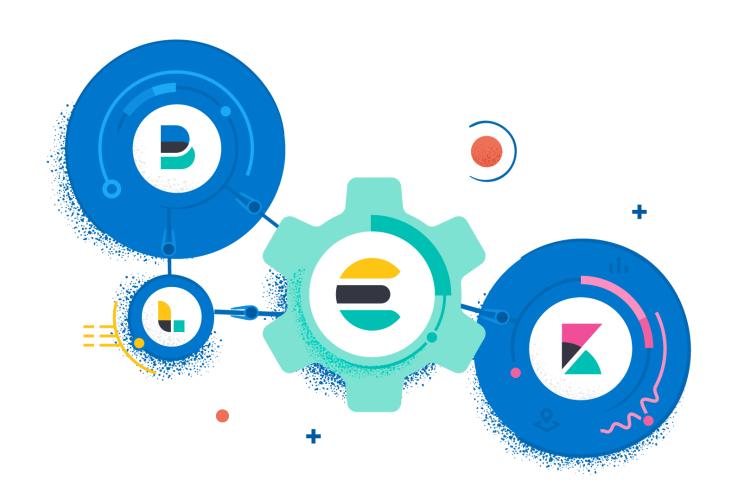
빅데이터 검색엔진 Elastic Stack



Elastic Stack

Elasticsearch, Logstash, Kibana 세가지 오픈 소스 프로젝트가 ELK Stack이라는 명칭으로 서비스가 제공되었고 여기에 Beats를 도입하여 Elastic Stack이라고 합니다.

https://github.com/elastic



Data Collection Data
Aggregation
& Processing

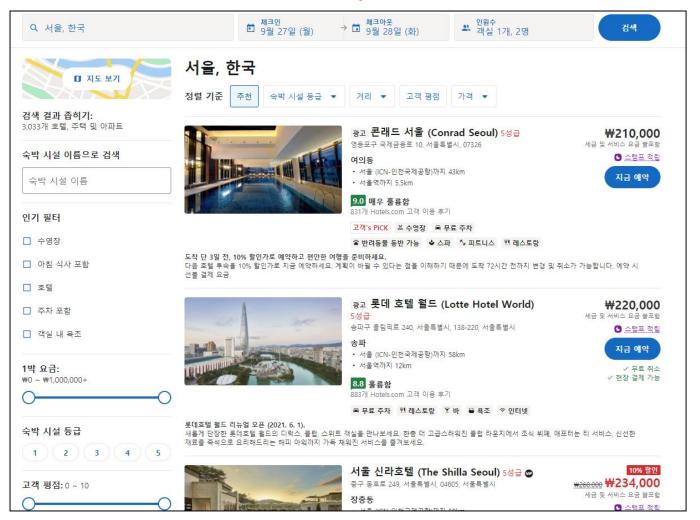
Indexing & storage

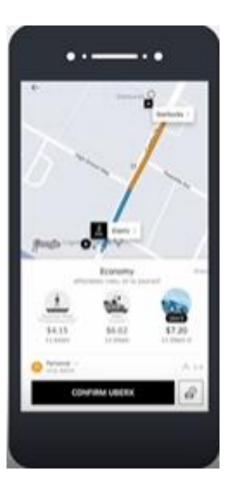
Analysis & visualization

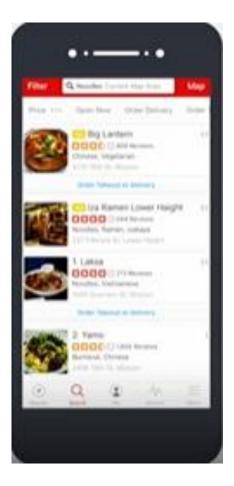
Elasticsearch

아파치 루씬 기반의 Full Text 검색이 가능한 오픈소스 분석엔진입니다.

주로 REST API를 이용해 처리하며, 대량의 데이터를 거의 실시간으로 저장, 검색 및 분석 할 수 있습니다.

