Elastic Stack-Day 01 (2019-06-08)

IT => DT

IT에서 DT(Data technology) 시대로

Big data

4v = volume(대량) + variety(다양) + velocity(빠르게) => value

빠르게 쌓이는 대량의 다양한 데이터로부터 가치있는 정보를 얻는다

NoSQL

= Not only SQL

=> (최근엔) No structured Query Language = 정형화되지 않은 데이터에 대한 질의 언어

Elastic Search

elastic은 사전상 융통성있는

=> 확장 가능하고, 어떤 Application에도 쉽게 적용 가능한

Elastic stack(ELK stack)

=> 이전엔 ELK stack이라 불렀으나 최근엔 Elastic stack 이라고 부름

Elastic search (검색 엔진)

betas (데이터 수집)

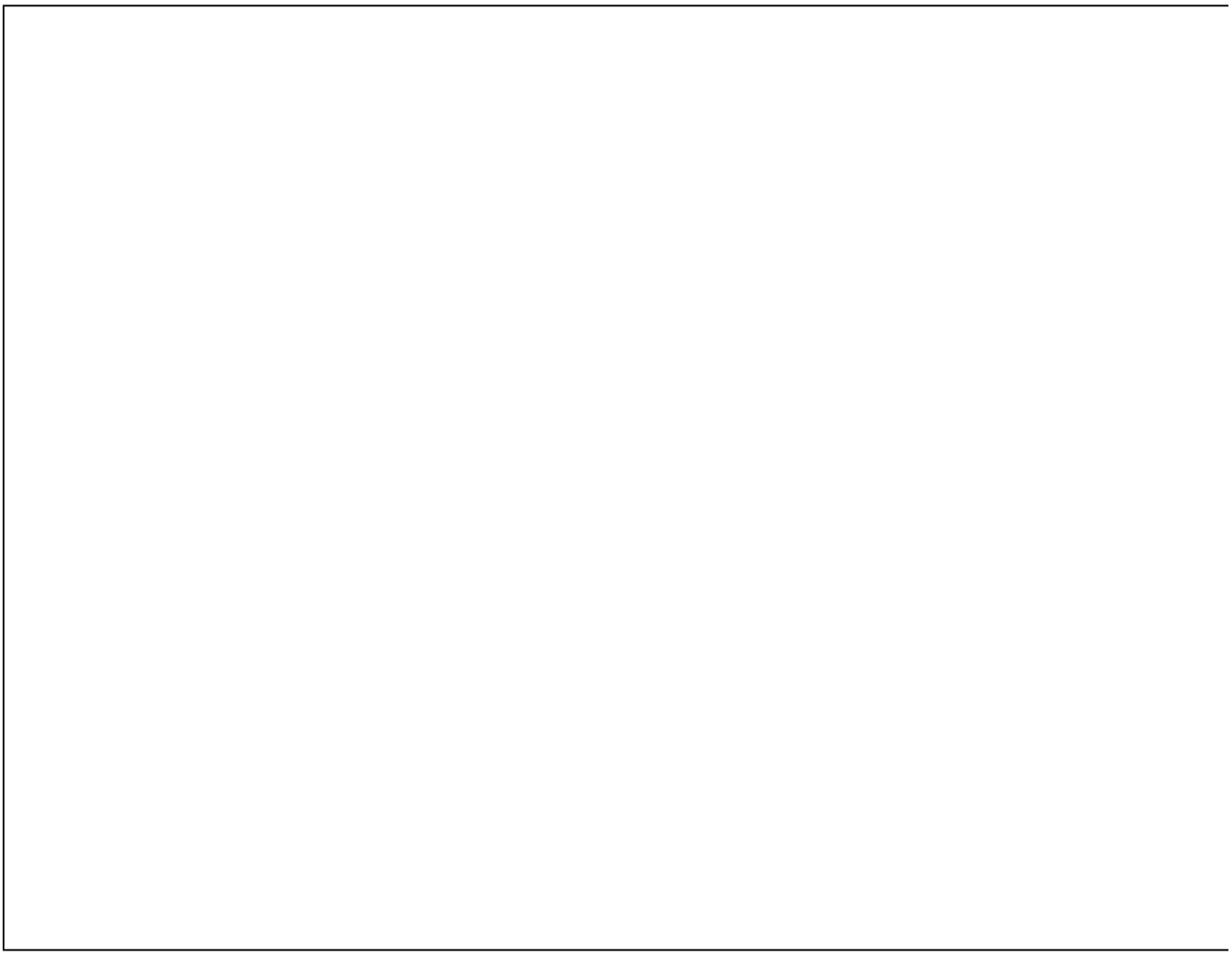
Logstash (데이터 수집 혹은 가공)

