ElasticStack

Elastic Stack

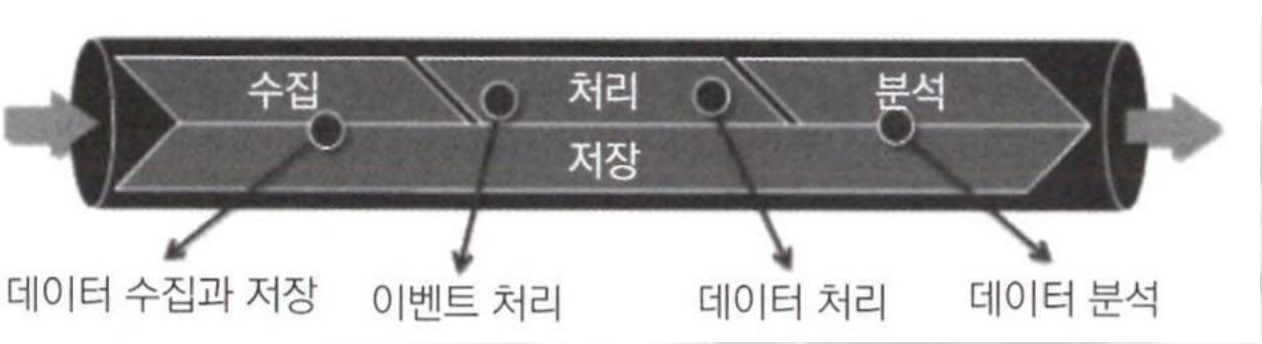
- Elastic Search 탄생
  - ✓ 엘라스틱이 등장했을 무렵에 사용자의 요구 사항은 너무나 명확했는데 바로 사이트 내에서 전문 검색(full text search) 기능을 제공하는 강력한 소프트웨어의 필요성
  - ✓ 인터넷 검색 서비스는 계속해서 등장하고 있었지만 구현체를 공개하지 않았기 때문에 일반 개발자는 검색 이면에 숨은 비밀에 접근하기가 좀처럼 쉽지 않았기 때문에 기업 내에서 대량의 데이터베이스나 파일 시스템을 검색할 수 있는 라이브러리나 프레임워크는 벤더가 제공하는 값비싼 솔루션에 의존할 수밖에 없었고 이는 기업용 검색 기능의 발전에 크나큰 걸림돌이 됨
  - ✓ 1999년 하둡(Hadoop)으로 유명해진 더그 커팅(Doug Cutting)은 제록스 PARC 와 익사이트(Excite)에서 검색 엔진을 만들던 역량을 이용해 검색 엔진 라이브러리를 만들게 되는데 더그 커팅 부인의 중간 이름을 딴 루씬(Lucene)
  - ✓ 루씬은 2001년도에 아파치 자카르타(Jakarta)에 합류했고 규모가 점점 확장됨에 따라 2005년 2월에 독자적인 아파치 프로젝트로 분리
  - ✓ 루씬은 전문 인덱싱과 검색 기능을 요구하는 애플리케이션에서 사용되기 시작했으며 더그 커팅은 2003년 6월에 100만 페이지를 가져오는 웹 크롤러인 아파치 너치(Apache Nutch)를 만들어 루씬의 위력을 증명
  - ✓ 나중에 아파치 프로젝트에 들어가면서 현재까지 개발이 이어지고 있는 너치는 맵리듀스 기능과 분산 파일 시스템을 기반으로 만들어졌는데 이것이 바로 하둡의 시초가 됨

Elastic Stack

- Elastic Search 탄생
  - ✓ 루씬은 라이브러리 형태로 되어 있었기 때문에 제대로 사용하기 위해 해야 할 작업이 무척 많음
  - ✓ 루씬 덕분에 완전 처음부터 작업할 필요는 없었지만 사용자 편의성을 제공하기 위한 상부 작업은 피할 수 없는 상황이었는데 이때 기술 잡지를 만들던 CNET 네트워크에서 2004년경 사내 웹사이트 검색을 위해 직접 제작한 검색 엔진의 소스 코드를 2006년에 아파치 재단에 기부하기로 결정했는데 솔라(Solr)라는 이름의 이 프로젝트는 상용 서비스에 적합한 수준의 높은 트래픽을 뒷받침하는 고성능의 검색 엔진으로 자리 잡으면서 많은 사람의 관심을 끌었고 2010년 3월에는 루씬과 통합되어 커미터들이 함께 개발을 진행했고 2021년 2월에는 루씬에서 분리되어 완전히 독립된 아파치 최상위 프로젝트로 자리 잡음
  - ✓ CNET 네트워크가 솔라를 개발하고 있었을 2004년 무렵에 요리 학교를 다니던 자신의 부인이 손쉽게 레시피를 검색할 수 있도록 루씬을 기반으로 컴파스(Compass) 프로젝트를 진행하던 샤이 배넌(Shay Banon)은 버전 3에 이르러서 많은 부분을 재작성해 확장 가능한 검색 엔진 솔루션을 만들면 좋겠다는 생각을 하게 되고 처음부터 분산 환경을 위한 솔루션, HTTP 상에서 JSON으로 인터페이스를 지원하는 솔루션, 자바 이외에도 다양한 프로그래밍 언어를 지원하는 솔루션을 표방하고 작업에 들어간 배넌은 2010년 2월에 엘라스틱서치를 공개

Elastic Stack

- Elastic Stack 으로 발전
  - ✓ 오픈 소스 검색 엔진 부문에서 먼저 출발한 아파치 솔라의 길을 그대로 따라갔다면 엘라스틱서치 역시 세간의 관심을 크게 끌지 못했을지도 모르지만 엘라스틱은 단순 검색 엔진에 머무르는 대신 플랫폼으로 발전하는 길을 택함
  - ✓ 데이터가 점점 늘어나는 상황에서 전문 검색은 대량의 데이터를 저장해서 처리하는 한 가지 방법에 불과했고 다양한 사용 사례를 유연하게 지원하기 위해서는 검색 엔진을 뛰어넘는 플랫폼이 필요해졌지만 유연성을 확보하려면 데이터를 수집, 가공, 저장, 분석, 시각화하는 일련의 파이프라인을 구성하기 위해 다양한 오픈 소스 소프트웨어를 조합해야 하는 불편함이 있었는데 엘라스틱은 바로 이 지점에서 사용자의 가려운 곳을 긁어주는 해법을 제시했으며 간편하게 빅데이터 파이프라인을 구성하고 싶은 기업의 요구 사항을 충족하기 위한 스택을 층층이 쌓아 올리기 시작
  - ✓ 일반적인 빅데이터 파이프라인