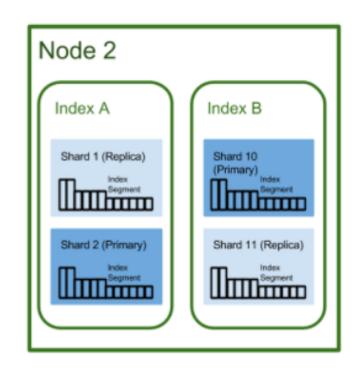






Elasticsearch 아키텍처

ElasticSearch Cluster Node 1 Index B Index A Shard 1 (Primary) Shard 10 (Replica) Segment Shard 2 (Replica) Segment Segment



- 역색인(Inverted Index)을 통한 빠른 검색
- 클러스터 구성을 통한 분산처리 및 고가용성
- Replica를 활용한 데이터 안정성 증대
- Shard 분배를 통한 선형적 확장(scale-out)
- RESTful API 지원
- Schemaless 지원
- 인덱스 기반의 타입 및 색인 방식 설정 지원

Elasticsearch 아키텍처

Cluster

- 하나 이상의 Elasticsearch 노드들로 구성된 노드의 집합
- 클러스터는 Elasticsearch 시스템을 구성하는 가장 큰 단위로 독립적으로 운용

Node

■ 노드는 Elasticsearch가 실행중인 인스턴스

Shard

■ 데이터를 분산하여 저장하기 위해 Index의 범위를 나눈 것

Replica

■ 분산 환경에서 데이터의 신뢰성을 높이기 위해 여러 노드에 데이터를 복제하여 저장하는 것

Elasticsearch 논리 아키텍처

Document

- 하나의 JSON 오브젝트로 elasticsearch 시스템에서 데이터를 구성하는 최소 단위
- 일반적인 row형 데이터베이스에서 하나의 row에 대응되는 개념

Field

- Document를 구성하는 하나의 key-value pair에 해당
- RDBMS의 열(column)에 대응되는 개념
- Elasticsearch는 schema-less 구조이므로, 하나의 Field에 다양한 타입의 데이터를 저장할 수 있음

Mapping

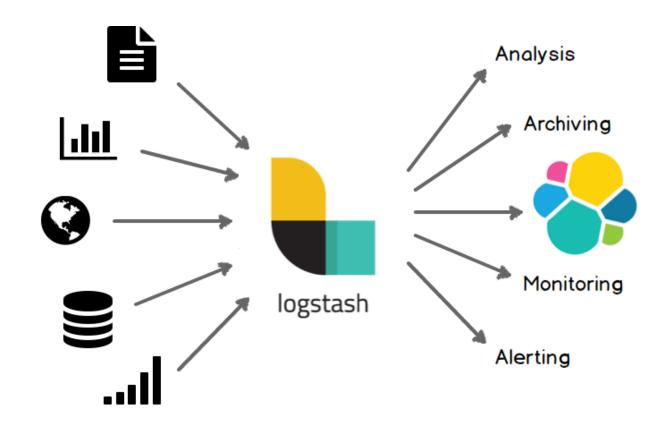
- 하나의 Document를 구성하기 위해 필요한 Field와 Field의 속성, 색인 방법을 정의하는 일련의 과정
- Mapping 과정은 RDBMS에서 스키마를 설계하고 인덱스를 설정하는 과정과 유사

Indices

- 여러 Type의 집합으로 RDBMS에서 Database에 대응되는 개념
- Index는 Shard와 Replica가 이루어지는 최소 단위

Logstash

다양한 플러그인을 이용하여 데이터 집계 및 보관, 데이터를 처리하며, 파이프라인으로 데이터를 수집하여 필터를 통해 변환 후 Elastic Search로 전송합니다.