Kibana (데이터 시각화)

ELK와 비슷한

Prometheus, Grafana (APM?)

=> Apache2 license



Jaeger

Istio

Docker에 ELK image

https://hub.docker.com/r/sebp/elk

DB Ranking

https://db-engines.com/en/ranking

Elastic search는 검색엔진 1위

시계열 DB 분야에도 상위랭크 (1위는 Influx DB)

Beats

경량 에이전트로 데이터를 Logstash or Elastic search로 전송

Beat	설명
PacketBeat	패킷 분석 로그, wireshark와 연동하여 보안 관련 패킷 분석 가능
FileBeat	Filebeat는 내부 모듈(auditd, Apache, NGINX, System, MySQL 등)을 통해 일반적인 형식의 로그 데이터를 단일 명령으로 간편하게 수집
MetricBeat	시스템의 CPU 사용률과 메모리, 파일 시스템, 디스크 IO, 네트워크 IO에 대한 통계와 시스템에서 실행되는 모든 프로세스에 대한 통계수집

Winlogbeat	Windows 이벤트 로그
Heartbeat	Heart beat에 대한 로그
Auditbeat	Linux 감사 프레임워크와 직접적으로 통신하여 auditd와 동일한 데이터를 수집

Logstash

아래 세 과정을 통해 데이터 가공 및 전송

- input
- filter
- output

Elastic Search

장점

- RESTful API 제공
- 여러 인덱스를 한꺼번에 조회할 수 있어 멀티테넌시 기능 제공
- 역색인
  => 텍스트 정보를 형태소 분석하여 검색 사전을 만들고 Inverted index 방식으로 데이터를 저장함
  => Apache solar보다 성능이 뛰어는 이유
  => 역색인이란 종이책의 마지막 페이지에서 제공하는 색인 페이지와 비슷
  ex)
  데이터 → 10p, 40p, 248p
  검색 → 11p, 72p, 891p

단점

- 이러한 역색인 과정과 커밋 플러시 등의 작업이 있어
  실시간이 아닌 준 실시간(Near Realtime)임
- 트랜잭션 롤백 미지원
- 데이터 업데이트가 아닌 삭제 후 재저장과 형태로 동작
  => Immutable 하게 데이터를 관리하므로 장점도 됨

ELK 설치

JDK 설치 (1.8 이상으로 설치 할 것)

JAVA_HOME, PATH, CLASSPATH 적용

ELK 설치

x-y-z 버전을 동일하게 다운로드하는 것이 좋음 (예제는 6-4-3)

CentOs의 경우 MACOS/LINUX 혹은 RPM 버전으로 다운로드 (예제는 RPM 버전)

Product	Download URL
ElasticSearch	https://www.elastic.co/kr/downloads/past-releases/elasticsearch-x-y-z
Filebeat	https://www.elastic.co/kr/downloads/past-releases/filebeat-x-y-z
Logstash	https://www.elastic.co/kr/downloads/past-releases/logstash-x-y-z
Kibana	https://www.elastic.co/kr/downloads/past-releases/kibana-x-y-z

Elastic stack server에 아래 rpm 설치 (root가 아닌 일반 사용자 계정으로 권한 에스컬레이션을 이용하여 설치할 것)

