Elastic 소개자료

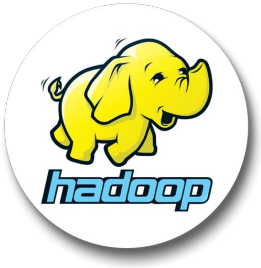
1. 엘라스틱 소개

Trend of Bigdata Analysis

RDBMS

BIGDATA

저장...



초고속 DB 엘라스틱서치

분산처리 기반,
고가용성 지원,
쉬운 확장성(Scale-out),
개발 친화적인 REST API,
전문검색(full-text search)
등의 특징을 가진
실시간 검색과 분석에 강점을 가진 데이터 플랫폼으로 발전

elastic

Why Elastic?

Elastic은 기존 오픈소스들의 약점을 해결하고, 새로운 시장의 요구를 수용 합니다.

기존 솔루션의 한계

- 기존 Database의
  비정형 데이터 수용과
  복잡한 JOIN의 문제

- Hadoop의
  대용량, 실시간 분석에 대한
  지원 부족

- NoSQL의 분석 기능 부족

- 개별 솔루션들의
  다양한 Use Case 대응의 어려움

기존 솔루션의 한계 해결

데이터 문제
해결을 위한
플랫폼,
"Elastic Stack"

시장의 요구 반영

새로운 시장의 요구

- 수평구조의 확장
- 실시간 분석
- 유연한 데이터 모델
- 신속한 쿼리 수행
- Schemaless
- 복잡한 요건에 대한 쿼리

elastic

Elasticsearch는 데이터를 수집, 검색(쿼리) 및 분석할 수 있는 데이터 플랫폼(초고속 DBMS)입니다
Elasticsearch를 이용한 연관 분석, 집계, 시각화를 통해 새로운 insight를 빠르게 얻을 수 있습니다

초고속DB

대용량 데이터(테라~페타급)도 실시간 수집/분석이 가능

- ✔ 초당 수백/수천만 건의 대용량 데이터를 실시간으로 수집, 분석하고, 초당 수천/수만쿼리까지 처리하여 시각화까지 가능한 플랫폼
- ✔ 분산처리 기반, 고가용성 지원, 쉬운 확장성(Scale-out), 개발 친화적인 REST API, 전문검색(full-text search) 등이 특징이며 모든 컴포넌트가 Scale-out구조로 되어있어 서버(노드)확장으로 퍼포먼스 유지가능

저장

세상의 모든 데이터 형태를 저장!

- ✔ 정형, 비정형 등 모든 데이터를 저장시 자동 인덱싱함으로써 실시간 쿼리/분석을 가능케 함
- ✔ 다양한 데이터 타입(숫자, 로그, Metric, 위치 정보, Text, 날짜등)에 대해 현존하는 최고의 압축/저장 기술로 저장되어, 가장 빠르게 쿼리 결과를 얻을 수 있음

분석

데이터를 유기적으로 조합하고 분석!

- ✔ 데이터를 유기적으로 조합하고 분석 및 시각화하여 새로운 Insight를 빠르게 얻음. 기존 무거운 DW대비, 36시간 걸리던 것이 불과 몇 초 만에 가능
- ✔ (예, 비지니스 분석/ 리스크 분석 / 실시간 트랜잭션 모니터링 / 트래픽 분석 / 로그 분석 / 트랜드 분석/ 보안 로그분석/ APM 등)

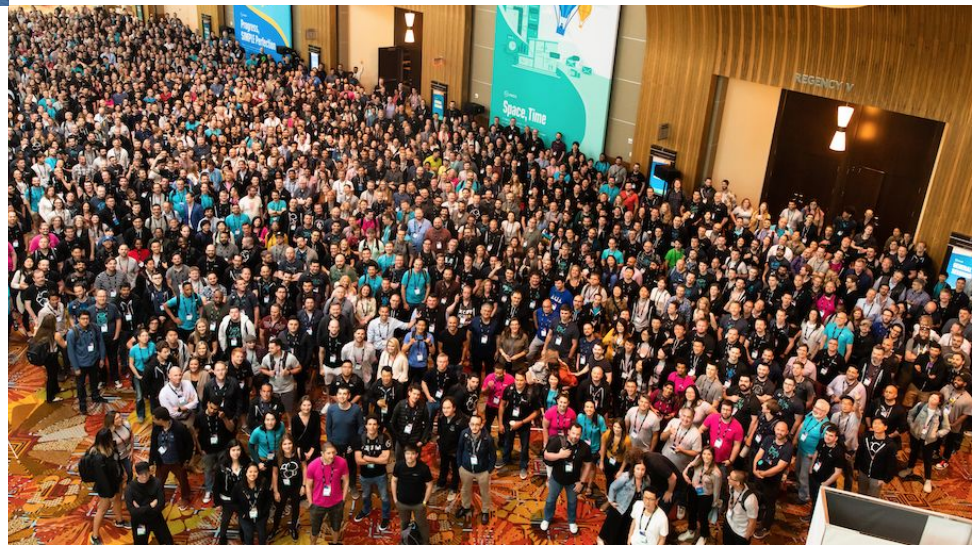
elastic

40여 개국에 분산된 직원

글로벌: 2000명

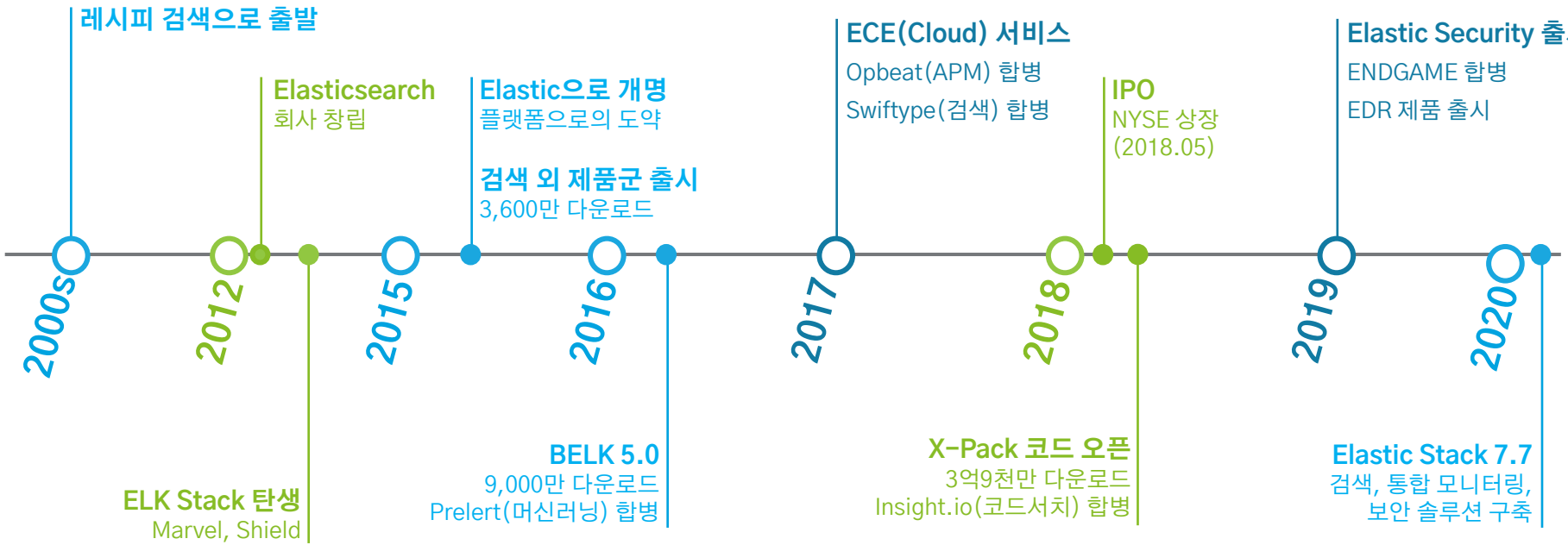
국내 20명(한국어 기술지원: 8명, 컨설턴트: 2명)

NYSE 상장 기업



Elastic 히스토리

검색엔진에서 시작하여 통합 모니터링 및 보안에 이르기까지,
기업의 엔터프라이즈 데이터와 관련된 모든 문제를 해결하기 위한 솔루션 기반의 플랫폼으로 나아갑니다



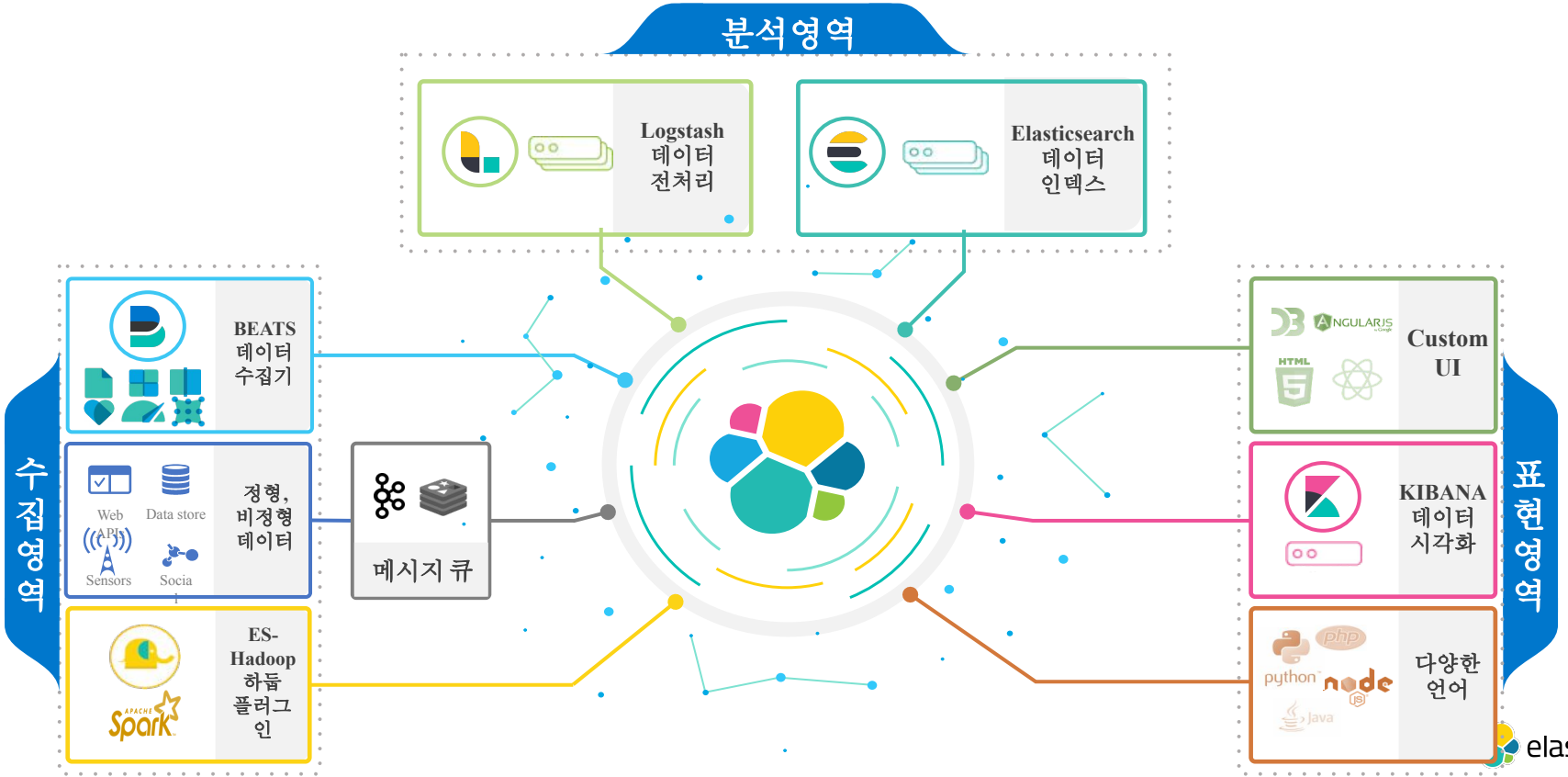
Index Ventures NEA BENCHMARK

2012~2014년,
3차례 걸쳐 104M(USD) 투자유치

elastic

엘라스틱 Data Platform

Elasticsearch는 단순한 NO-SQL DB나 일반적인 검색 엔진이 아닙니다.
Elasticsearch는 수집/저장/시각화 모듈을 모두 포함한 데이터 플랫폼입니다.



Global Subscription 고객사

TECHNOLOGY	FINANCE	TELCO	CONSUMER	HEALTHCARE	PUBLIC SECTOR	AUTOMOTIVE / TRANSPORTATION	RETAIL
Adobe	BARCLAYS	orange	Uber	UnitedHealth Group		RENAULT	Walmart
cisco	DISCOVER	COX	instacart	UCLA Health	United States Census Bureau	AIRBUS	The Home Depot
workday	USAA	SoftBank	tinder	Yale NewHaven Health	GSA	FUJITSU	ebay
Microsoft	collector bank	verizon	Activision Blizzard	MAYO CLINIC	JPL Jet Propulsion Laboratory	SAMSUNG	Kroger
IBM	ING	Sprint	lyft	MERCK	WILSON NORTH CAROLINA		Walgreens

elastic

엘라스틱 활용 범위

비지니스 분석에서 통합 로그(로그/메트릭/APM/Uptime) 그리고 보안(SIEM + EDR)까지 단일 플랫폼에서 분석이 가능
Elasticsearch는 보안, 비즈니스, 포털, 금융, 커머스 등 다양한 분야에서 데이터 분석 도구로 사용하고 있습니다.

보안 분석 (SIEM+EDR)

✔ 각각 다른 보안 솔루션 로그를 수집하여,
통합 분석을 통한 위협 탐지 용도로 활용

slack, cisco, BARCLAYS, HYUNDAI, KIA MOTORS, FireEye, 위메프 WEMAKEPRICE, SAMSUNG

로그/성능(metric)/APM

✔ 다양한 장비 및 Container 로그들을
수집하여, 머신러닝을 통한
장애 탐지 및 통계분석에 활용

NAVER, B kakaobank, citi, verizon, NH농협은행, pay

마케팅/비지니스 분석

✔ SNS, 기사, 웹 문서 와 같은
여러 형태의 정보를 수집/분석하여,
마케팅 분야의 새로운 가치 창출로 활용

ebay, Walmart, KB국민카드, Uber, Goldman Sachs, Sprint

검색

✔ 기업내부용 Enterprise Search부터
서비스용 검색 서비스로 활용

posco, TMON, ebay, AIRBUS, WPS, AMOREPACIFIC CORPORATION (주)아모레퍼시픽

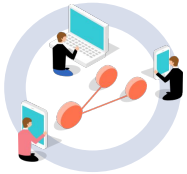
왜 Elasticsearch 인가?