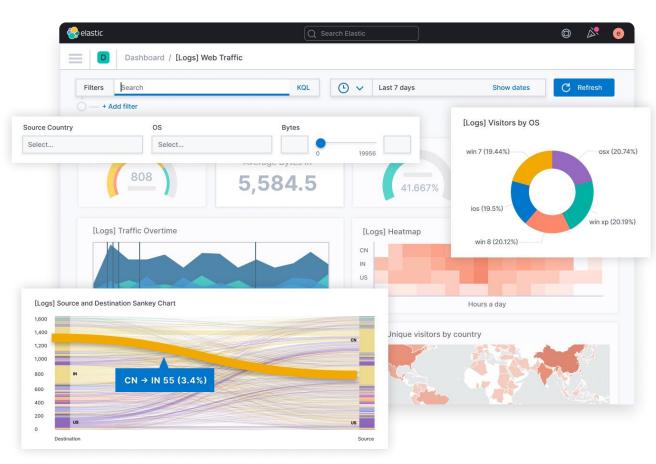


https://www.elastic.co/kr/logstash/

- 입력: 다수의 소스에서 동시에 이벤트를 가져오는 다양한 입력을 지원. 로그, 메트릭, 웹 애플리케이 션, 데이터 저장소 및 다양한 AWS 서비스에서 스트리밍 방식으로 수집
- **필터** : 형식이나 복잡성에 상관없이 설정을 통해 데이터를 동적으로 변환
- **출력** : Elastic Search, Email, ECS, Kafka 등 원하는 저장소에 데이터를 전송

Kibana

Elasticsearch에서 색인된 데이터를 검색하고 시각화 하는 기능을 제공합니다.



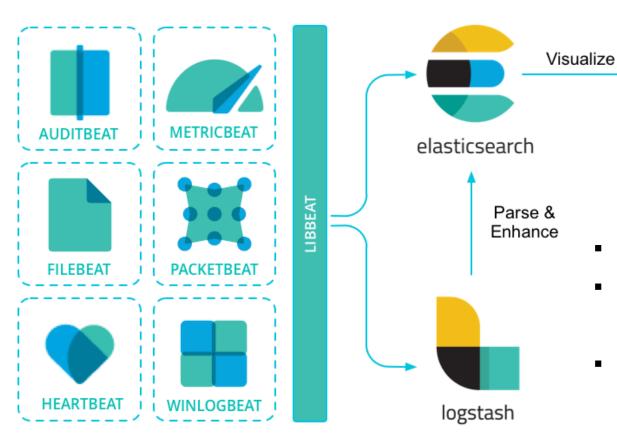
히스토그램, 선 그래프, 원형 차트, 선버스트 등
 다양한 시각화 요소 제공

차트, 지도, 필터를 모두 가져와 한 곳에서 데이터
 의 전체를 볼 수 있는 대시보드를 빠르게 생성

https://www.elastic.co/kr/kibana/

Beats

경량 데이터 수집기로 에이전트로 설치되어, 데이터를 Logstash 또는 Elastic Search로 전송하는 도구입니다.