- Elastic search
- Logstash
- Kibana

SSL 통신을 위한 인증서 생성을 위한 환경설정

- openssl을 통해 키 생성

방화벽 해제

- 9200 포트, 5443포트에 대해 방화벽 해제
- 방화벽 reload
- firewall-cmd --list-all 명령어로 방화벽 해제 확인

echo $JAVA_HOME에 설정된 Java home을 elastic search 설정파일에 입력

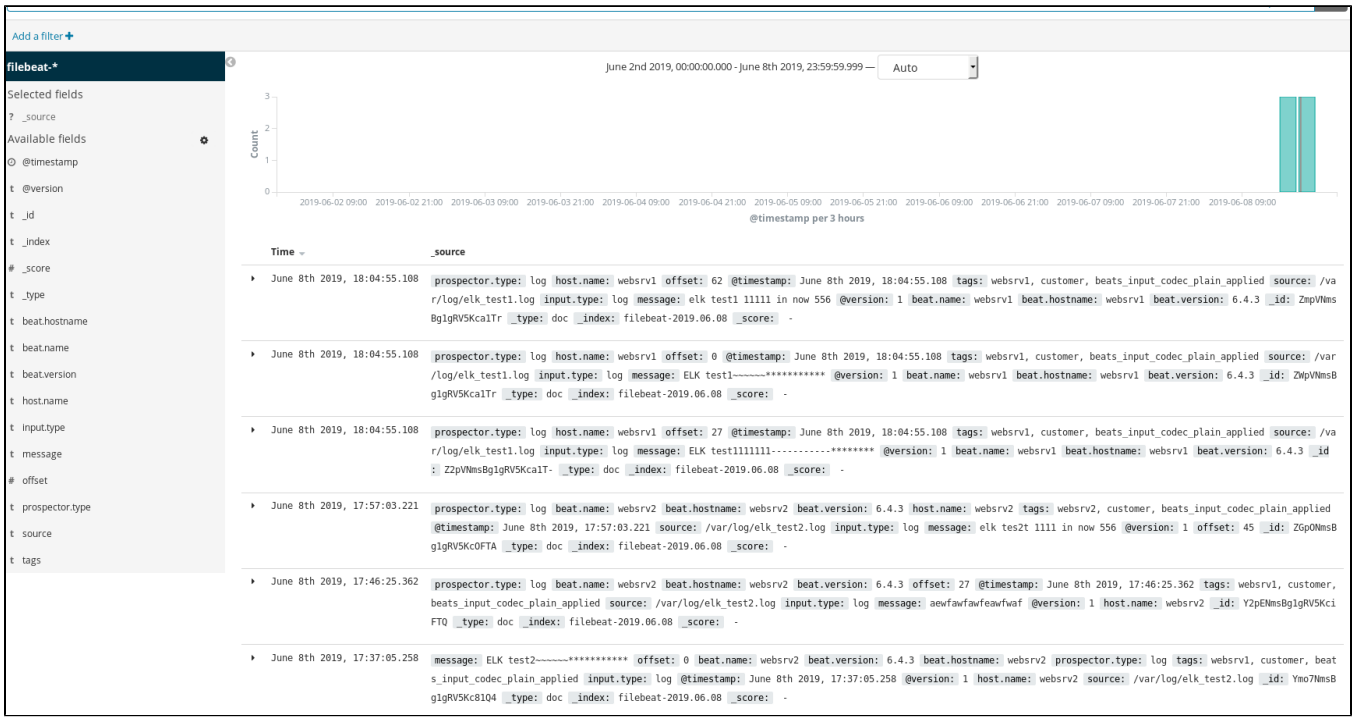
아래 각 명령어로 상태 확인 (문제 생길 경우 rpm으로 지웠다가 다시 설치)