Elasticsearch는 빠르고, 유연하며, 안정적으로 확장 가능합니다. 다양한 데이터를 수집하고 통합 분석할 수 있습니다.

특장점

속도

전문 검색(Full-text)을 위한
역 인덱스(inverted index) 구현

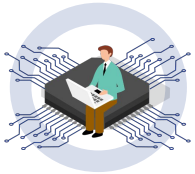


확장성

노드 구성의
수평적 확장 용이

유연성

숫자, 텍스트, 위치정보, 정형,
비정형 모든 유형의 데이터를 수용



복원력

분산 저장 설계로 하드웨어
오류나 네트워크 단절에서 안전

다양한 기능

Monitoring

앱 및 서비스가 사용자에게
영향을 미치기 전에
가용자원을 체크할 수
있습니다

METRICS

여러 데이터 메트릭
소스를
통합 분석합니다

SECURITY

인증로그, 감사데이터,
서버의 대화형 데이터로그를
분석하여 위협에 대응합니다

APM

미들웨어, 어플리케이션
로그를
수집하여 추이를 관찰/알람
합니다.

SEARCH

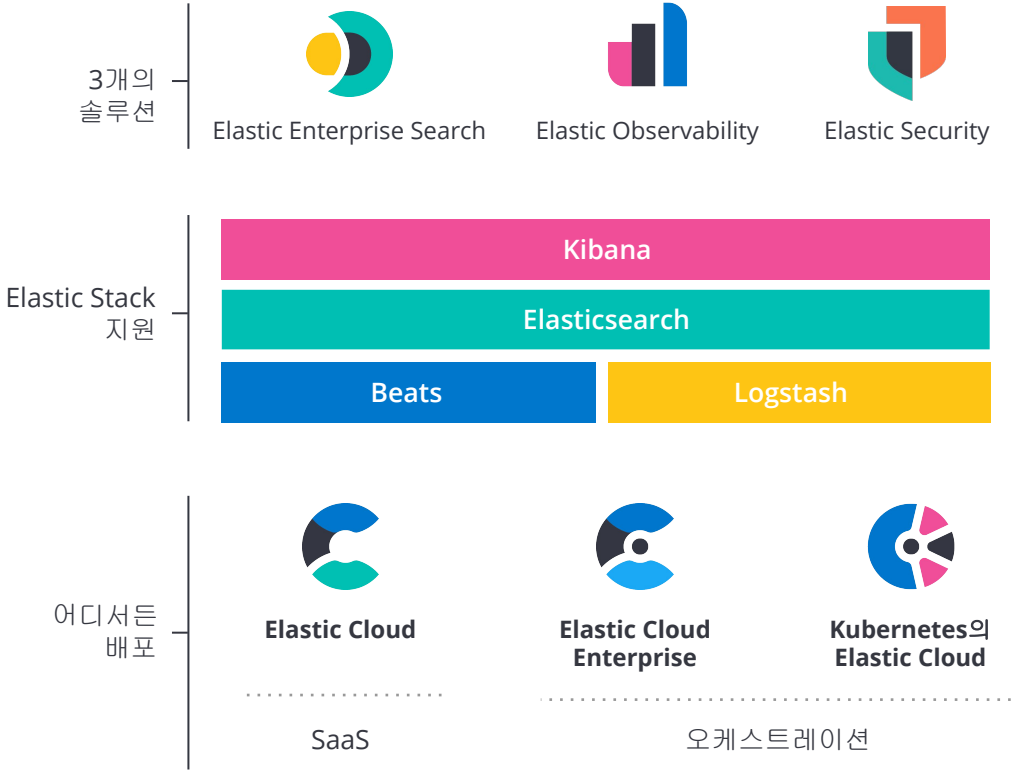
알고리즘 기능, 자동 완성, 필터,
랭킹 등을 이용하여 최적의
검색 결과를 확인 할 수
있습니다

elastic

2. 엘라스틱 스택 기능

- Elastic Stack 기본 구성
- Beats Overview
- Logstash Overview
- Kibana Overview

Elastic 한 눈에 보기



elastic

Elastic 기본 구성

데이터 [Source]

- 신용정보
- 거래 정보
- 회원정보
- 상품 정보

수집 [Shipper]

Beats

경량수집기

전처리 [Parser]

Logstash

ET
L

인덱싱 [Indexing]

Elasticsearch

분석엔진

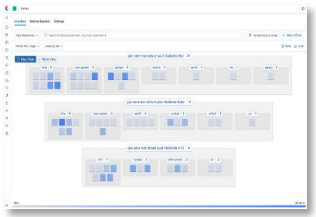
시각화 [Visualization]

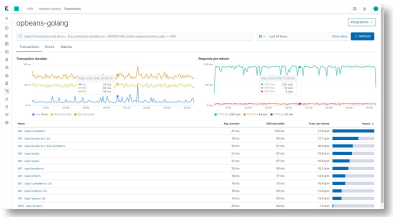
Kibanaa

UI

시각화

관리자





x-pack

Extensions for the Elastic Stack

security

Alerting

reporting

monitoring

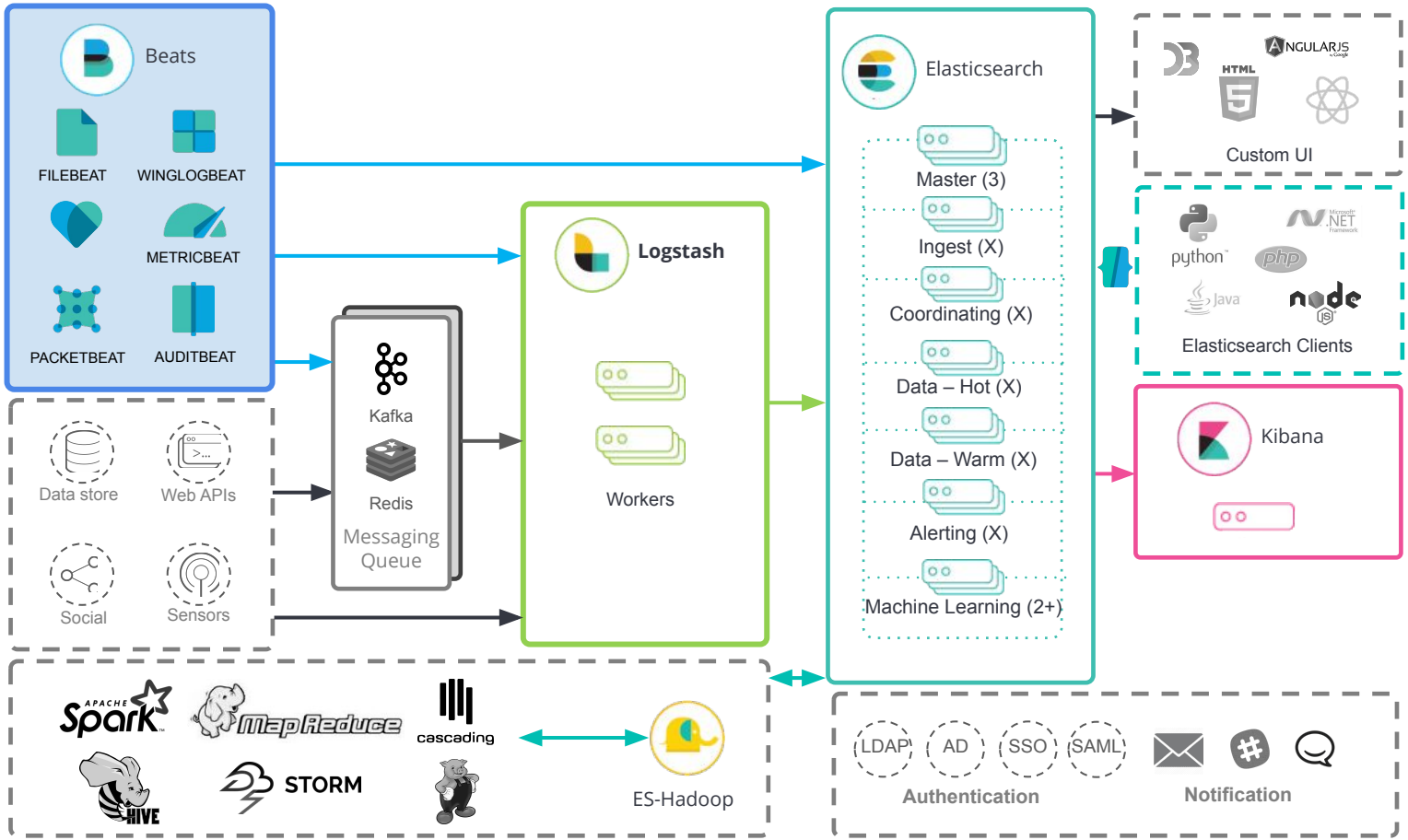
graph

machine learning

elastic

Beats

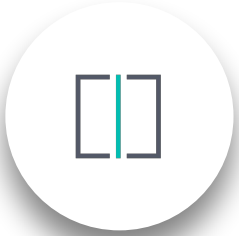
- 소스로부터 파일, 시스템 메트릭, 네트워크 패킷 등을 수집해서 Elasticsearch, Logstash, 및 메시지 큐로 전송



Beats 제품군

수집할 데이터의 종류에 따라 각기 다른 Beats 들이 제공됩니다.