https://www.elastic.co/kr/beats/

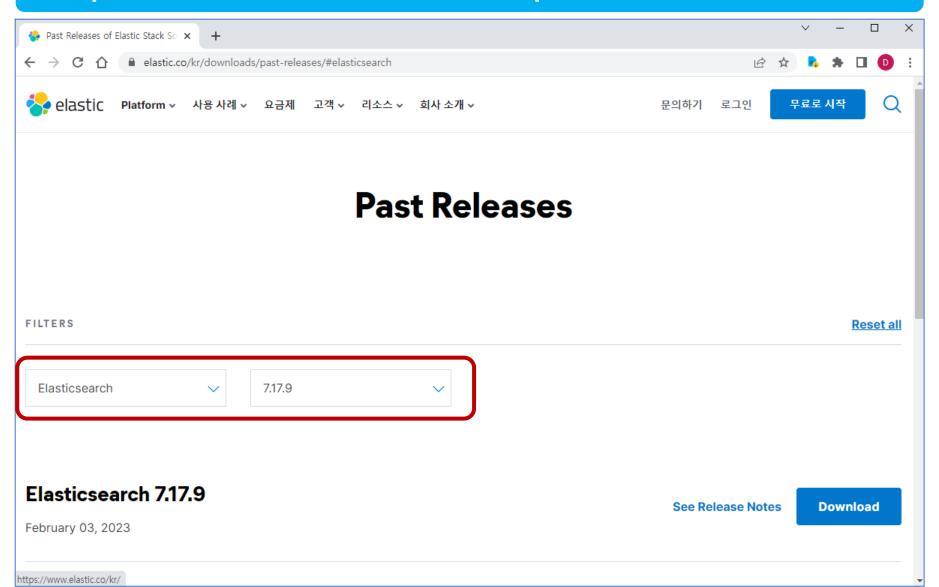
■ Filebeat : 서버에서 로그 파일을 제공

kibana

- Metricbeat : 서버에서 실행중인 운영 체제 및 서비스에서
 메트릭을 주기적으로 수집하는 서버 모니터링 에이전트
- Packetbeat : 응용 프로그램 서버간에 교환되는 트랜잭션에 대한 정보를 제공하는 네트워크 패킷 분석기
- Winlogbeat : Windows 이벤트 로그를 제공