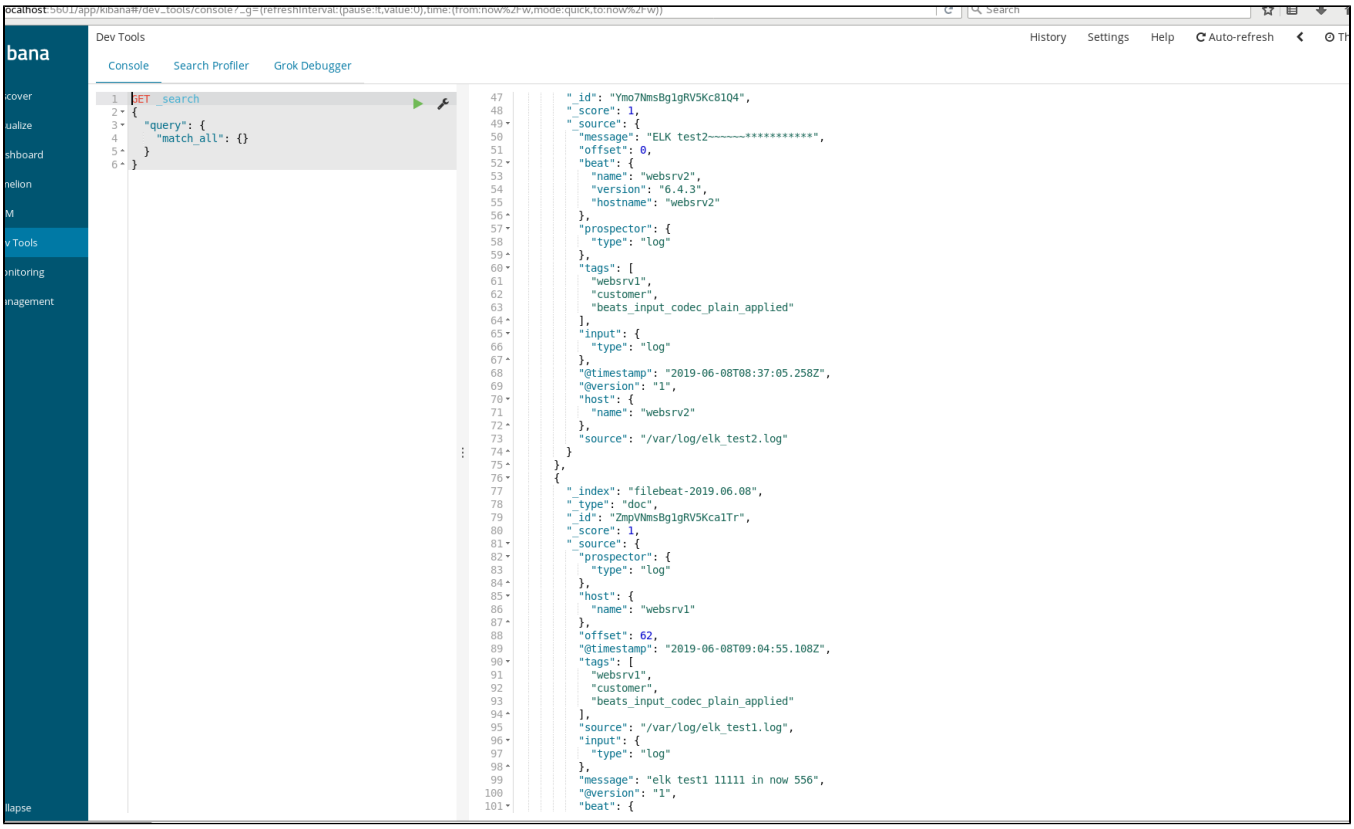
- systemctl start elasticsearch
    - systemctl status elasticsearch
- systemctl start logstash
    - systemctl status logstash
- systemctl start kibana
    - systemctl status kibana

대상 서버에 아래 rpm 설치

- filebeat

filebeat가 설치된 곳에서 대상 로그 파일 수정 후 kibana에서 확인





후기

데모 자료에 정리된 스크립트대로 설정을 해도 제대로 동작하지 않는 케이스가 많음

=〉 시연자가 스크립트를 실행하는 계정을 혼동하여 발생 (root로 실행하지 않고 일반 사용자 계정으로 권한 상승하여 실행해야 하는 경우)

=〉 데모 자료에 정리된 스크립트나 설정파일 변경 등의 작업 내용을 어느 서버에서 실행해야하는지 제대로 정리가 되어있지 않아 교육 참석자들의 혼동으로 수업 진도가 늦음