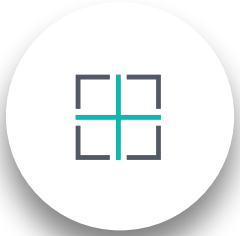
Packetbeat
네트워크 데이터



Auditbeat
감사 데이터

Metricbeat
시스템 메트릭

Filebeat
로그 파일



Winlogbeat
Windows 이벤트 로그

Heartbeat
가동시간 모니터링

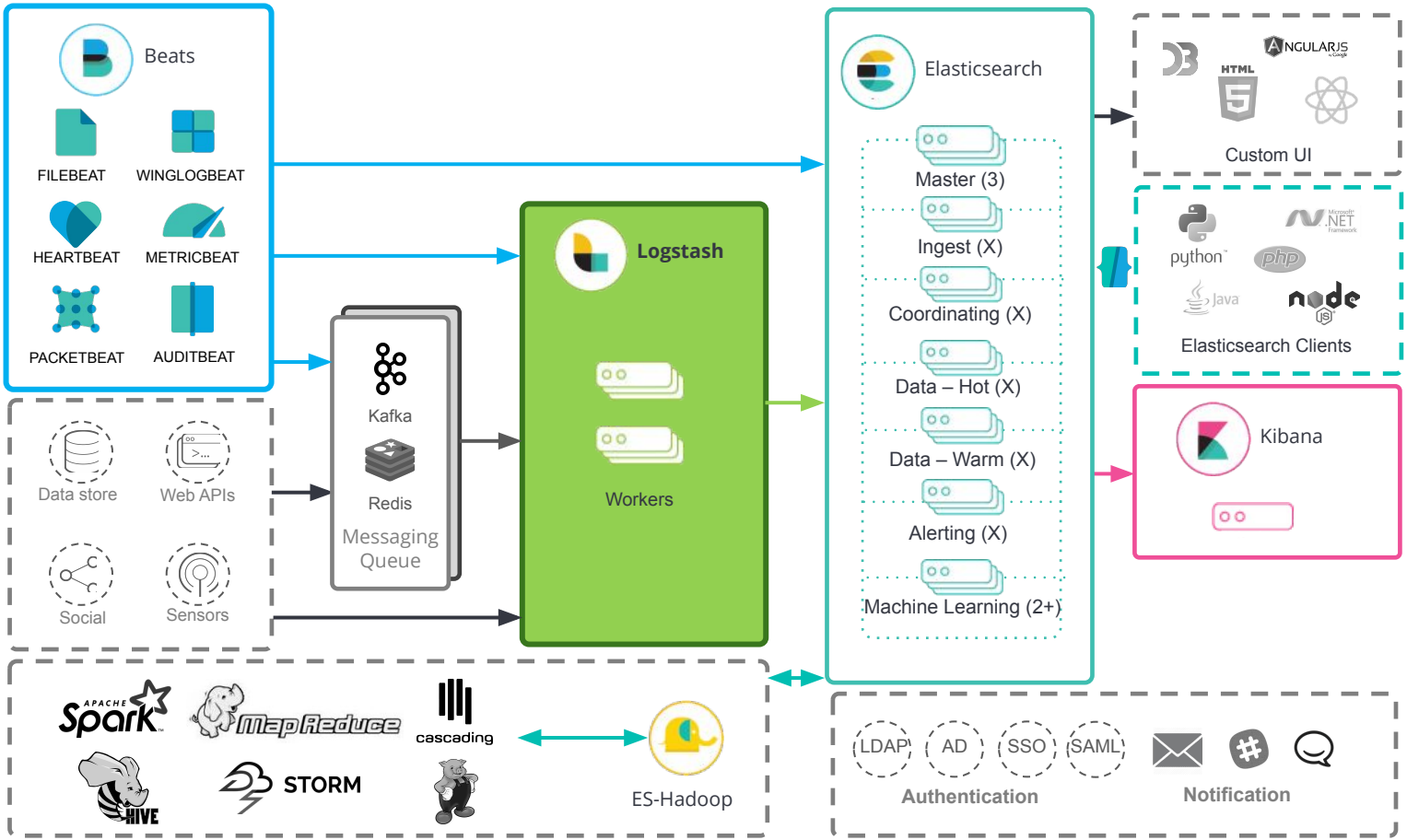
Functionbeat
클라우드 시스템의
서버리스 데이터

그 외에도 커뮤니티에서 만든 90여 개의 오픈소스 커뮤니티 Beats 가 있습니다

elastic

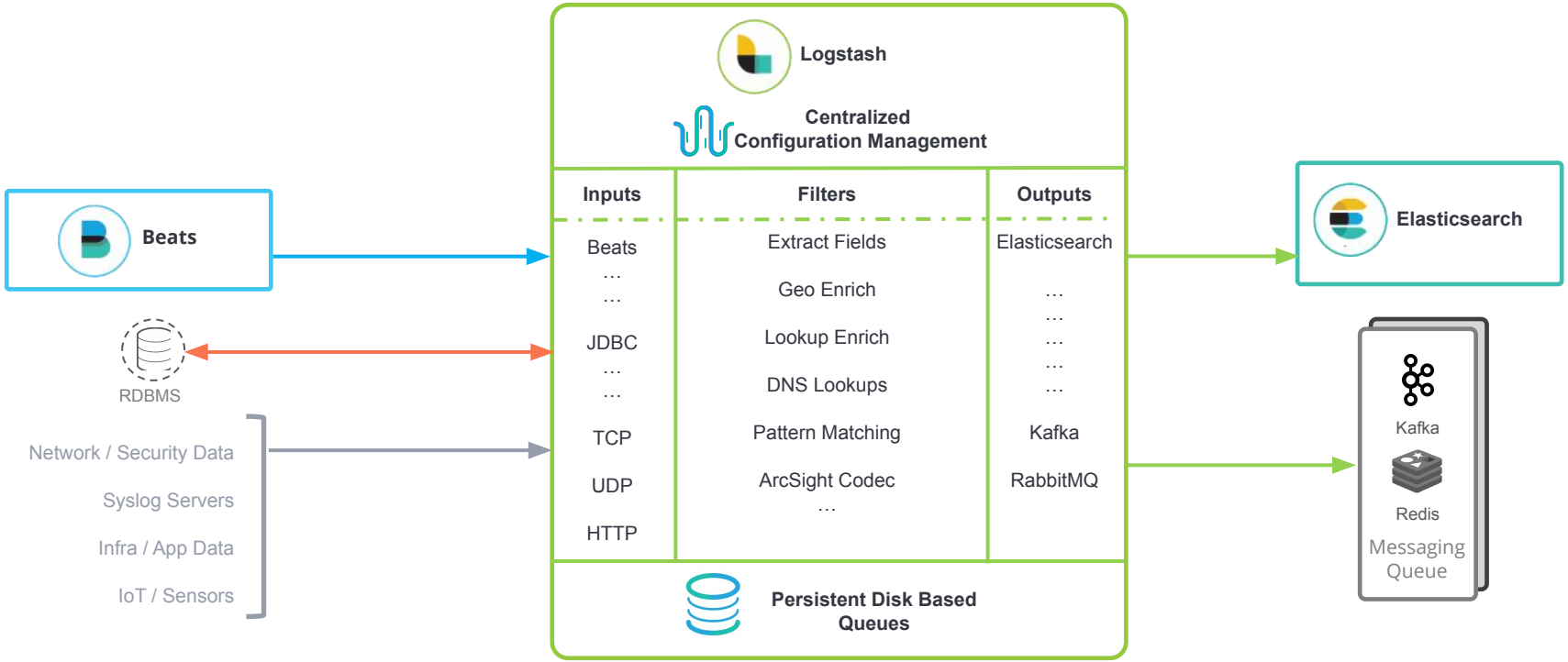
Logstash

- 파이프라인을 통해 입력 받은 데이터를 변환 & 확장 하고 Elasticsearch 를 비롯한 다양한 시스템으로 전송



Logstash 구조

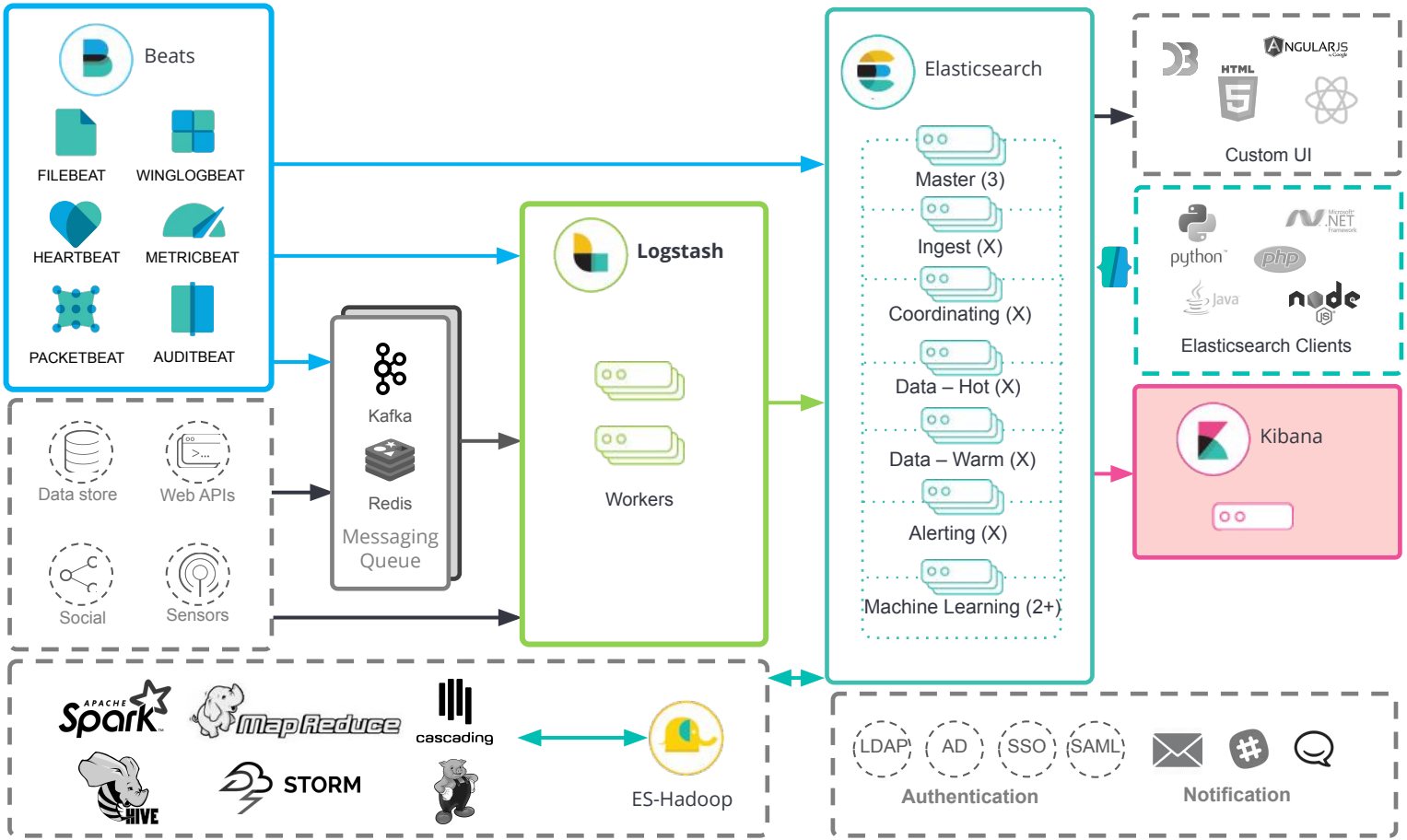
지원 Input : Beats, JDBC, TCP, HTTP, File, 로컬 파일, 각종 SNS API, 깃헙, 세일즈포스 등의 서드파티 시스템 데이터 등
지원 Filter : 필드 파싱, 패턴 매핑, Plane Text → JSON 형식으로 변환 IP 주소로부터 로컬정보 확장 등
지원 Output : Elasticsearch, Kafka, RabbitMQ, 원격 파일 시스템, 아마존 S3 버킷 등



elastic

Kibana

- Elasticsearch에 저장된 모든 데이터를 시각화 & Dashboard 생성하여 다양한 방법으로 분석

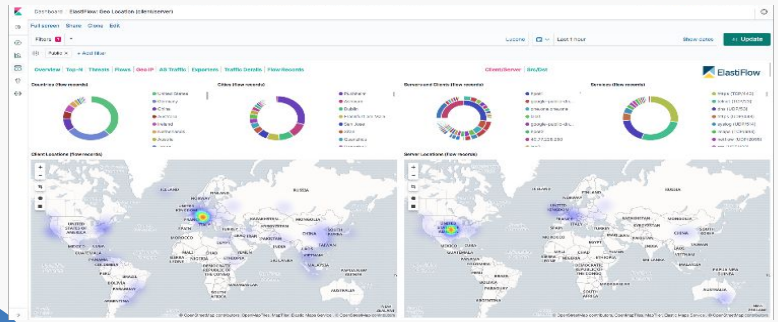


Kibana 활용

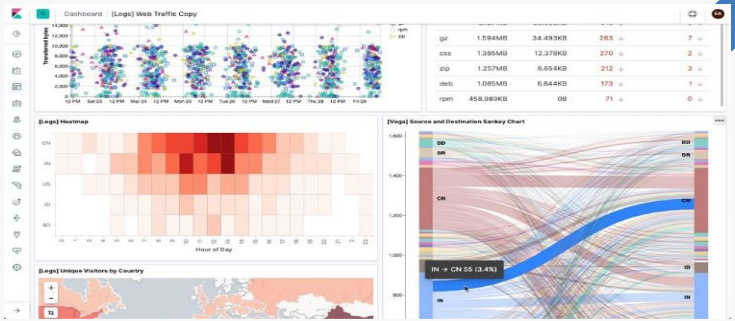
다양한 분석 및 Self BI 환경 제공

수집된 데이터를 KIBANA를 통해 다양한 방법으로 분석할 수 있다.

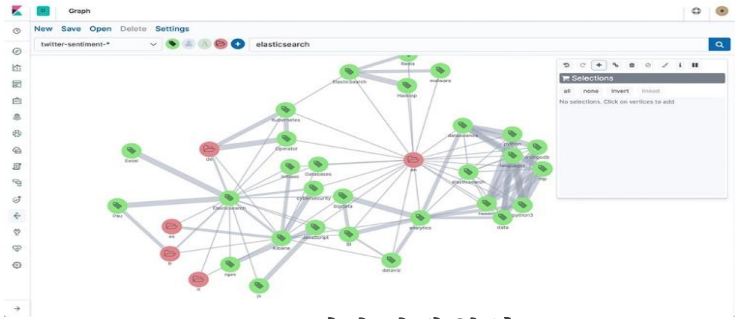
▲ 시계열 분석



▲ Netflow 및 위치 분석



▲ Heatmap 및 Vega 분석



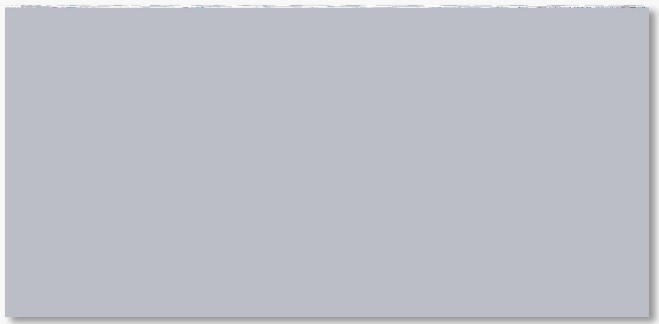
▲ 연관 관계 분석

Kibana 활용

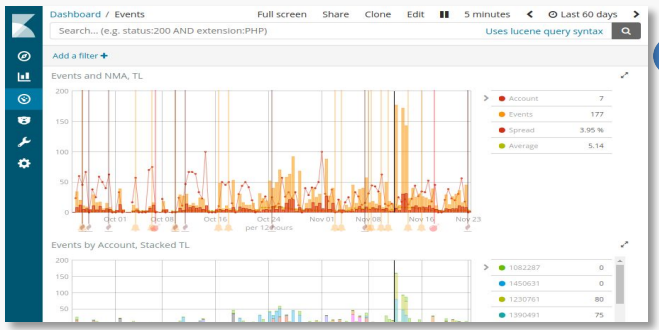
다양한 시각화

수집된 데이터를 KIBANA를 통해 다양한 방법으로 분석할 수 있다.

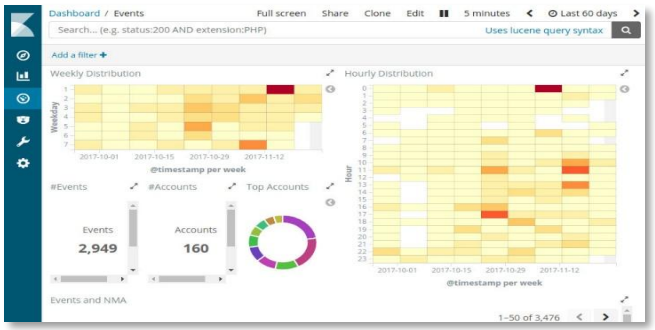
▲ 시스템 모니터링 시각화



▲ 통계 지표 시각화



▲ 시계열 데이터 시각화



▲ 데이터 간의 연관관계 시각화

Elastic Solution

Elastic Enterprise Search

Site Search
No query-based pricing

App Search
No docs-based pricing

Workplace Search
No user-based pricing

Elastic Observability

APM
No agent-based pricing

Logs
No ingest-based pricing

Metrics
No host-based pricing

Elastic Security

SIEM
No seat/ingest-based pricing

Endpoint Security
No endpoint-based pricing

elastic

3. 실시간 빅데이터 플랫폼 Elastic

- Elastic vs Hadoop 기술 비교
- 실시간 빅데이터 플랫폼 요구 사항
- Elastic 기반 플랫폼 구축 기대 효과 및 특징 요약
- 실시간 빅데이터 활용 레퍼런스 요약(국내)

Elastic vs. Hadoop 기술 비교

주요 차이점	하둡	엘라스틱서치
한줄 요약	**대규모 분산 저장소**	**대규모 분산 검색 데이터베이스**
실시간성	배치 중심, Spark 등의 추가 기술을 통해 실시간성 보완	실시간 중심 (예, 구글 검색)
응답성	수 시간, 수일 단위	수 밀리초, 수초 단위
처리 단위	파일 단위 처리 (맵-리듀스)	문서 단위 처리 (JSON 형식)
배포 복잡도	**상대적으로 복잡. 데이터의 "처리"에 다수의 오픈소스를 조합하여 사용하므로 자유도와 복잡도가 모두 높음**	수집, 저장/검색, 시각화 3가지로 구성으로 단순. JSON 형식의 DSL 쿼리 기반으로 상대적으로 쉽고 강력함.
데이터 스키마	주로 키=값 형태로 만들어진 파일을 업로드하여 처리. 특별한 스키마 제약이 없지만 함께 사용하는 기술에 따라 특정 스키마로의 재가공 필요.	정형, 반정형, 비정형 데이터를 JSON 형태로 변환하여 저장. 텍스트 형태가 아닌 파일(MP3, 비디오 등)은 저장 및 처리가 어려움.
분석 활용성	MR, Pig, Hive 전문가가 필요하며 R이나 SAS같은 통계분석 시스템과 연계 개발이 이루어져야하고 많은 전문가/인건비가 필요	통계 및 고급 분석이 가능한 쿼리 API가 제공되며 데이터의 검색, 집계 및 역할별 시각화를 손쉽게 활용 가능. Rest API를 통한 분석 시스템간 연계 분석 가능
지원 언어	하둡은 다양한 언어를 제공하지 않음	엘라스틱은 Ruby, Lua, Go 등 다양한 언어를 지원하며 이 언어들은 하둡에선 지원되지 않음

