

# 이성준

대한민국 성남시 분당구 / [sungjunlee127@gmail.com](mailto:sungjunlee127@gmail.com)

## 주요 직무 경험 및 핵심 역량

- 풍부한 대용량 데이터 수신, 처리 및 가공, 전달을 담당하는 각종 데이터 파이프라인 구축 경험
- 대용량 색인, 검색, 시각화를 위한 Elastic Stack 구축 및 다년간의 운영 경험
- Hadoop Ecosystem 기반의 빅데이터 ETL 처리 경험 다수
- 컨테이너 환경에서의 애플리케이션 배포 및 운영 경험 다수

## Skill & Tools

- Language: Java, Kotlin, Python, SQL
- OS: Ubuntu, CentOS
- DB: Elasticsearch, PostgreSQL, Oracle, MongoDB, Redis
- Framework: Spring, Hadoop
- Container Platform: Kubernetes, Docker
- MQ: Kafka
- CI / CD: Jenkins, Airflow
- BI Tool: Kibana
- Data Pipeline: Logstash

## 경력 Detail

### NAVER Shopping

Software Engineer Intern, 2017.12 - 2018.02

Software Engineer, 2018.03 - **재직중**

#### Elastic Stack 기반 통합 모니터링 및 검색 시스템 구축

- Elastic Stack 으로 상품 수신부터 검색 서비스 노출까지 모두 확인 가능한 통합 모니터링 & 검색 시스템 구축
- 5개 클러스터 내 95개의 노드에 걸쳐 400억 건 이상의 docs 저장
- Curator on Airflow 방식으로 인덱스 최적화 및 관리

#### 상품 수신 파이프라인 개발 및 운영

- 외부 물의 상품 데이터를 검색에 노출시키기 위해 수신, 처리하는 stream & batch 기반의 상품 연동 파이프라인 개발 및 운영
- 매일 수십만 개의 물으로부터 10억개 이상의 상품을 받아 처리
- 실시간 처리율 향상을 위해 일부 batch를 kafka기반의 stream 구조로 개편

#### 리뷰 수신 플랫폼 개편 및 운영

- 기존의 거대한 On-premise monolithic batch 기반 리뷰 수신 플랫폼을 Kubernetes MSA stream 기반으로 개편
- 1일 1회 수신 구조에서 실시간 처리 구조로 개선해 처리시간 50% 이상 단축 및 처리량 증가
- 실패 시 원인 파악 및 추적, 특정 부분 자동 재처리가 가능하도록 구성하여 fault-tolerance 확보

#### 메인 DB 전환에 따른 플랫폼 호환성 확보

- 기존 Oracle 을 Citus 기반의 PostgreSQL로 전환하는 작업이 조직 전체에 걸쳐 진행됨에 따라 Oracle 에 맞춰진 쿼리, 프로시저, 코드 등을 분산 Citus 환경에서도 잘 동작하도록 변경하고 최적화 진행
- Bulk write 시 Shard lock 방지를 위한 Shard Id 기준 그룹화 처리, 성능 향상을 위한 unnest & split 방식 적용, distributed table 에 맞게 쿼리 변경

#### **연관상품 추가 수신 및 수신 방식 개편 및 운영**

- 특정 상품과 연관된 상품을 수신, 처리하는 파이프라인에서 신규 타입 상품 추가 수신 개발 및 비효율성 제거
- 상품 Full dump 수신 방식을 기존 HBase random read 에서 Map Reduce 방식으로 변경해 부하 감소
- 처리 구조를 개선해 일부 불필요한 테이블 사용 제거해 RDB dependency 축소
- 전반적으로 기존 대비 처리시간 약 80% 단축 & 추가 처리 상품 약 40% 증가

## 판매지수 수신 개편 및 운영

- 기존 On-premise batch 기반 판매지수 수신 파이프라인을 Kubernetes stream 기반으로 개편
- 수신 & 처리에 실패한 부분을 자동으로 재 수신 및 보정 되도록 개선하여 전반적으로 fault-tolerance 보장
- 처리 시간 레거시 시스템 대비 65% 이상 단축, 제휴 업체들로부터 문의 인입 대폭 감소

## 체크아웃 리뷰 카운트 수신 개편 및 운영

- 제휴 업체로부터 리뷰 카운트를 수신 받아 연동하는 파이프라인의 처리 및 구동 방식 개편
- 기존 Java batch를 Kotlin coroutine 기반으로 개편하여 처리시간을 1시간에서 10분으로 약 83% 단축
- 처리 결과 File을 생성 후 rsync로 전달하던 부분을 Kafka로 변경하여 서버 의존성 제거
- On-premise 방식에서 Kubernetes operator on Airflow 로 전환하여 맥등성 및 안정성 확보

## NAVER D2 기고

- Elastic Stack 으로 통합 모니터링 시스템을 구축한 경험을 바탕으로 네이버 기술 블로그인 D2에 'Elastic Stack과 Lambda 를 활용한 모니터링 시스템 구축'으로 게시글을 기고
- 링크: <https://d2.naver.com/helloworld/9878588>

## 인턴십 멘토링 활동

- 네이버 해데이 채용연계형 인턴십 멘토로 참여해 인턴 과제 선정, 업무 가이드, 회사생활 지원 등의 멘토링을 진행하여 인턴 사원의 성공적인 정규직 전환에 도움을 줌

## NAVER Engineering Day 2019 발표

- 네이버 엔지니어링 데이에 연사로 참여해 'Kafka Lag 감지를 통한 Kubernetes Auto-scaling 적용기' 발표
- 링크: <https://blog.naver.com/occidere/221758990374>

## 학력

---

국민대학교 컴퓨터 공학 전공, 2012.03 – 2018.08, 학점 3.74 / 4.5

## 기타 개인활동 링크

- 
- GitHub: <https://github.com/occidere>
  - LinkedIn: <https://www.linkedin.com/in/sungjun-lee/>
  - Blog: <https://blog.naver.com/occidere>
  - Stackoverflow: <https://stackoverflow.com/users/7110084/occidere>
  - Elastic Bronze Contributor 2021: <https://www.credential.net/557b8611-e161-4a5b-92bb-6b7ea5a0af7d>
  - Leetcode: <https://leetcode.com/occidere/>
  - Baekjoon: <https://www.acmicpc.net/user/occidere>