Elasticsearch 설치

https://www.elastic.co/kr/downloads/past-releases/#elasticsearch



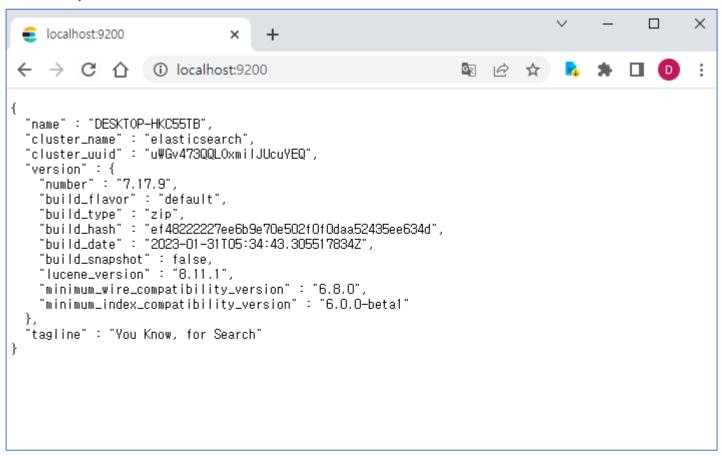
Elasticsearch 설치

■ 실행

- cd bin
- elasticsearch

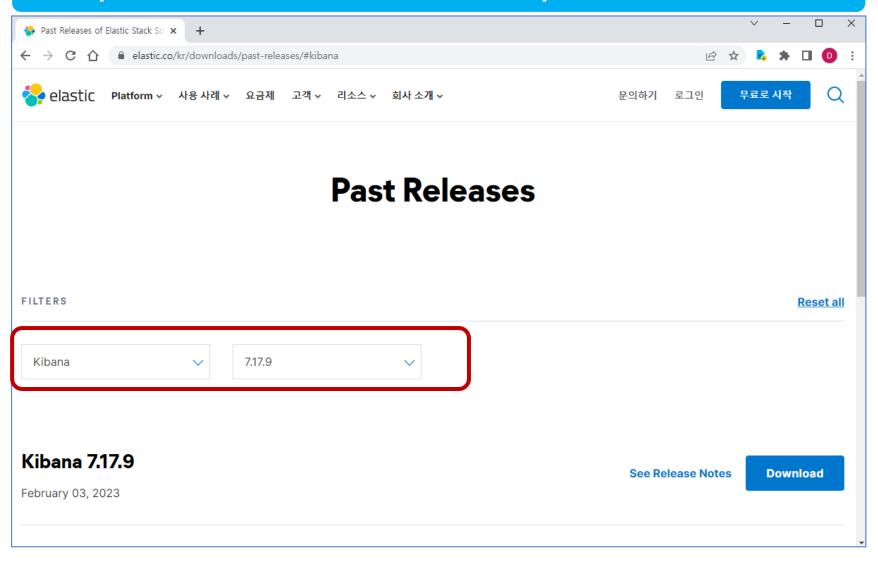
■ 실행 확인

- curl localhost:9200
- http://localhost:9200/ 접속



Kibana 설치

https://www.elastic.co/kr/downloads/past-releases/#kibana



Kibana 설치

kibana.yml

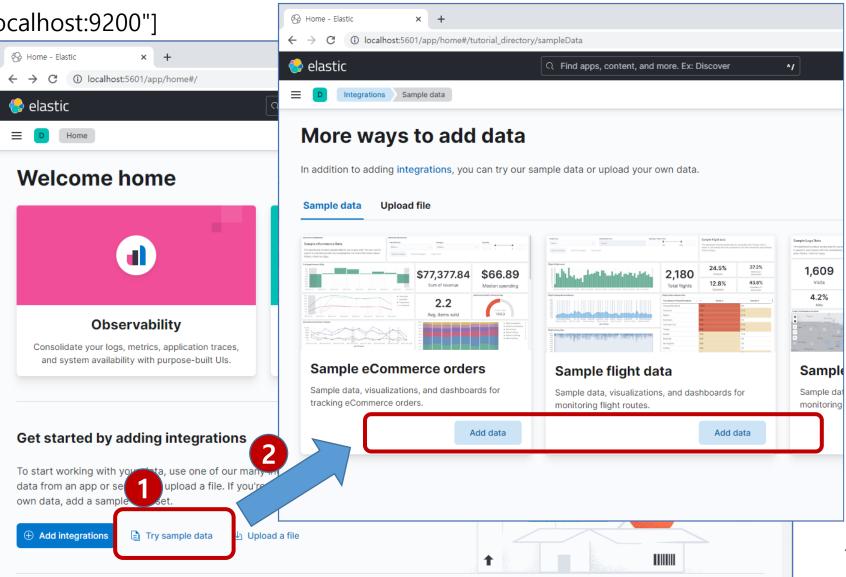
elasticsearch.hosts: ["http://localhost:9200"]