실시간 빅데이터 플랫폼 요구사항

- **안정적이고 확장 가능한 완전한 데이터 플랫폼 구축**
- 운영 시스템 자원에 부담을 주지 않는 데이터 수집
- 수집 데이터의 활용이 용이하도록 수집 중 데이터 강화(Security, 압축 등)
- 원본 데이터의 유실 또는 변조 없이 보관
- **노드 장애 대응: 한대의 서버 중단 상황에서도 완전한 검색 및 데이터 수집**
- 실시간 데이터 추출 및 활용을 위한 검색 성능 보장

- **주요 항목**
  - 데이터 수집 용이성(Data Source 추가 용이성 확보 - API, Log 수집, DB연계 등)
  - 수집 대상 시스템에 대한 간섭 최소화
  - 데이터 누락 방지(Data Integration 솔루션 제공 - Beats, Logstash, Watcher 등)
  - 플랫폼의 SLA 보장
  - 안정적 기반인프라 구축을 통한 Data 플랫폼 확장성 보장

elastic

elastic Stack 기반 빅데이터 플랫폼 구축 기대 효과

- 실시간 우선 플랫폼 구축 용이
- 즉시적 분석 모델 반영
- 개발 기간 단축(수집 Agent 활용 및 Plug-in 적용 - 솔루션 공식지원)
- 운영 편의성, 유지보수 용이성 향상(플랫폼 통합 및 보안 적용)
- **추가기대 효과 - 빅데이터 플랫폼에 대한 통합 인프라 모니터링 제공(SLA)**

  - 분석가 환경 별 모니터링

  - 운영 환경 모니터링

  - 부서/영업 채널별 환경 구성 및 모니터링 등

elastic

elastic 특징 요약

- 선택적 확장성: 데이터 증가로 인한 라이센스 추가 비용이 발생하지 않고 오직 노드 수에 따라 과금. 성능과 비용 간의 균형 있는 조정이 가능하여 라이센스 비용에 자율성 확보 가능함
- 실시간성: 실시간 수집, 실시간 검색 외에 검색 범위 전체에 대한 응답성 또한 실시간성 범주에 포함 됨. 5분/10분 단위의 중간 집계를 통한 결과는 실시간 집계가 될 수 없음.
- 스코어 기반 연관도 검색 결과: 다수 키워드의 검색 결과를 단순히 시간 역순으로 제공하는 것을 넘어 연관도에 따른 점수를 기반으로 결과를 제공하여 보다 정확한 검색 결과 제공
- 유사 검색 결과: 반드시 일치하는 검색 키워드만 반환하는 것으로는 부족, 로그 내용중 메시지에 유사 키워드에 대한 결과까지 반환
- 연동 용이성: 데이터 플랫폼의 일부가 아닌 모든 기능을 Rest API로 제공하여 기존 시스템과 연동할 수 있어야함. 시각화 요소의 기존 시스템 내에 내포(embed)가 가능함 (Canvas, Visualization, Map 등)
- **데이터안정성: 데이터 노드 장에 발생시 복구시까지 데이터의 수집 및 검색에 영향을 주지 않음**

실시간 빅데이터 활용 레퍼런스 요약(국내)

- **실시간 빅데이터 활용 사례**

웹툰	클릭스트림 기반 웹툰 내 PPL 효과 분석(효과 분석 Dashboard 구축)	S 금투	고객 투자 성향 분석 및 이상거래 감지(ML)
Logistics	택배 송장데이터 기반 제품 별 마켓쉐어 분석 및 데이터 마켓플레이스 구축	K 증권	고객별 일/시간단위 거래 분석 이상거래 감지
모빌리티	혼잡도 분석 및 실시간 배차 및 실시간 금액 산정	K 카드	위치기반 Payment 마케팅 연동 SNS Crawling 데이터 기반 평판 분석
커머스	클릭스트림 기반 상품 노출 순위 선정 및 고객별 선호 제품 색상 배치	S 은행	장애 이슈 트래킹(통합 인프라 모니터링 및 장애 예측)
IPTV	IPTV 셋탑 로그 기반 고객별 VOD 추천 및 메뉴 구성	B 카드	재택근무 VDI 사용현황 모니터링
OTT	연령, 성별, 지역별 실시간 컨텐츠 추천 및 인기순위 추출	K 은행	플랫폼 내 데이터 기반 신용등급 산정 및 상품 추천

추가 자료

	Open Source	Basic	Gold	Platinum
Elastic Stack	Elasticsearch, Kibana, Beats, Logstash	Elasticsearch, Kibana, Beats, Logstash	Elasticsearch, Kibana, Beats, Logstash	Elasticsearch, Kibana, Beats, Logstash
X-Pack				
Monitoring		✔	✔ + Multicluster + configurable data retention	✔ + Multicluster + configurable data retention
Security			✔	✔ + field/doc level security + custom realms
Alerting			✔	✔
Graph				✔
Reporting			✔	✔
Machine Learning				✔
Dev Tools		✔ + Search Profiler	✔ + Search Profiler	✔ + Search Profiler
Tile Service	✔ Zoom Level 10	✔ Zoom Level 18	✔ Zoom Level 18	✔ Zoom Level 18
Support			Business Hour 6 Support Contacts 4 Hour Critical Response	24/7/365 8 Support Contacts 1 Hour Critical Response Emergency Patches

고객을 위한 지원 (유료) 서비스

Elastic은 서브스크립션을 통한 기술 지원, 컨설팅 서비스, 교육 서비스를 제공합니다.

기술 지원

- 글로벌 기술 지원 체계
- 전담 연락 엔지니어 배정

컨설팅 서비스

- On-site 컨설턴트 방문
- 비즈니스 단계별 맞춤 컨설팅

교육

- 고객사 역량 강화
- 자격증 취득 지원

elastic

기술 지원

온라인 / 음성 및 화상 전화

24/7/365 SLA 기반 신속 대응

장애 발생 시 1시간 이내 정상화 복구
오류나 버그에 대한 긴급 패치 적용
노드 장애 대응: 한 대의 서버 중단 상황에서도 완전한 검색 및 데이터 수집
다중 데이터 센터 간 DR 및 완전 백업 불필요
단순 장애 현황 대시보드 외 장애 징후에 대한 사전 경고 및 알림

전담 Certified 리드 엔지니어 배정

영업 시간 내 한국어 지원
기술 지원 문의 개수 제한 없음

Elastic 공식 진단 툴을 통한 정확한 원인 분석 및 해결

아키텍처 / 인덱스/ 샤드 등 시스템 최적화 가이드
클러스터 관리 및 검색 성능 튜닝
쿼리 성능 최적화
운영으로 전환 시 마이그레이션 및 업그레이드 지원
Elastic Stack 및 상용 기능 Best Practice 제공

elastic

컨설팅 서비스

컨설팅 서비스 영역

아키텍처
엘라스틱서치 클러스터, 엘라스틱 스택 설계
시스템 통합 설계

SIEM
Elastic Stack 기반 사내 보안 솔루션 구축

Machine Learning
데이터 분석 및 결과를 업무 프로세스에 적용

검색
검색엔진 교체
검색 성능 평가, 진단, 해결방안 제시

시각화
Canvas로 구현하는 인포그래픽
시각화 라이브러리 Vega

ECE
다중 엘라스틱서치 클러스터의 오케스트레이션 환경

컨설팅 프로젝트 형태

Proof of Concept (PoC)
짧은 시간 내 (기술 학습 시간 단축, Trial & Error 최소화)
최소의 인원으로 (컨설턴트의 고객사 상주)
최상의 퀄리티 (컨설팅 이전 케이스 활용)

본 구축
프로젝트 목적에 맞는 데이터 설계
데이터의 확장에도 성능 문제 없는 클러스터 설계
운영에 필요한 핵심 기능 도입
컨설팅(단기적/현장)과 기술지원(장기적/원격)의 협력형 지원체계 구축으로 최적의 프로젝트 구축 및 운영
엘라스틱 도입 영역 확장

문제 해결 형
자체적으로 구축하고 운영하던 엘라스틱 스택의 문제 발생 시
문제 원인 진단 및 설명 (진단 방법까지 전달)
Elastic를 개발한 개발팀이 전하는 해결 방안 가이드

Best Practices

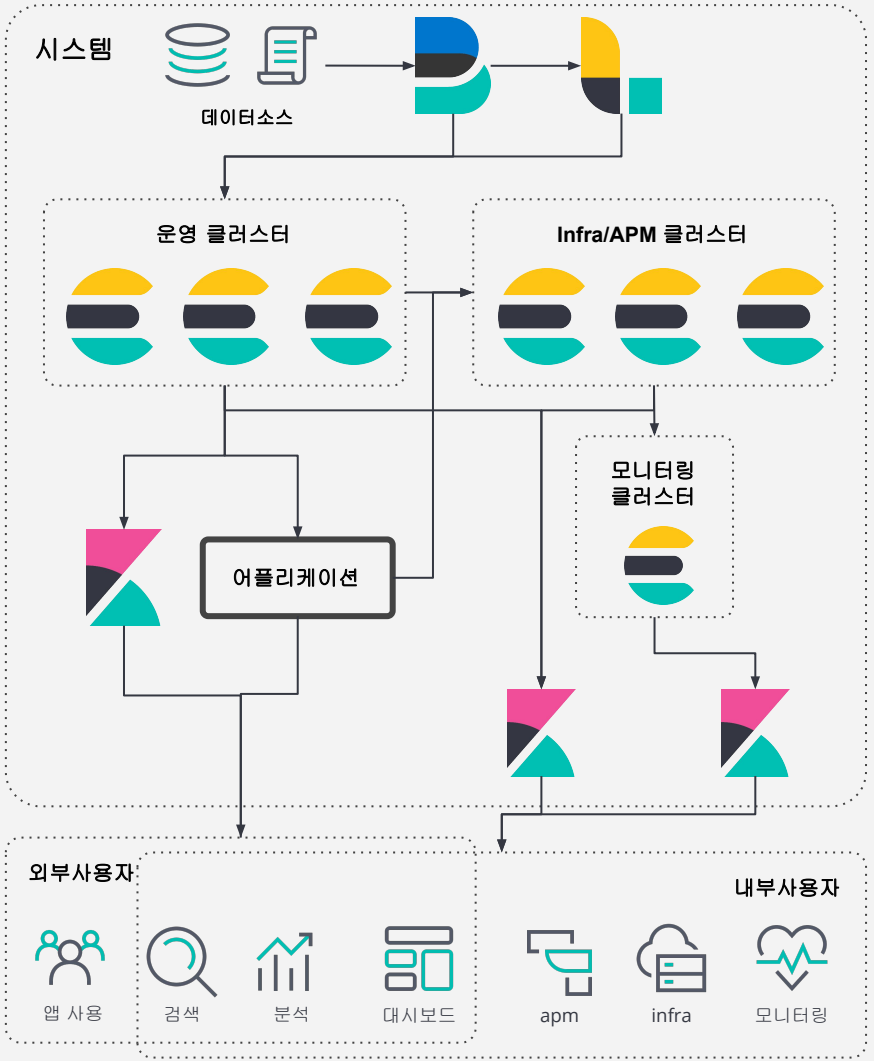
Elastic Stack deploy

클러스터 설계 및 운영
- 엘라스틱서치 클러스터를 위한 HW 스펙 / 설정 추천
  엘라스틱 스택 아키텍처 추천
- 엘라스틱서치 클러스터 사이징 방법론
- 엘라스틱 스택 아키텍처 모범사례
- 엘라스틱서치 클러스터 운영전략 모범사례

인덱스 Setting/Mapping/운영
- 인덱스 setting / mapping 추천
- 데이터 모델링 모범사례
- 데이터 색인의 내부 로직

엘라스틱서치 기능
- 고객의 유즈케이스에 유용한 기능 추천 및 적용 방안 안내

모니터링
- 엘라스틱 스택 / APM / Infrastructure 모니터링 방안



Best Practices

Elastic Stack for security

데이터 중앙화

다양한 소스에서 발생하는 데이터를 ES 클러스터에 색인
- Packetbeat, Winlogbeat/Auditbeat
- 기성 보안 솔루션에서 수집 된 로그
- 사내 보안 솔루션에서 수집 된 로그

데이터 고도화(Enrichment)

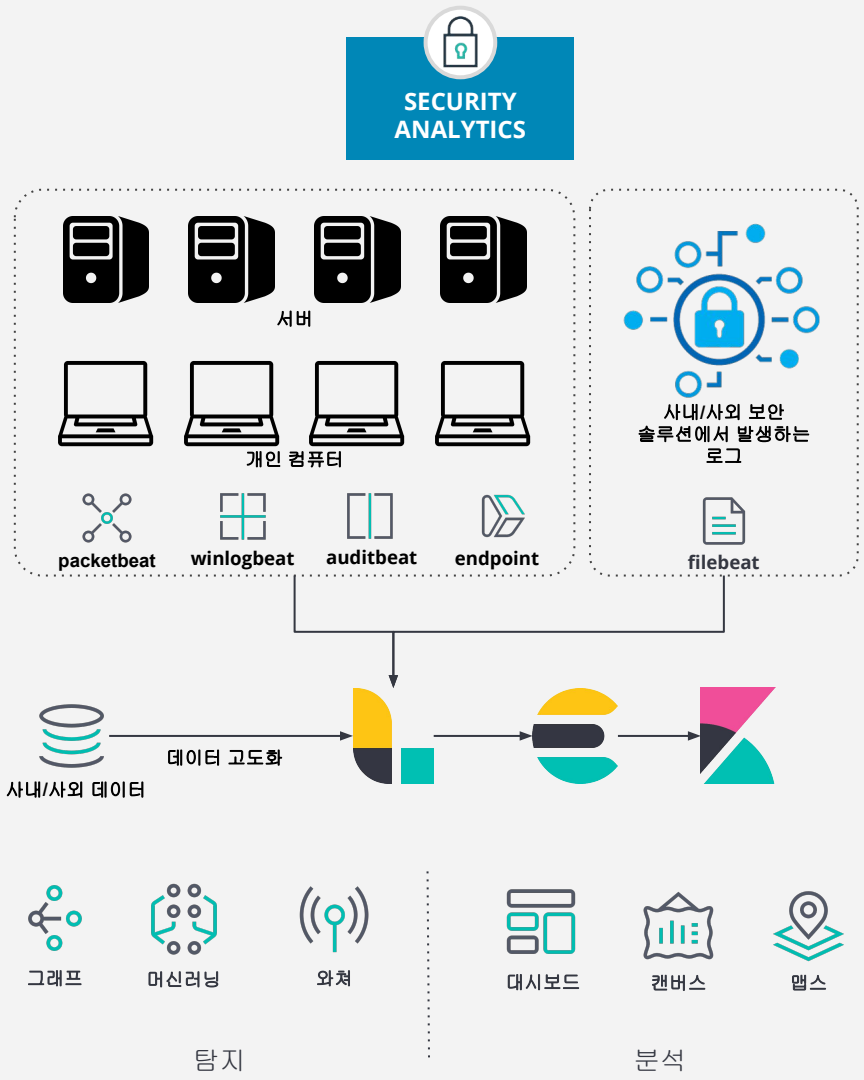
수집된 로그에 정보를 추가하여 고도화 함
- 내부) 사내 보유 데이터베이스
- 외부) 오픈데이터, 외부 솔루션에 대한 요청 결과