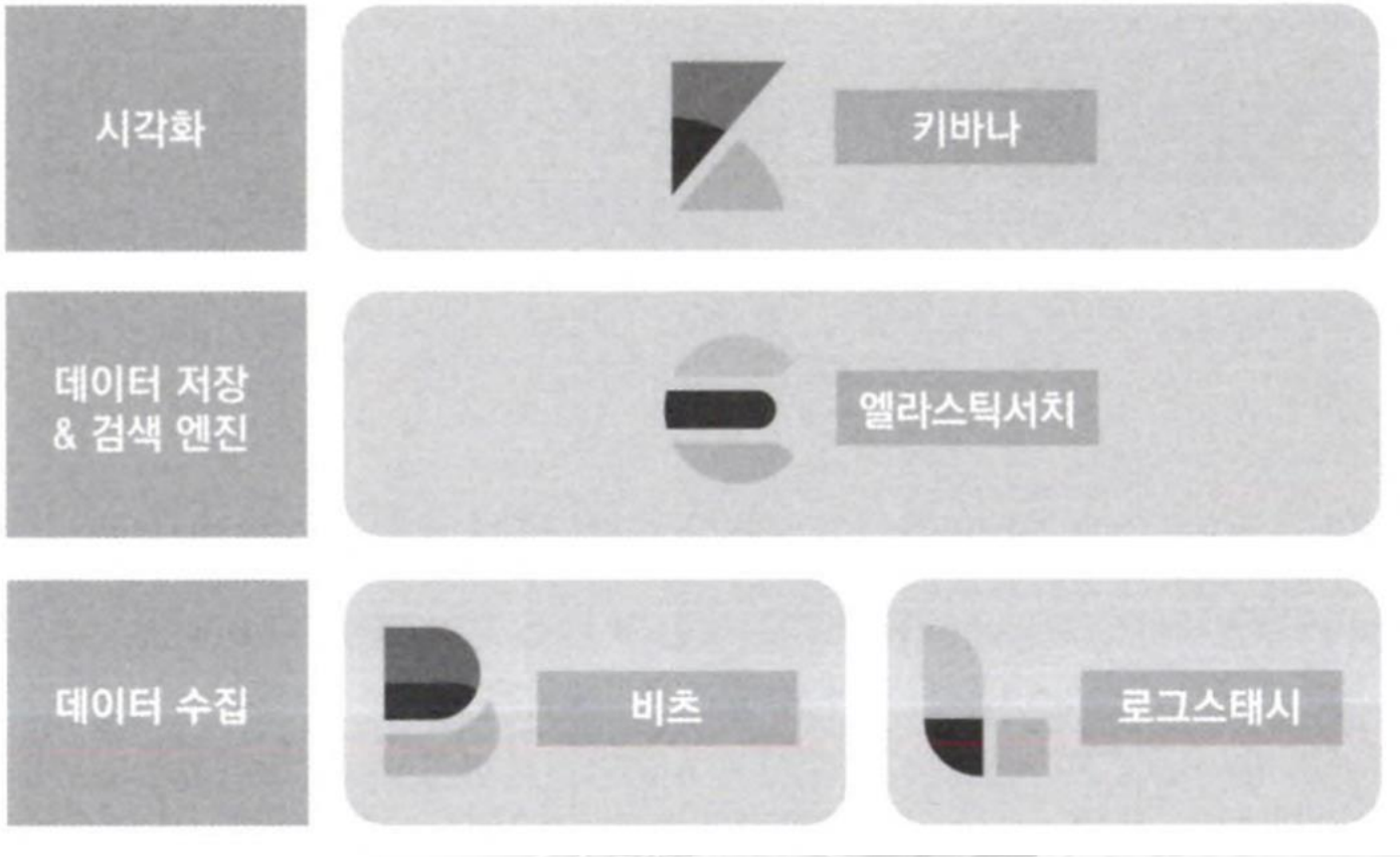
Elastic Stack

- Elastic Stack 으로 발전
  - ✓ 엘라스틱서치가 개발될 무렵 두 가지 오픈 소스 프로젝트가 추가로 진행되고 있었는데 요르단 시셀(Jordan Sissel)은 사용자가 선택한 보관함(Stash)으로 로그를 보내는 오픈 소스 데이터 수집 도구인 로그스태시(Logstash)를 개발하고 있었고 라시드 칸(Rashid Khan)은 오픈 소스 시각화 UI인 키바나Kibana를 개발하고 있었음
  - ✓ 샤이 배넌은 두 사람과 친분이 있었고 각자 진행 중인 프로젝트에 대해서도 잘 알고 있었기에 2012년 함께 모여 팀을 결성하기로 하고 엘라스틱서치, 로그스태시, 키바나를 합친 엘라스틱 스택(Elastic Stack)을 개발
  - ✓ 초창기에는 엘라스틱 스택이 느슨한 형태로 개발되었으며 각자 원하는 버전을 원하는 시점에 출시하는 형태였지만 사용자 요구에 맞춰 제품은 점차 개선을 거듭하면서 새 플러그인과 확장이 추가되기 시작했는데 스택을 구성하는 요소에 버전 번호가 제각각 붙었기 때문에 호환성을 유지하는 과정에서 큰 혼란이 벌어짐
  - ✓ 정상적인 동작을 위해 엘라스틱 스택을 구성 하는 각 구성요소의 어떤 버전을 함께 사용해야 할지 파악하기가 쉽지 않았으며 플러그인과 확장도 버전 별로 조합을 확인하는 과정이 고통스러웠는데 이를 해소하기 위해 2015년 10월에는 버전을 동일하게 맞춘 제품을 동시에 출시해서 안정적인 기반을 확보
  - ✓ 2015년에 베를린의 패킷비트(Packetbeat)라는 회사가 네트워크 데이터를 엘라스틱서치로 보내는 방법을 연구해서 제품을 개발하고 있다는 사실을 엘라스틱 개발팀이 알게 되자 엣지 장비(edge device)에서 로그스태시 와 엘라스틱서치로 네트워크 데이터, 로그, 메트릭, 성능 진단 데이터를 보내는 경량 데이터 수집기의 가능성을 고려하게 되고, 패킷 비트와 연합해 비츠(Beat)가 탄생
  - ✓ 이런 과정을 거치면서 여러 오픈소스 소프트웨어가 합쳐진 결과 엘라스틱 스택이 완성

Elastic Stack

❖ Elastic Stack 으로 발전

✓ 엘라스틱 스택 구조



Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 엘라스틱 스택은 일반적인 빅데이터 파이프라인을 구성하기 위한 데이터 수집, 가공, 저장, 분석, 시각화에 필요한 모든 소프트웨어를 갖추고 있음
  - ✓ 비츠와 로그스태시는 데이터를 수집하고 가공하는 역할을 맡으며 엘라스틱 서치는 저장하고 분석하는 역할을 담당하고 키바나는 엘라스틱 서치에 저장된 데이터를 시각화하고 모니터링하는 역할을 수행
  - ✓ 각각을 결합하면 더 강력해지지만 필요에 따라서는 키바나를 제외한 엘라스틱 스택의 나머지 구성 요소만으로도 개별적인 빅데이터 파이프라인의 일부로 사용될 수 있음
  - ✓ 상용 라이선스를 구입할 경우에는 그래프 분석이나 머신러닝 같은 고급 기능도 사용 가능

Elastic Stack

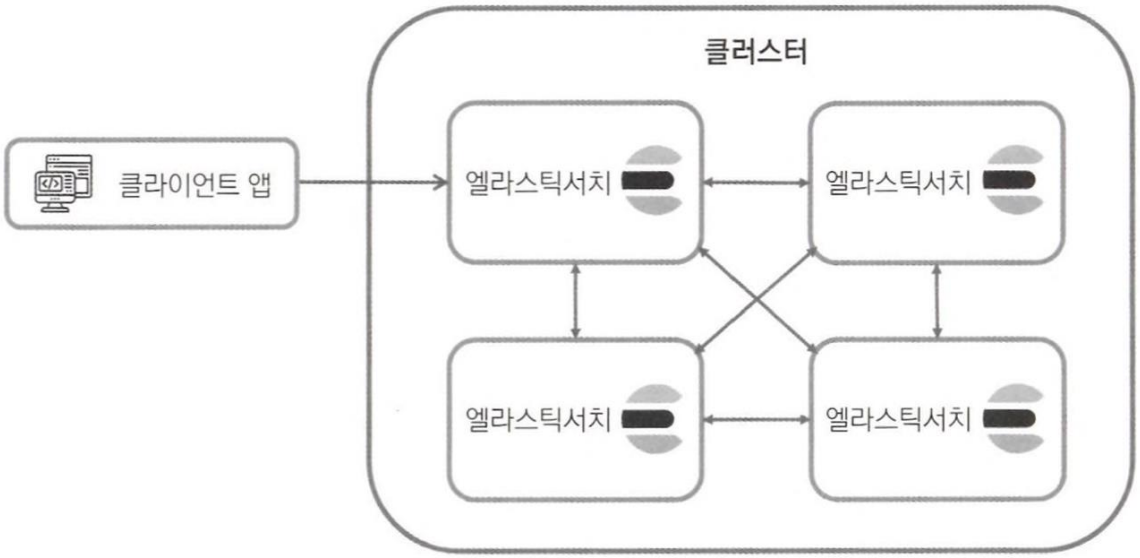
- Elastic Stack 구성 요소
  - ✓ 엘라스틱 서치(Java): 분산 검색 엔진
    - 엘라스틱 서치는 검색 엔진(search engine)이지만 구글이나 네이버 같은 포털 서비스와는 다름
    - 검색 엔진은 내부적으로 각 도큐먼트를 인덱싱(Indexing)하고 빠르게 검색하는 데 사용하는 기술로 검색 엔진을 이용해 상위에서 구글, 네이버 같은 서비스를 만들 수 있음
    - 엘라스틱 서치는 모든 레코드를 JSON 도큐먼트 형태로 입력하고 관리하고 있으며 일반적인 데이터베이스와 마찬가지로 쿼리한 결과에 대해 일치하는 원본 도큐먼트를 반환
    - 엘라스틱 서치는 텍스트 외에도 숫자, 날짜, IP 주소, 지리(geo), 정보 등 다양한 데이터 타입에 대해 최적화되어 있음
    - 검색 엔진이면서 데이터베이스이기도 하고 텍스트 외에 다양한 데이터 타입을 지원한다는 점이 엘라스틱 서치라는 이름과 쉽게 매칭되지 않지만 엘라스틱 서치를 일종의 NoSQL 데이터베이스 와 유사
    - 엘라스틱 서치는 텍스트나 도큐먼트의 경우 인덱싱 시점에 분석을 거쳐 용어 단위로 분해되고 역인덱스 사전을 구축하는데 숫자나 키워드 타입의 데이터들은 엘라스틱 서치 집계를 위해 집계에 최적화된 컬럼 기반 자료구조를 저장
    - 엘라스틱 서치는 이렇게 최적화된 자료구조들을 바탕으로 병렬 처리나 분산 처리를 할 수 있음

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 엘라스틱 서치: 분산 검색 엔진
    - 엘라스틱 서치가 관계형 데이터베이스는 물론 다른 NoSQL 제품이 상상하기 어려운 빠른 검색과 집계 성능을 실현하는 이유는 검색 엔진인 동시에 데이터베이스이기 때문
    - 이론적으로는 충분한 크기의 엘라스틱 서치 클러스터가 구성되어 있다면 데이터의 양과 무관하게 1초 이내의 응답 속도를 기대할 수 있음
    - 엘라스틱 서치에 대한 기본적인 사용 경험은 몽고디비 같은 도큐먼트 기반 NoSQL 과 유사하지만 다른 NoSQL 제품을 압도하는 검색 기능과 성능이 엘라스틱 서치의 가장 큰 특징이자 활용 목적
    - 검색 엔진으로서 엘라스틱 서치의 중요한 특징 중 하나는 스코어링(scoring) 즉 연관도에 따른 정렬로 단순히 필드값을 기준으로 한 정렬은 어떤 데이터베이스에서도 제공되지만 엘라스틱 서치는 검색어에 대한 유사도 스코어를 기반으로 한 정렬을 제공하는데 특히 복잡한 문자열 콘텐츠에서 검색을 수행할 때 큰 효과를 보임
    - 이 외에도 다양한 스코어링 방법을 포함해 사용자가 정렬 방식을 다양하게 정의할 수 있다는 특성은 용도에 따라 큰 장점으로 작용
    - 검색 엔진으로서의 강력한 기능들은 엘라스틱 서치의 기반이 되는 아파치 루씬(Apache Lucene)에서 비롯됐었는데 루씬은 로우 레벨의 검색 엔진으로서 코드 상에서 자바 라이브러리 형태로 참조해 사용이 가능
    - 엘라스틱서치가 제공하는 데이터 타입 별로 최적화된 자료구조 지원이나 스코어링 등은 루씬의 특징으로 볼 수 있지만 엘라스틱 서치는 이를 분산 시스템으로 확장하며 성능, 유연성, 활용성 면에서 극적인 개선을 이룸

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 엘라스틱 서치: 분산 검색 엔진
    - 엘라스틱 서치 분산 시스템 구성



