Elasticsearch(ElasticDB) 개념

1. Elasticsearch란?

Elasticsearch는 분산형 검색 및 분석 엔진으로 데이터를 분산 관리하고 빠르게 검색할 수 있는 기능을 제공함. 여러 서버를 연결해 클러스터를 구성하는 것이 특징이며, 주로 대량의

주로 로그 분석, 실시간 검색, 위치 기반 데이터 검색 등에서 널리 사용됨

2. Elasticsearch 주요 개념

1-1 노드(서버)

- Elasticsearch 클러스터를 구성하는 개별 서버 또는 인스턴스를 의미함
- 하나의 Elasticsearch 인스턴스를 실행하면, 그 인스턴스가 하나의 노드로 동작
- 노드는 각각 데이터 저장, 검색 요청 처리 등의 역할을 수행하며 신속한 검색을 제공

1-2 클러스터

- 클러스터는 Elasticsearch에서 여러 노드가 모여 하나의 시스템처럼 작동하는 집합을 의미함
- 문서, 인덱스, 샤드, 복제본을 생성하고 관리하는 전체 시스템을 의미함

2-1 문서

- 기본 데이터 단위로, JSON 형식으로 저장
- 하나의 문서는 사용자 정보나 제품 데이터 같은 개별 데이터

2-2 인덱스

- 문서들의 집합체로 여러 문서가 모여 하나의 인덱스를 형성
- 인덱스를 사용하여 데이터를 빠르게 검색하고 관리 가능

2-3 샤드

- 데이터가 많아질 경우 하나의 서버에 모든 데이터를 담을 경우 부하가 심해짐
- 샤드를 이용하여 인덱스를 여러 서버에 나누어 저장하는 기능을 제공
- 인덱스가 샤드로 나뉘어있어도 데이터를 검색하는 데 문제가 없으며, 검색 속도가 더 빨라짐
- Elasticsearch는 백그라운드에서 각 샤드에 분산된 데이터를 자동으로 모아서 결과를 하나의 인덱스처럼 보여줌
- 물리적으로는 여러 서버에 샤드(문서 단위로) 분산 되어 있지만 논리적으로는 모든 샤드들이 하나의 인덱스를 구성함

2-4 복제본

- Elasticsearch는 인덱스를 생성할 때 기본적으로 샤드의 복제본이 하나씩 자동으로 생성됨
- 복제본 샤드는 원본 샤드와 동일한 데이터를 가지고 있으므로 읽기 요청을 복제본에도 분산할 수 있어 읽기 속도가 빨라지며 고가용성을 보장함.

- [illegible]

Revision #31
Created 8 November 2024 08:06:00 by [illegible]
Updated 9 November 2024 13:59:47 by [illegible]