이상징후 탐지

- 수집된 로그에서 이상징후 탐지
- 탐지된 시점에만 이메일, 슬랙 등으로 알람 생성

데이터 분석

다양한 시각화 메뉴를 사용하여 이상징후의 근본 원인을 인터렉티브하게 분석



Best Practices

Analyze data with ML

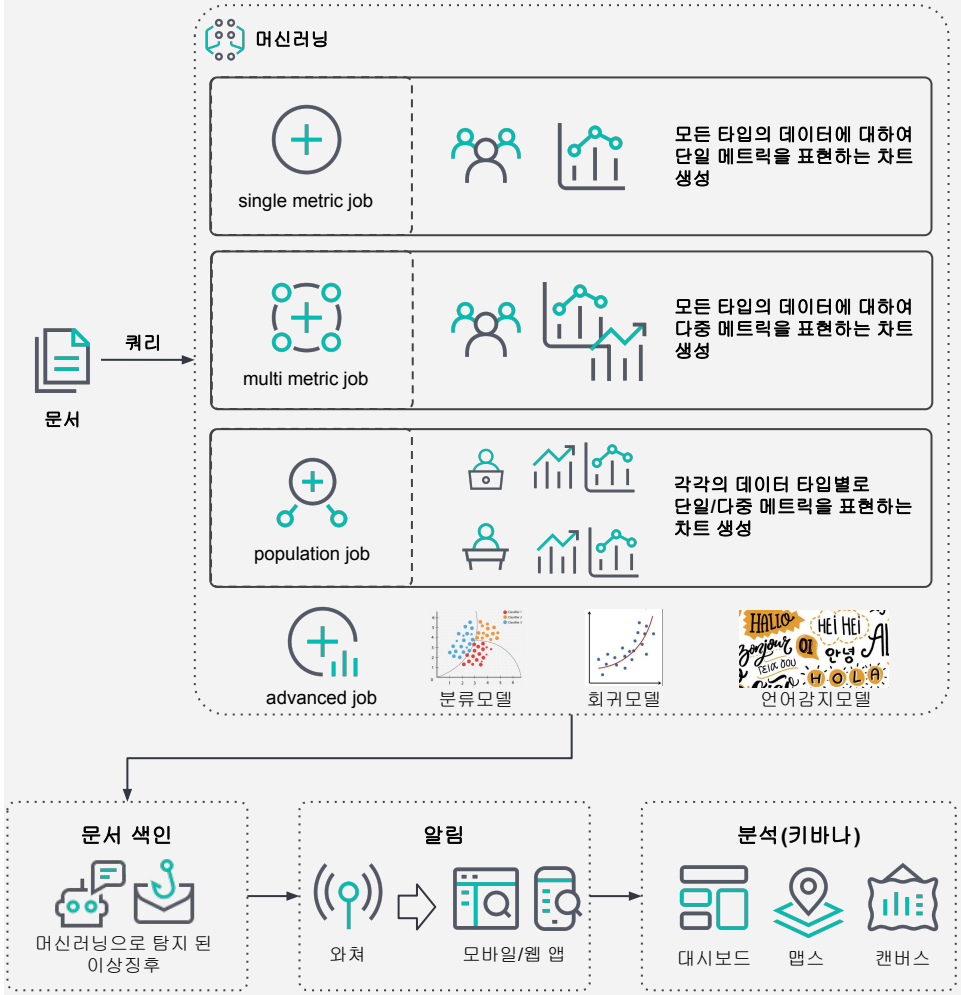
데이터의 패턴 분석

이상징후 탐지란?

- 데이터로 생성한 그래프가 과거의 패턴과 상이한 시점 감지
- 7 days / 24 hours 모니터링 필요
- 많은 노동력이 필요하거나 데이터가 많은 경우 불가능함
- 엘라스틱 머신러닝을 통해 세계에서 가장 뛰어난 이상징후 탐지 알고리즘 제공
- 데이터의 패턴 학습
- 데이터 타입/시간 별 이상징후 임계치의 동적 변경

업무 효율 개선

- 이상징후가 탐지 되었을 때만 업무 수행
- 이메일, 슬랙, 티켓 등으로 알림 수신
- 알림에 포함된 링크를 통해 머신러닝 결과 페이지 연동
- 머신러닝과 키바나 대시보드 연동
- 대시보드에서 필터/쿼리를 이용해 근본원인 분석



7d, 24h 모니터링 **불필요** / 키바나를 통해 **인터렉티브하게** 근본원인 분석

Best Practices

Elasticsearch for search

검색엔진 교체

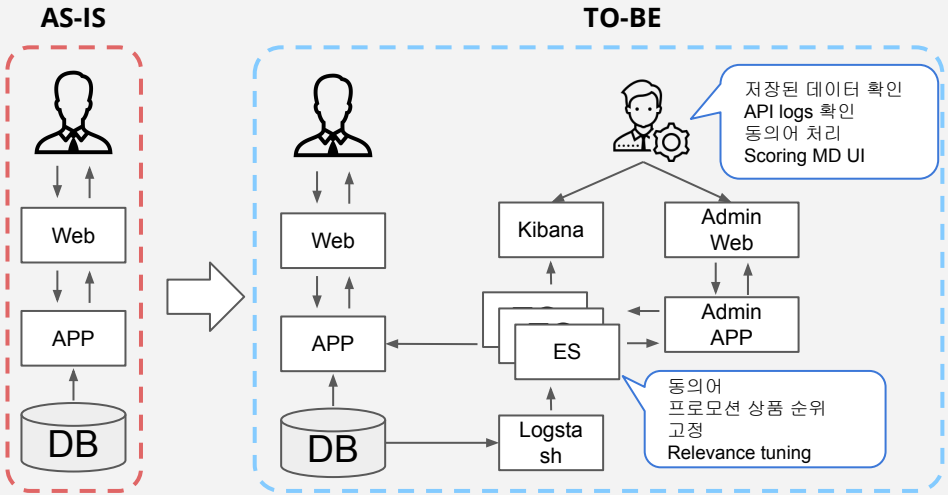
- Elasticsearch cluster 사이징
- 데이터 수집 프로세스 설계/구축/성능평가
- 인덱스/샤드 디자인
- 형태소 분석기, 동의어, 불용어
- 프로모션 상품 순위 고정, 검색 결과 조정
- Admin 페이지

검색 성능 개선

- 성능평가를 통한 파라미터 파인 튜닝
- 검색 성능 개선을 위한 업무 프로세스 가이드
  쿼리 최적화, 문제적 쿼리 선별

Elastic Enterprise Search

- Workplace Search
- App Search
- Site Search



		1	2	3	4	5	6	7	8	9
백엔드	App) ES client 추가 및 쿼리 추가		■							
	Admin App) 서버 로직 개발			■	■	■	■	■		
프론트	Web) Search box endpoint 수정		■							
	Admin Web) UI 디자인	■	■	■						
	Admin Web) UI 개발			■	■	■	■	■		
엘라스틱	AWS instance 구축 및 Elastic Stack 설치/설정	■								
	Search query 개발		■	■						
	Admin 기능용 request 개발				■	■	■	■	■	
	운영 환경 셋업(metricbeat, APM, Alert, ...)									■

Best Practices

Advanced Visualization

Canvas

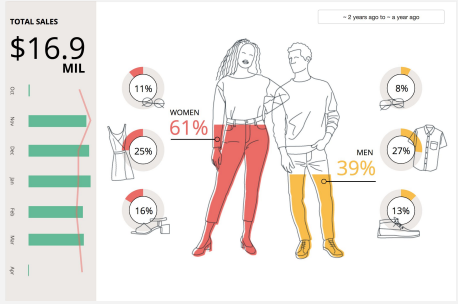
- Elasticsearch에 저장된 데이터가 실시간으로 반영되는, CSS로 디자인한 인포그래픽 생성
- PDF 리포트 생성 및 키오스크 모드 제공
- ES SQL 지원

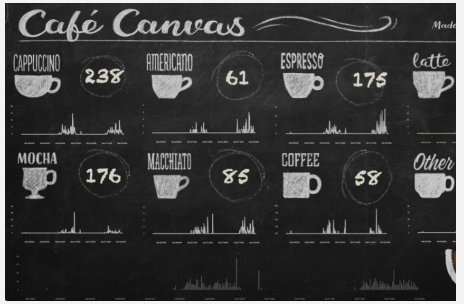
Vega/Vega-lite

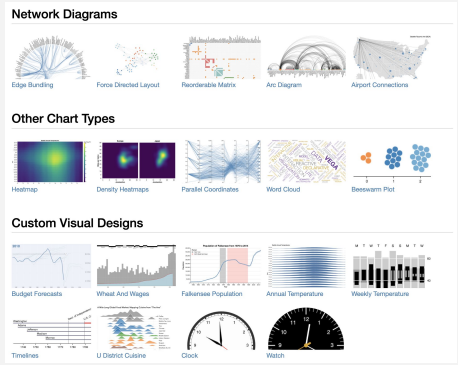
- Vega/Vega-lite에서 지원하는 그래프를 Elasticsearch에 저장된 데이터를 기반으로 Kibana상에서 사용 가능
- Vega/Vega-lite를 Kibana로 포팅하는 가장 효과적인 업무 프로세스 가이드

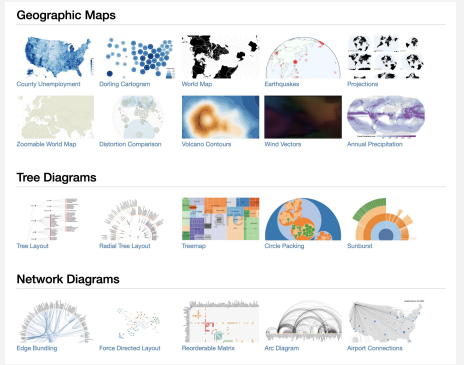
Maps

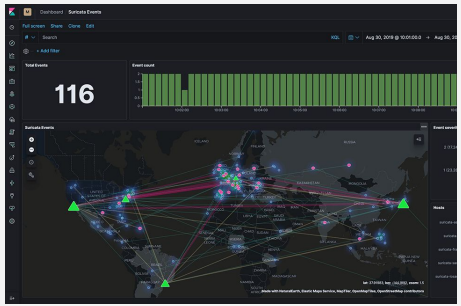
- 위치 기반 데이터의 분석환경 제공
- Rightmove(부동산 임대 및 판매업) Geo distance, polygon/shape, geo-bounding box searches를 사용하여 개인에게 이상적인 주택 추적











Best Practices

Best approach for infrastructure

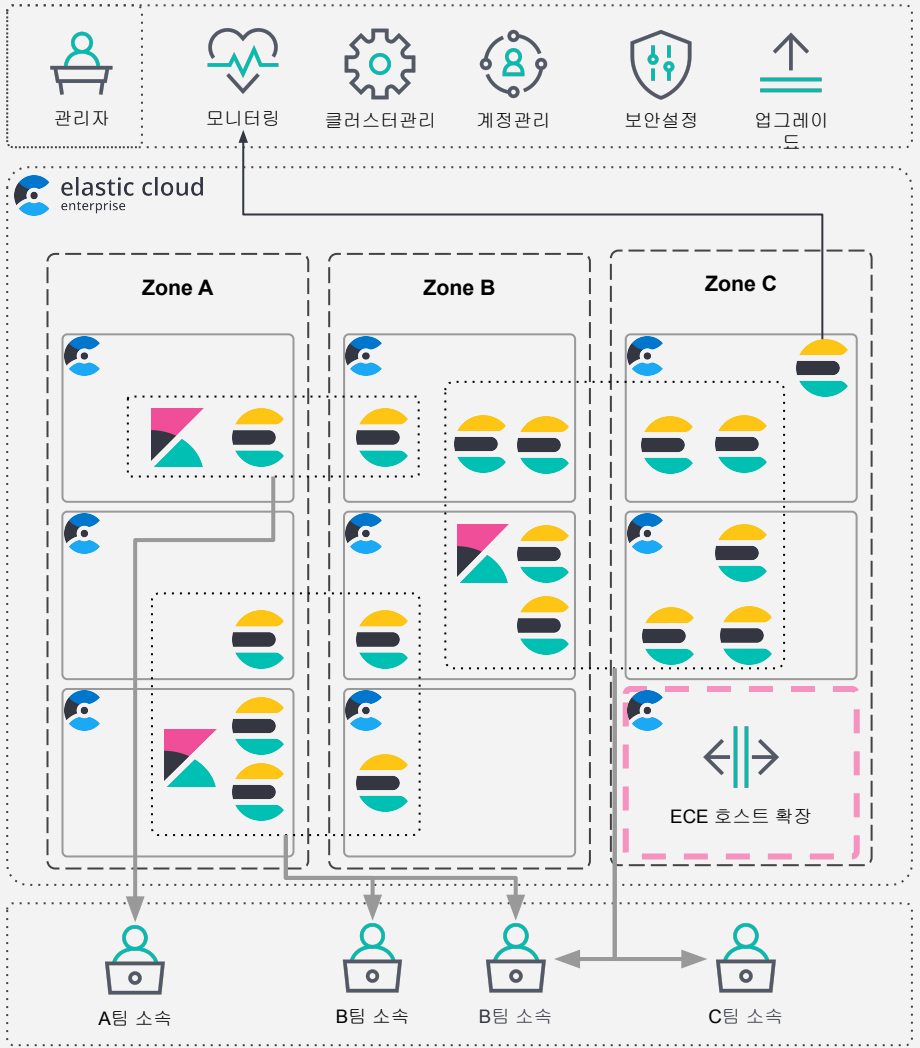
다수의 엘라스틱서치 클러스터 관리

- 쉬운 클러스터 생성/수정/삭제
  어드민 UI 제공, 도커 컨테이너 기반 클러스터 배포
- ECE 클러스터 / 엘라스틱서치 클러스터 모니터링
- 다중 존으로 높은 가용성 제공
- 엘라스틱 클라우드와 동일한 운영 환경(안정적 시스템)
- 엘라스틱 클라우드 운영 노하우 기반 기술지원
- 손쉬운 ECE 클러스터 확장

업무 효율 개선

인프라팀

- 메뉴얼적 클러스터 생성/수정/삭제 불필요
- 모니터링 시스템 기본 제공
- 엘라스틱서치 클러스터 사용팀
- SaaS 고객처럼 UI를 통한 클러스터 생성/수정/삭제로
  업무 프로세스 단축



Elastic 공인 교육

training.elastic.co

Elasticsearch Engineer I, II

클러스터 가동에서 고급 관리 기술 마스터링에 이르기까지 Elastic Certified Engineer 코스를 통해 Elasticsearch의 전문가가 되실 수 있습니다.