Elastic Stack

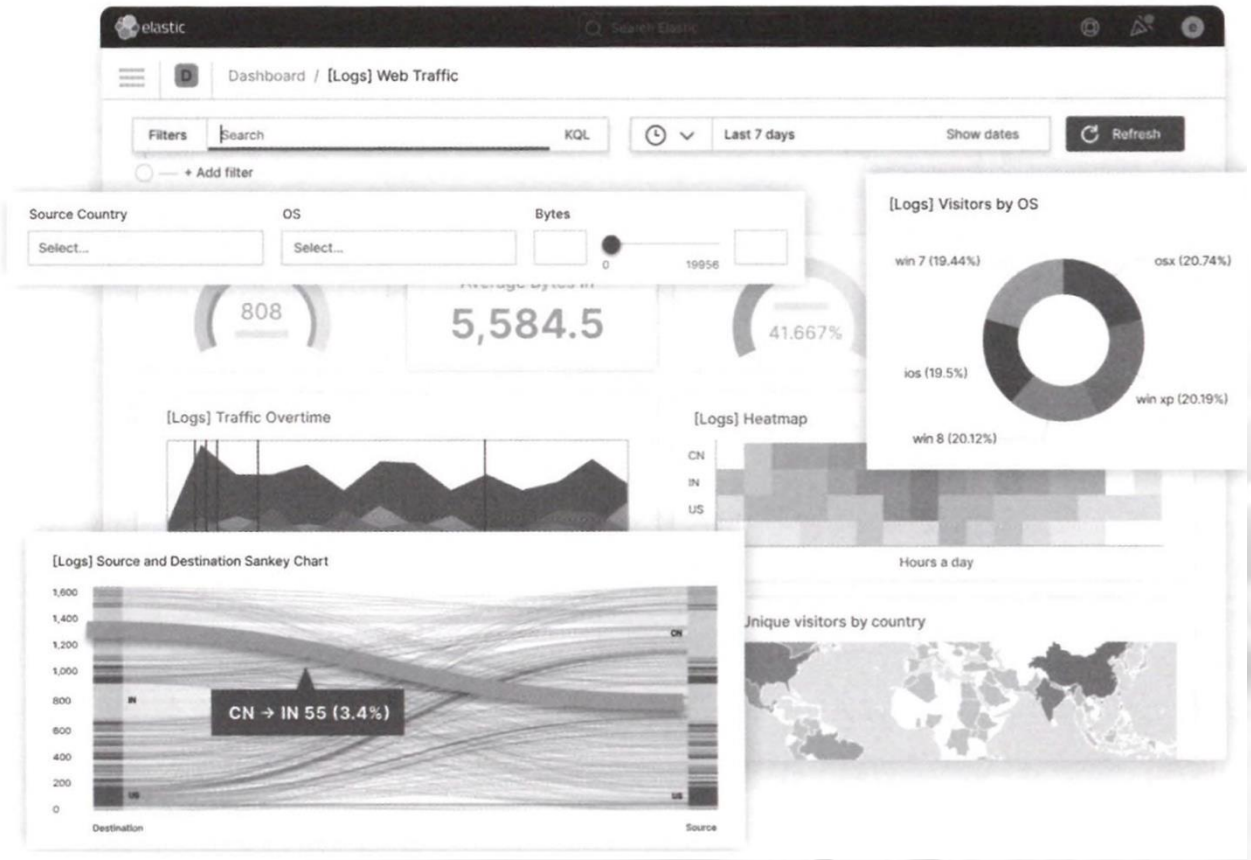
- Elastic Stack 구성 요소
  - ✓ 엘라스틱 서치: 분산 검색 엔진
    - 분산 시스템으로서 엘라스틱 서치는 복수의 루씬 인스턴스를 병렬로 배치하고 분산 처리해 검색 속도를 무한히 확장 할 수 있게 했고 노드 간 복제 기능을 통해 일부 노드가 다운되더라도 정상적으로 서비스를 지속할 수 있게 함
    - 무엇보다 편리한 점은 모든 통신을 REST API를 이용하도록 만들어 프로그래밍 언어와 무관하게 사용자가 쉽게 접근할 수 있도록 활용성을 높였다는 것
    - 엘라스틱 서치에도 단점은 있는데 저장 공간이 크게 압축되지 않고 시스템 리소스를 많이 사용
    - 엘라스틱 서치는 DSL(Domain Specific Language) 쿼리를 채용하는데 JQIN 쿼리가 사실상 어렵기 때문에 반정규화를 기본으로 모델링해야 하고 인덱스가 불변의 자료구조이기 때문에 도큐먼트를 수정하거나 삭제할 경우에 비용이 저렴하지 않지만 이러한 단점들은 검색 성능을 끌어올리기 위해 어느 정도 트레이드오프가 이뤄진 것으로, 엘라스틱 서치가 필요한 수준의 대량 데이터를 처리할 때는 일반적으로 용인되는 제약들
    - 엘라스틱서치는 대용량 데이터에 대한 빠른 검색과 집계가 필요할 경우 우선적으로 고려해야 하는 데이터베이스 중 하나라고 할 수 있음

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 키바나(JavaScript): 시각화와 엘라스틱 서치 관리 도구
    - 엘라스틱 서치는 사용자의 모든 입력을 REST API 형태로 받아들이기 때문에 별도의 드라이버 라이브러리가 없더라도 웹 브라우저나 curl 명령 등 친숙한 도구들을 이용해 기능을 활용할 수 있음
    - 엘라스틱 서치는 REST API가 잘 설계된 제품이긴 하지만 복잡한 요청을 일일이 작성하기에는 다소 불편할 수 있음
    - 키바나는 바로 이런 불편함을 해소해주는데 키바나는 엘라스틱 스택 의 UI를 담당
    - 엘라스틱 서치에 대한 대부분의 관리 기능, API를 실행할 수 있는 콘솔, 솔루션 페이지들, 그리고 스택의 각 구성 요소들을 위한 모니터링 페이지 등이 모두 키바나에 포함되어 있지만 엘라스틱 서치 기반의 시각화 도구로서 키바나의 가장 중요한 기능은 시각화와 대시보드라 할 수 있음
    - 키바나는 일반적으로 많이 사용되는 라인 차트, 파이 차트 등과 테이블, 지도 등의 다양한 시각화 요소들을 클릭 몇 번으로 쉽게 구성할 수 있게 해줌
    - 시각화 요소들을 한 화면에 마음대로 배치하고 실시간으로 업데이트할 수 있는 대시보드 기능은 엘라스틱 스택의 확산과 보급에 지대한 공헌을 함

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 키바나(JavaScript): 시각화와 엘라스틱 서치 관리 도구
    - 키바나의 커스텀 대시보드



Elastic Stack

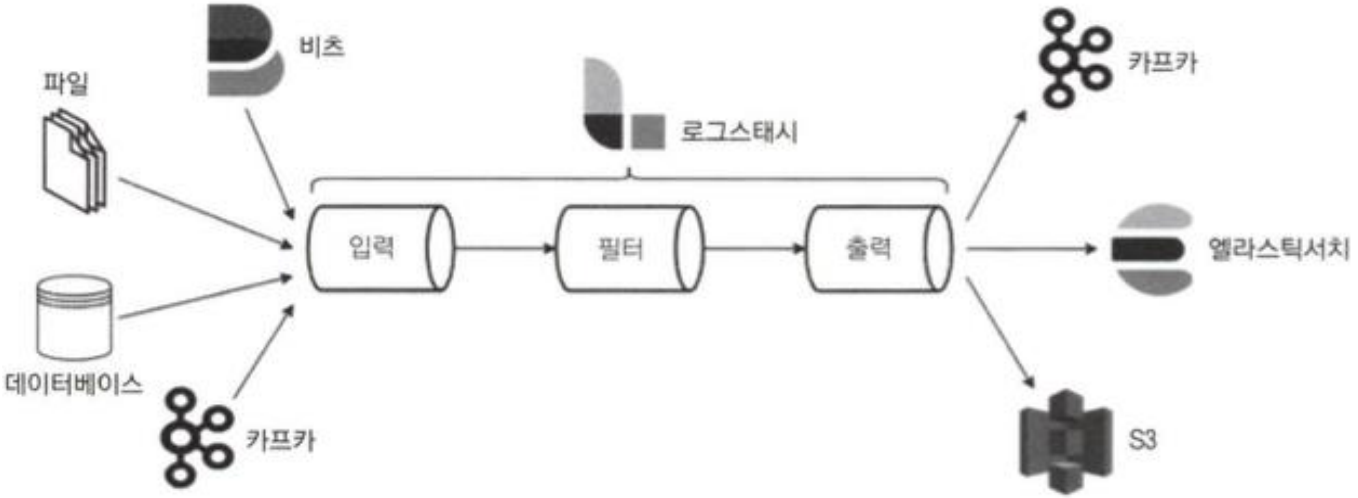
- Elastic Stack 구성 요소
  - ✓ 키바나(JavaScript): 시각화와 엘라스틱 서치 관리 도구
    - 대시보드에서 실시간 모니터 링이 가능하고 데이터 분석도 가능
    - 요즘은 프레젠테이션을 만들듯이 시각화를 구성할 수 있는 캔버스, 서버에 접속해 tail 명령을 사용하듯 실시간으로 인덱싱되는 로그를 지속적으로 확인하고 검색할 수 있는 로그, 애플리케이션의 성능 모니터링을 위한 APM과 보안 이벤트를 관제할 수 있는 SIEM 등 강력한 솔루션들이 포함되어 있어 때로는 키바나를 사용하기 위해 엘라스틱 서치를 고려하는 경우도 있음
    - 키바나는 새로운 버전이 나올 때마다 많은 기능이 추가되는 편이므로 구 버전의 키바나를 사용해본 적이 있다면 한 번쯤 최신 버전의 엘라스틱 스택을 설치해 최신 버전의 키바나를 체험해보는 것이 좋은데 다양한 기능을 체험해볼 수 있도록 샘플 데이터를 기본으로 제공하기 때문에 어렵지 않게 기능을 둘러볼 수 있음

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 로그스태시(Jruby): 이벤트 수집과 정제를 위한 도구
    - 대량의 데이터를 검색하기 위해 가장 먼저 선행돼야 할 작업은 데이터를 적재하는 것인데 이는 빅데이터를 다룰 때 가장 손이 많이 가는 작업이며 변경되는 상황에 맞춰 꾸준히 모니터링을 해야 하고 개선도 필요한 부분
    - 데이터 수집과 가공 기능을 제공하는 로그스태시를 사용하면 로그, 메트릭, 웹 애플리케이션 등 다양한 소스로부터 로그를 수집할 수 있고 필터 기능을 이용해 비정형이나 반정형 데이터를 분석하기 쉬운 형태로 정제할 수 있고 엘라스틱서치 외 에도 다양한 플랫폼으로 정제된 데이터를 내보낼 수 있음
    - 로그스태시는 형식에 무관하게 데이터를 동적으로 수집, 변환, 전송하는 구조로 되어 있음
    - 비구조적인 데이터에서 구조를 도출하며 IP 주소에서 위치 정보 좌표를 해독하고 민감한 필드를 익명화하거나 제외시키는 등의 전반적인 작업을 쉽게 해줌

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 로그스태시(Jruby): 이벤트 수집과 정제를 위한 도구



Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 로그스태시(Jruby): 이벤트 수집과 정제를 위한 도구
    - 로그스태시는 별도의 코딩 없이 간단한 설정만으로 로그를 가공할 수 있음
    - 확장 가능한 200개 이상의 플러그인 덕분에 설정의 대부분은 플러그인 사용 방법인데 그 사용법이 크게 어렵지 않지만 로그스태시가 익숙하지 않다는 이유로 로그스태시가 아닌 커스텀 애플리케이션을 작성해 소스 데이터를 수집하려는 경우도 있음
    - 로그스태시의 진면목은 단순 소스 데이터 정제가 아니고 엘라스틱 서치의 인덱싱 성능을 최적화하기 위한 배치 처리와 병렬 처리가 가능하며 영속적인 큐를 사용해 현재 처리 중인 이벤트의 최소 1회 전송을 보장해줄 뿐만 아니라 유동적인 처리 방식으로 인해 수집 중인 데이터 양이 급증하는 부하 상황에서도 안정성을 보장해주는데 이를 커스텀 애플리케이션으로 구현할 경우 신경써야 할 부분이 많고 문제 발생 시 디버깅이나 튜닝이 쉽지 않다는 점에서 따라오기 어려운 로그스태시의 강점이라 할 수 있고 로그스태시를 사용하면 공동체에서 유사한 요구 사항을 해결한 다양한 사람들의 도움을 기대할 수 있으며 잘 알려진 제품들의 경우 이미 만들어진 플러그인이나 모듈을 활용할 수도 있음
    - 데이터 수집단에서 커스텀 애플리케이션 과 로그스태시를 조합해 사용하는 경우도 있는데 가급적 로그스태시만으로 구성하는 방식이 관리나 안정성 측면에서 장점이 많음

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 비츠(Golang): 엣지단에서 동작하는 경량 수집 도구
    - 로그스태시의 기능은 충분히 강력하지만 이벤트 정보를 수집하기 위해서는 실제 서비스가 동작하는 호스트에 수집기를 설치해야 하는 경우가 많은데 로그스태시는 다양한 필터와 설정을 지원하는 만큼 무겁기 때문에 이러한 목적으로는 활용도가 떨어질 수 있음
    - 엘라스틱 스택 에는 이를 위해 파일비트, 메트릭비트 등 비츠라고 부르는 경량 수집기가 포함되어 있음
    - 각 비트는 로그 수집, 시스템 지표 수집 등 특정 목적에 최적 화된 에이전트이며 가볍기로 유명한 고 프로그래밍 언어로 작성되었고 로그스태시 수준의 복잡한 이벤트 가공은 지원하지 않아 가벼우므로 각 서비스 호스트에 비교적 부담 없이 설치할 수 있음
    - 이러한 양쪽의 장점을 활용하기 위해 비츠와 로그스태시를 혼합해 많이 사용
    - 비츠에서 각 서비스 호스트의 정보를 수집하며 로그스태시에서 이를 취합하고 가공해 엘라스틱서치로 전송하는 형태의 아키텍처를 많이 사용
    - 비츠는 데이터 수집을 위해 용도별로 최적화된 경량 에이전트로 제공되며 특성과 성격에 맞는 다양한 비트들이 존재하고 특정 비트는 잘 알려진 다수의 서비스에 대한 모듈을 제공
    - 모듈에는 수집 설정 프리셋, 샘플 대시보드 등이 포함되므로 별도의 설정 과정 없이도 손쉽 게 이벤트를 수집할 수 있음

Elastic Stack

- Elastic Stack 구성 요소
  - ✓ 기타 솔루션
    - 엘라스틱 스택에는 애플리케이션 성능 모니터링을 위한 APM, 보안 이벤트 분석을 위한 SIEM, 컨테이너나 다수의 서비스를 쉽게 모니터링하기 위한 인프라 모니터링 등의 솔루션이 포함되어 있음
    - 비츠 같은 대표적인 사례를 보면 알겠지만 엘라스틱 스택의 버전이 올라가며 엘라스틱 스택을 사용해 사용자들이 손수 구현해오던 많은 시스템들이 솔루션으로 추가되고 있음

Elastic Stack

- Elastic Stack 용도
  - ✓ 전문 검색 엔진
    - 엘라스틱 스택의 가장 기본적인 사용처는 바로 전문 검색 엔진 구현
    - 전문(Full Text)은 단순한 문장부터 뉴스 기사나 논문 등 다양한 글의 전체 내용을 의미
    - 대상 도큐먼트가 많지 않다면 일반적인 관계형 데이터베이스의 LIKE 질의만으로도 충분히 검색이 가능하지만 도큐먼트 수가 조금만 늘어나도 인덱스의 도움이 없이 빠른 검색은 불가능에 가까움
    - 구글에서 엘라스틱서치 사용해야 하는 이유로 검색을 하면 관련된 문서를 찾아주는데 이런 검색이 바로 전문 검색
    - 전문을 빠르고 정확하게 검색하기 위해 전문을 용어(terms) 단위로 분석해 인덱싱해두고 이를 기반으로 검색을 수행하는 역인덱싱(inverted indexing) 기법이 많을 활용
    - 전문 검색은 관계형 데이터베이스와 커스텀 애플리케이션을 이용해 직접 개발할 수 있지만 완성도와 정확도가 떨어지며 불필요한 개발 공수가 추가되기 때문에 추천하지 않음
    - 특히 빅데이터에서 역인덱싱을 이용해 전문 검색을 해야 한다면 엘라스틱 서치 같은 검색 엔진을 활용하면 됨
    - 엘라스틱 스택은 엘라스틱서치, 로그스태시, 키바나, 비츠 등을 쉽게 사용할 수 있게 온라인 문서를 제공
    - 특정 단어를 검색하면 엘라스틱 서치가 온라인 문서를 대상으로 검색 엔진을 통해 찾은 결과를 보여줌

Elastic Stack

- Elastic Stack 용도
  - ✓ 전문 검색 엔진
    - 역인덱싱이라는 기술 자체는 잘 알려져 있고 단순하게 보면 구현이 크게 어렵지는 않지만 원문을 어떻게 수집하고 어떤 방식으로 분석해 용어를 만들고 검색 시에 어떤 기준으로 유사도를 측정하고 어떤 문서를 결과 상단에 위치시킬지 와 같은 작업은 단순하지 않음
    - 실제 검색 서비스를 만들기 위해서는 고려해야 할 사항 이 한두 가지가 아님
    - 엘라스틱 스택에는 용어 분석을 위한 다양한 언어별 분석기가 준비되어 있고, 유사도 스코어링을 위해서도 다양한 방법을 제공할 뿐 아니라 로그스태시를 이용해 도큐먼트 수집 단계에서부터 도움을 주며 키바나를 이용해 별도의 UI 없이도 검색 동작을 테스트해볼 수 있음
    - 전문 검색 엔진 외에도 레거시 관계형 데이터베이스에 대한 외부 검색 엔진으로 활용되기도 하는데 이 경우 엘라스틱 서치에는 원문을 저장할 필요가 없기에 큰 용량을 차지 하지는 않으면서도 빠른 검색 성능을 기대할 수 있다는 장점이 있지만 굳이 관계형 데이터베이스에 원문을 유지할 필요가 없다면 하나만 사용하는 방식이 관리상 편리하기 때문에 엘라스틱 서치로 완전히 전환하는 경우가 많음
    - 검색 엔진 용도로 엘라스틱 서치를 활용할 때는 일반적으로 도큐먼트의 양이 다른 용도에 비해 많지 않기 때문에 상대적으로 작은 클러스터 크기로도 충분한 활용이 가능하며 레플리카 수를 늘려 동일 인덱스에 대한 검색 성능을 높일 수도 있음

Elastic Stack

- Elastic Stack 용도
  - ✓ 로그 통합 분석
    - 애플리케이션에서 발생되는 로그는 디버깅에 있어 중요한 요소 중 하나
    - 애플리케이션이 단순하고 발생되는 로그의 양이 많지 않다면 일반적인 텍스트 에디터나 간단한 셸스크립트만으로도 발생되는 로그를 충분히 파악할 수 있지만 애플리케이션이 복잡해져서 로그의 발생량 이 많아지고 복수의 서비스가 연계되어 동작한다면 로그의 발생 위치를 모두 기억해야 할 뿐 더러 발생 순서를 추적하기도 어렵고 이런 로그는 커스텀 애플리케이션 외에 운영을 위해 사용되는 여러 서비스와 기반 시스템에서도 발생되기 때문에 서비스가 조금이라도 커지면 단순히 여러 창에 유닉스 tail 명령을 이용해 로그를 추적하는 방식은 한계에 부딪히게 됨
    - 엘라 스틱 스택은 여러 장비와 서비스에서 발생하는 로그들을 통합하고 검색하는 데 최적화된 솔루션이라 할 수 있는데 시스템과 호스트 쿠버네티스, 아파치, MySQL, 윈도우 같은 다양한 환경에 서 생성되는 로그를 별도의 복잡한 구성 없이도 바로 수집 가능
    - 비츠를 사용하면 적은 리소스로 각 장비의 로그들을 빠르게 수집할 수 있고 로그스태시는 다양한 필터를 통해 일원화된 형태로 가공을 도우며 엘라스틱 서치의 대용량 로그에 대한 빠른 인덱싱 성능과 텍스트 검색 능력은 여러 곳에 흩어진 서비스 로그들을 통합해서 연관 분석을 지원
    - 키바나의 로그 UI 나 대시보드는 시계열로 발생하는 로그들을 직관적으로 모니터링할 수 있도록 도와줌

Elastic Stack

- Elastic Stack 용도
  - ✓ 로그 통합 분석
    - 키바나에서 제공하는 로그 모니터링 화면

Elastic Stack

- Elastic Stack 용도
  - ✓ 로그 통합 분석
    - 로그 UI는 각 서버에서 수집된 로그를 한 페이지에서 실시간으로 추적할 수 있게 도와줌
    - 여러 곳에 흩어진 서버의 로그를 한 페이지로 본다는 사실에 주목해야 하는데 운영 업무에서는 이런 형태의 시각적인 모니터링 도구가 중요
    - 로그 모니터링을 위해 엘라스틱 스택에서 할 수 있는 업무
      - 먼저 파일비트는 내장된 모듈을 이용해 별도의 설정 없이도 빠르게 로그를 수집할 수 있게 해주는데 파일비트의 모듈은 알려진 서비스들을 정제하기 위한 파이프라인 설정, 샘플 대시보드 등을 포함하고 있기 때문
      - 커스텀 애플리케이션에서 발생한 로그의 경우 로그스태시의 필터 기능이나 엘라스틱 서치의 수집 파이프라인 기능으로 정제가 가능
      - 패턴을 기반으로 원문을 가공해 반정형의 텍스트 메시지로부터 정형화된 수치 등의 정보를 뽑아내어 인덱싱하면 훨씬 더 다양한 방식으로 로그 분석이 가능
      - 엘라스틱서치에서는 로그 원문의 빠른 검색 후 하이라이트가 가능할 뿐 아니라 날짜나 수치에 대한 범위 조회, IP 대역 조회 등은 로그 분석 시 큰 도움이 됨

