 2 falhar, o P1 fica indisponível. O cluster automaticamente promove a réplica R1 (que está no Nó 1) para se tornar o novo primário. Nenhuma requisição é perdida e o sistema continua funcionando.