Elasticsearch Engineer I 커리큘럼

Elasticsearch Engineer II 커리큘럼

Data Analytics

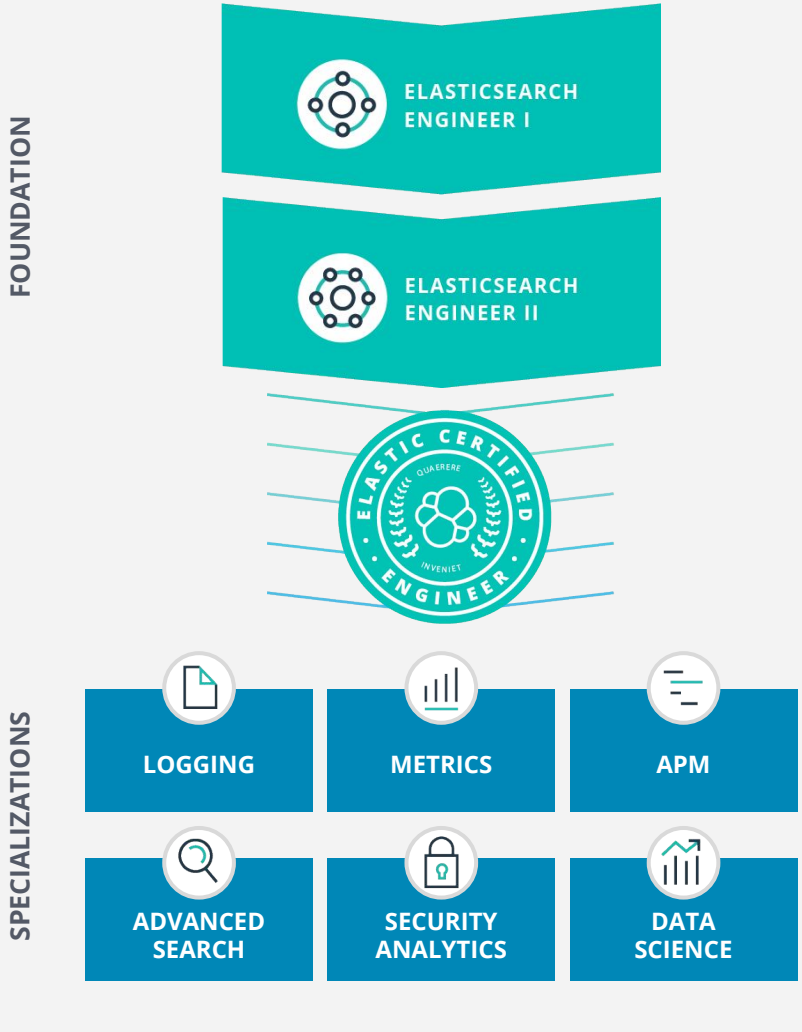
Kibana 입문 교육으로 데이터 시각화 및 고급 분석 기술을 활용하실 수 있습니다.

Data Analysis with Kibana 커리큘럼

Observability Engineers

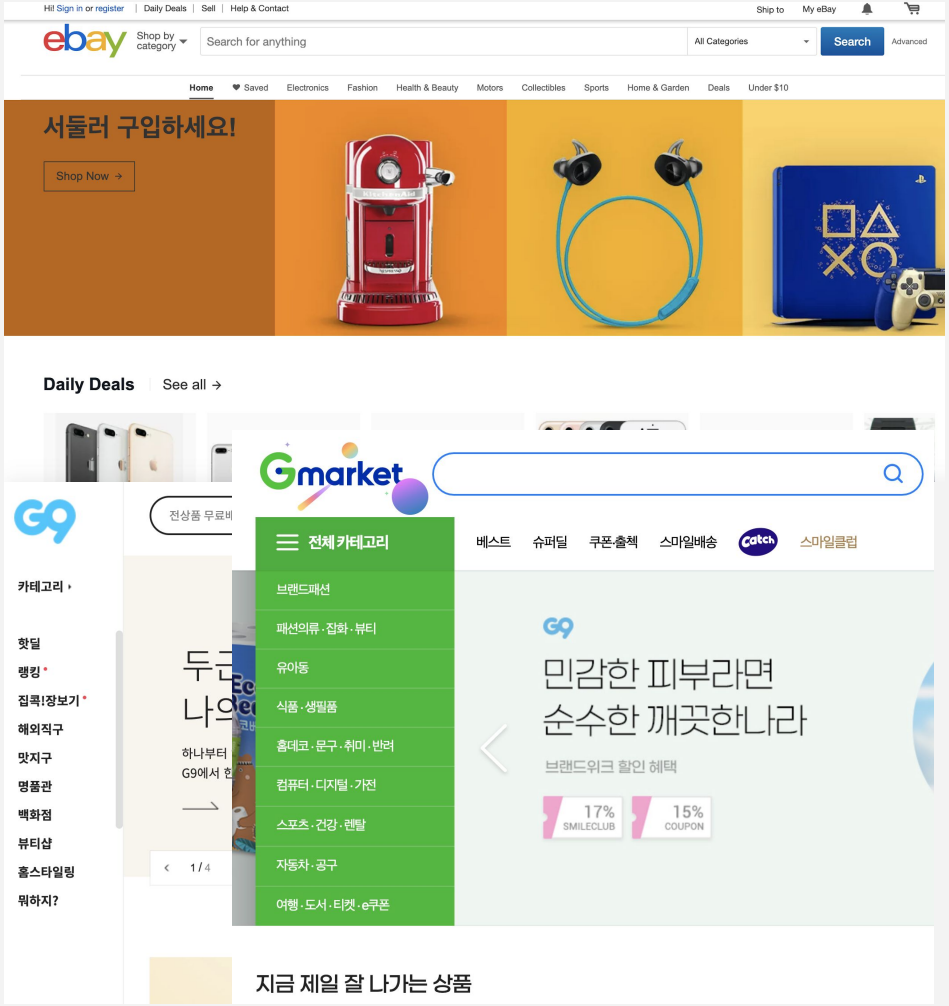
Elastic Stack을 활용해 Observability 데이터 (로그, 메트릭, APM, Uptime)를 통합하는 방법을 배우실 수 있습니다. 대시보드, Alerting, Machine Learning 등의 기능을 통해 통합적인 관측성을 학습합니다.

Elastic Observability Engineer 커리큘럼



Customer References

더 많은 Use Cases 보러가기: https://www.elastic.co/use-cases/

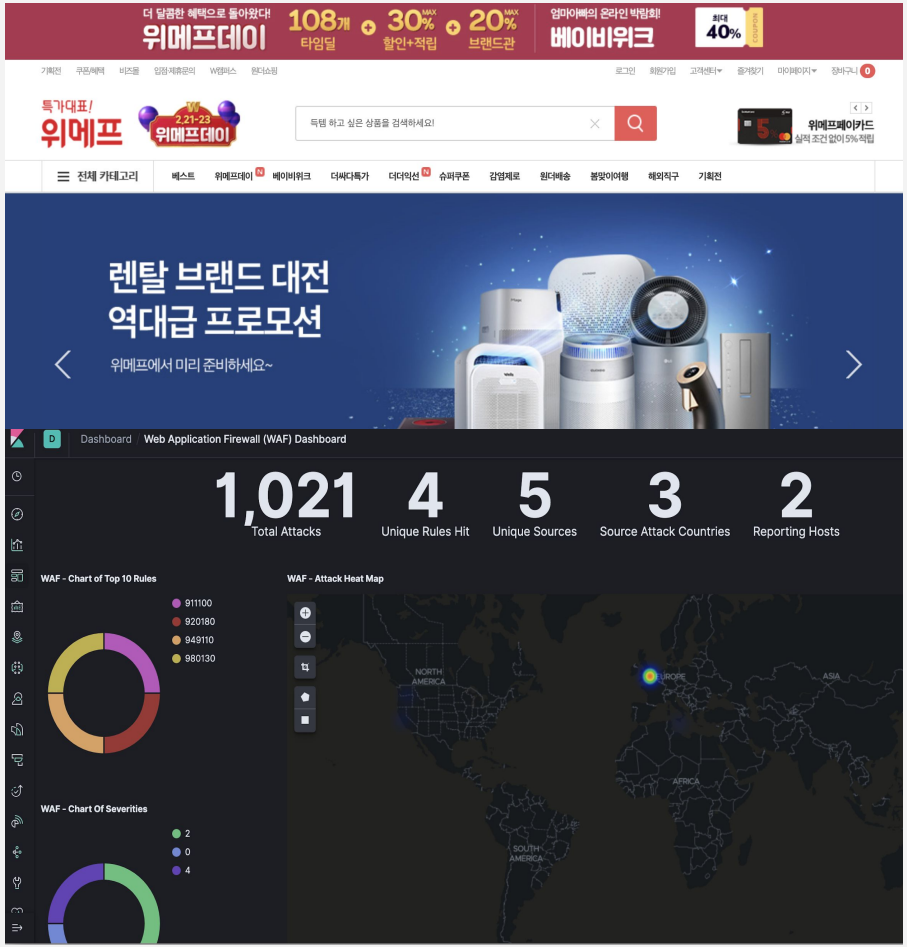


이베이HQ & Korea

- 이베이HQ, 이베이Korea 운영로깅
- 이베이HQ:2015년기준 1.62억명 유저,
  약 8억개의 상품리스트를 1초 이내 검색
  - ❑ 사이트 검색
  - ❑ 구매 및 이상징후, 운영로깅 분석
  - ❑ 보안분석 및 위협탐지
  - ❑ 모바일 로깅 데이터
  - ❑ 2014년부터 약 30여개 프로젝트진행
- 이베이Korea: 운영로깅 이상징후분석
  Auction, G9, Gmarket 통합주소 G9: 상품검색

" This is not SMALL data"

elastic

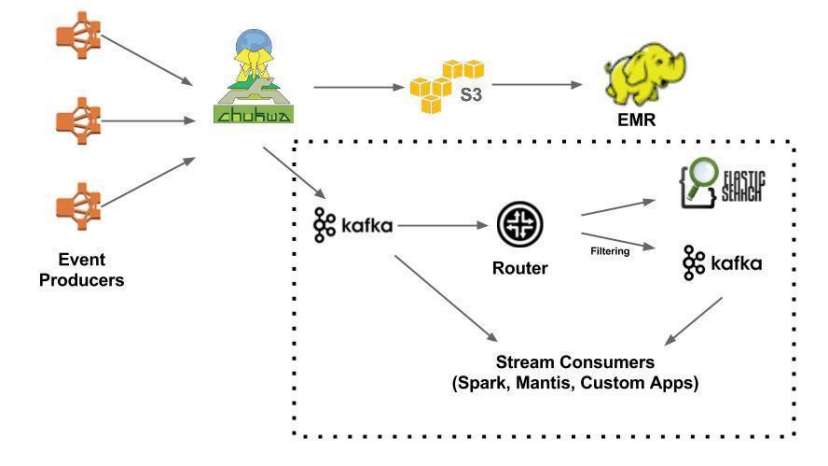


위메프

- 위메프 보안분석: SIEM, 이상징후 ML탐지
- 위메프 상품검색
  - 3억건, 일 증가수 200만건
  - 전체 상품검색 엘라스틱 구성

"보안은 고객의 신뢰를 좌우해요"

elastic

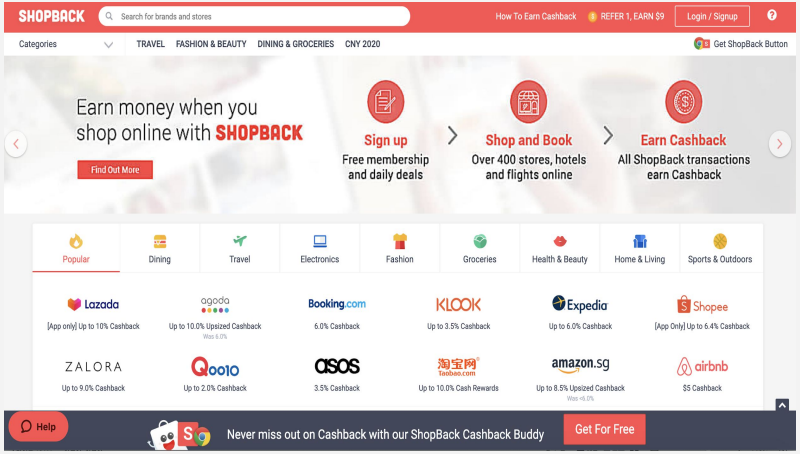


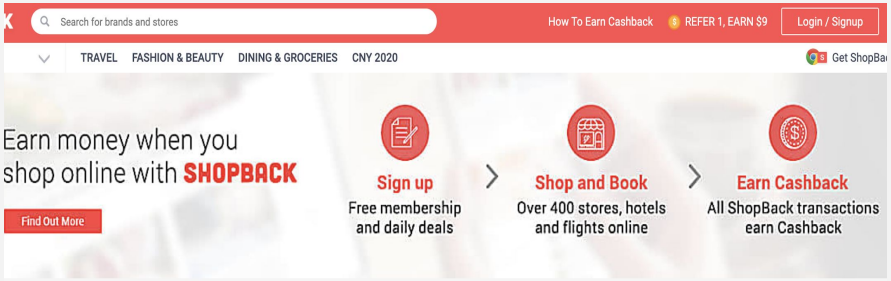
NETFLIX

- 5.8 페타 바이트의 DB: 1초 이내에 검색
- 개인의 시청시간, 이력, 클릭수 등을 학습하여 개인화된 추천 서비스 제공

"넷플릭스는 로깅 회사 입니다
그리고 영화 스트리밍도 제공하죠!"