Elastic Stack

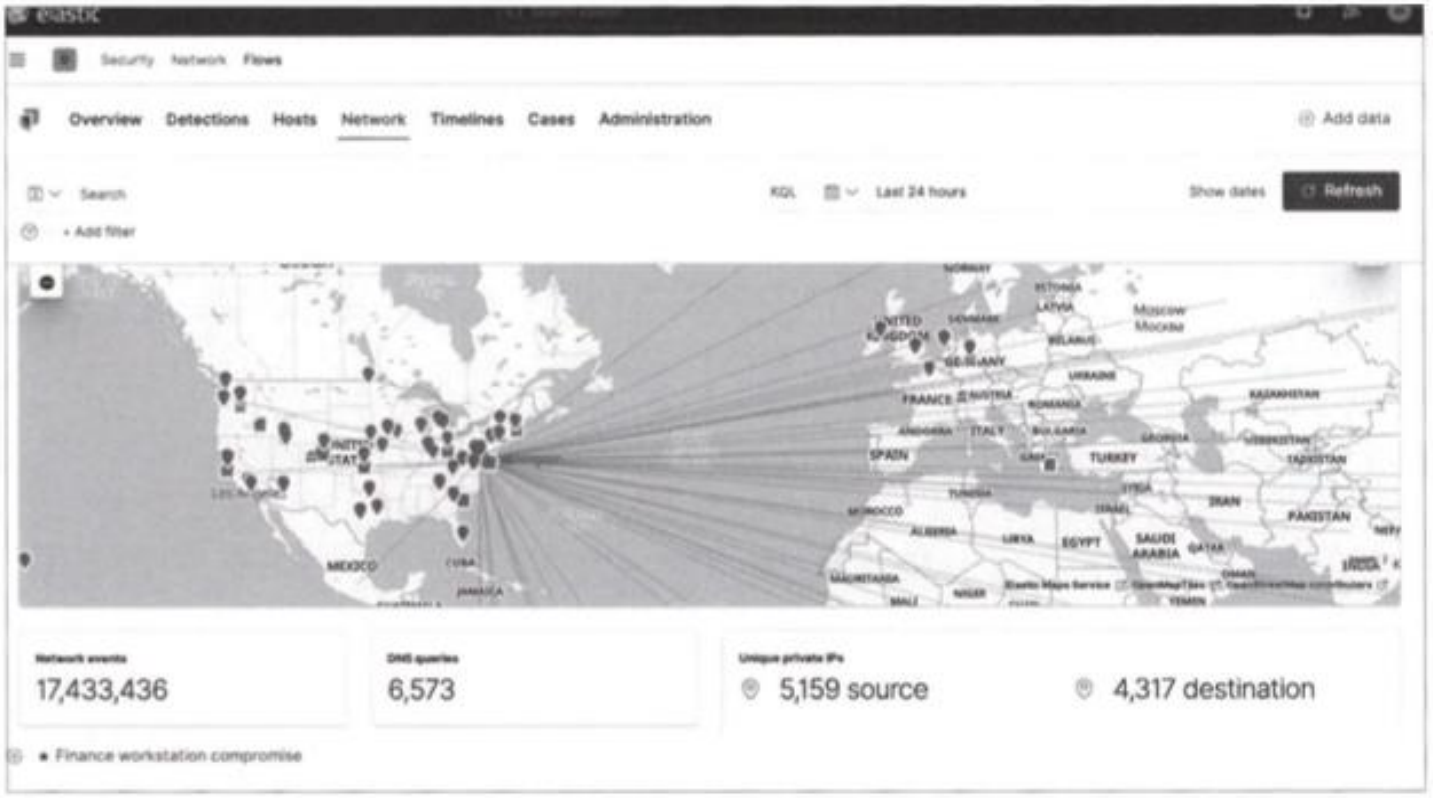
- Elastic Stack 용도
  - ✓ 로그 통합 분석
    - 로그 모니터링을 위해 엘라스틱 스택에서 할 수 있는 업무
      - 엘라스틱 서치에 저장된 데이터들은 손쉽게 교차 분석이나 연관 분석이 가능
      - 관계형 데이터베이스와 다르게 엘라스틱 서치는 인덱스 패턴을 이용해 한 번에 여러 인덱스에서 동시에 조회가 가능한데 이는 복잡한 JOIN 쿼리 없이도 각기 다른 서비스에서 동시간 대에 어떤 로그가 발생했는지 탐색하고 공통된 요소를 찾아 문제의 원인을 분석할 수 있게 해줌
      - 비츠의 모듈을 사용할 경우 엘라스틱에서 제공하는 ECS(Elastic Cmmon Schema) 구조에 맞춰 로그를 인덱싱하는데 이 기능은 다른 서비스들을 같은 스키마에 매핑함으로써 연관 분석을 용이하게 해줌
      - 2개의 다른 서비스가 각각 srcIP라는 필드명과 source_ip라는 필드명으로 소스 IP 정보를 매핑했다고 가정하면 일반적으로 각 로그에 대한 쿼리를 별도로 작성해야 하겠지만 표준 스키마에 맞춘다면 필드명이 동일하므로 하나의 쿼리로 검색이 가능
      - 커스텀 애플리케이션의 로그를 수집하더라도 ECS 포맷에 맞춘다면 엘라스틱 스택의 각종 기능들과 손쉬운 연계가 가능

Elastic Stack

- Elastic Stack 용도
  - ✓ 보안 이벤트 분석

Elastic Stack

- Elastic Stack 용도
  - ✓ 보안 이벤트 분석



Elastic Stack

- Elastic Stack 용도
  - ✓ 보안 이벤트 분석
    - SIEM(Security Information and Event Management)이나 ESM(Enterprise Security Management) 등으로 불리는 솔루션은 조직 내에 속한 다양한 장비들로부터 보안 이벤트를 수집하고 분석할 수 있게 하려는 목적으로 만들어졌음
    - 보안과 관련된 이벤트는 다양한데 보안 장비, 웹 서버, 애플리케이션, 각 사용자 의 PC 등에서 실시간으로 발생하는 이벤트의 양은 다른 어떤 로그와 비교해도 무시할 수 없는 수준이지만 공격 방식의 다변화에 따라 이 이벤트들은 실시간으로 수집됨과 동시에 실시간 검색과 필터링을 통한 연관 분석과 머신러닝을 이용한 이상 징후 탐지까지 연결돼야 함
    - 엘라스틱 SIEM은 비츠의 모듈을 이용해 애플리케이션, 엔드포인트, 인프라스트럭처, 클라우드, 네트워크 등 다양한 소스에서 수집한 이벤트를 기반으로 엔드포인트 활동, 인증 로그, DNS 트래픽, 네트워크 플로우 에서 이상 징후, 불법적인 로그인 시도, 사용자 접근 패턴 등의 문제를 빠르게 찾아낼 뿐만 아니라 SIEM UI에서는 분석을 비롯한 고유한 탐지 규칙 관리 또한 가능하고 미리 빌드된 머신러닝 모델과 즉시 사용 가능한 알고리즘을 통해 알려지지 않은 위협도 밝혀낼 수 있음
    - 엘라스틱 서치의 실시간 검색 성능, 비츠의 이기종 장비들에 대한 각종 이벤트 수집 기능, ECS에 의해 통일된 스키마에 기반한 원활한 연관 분석, 키바나에서 제공되는 SIEM, 지도, 그래프 등의 각종 지원 UI로 인해 보안 이벤트 분석은 엘라스틱 스택의 가장 대표적인 사용 사례 중 하나라 할 수 있음

Elastic Stack

- Elastic Stack 용도
  - ✓ 애플리케이션 성능 분석

Elastic Stack

- Elastic Stack 용도
  - ✓ 애플리케이션 성능 분석
    - 애플리케이션을 운영하는 데 있어 가장 중요한 것 중 하나는 안정적인 서비스 유지를 하는 것인데 이를 위해서는 애플리케이션의 상태를 지속적으로 모니터링해야 함
    - 단순히 하나의 애플리케이션뿐만 아니라 연계된 다른 모든 서비스의 성능 정보를 모니터링해야 함
    - 엘라스틱 스택의 애플리케이션 성능 모니터링 도구인 APM은 프로그래밍 언어별 에이전트를 통해 성능 지표 수집을 돕고 분석을 위한 이를 제공할 뿐 아니라 메트릭비트와 패킷비트를 사용하면 시스템을 비롯해 여기 연계된 다양한 서비스들의 성능 정보를 수집할 수 있게 도와줌
    - 엘라스틱 APM은 로그, 인프라 모니터링 등 다른 UI 와 연계했을 때 효과가 배가되며 서비스 지도를 통한 서비스 연결 관계 파악, 트랜잭션의 분산 추적, 경보 기능을 사용한 즉각적인 대응, 자바, 노드, PHP, 닷넷 같은 다양한 프로그래밍 언어 지원, 머신러닝을 사용한 비정상적인 응답 시간 탐지와 같은 APM의 핵심적인 기능을 제공하므로 단순 모니터링을 넘어 문제 원인 분석과 서비스의 상태 파악을 위한 관측성 강화에 이바지하고 있음
    - 다양한 경로로 수집된 정보 들은 컨테이너나 호스트 집합의 상태를 한눈에 확인할 수 있는 인프라 UI. 메트릭비트에서 모듈 별로 제공하는 샘플 대시보드, 트랜잭션이나 오류 스택 트레이스 등을 빠르게 확인할 수 있는 APM UI 등 키바나에 제공되는 이를 통해 분석이 가능할 뿐 아니라 다른 도큐먼트와 마찬가지로 인덱스 형태로 저장되기 때문에 일반적인 쿼리와 집계가 가능하고 키바나의 시각화와 대시보드를 활용할 수도 있음