■ 실행

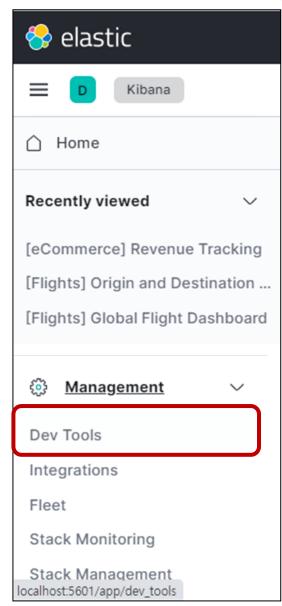
- cd bin
- kibana

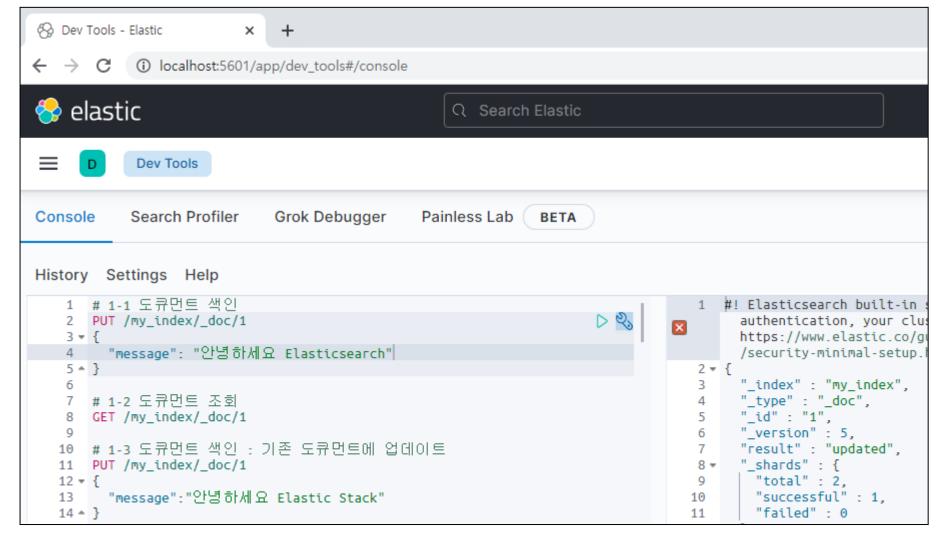
■ 사용

■ http://localhost:5601/ 접속



Elasticsearch 실습

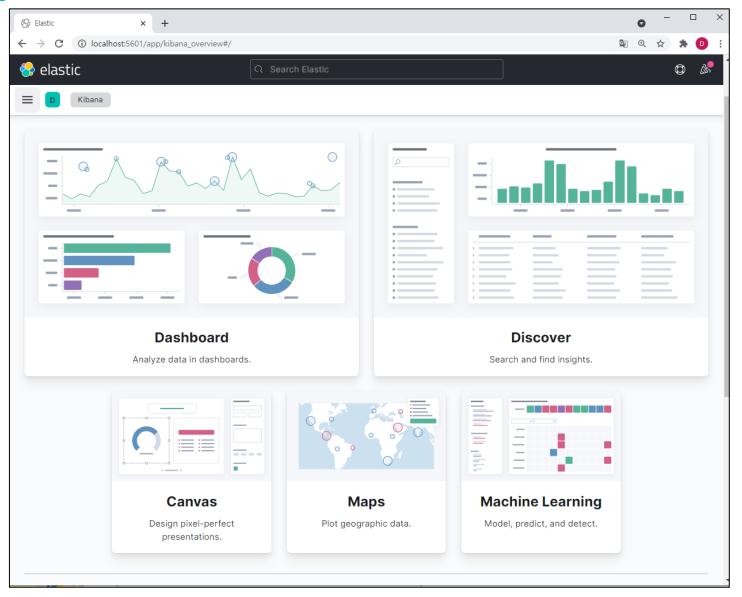




교육동영상: https://www.elastic.co/kr/webinars/getting-started-elasticsearch

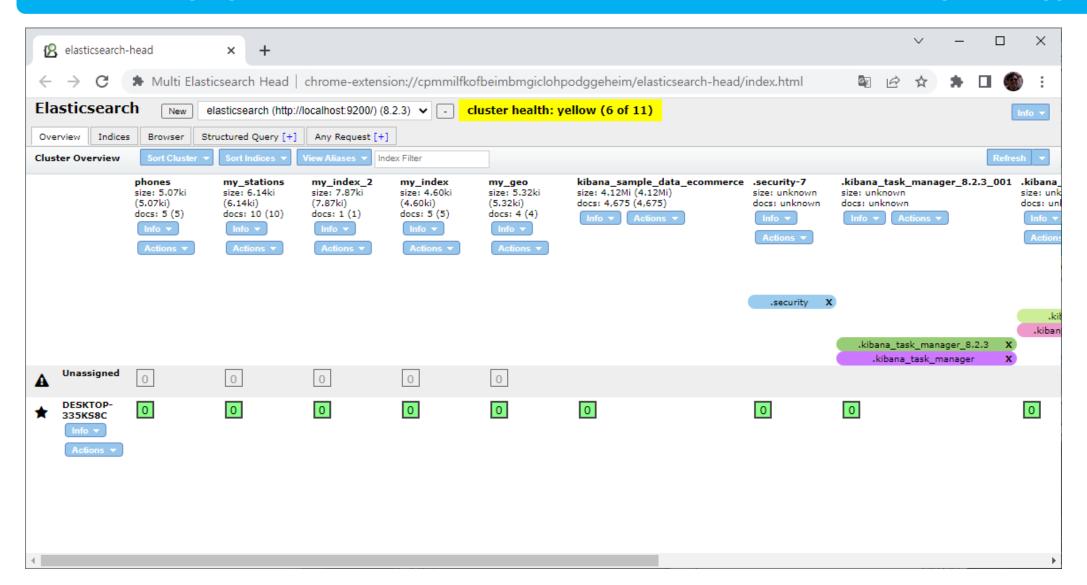
실습코드: https://github.com/kimjmin/elastic-demo/blob/master/demos/get-started/elasticsearch-7x.md

Kibana 실습



ElasticSearch 클러스터 관리도구

https://chrome.google.com/webstore/detail/multi-elasticsearch-head/cpmmilfkofbeimbmgiclohpodggeheim



Thank you