elastic



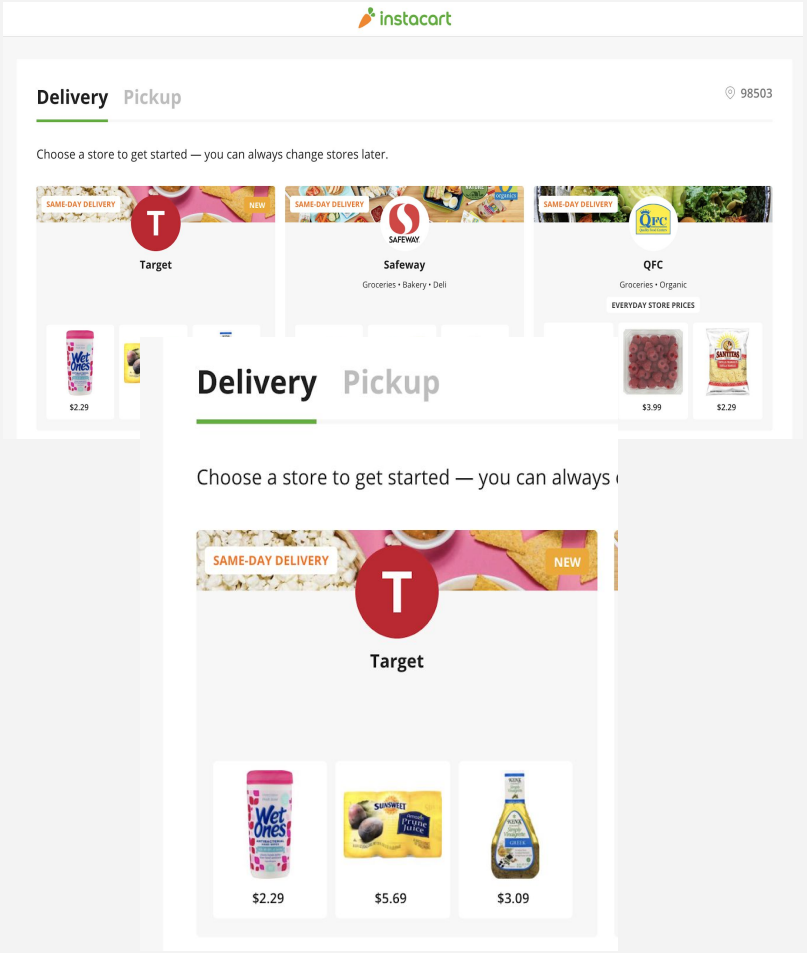


ShopBack

- 30개의 리스트가 2초 -> 0.1초로
- 개인화 쇼핑 추천기능
- 다국어 검색 지원
- 1300백만 상품리스트 인덱싱
- 200백만 카테고리 분류
- 6개국 동시 서비스
- 하루 2.7GB 로깅, 메트릭 수집

"고객의 관점에서, 빠르고 쉽게"

elastic



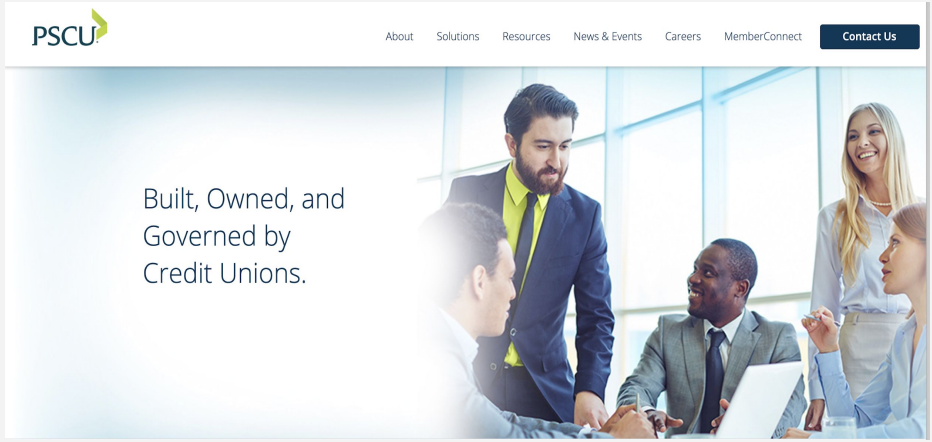
Instacart

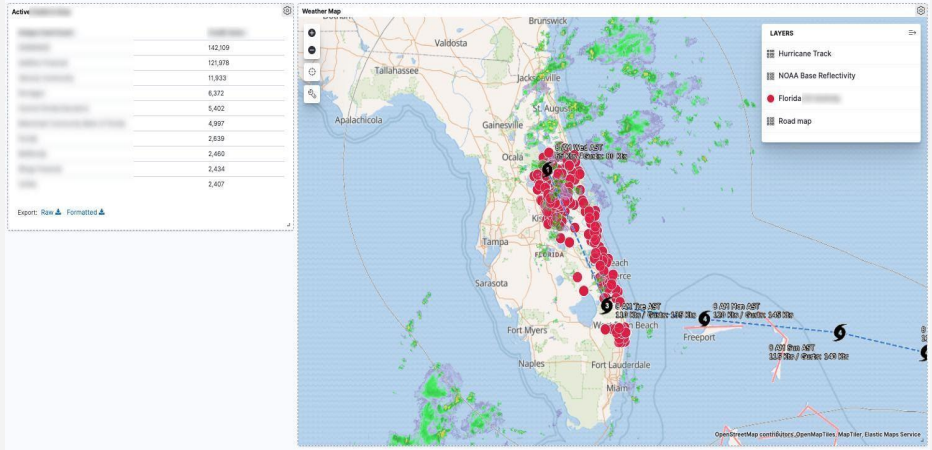
- 5억개의 리스트를 2만개의 지역과 연계해서 검색
- 카탈로그 검색(슈퍼->온라인)
- 구매패턴 학습으로 개인화 추천
- 구매를 위한 **다이내믹 장바구니 제공**
- 개인화 상품 랭킹, 검색 기능 제공
- 개인화 **변동형 가격 제공**

"온전히 한명만을 위한 사이트처럼"

elastic

미국신용조합 PSCU



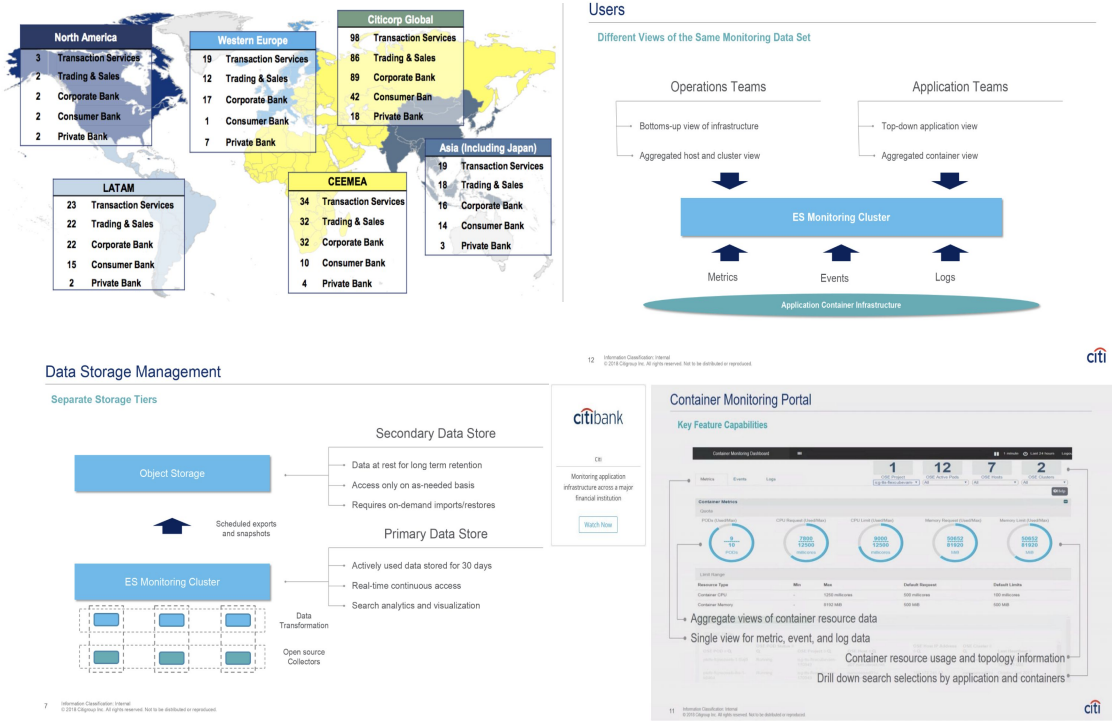


- 적재만 24시간에서 실시간 탐지로 변경
- 99% MTTK, TTD 감소(보안위협탐지)
- 다양한 데이터를 추가 적재하여 데이터 인사이트 확장
- 상관분석을 통한 이상징후 인텔리젼스 구현

"18개월 동안, 41억 이상거래탐지!!!, 브라보"

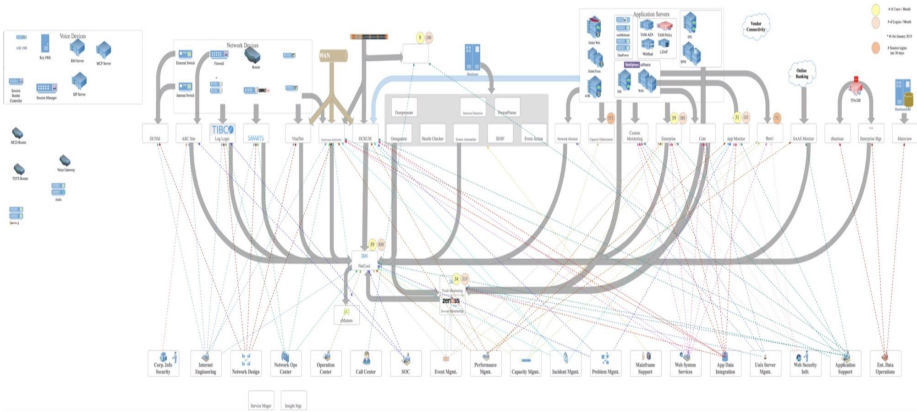
elastic

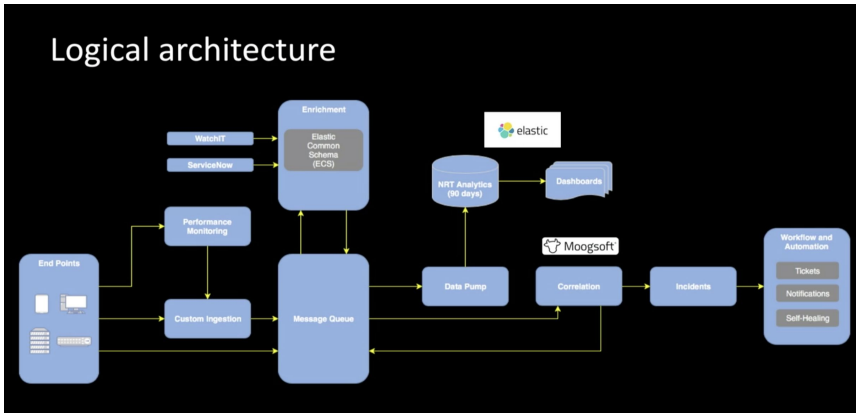
CITI GROUP (Elastic을 사용한 통합 로그/메트릭 모니터링)



- 100여 개국에 진출해 있고 160여 개국에서 클라이언트를 보유하고 있는 CitiGroup 전체 시스템 로그/메트릭 데이터를 통합하여 엘라스틱에 저장(30일 보유)후 30일 지나면, 저렴한 secondary 파일시스템에 자동으로 보관
- 엘라스틱 Alerting(Rule-Engine)(이전의 Watcher)을 사용해 컨테이너 인프라 모니터링을 설정함으로써, 모든 부서를 대상으로 컨테이너 메트릭, 이벤트, 로그를 수집하는 시스템을 개발할 수 있었고, 이 정보는 감사 발견사항, 알림 관리, 티켓 발행 등 포함하는 다양하게 활용함
- Kibana를 사용해, 실시간으로 애플리케이션 상태의 집계 보기, 메트릭과 로그의 개별 보기, 컨테이너 리소스 사용 및 토폴로지 정보를 제공하는 대시보드를 생성, 사용자는 애플리케이션이나 컨테이너로 세부적으로 들어가 그 소스에서 문제를 발견할 수도 있습니다. 운영팀과 애플리케이션팀은 데이터를 보는 방법을 뒤집어, 도구 성능을 모니터링하고 실제 도구의 효과성 개요를 보여주는 보고서를 실행하고 있음

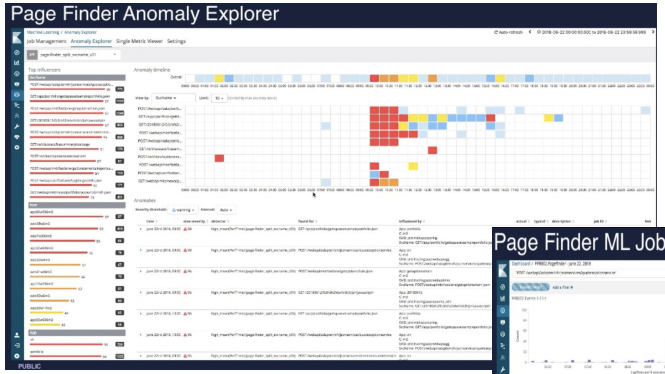
KEY BANK (End to End Enterprise Monitoring구축)