Elastic Stack

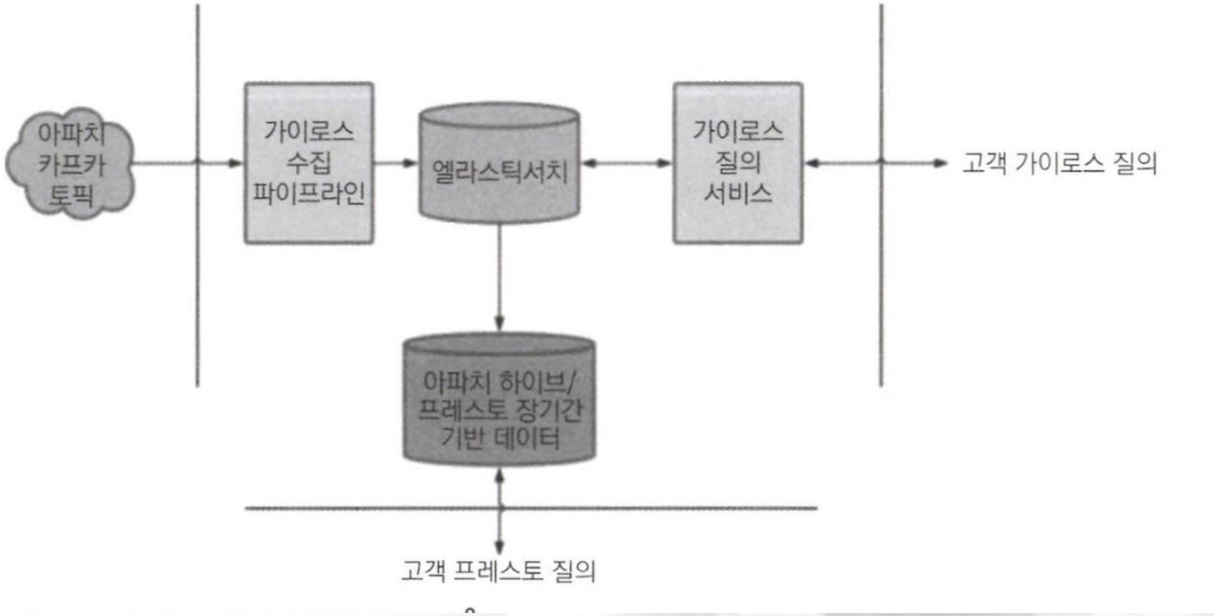
- Elastic Stack 용도
  - ✓ 애플리케이션 성능 분석
    - 사용자는 수집된 지표들을 활용해 고유한 대시보드를 만들어 자신의 애플리케이션 특성에 따른 맞춤형 모니터링이 가능하고 로그 통합 분석 과 함께 사용할 경우 문제 시점의 로그를 빠르게 확인하고 원인을 파악할 수 있음
    - 유료 서브스크립션에 포함된 기능인 와처(Watcher) 나 머신러닝 등의 기능과 연계하면 장애나 이상 상황을 빠르게 파악하고 원인을 효과적으로 분석할 수 있음
    - APM 전문 솔루션 대비 가장 큰 장점은 이러한 확장성
    - 기본적으로 제공되는 기능은 전문 솔루션에 비해 부족할 수 있으나 커스터마이징을 통해 자신의 비즈니스 로직에 최적화된 모니터링 솔루션을 구축할 수 있음

Elastic Stack

- Elastic Stack 용도
  - ✓ 머신러닝
    - 보안 이벤트 분석과 애플리케이션 성능 분석에서도 잠깐 등장한 머신러닝은 엘라스틱 스택에서 유료 라이선스를 구매할 경우 사용할 수 있는 강력한 기능 중 하나
    - 데이터를 엘라스틱 서치에 넣은 다음에 비지도 머신러닝 기법을 활용해 데이터에서 패턴을 발견할 수 있음
    - 시계열 모델링은 시간과 관련된 데이터에서 이상 징후를 탐지하고 과거 데이터를 기반으로 동향을 예측하게 도와주는데 이런 기능을 활용해 SEM 통합, APM 통합, 로그 통합 부문에서 비정상적인 상황을 특별한 규칙 지정 없이도 빠르게 발견할 수 있음

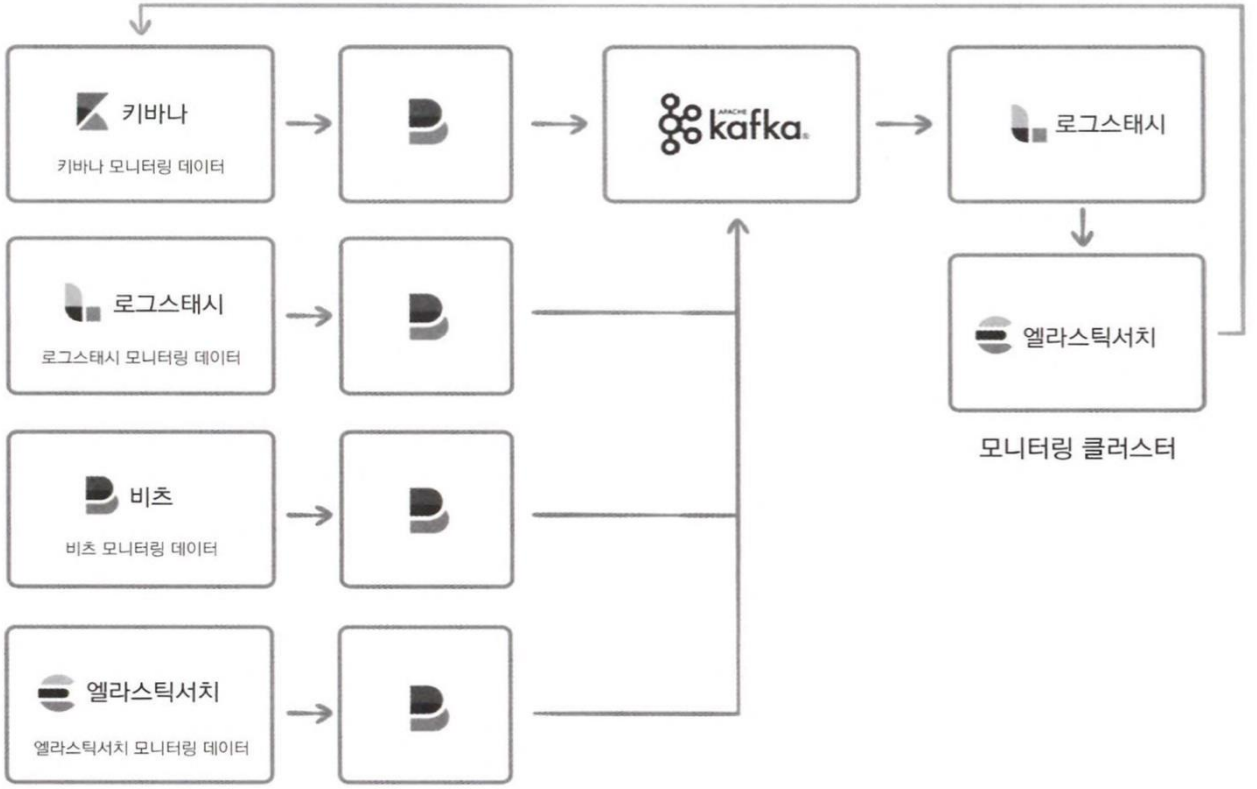
Elastic Stack

- 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 우버는 대규모 실시간 데이터 통찰 플랫폼인 가이로스(Gairos)를 구축해 가격 책정, 최대 ETA(도착 예정 시간) 계산, 수요/공급 예측과 같이 정보에 입각한 결정을 내리고 있는데 가이로스는 아파치 카프카(Kafka)와 아파치 하둡 생태계(HDFS/Hive/Presto)를 연계해 실시간 분석과 장기 분석을 지원하며 동적인 가격 정책과 환경 친화적인 운전 경로를 운전자에게 제공



Elastic Stack

- 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 엔터프라이즈 데이터 버스인 카프카와 연동
    - 카프카를 이용한 안정적인 시스템 구성

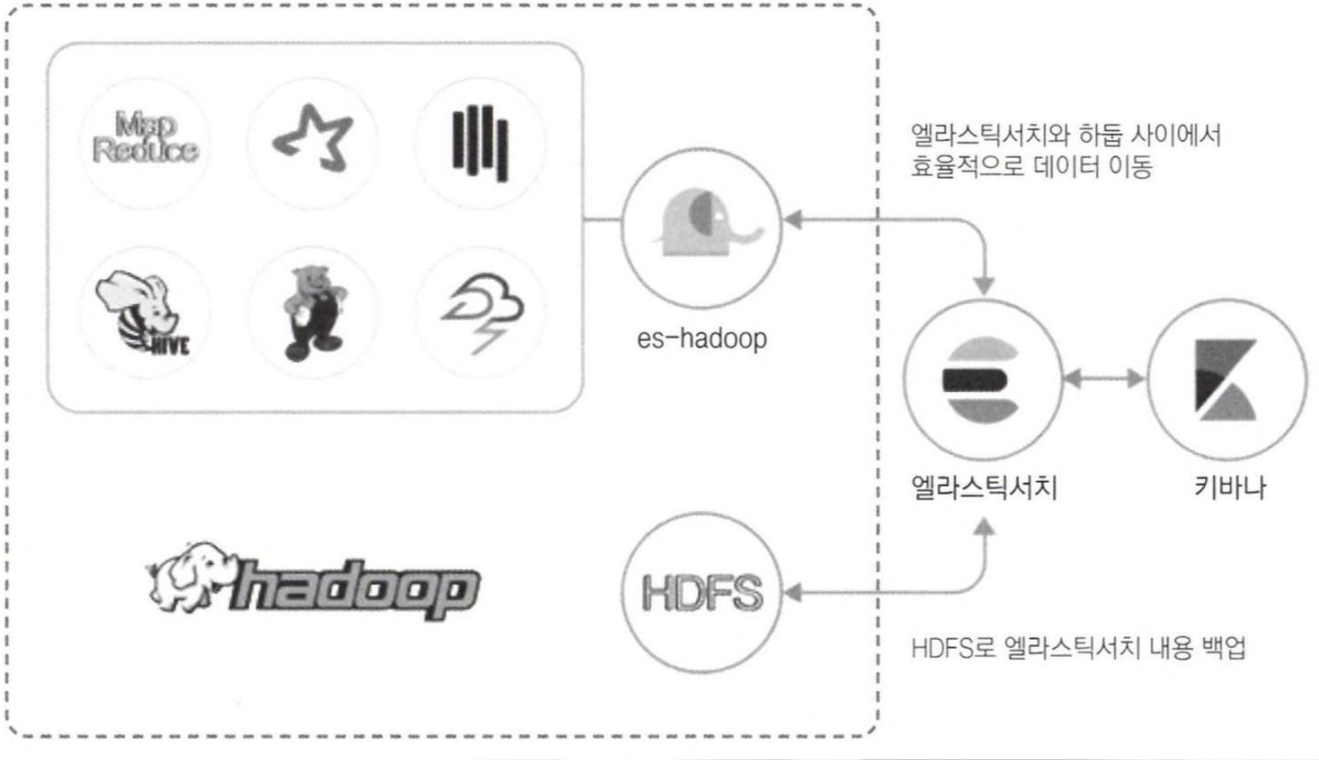


Elastic Stack

- ❖ 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 엔터프라이즈 데이터 버스인 카프카와 연동
    - 아파치 카프카는 분산 데이터 스트리밍 플랫폼
    - 대량의 데이터를 실시간으로 배포하는데 최적화되어 있어 많은 빅데이터 플랫폼에서 기본적으로 채택하고 있음
    - 카프카와 엘라스틱 스택은 데이터 수집 시점에서 많이 연계되어 사용
    - 카프카를 연계해 비츠, 로그 스태시, 엘라스틱 서치와 통신하는 아키텍처
    - 비츠에서 수집한 각 장비의 이벤트를 카프카로 전송하고 이를 로그스태시로 다시 읽어 들이거나 카프카에 저장된 다른 시스템의 이벤트를 엘라스틱 서치로 읽어 들이는 등의 구성으로 자주 사용
    - 카프카를 사용하면 엘라스틱 서치의 인덱싱 성능이 순간적으로 충분하지 않거나 로 그스태시나 엘라스틱 서치의 불안정으로 정상적인 인덱싱이 불가능할 때도 데이터의 유실이 방지되며 타 시스템과 데이터 연계가 손쉬워진다는 장점이 있

