


캡스톤 디자인 I

종합설계 프로젝트

프로젝트 명	대용량 트래픽 병렬처리 서비스 연구
팀 명	API Rail
문서 제목	중간보고서

Version	1.0
Date	2022-04-02

팀원	강동호 (조장)
	최주원
지도교수	이창우 교수

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

CONFIDENTIALITY/SECURITY WARNING

이 문서에 포함되어 있는 정보는 국민대학교 소프트웨어융합대학 소프트웨어학부 및 소프트웨어학부 개설 교과목 캡스톤 디자인I 수강 학생 중 프로젝트 “대용량 트래픽 병렬처리 서비스 연구”를 수행하는 팀 “API Rail”의 팀원들의 자산입니다. 국민대학교 소프트웨어학부 및 팀 “API Rail”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

문서 정보 / 수정 내역

Filename	중간보고서-대용량 트래픽 병렬처리 서비스 연구.doc
원안작성자	강동호, 최주원
수정작성자	강동호, 최주원

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2022-04-02	최주원	1.0	최초 작성	
2022-04-02	강동호	1.1	내용 추가	카프카 및 프로세싱 관련 내용 추가
2022-04-03	최주원	1.2	내용 수정	문서 키워드 체크 및 로직 내용 추가

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

목 차

1	프로젝트 목표	4
2	수행 내용 및 중간결과	5
2.1	계획서 상의 연구내용	5
2.2	수행내용	5
3	수정된 연구내용 및 추진 방향	6
3.1	수정사항	6
4	향후 추진계획	7
4.1	향후 계획의 세부 내용	7
5	고충 및 건의사항	8

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

1 프로젝트 목표

현재 (주)스토어링크에서는 쇼핑몰 별로 키워드를 검색할 때, 각종 후보군을 통해 순위를 예측하고, 관리해주는 서비스를 운영하고 있다. 가령, 어떤 키워드를 통해 유료광고와 마케팅을 진행해야 더 나은 순위를 보장할 수 있는지 자동화된 시스템을 통해 예측하고, 이에 대한 솔루션을 제공한다. 이때, 활용되는 데이터는 쇼핑몰별 순위나 리뷰, 키워드 그리고 매출액을 수집하여 진행되는데, 날이 갈수록 사용자들이 자주 입력하는 상위 10%의 키워드가 증가하고 있다. 이에 따라서, 서비스에서 데이터를 수집할 때 더 나은 안정성과 시간을 단축을 위해서 프로젝트를 진행하게 되었습니다.

 <div> 국민대학교 소프트웨어학부 캡스톤 디자인 I </div>	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

2 수행 내용 및 중간결과

2.1 계획서 상의 연구내용

스토어링크는 마케팅을 수집한 데이터를 바탕으로 분석하고 매출 상승을 위한 솔루션을 제공하는데 필요한 데이터가 누락되고 느리게 수집되는 현상이 있었습니다.

이 때 기존의 데이터 수집기가 Spring Scheduler 로 순차적으로 수집하며 Crawling API를 호출하게 되고 얻어진 데이터를 영속성으로 전환해 ORM으로 데이터베이스에 저장합니다.

이 과정에서 발생하는 API의 오류나 데이터베이스의 트랜잭션 누락이 간혹 발생하고 있었고 문제 상황을 체크하고 다시 처리를 시도하는 과정과 처리량을 제어하기 위한 Back-pressure 기능을 도입하기 위해 Kafka를 사용하기로 결정했습니다.

2.2 수행내용

우선 코드를 공개하지 못하는 사내 API와 외부 API로 구분합니다.

우선 API들이 동기 구조로 돌아가고 있는 사내 내부 API셋을 비동기로 전환 완료하여 처리되지 않는 API가 호출되더라도 지속적으로 처리량을 받을 수 있는 Netty Client 로 전환하였습니다.

Docker Compose를 활용하여 Zoo Keeper와 Kafka Manager를 실행시켰을 때 카프카가 실행되어 실행의 우선순위를 지정하여 각 서비스들의 의존성을 정의하여 실행단계를 유지시켰습니다.

Meta에서 제공되는 오픈소스 개발 문서 프레임워크인 Docusaurus 를 통해 React.js 컴포넌트를 이용한 사내 개발문서를 작성하였고 매주 회의를 진행하여 사이트에 회의록도 함께 배포될 수 있도록 구성하였습니다.


Spring WebFlux에서 Kafka Reactive 모듈을 활용하여 스토어링크 API를 카프카 토픽을 Consume 하여 호출할 수 있도록 구성했습니다.

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

3 수정된 연구내용 및 추진 방향

3.1 수정사항

기존에 어떤 시스템 아키텍처와 인프라 스트럭처로 구성을 해야 할지 테크 스택에 대해 많은 논의를 거치고 시작하였기 때문에, 현재 시점까지의 수정된 연구내용은 없습니다.

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

W

4 향후 추진계획

4.1 향후 계획의 세부 내용

현재는 한번에 얼마만큼의 패킷이 전달되는지 검증 과정에 있고 API 마다 처리속도의 차별점을 뒤야 하기 때문에 처리 속도량에 따른 토픽을 두어 빠른 연산과 느린 연산 보통 연산에 대해 각각 구분하는 스트림을 정의합니다.

Docusaurus Front-end 에서 API 토픽 테스트 시뮬레이터 작업을 진행하여 Swagger 에서만 볼 수 있던 API문서를 React.js 프론트엔드에서 확인하여 테스트를 즉각 진행할 수 있는 환경 구성하여 Swagger 가 구성되어 있는 서비스는 바로 API로 사용할 수 있게 구성합니다.

API 호출 시각화를 적용하여 현재의 카프카에 얼마만큼의 메시지가 오가는지 볼 수 있는 차트 생성하여 제공할 예정입니다.

에러로 인해 처리되지 못한 개수와 Retry 요청 그리고 처리가 완료된 개수를 수치로 환원하여 모니터링 가능한 API 스트림 모니터링 툴을 제작할 예정입니다.

 국민대학교 소프트웨어학부 캡스톤 디자인 I	중간보고서		
	프로젝트 명	대용량 트래픽 병렬처리 서비스 연구	
	팀 명	API Rail	
	Confidential Restricted	Version 1.2	2022-APR-03

5 고충 및 건의사항

1. 사내 시니어 부재로 인해 제대로 메시지 브로커 등을 사용하고 있는지에 대해 확인받을 수가 없는게 아쉽습니다.