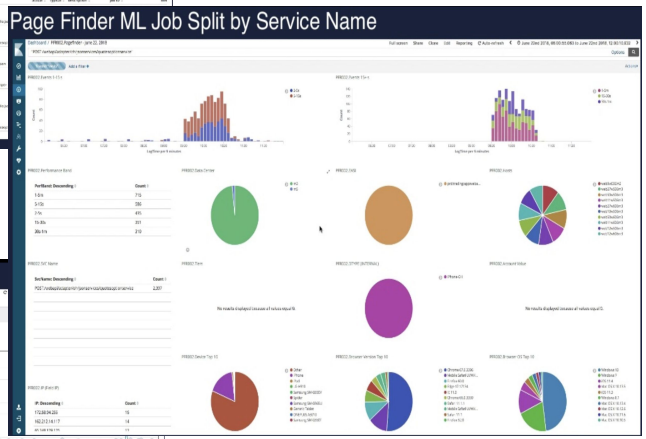


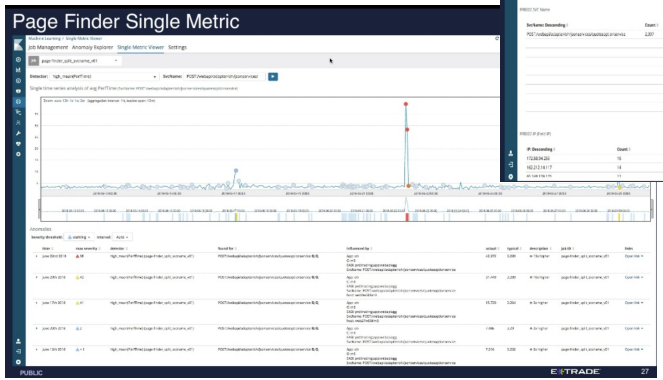


- KEY BANK는 미국 대형 은행중 하나로, 직원수 18,000명, 15개주에서 1,100의 지점과 1,400 ATMs 을 운영함
- 왼쪽 아키텍쳐는 엘라스틱으로 전체 모니터링을 통합하기전 모습으로, 개별 모니터링tool만 21개를 갖고 있었고, MTTR(원인을 분석하여 문제를 해결하는데 소요된 시간)은 더욱더 오래걸리고, 원인을 조사하는데 필요한 로그나 메트릭(성능값)조차 충분히 제공되지 못했음
- 엘라스틱을 도입후, 전체 지점의 수천대의 워크스테이션에 윈도우로그비츠와 메트릭비츠를 설치하여 데이터를 수집하기 시작했고, 그동안의 문제가 되었던 부분에 대해 3대 주요 원인을 확인할 수 있었음: consistently high disk I/O, low available memory, and a completely saturated network
- 기존 개별 모니터링 시스템을 통합하면서 약 50억의 비용을 절감(SW비용 및 각 시스템의 전문 인력 운영/유지 비용등 포함)

E-Trade (엘라스틱 머신러닝을 통한 시스템 장애 사전대응)





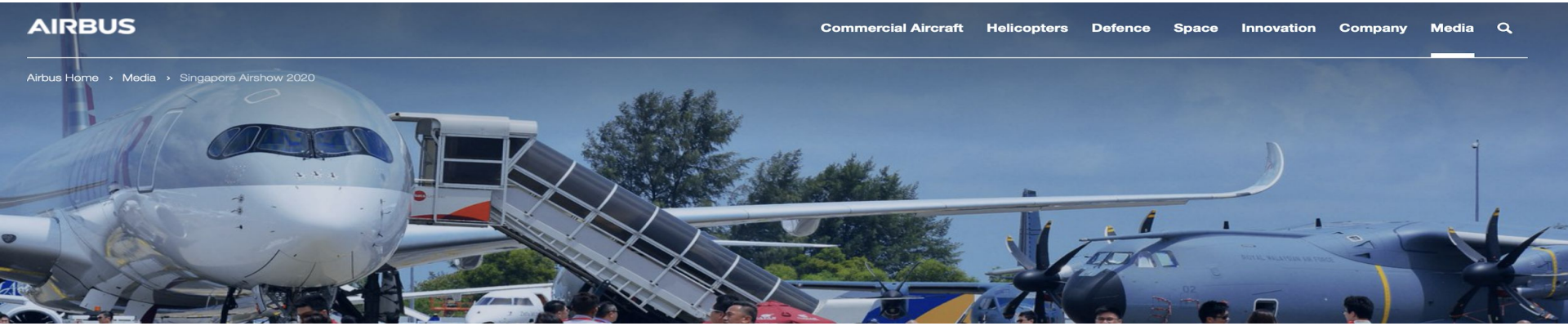


- E-Trade는 미국 최대규모의 온라인 증권사이며 1초에 백만건의 다큐먼트를 실시간 저장하고 있으며 전체 170노드(이중 데이터 노드는 100개)에서 115TB의 데이터를 기반으로 엘라스틱 머신러닝을 통해 이상징후를 탐지하여 시스템에 문제가 심각해지기 전에 문제 해결이 가능해짐

- 이상징후 탐지에 2개의 job을 동시에 적용하는데,
  첫번째는 서비스 performance시간(Perf_time)이 길어지는것과 단기간에 급증하는 count에 대한 부분을 catch하는 머신러닝 Job을 만들어서 Spike(단기간 급증)를 발견하고, 엘라스틱 키바나를 통해 drill-down하면서 그에 대한 원인을 빠르게 파악하고 있으며, 실사례로 서비스 퍼포먼스문제가 특정 한 데이터센터에서만 발생한것을 빠르게 파악하여 그쪽에 연결된 서비스를 다른 데이터센터로 돌린후 장애를 해결하고 다시 원상복귀함으로써 실제 고객은 장애를 격지않도록 조치함

Customer Reference

AIRBUS



신속한 쿼리 응답:

Airbus는 Elasticsearch를 통해 쿼리 응답 시간을 2초 이하로 단축했습니다.

ADNS 플랫폼의 목표는 앞으로도 계속 증가할 **6TB** 문서에 쿼리 작업을 실시하며 모든 사용자 요청에 **2초 이내 응답**하고, **분당 3000개의 쿼리**를 처리한다. 요청 내용과 요청자의 프로필에 따라 적절한 사용자 **액세스 권한**을 제공

Near Real-Time 액세스:

Kibana를 통해 Elastic 플랫폼에 통합된 시각 모니터링 기능으로 최적의 기능을 보장합니다.

항공기 기술 문서에 대한 **Near Real-Time** 액세스를 지원 **20억 개 이상**의 정형/비정형 데이터 블록으로 이루어진 데이터베이스를키워드만으로검색할 수 있도록 지원

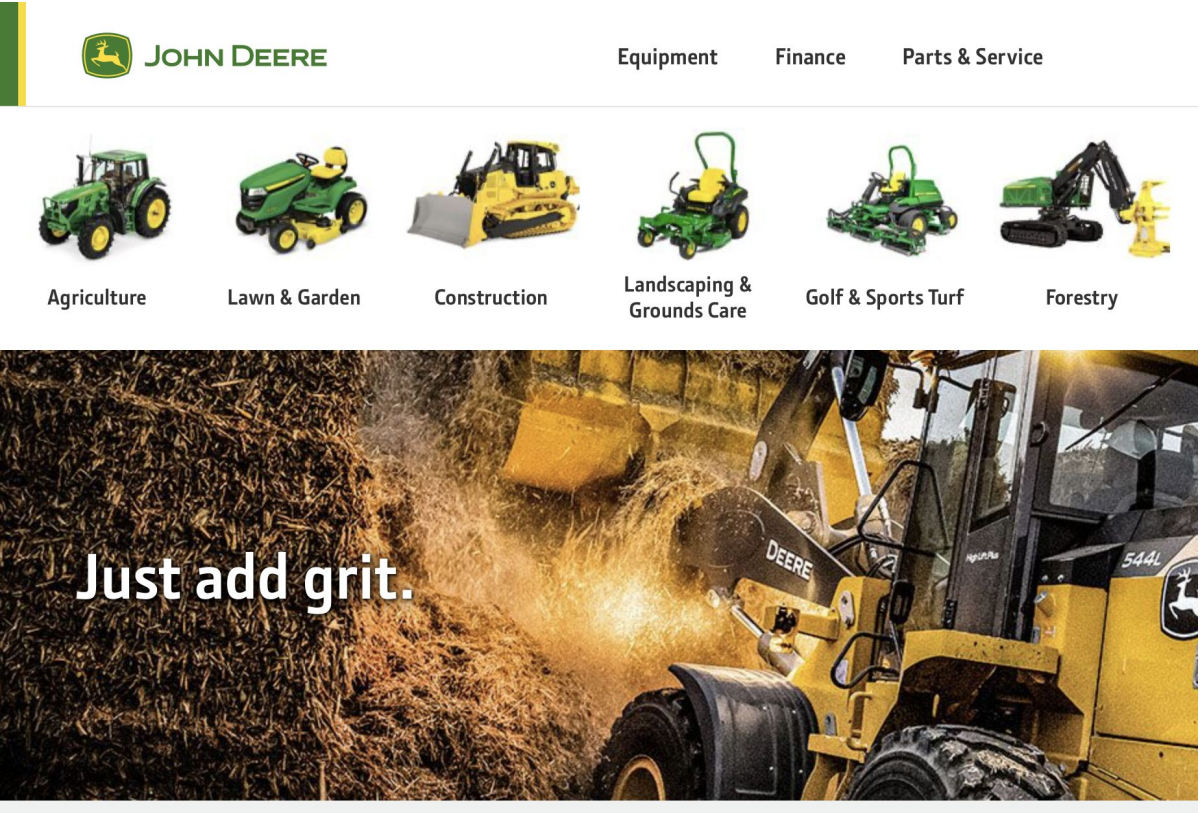
제품 개발 기간 단축:

Airbus는 Elasticsearch로 ADNS 솔루션의 개발 기간을 단축했습니다.

디지털 문서 검색 기능과 컨설팅 플랫폼을 이용해 서비스 수준 협약 **SLA**(Service Level Agreement)을 성공적으로 이행
운영 인프라는 두 개의 내부 가동/가동 **이중화 데이터 센터**로 구성되어 있으며 각 데이터 센터에는 **Elasticsearch 서버 25개의 노드**로 구성

elastic

Customer Reference



Logcentral@Deere

25개의 노드로 구성
- Index rate 20,000 events/s
- 180억건의 문서
- 11TB storage
- 14개의 Applications Logging

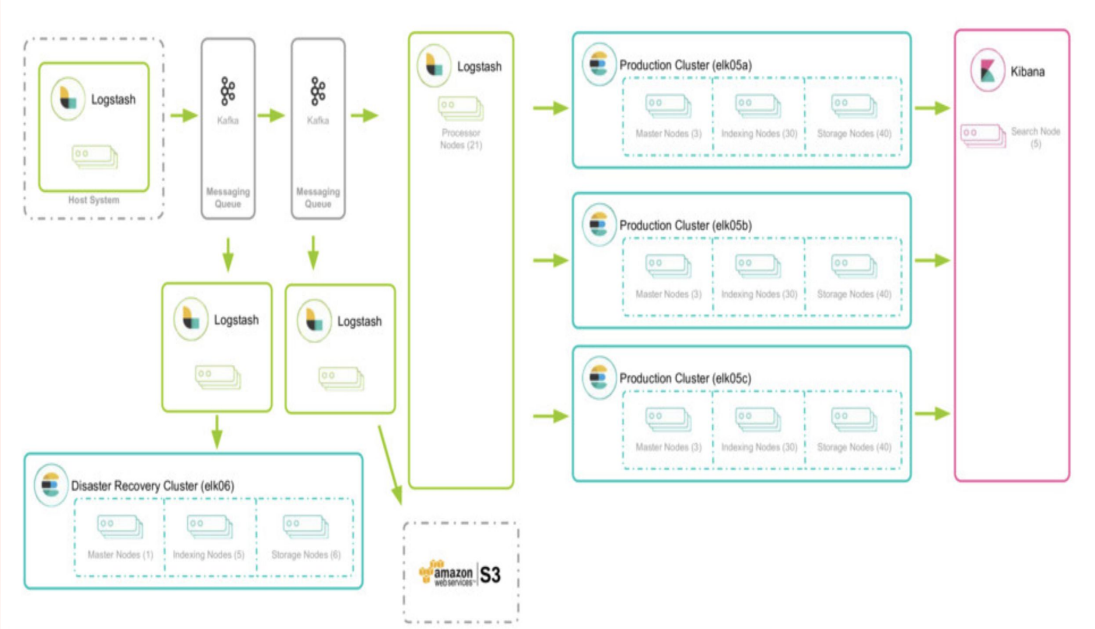
Logcentral@Cloud

8개의 노드로 구성
- Largest cluster has 51 Nodes
- 500억건의 문서
- 200TB storage
- 100+ Applications Logging

Customer Reference

fitbit

fitbit PRODUCTS TECHNOLOGY MOTIVATION



- Log analytic Platform Architecture -

Scaling Log Aggregation At Fitbit

- 약 **300 노드**를 통해 Logging 플랫폼 운영
- **초당 300,000건**의 로그 처리
- **일당 250~300억건**의 로그 메시지 처리
- 로그 타입에 따라 **7 ~ 90일간 로그 저장**
- 디스크 상에 **평균 1700 ~ 1800억건**의 로그 저장
- Kafka를 통해 **44~48 TB**의 데이터 전송
- 효율적인 **DR(Disaster Recovery) 시스템** 구성

elastic

Customer Reference

On/Offline 시설 차세대 보안 플랫폼(NGAV) 구축

- **목적 및 요구사항**
  - 에너지 회사 시설의 손상과 손실 전에 중요한 자산을 보호
  - 솔루션 운영 및 관리시간 단축
  - 표적 공격, 맬웨어 및 그 이상을 차단
  - 최소한의 성능 영향으로 24x7 온라인 및 오프라인 보호
  - 오 탐지 최소화로 정확한 알림 제공
  - 경고를 신속하게 심사하고 수정

- **적용 솔루션**
  - Endpoint Security

- **데이터 표현**
  - 시각화 및 대시보드 도구 : Kibana 외

- **기대효과**
  - Elastic Endpoint Security 보호 기능을 통해 데이터 손실 전에 시스템을 악용 할 수있는 악용, 맬웨어, 피싱, 파일없는 공격, 랜섬웨어 공격 차단
  - 파일리스 공격 기술이 쉘 코드 주입 및 DLL 주입과 같은 기술 차단
  - 가벼운 자율 에이전트를 통해 Offline 시에도 고가 자산을 24x7 보호
  - AI 기반 NLU (Natural Language Understanding) 기반 챗봇인 Artemis® 활용
  - 분석가가 우선 순위를 지정할 수있는 직관적 인 시각화 엔진 인 Resolver를 통해 회사의 기존 보안 리소스를 보강

Pacific Gas and Electric Company

Pacific Gas and Electric Company®

elastic