Elastic Stack

- 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 하둡 생태계 와의 연동
    - es- hadoop을 이용한 하둡 과 엘라스틱 서치의 연계

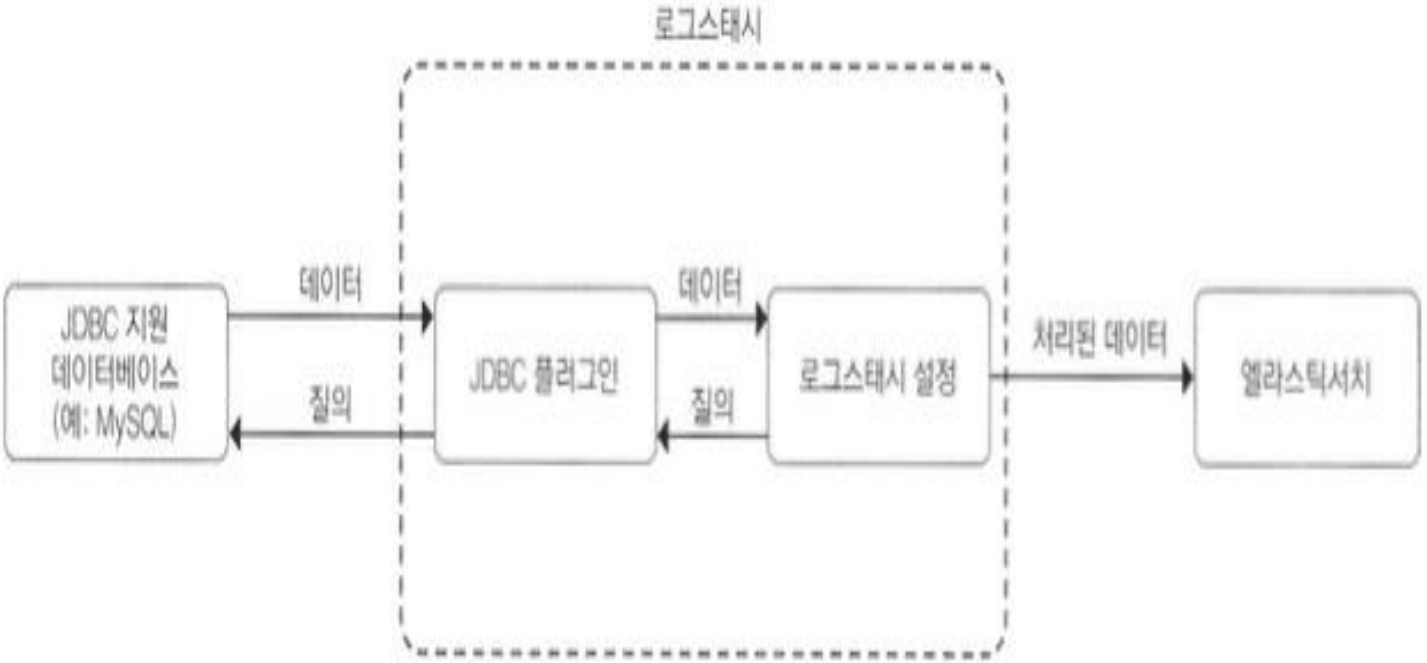


Elastic Stack

- ❖ 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 하둡 생태계 와의 연동
    - 우버에서는 이후에 시스템이 확장됨에 따라 빠른 처리를 위해 아파치 스파크(Spark)까지 도입
    - 엘라스틱 스택에서 es-hadoop 모듈을 사용하면 스파크에서 엘라스틱 서치 API를 사용해 도큐먼트를 읽어 들이고 반대로 인덱싱하는 등 서로 상호 작용을 수행할 수 있음
    - 엘라스틱 서치는 이미 인덱싱된 결과에 대해 빠른 검색을 제공하고 스파크는 기존에 인덱싱되지 않은 데이터에 대한 빠른 일괄 처리를 제공하기 때문에 엘라스틱 서치를 중간 집계 엔진이나 결과를 저장해 재활용하는 용도 등으로 활용할 수 있음
    - 두 솔루션에 모두 익숙해진다면 서로의 단점을 보완해 가장 빠르게 원하는 결과를 도출해낼 수 있을 것

Elastic Stack

- 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 관계형 데이터베이스와 연동



Elastic Stack

- 빅데이터 플랫폼의 일부로 동작하는 엘라스틱 스택
  - ✓ 관계형 데이터베이스와 연동
    - 관계형 데이터베이스에는 주로 레거시 데이터나 신뢰성이 중요한 데이터를 저장
    - 로그스태시는 JDBC 입력 플러그인, 필터 등 관계형 데이터베이스와 연계할 수 있는 다양한 방법을 제공
    - 기존의 관계형 데이터베이스에 저장된 데이터를 인덱싱하거나 입력받는 이벤트에 정보를 주입하는 등 여러 용도로 사용
    - JDBC 입력 플러그인을 사용해 관계형 데이터베이스에 저장된 데이터를 엘라스틱 서치로 이전할 경우 엘라스틱 스택의 우수한 성능을 이용해 관계형 데이터베이스로는 처리하기 어려운 집계도 빠르고 정확하게 처리 할 수 있고 텍스트 데이터의 경우 엘라스틱 서치에서는 색인만 수행하고 원문을 저장하지 않게 설정하는 방식을 이용해 저장소 용량을 아껴 쓰면서 검색 엔진의 전문 검색 기능만 활용하는 경우도 있음
    - JDBC 필터를 사용할 경우 관계형 데이터베이스에서 실 데이터를 불러와 매핑 하는 용도로 이벤트에 ID 형태로만 기록되는 정보를 활용할 수 있음

Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 엘라스틱 서치의 유사 제품군
    - 엘라스틱 서치는 검색 엔진이자 데이터를 저장할 수 있는 NoSQL 데이터베이스로도 볼 수 있음
    - 요즘은 NoSQL 데이터베이스 와 관계형 데이터베이스 간에도 점점 경계가 약해지는 추세이긴 하지만 NoSQL 데이터베이스는 범용으로 사용되는 관계형 데이터베이스 와 비교해 특정한 목표에 맞춰 최적화되는 경우가 많으며 강한 트랜잭션 기능을 포기하는 대신 확장성과 사용성에 초점을 맞춤
    - 최근에는 MySQL이나 PostgreSQL 같은 오픈 소스 관계형 데이터베이스에서도 JSON 형식을 지원하기 때문에 엘라스틱 이나 몽고디비 같은 도큐먼트 저장 목적으로도 일부 사용이 가능한 상태로 발전하고 있음
    - 엘라스틱 서치의 경우 근본이 데이터베이스가 아닌 검색 엔진이기 때문에 데이터를 빠르게 검색 하고 집계하는 데 최적화되어 있으며 그 대신 더 많은 시스템 리소스를 사용하기 때문에 일반적인 범용 데이터베이스처럼 사용하기에는 무리가 있음
    - 특히 저장된 데이터를 부분적으로 수정/삭제하는 작업은 내부적인 관점에서 보면 효율적이지 못하므로 시간에 따라 데이터가 발생하며 이미 발생한 데이터는 변하지 않는 immutable 시계열 데이터 유형을 다루기에 좋음
    - 많은 데이터베이스가 자체적으로 인덱스를 지원해 빠른 검색을 가능케 하지만 엘라스틱 서치의 인덱스 기술이나 지원되는 쿼리의 종류는 여타 데이터베이스와는 큰 차이가 있음

Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 엘라스틱 서치의 유사 제품군
    - 솔라는 엘라스틱 서치와 마찬가지로 루씬을 기반으로 만들어졌기에 상당히 유사
    - 솔라는 엘라스틱 스택 같은 통합된 플랫폼으로 발전하지 못했기에 최근에는 엘라스틱이 압도적으로 앞서나가는 상황
    - 엘라스틱은 스택의 다른 구성 요소들의 도움을 받아 로그 분석, 모니터링, 위치 기반 데이터 분석과 시각화에 적합하고 솔라는 정적 데이터를 처리하기에 적합
    - 엘라스틱은 솔라에 비해 월등히 많은 데이터 소스와 쿼리 DSL은 물론이고 다양한 쿼리 파서를 지원하고 있으므로 데이터 수집과 분석에 유리하며 주키퍼(ZooKeeper)에 의존하지 않으므로 분산 환경 구축과 운영 측면에서 우위에 서 있음

Elastic Stack

❖ 유사 제품 과의 비교

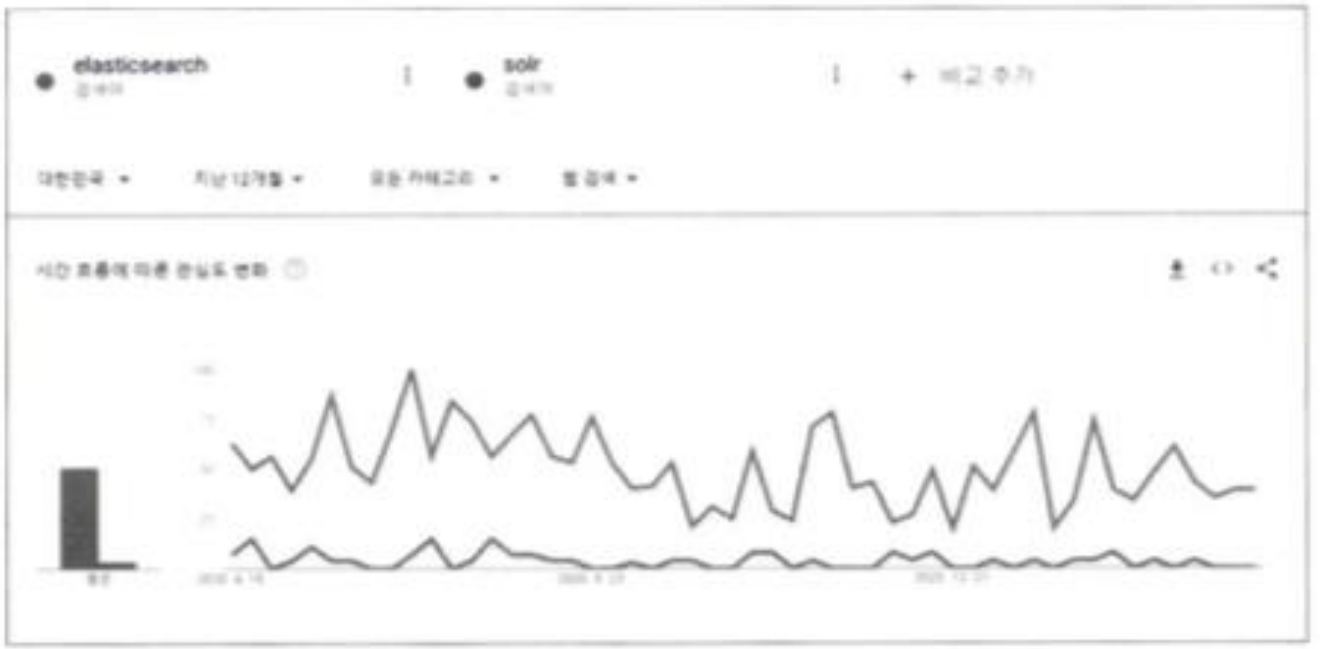
✓ 엘라스틱 서치의 유사 제품군

● 엘라스틱 서치 와 데이터베이스 비교

	엘라스틱서치	몽고디비	MySQL
분류	검색 엔진	도큐먼트 저장소	관계형 데이터베이스
최초 릴리스	2010년	2009년	1995년
스키마	자동 생성	자유	필요
인터페이스	REST API	전용 프로토콜	JDBC/ODBC
분산 저장	샤딩	샤딩	별도 제품으로 지원
트랜잭션	미지원	부분 지원	지원
JOIN	미지원	최근 도큐먼트 단위로 지원 시작	지원
특징	다양한 데이터 유형에 대한 빠른 검색/집계	높은 활용성/범용성	뛰어난 데이터 무결성

Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 엘라스틱 서치의 유사 제품군
    - ● 엘라스틱 서치 와 솔라 검색 추이



Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 로그스태시/비츠의 유사 제품군
    - 엘라스틱 스택에는 로그스태시 와 비츠라는 훌륭한 수집 기들이 포함되어 있는데 로그스태시는 상대적으로 무겁지만 다양한 입력, 출력, 필터 플러그인을 갖춰 설정만으로 거의 모든 형태의 가공을 수행할 수 있고 비츠는 각 비트의 기능은 제한적인 대신 본연의 기능에 충실해 빠르고 가볍게 이벤트를 수집할 수 있음
    - 엘라스틱 스택 내에서는 이 두 가지 수집기를 조합해 각 수집 대상으로부터 비츠를 이용해 빠르게 이벤트를 수집하고 로그스태시로 전달한 다음에 가공을 거쳐 엘라스틱 서치로 적재하는 식으로 유연한 활용이 가능
    - 데이터 플랫폼에 사용되는 유사한 제품으로 트레저 데이터(Treasure Data)에서 개발한 플루언트디(Fluentd) 와 비교할 수 있는데 플루언트디는 로그스태시와 설정의 형태만 다를 뿐 기능상으론 거의 동일한 역할을 수행할 수 있지만 엘라스틱 스택에 한정지어 사용한다면 로그스태시는 엘라스틱 서치, 키바나와 연계해 모니터링, 중앙 관리 등의 기능을 제공하는 한편 플루언트디는 이러한 지원을 받을 수 없기 때문에 되도록 로그스태시와 비츠의 조합을 사용하길 권장
    - 플루언트디는 CNCF(클라우드 환경에서 애플리케이션을 배포하고 작은 서비스를 패키징하는 기술) 스택이며 쿠버네티스나 오픈트레이싱(OpenTracing)을 사용하는 경우에 사실상 필수 구성요소로 볼 수 있기때문에 로그스태시와 플루언트디는 자신만의 영역을 점유하고 있음

Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 키바나의 유사 제품군
    - 키바나는 실시간 대시보드, 프레젠테이션 슬라이드처럼 시각화를 수행할 수 있는 캔버스 기능 등 엘라스틱 스택의 강력한 시각화 도구일 뿐만 아니라 엘라스틱 스택의 전반적인 모니터링, 관리 기능을 아우르는 유저 ui로서의 역할도 수행
    - 엘라스틱 서치만을 사용한다면 키바나는 선택이 아닌 필수라 해도 과언이 아님
    - 비교되는 시각화 도구로는 키바나를 포크해서 만든 그라파나가 대표적인데 전반적으로 키바나보다 기능은 부족하다고 볼 수 있으나 데이터 소스로 엘라스틱 서치 외의 다양한 시계열 데이터베이스를 활용할 수 있다는 장점이 있음
    - 타 데이터베이스의 데이터를 엘라스틱 서치로 이전한다면 키바나를 사용해 더 다양한 기능을 활용할 수 있겠지만 데이터 플랫폼 내의 여러 데이터베이스들을 ui 수준에서 통합해보고 싶다면 키바나 와 그라파나를 함께 사용하는 것도 좋은 선택이 될 수 있음
    - 태블로같은 비즈니스 분석 도구들은 용도가 조금 다른 데 실시간 모니터링에 초점을 맞춘 키바나 나 그라파나 와는 달리 엑셀이나 관계형 데이터베이스, 기타 형식의 파일 등 주로 정적인 파일을 분석해서 시각화할 때 많이 사용되므로 목적에 맞는 도구를 선택할 필요가 있음

Elastic Stack

❖ 유사 제품 과의 비교

✓ 키바나의 유사 제품군

● 시각화 도구 비교

	키바나	그라파나	태블로
실행 환경	웹 서비스	웹 서비스	설치/웹 서비스 모두 제공
데이터 소스	엘라스틱서치	엘라스틱서치, 그라파이트(Graphite), 몽고디비, 프로메테우스(Prometheus) 등	관계형 데이터베이스, 엑셀 파일, JSON 파일, 드롭박스 등
특징	실시간 대시보드 외에도 캔버스, 지도, 비츠/로그스태시와 연계되어 제공되는 솔루션 등 엘라스틱서치의 기능을 십분 활용한 다양한 시각화 세트 제공	라인, 바 차트 등 기본적이지만 인프라 모니터링에 부족함이 없는 시각화 요소들을 포함한 실시간 대시보드 제공	실시간보다는 다양한 시각화 도구를 이용한 분석에 특화되어 있음. 다양한 데이터 소스를 활용한 시각화 가능
기타 기능 지원	스택 관리, 모니터링 등 엘라스틱스택 특화 기능 지원. 알림이나 머신러닝 등 유료 기능 존재	알림 기능 지원	

Elastic Stack

- ❖ 유사 제품 과의 비교
  - ✓ 엘라스틱 스택의 유사 제품군
    - 앞서 소개한 제품들이 스택의 각 구성 요소들과 대응된다면 이와 다르게 스플렁크(Splunk)는 엘라스틱 스택 그 자체와 대응되는 제품이라 할 수 있음
    - 일반적으로 함께 사용되기보다는 둘 중 하나를 선택해 사용
    - 스플렁크는 대용량 로그 분석을 위한 엔터프라이즈 제품으로 전용 로그 수집기, 고유한 인덱싱 기술, 대시보드, 각 벤더들에서 제공하는 수많은 앱을 통해 엘라스틱 스택과 마찬가지로 이 기종의 로그를 쉽게 수집하고 시각화하는 데 도움을 줌
    - 엘라스틱 스택과 비견되는 가장 큰 장점은 엔터프라이즈 제품답게 제공되는 앱들이 다양하고 수집기나 대시보드 등이 더 잘 만들어져 있어 완성되어 있는 솔루션이라는 느낌을 준다는 것이지만 엘라스틱 스택은 스택 내 구성 요소들을 자유롭게 설계하고 기본 제공되는 대시 보드조차도 커스터마이징이 가능하다는 장점이 있음
    - 이러한 특성을 이용해 엘라스틱 스택은 커스텀 애플리케이션의 백엔드로 사용하거나 조직이나 서비스의 특성에 최적화된 시각화를 구축하기 좋고 비용적인 측면에서도 차이가 남
    - 엘라스틱 스택도 유료 서브스크립션을 제공하지만 기본적으로는 오픈소스이고 많은 기능이 추가되는 베이직 라이선스 구조이고 자사의 플랫폼으로 활용할 때는 무료로 제공되기 때문에 유료 옵션만 존재하는 스플렁크와는 차이가 있음

Elastic Stack

- 체계적인 문서화와 활발한 공동체 지원
  - ✓ 오픈 소스 소프트웨어의 경우 문서화가 약하고 기술 지원을 받을 통로가 마땅치 않은 경우가 많지만 엘라스틱 스택은 문서화가 상당히 잘되어 있으며 활발한 공동체가 기술 지원을 대신하고 있기에 오픈 소스 세계의 모범 사례로 불릴 만 함
  - ✓ 엘라스틱 스택과 제품에 관한 공식 문서는 https://www.elastic.co/guide/index.html 에서 확인할 수 있으며 엘라스틱 스택, 엘라스틱서치, 엘라스틱 클라우드, 키바나, 엔터프라이즈 서치, APM, 엘라스틱 보안, 로그스태시, 비츠에 대한 자세한 매뉴얼을 제공
  - ✓ 한국어를 사용하는 개발자들을 위해 엘라스틱, 로그스태시, 키바나 문서는 한국어로 번역해서 제공
  - ✓ 뉴스나 새로운 기술 설명은 공식 블로그인 https://www.elastic.co/kr/blog/에서 확인할 수 있음
  - ✓ 스택오버플로우(https://stackoverflow.com)는 가장 활발한 개발자 커뮤니티로 엘라스틱 제품 군의 많은 질문과 답변을 찾을 수 있으며 엘라스틱과 관련해 궁금한 내용을 검색하면 누군가가 이미 비슷하게 질문해서 해결책을 얻은 과정을 확인할 수 있으므로 막히는 부분의 문제들 대부분은 어렵지 않게 답변을 찾을 수 있을 것
  - ✓ 엘라스틱 포럼(https://discuss.elastic.co)은 좀 더 전문적 인 토론과 답변을 얻을 수 있는데 간단한 궁금증은 스택오버플로우에서 해결하고 전문적인 토론이나 소식들은 엘라스틱 포럼을 이용