44 멀티 클라우드 환경에서의 LLM 기반 로그 분석 및 이상 탐지 시스템

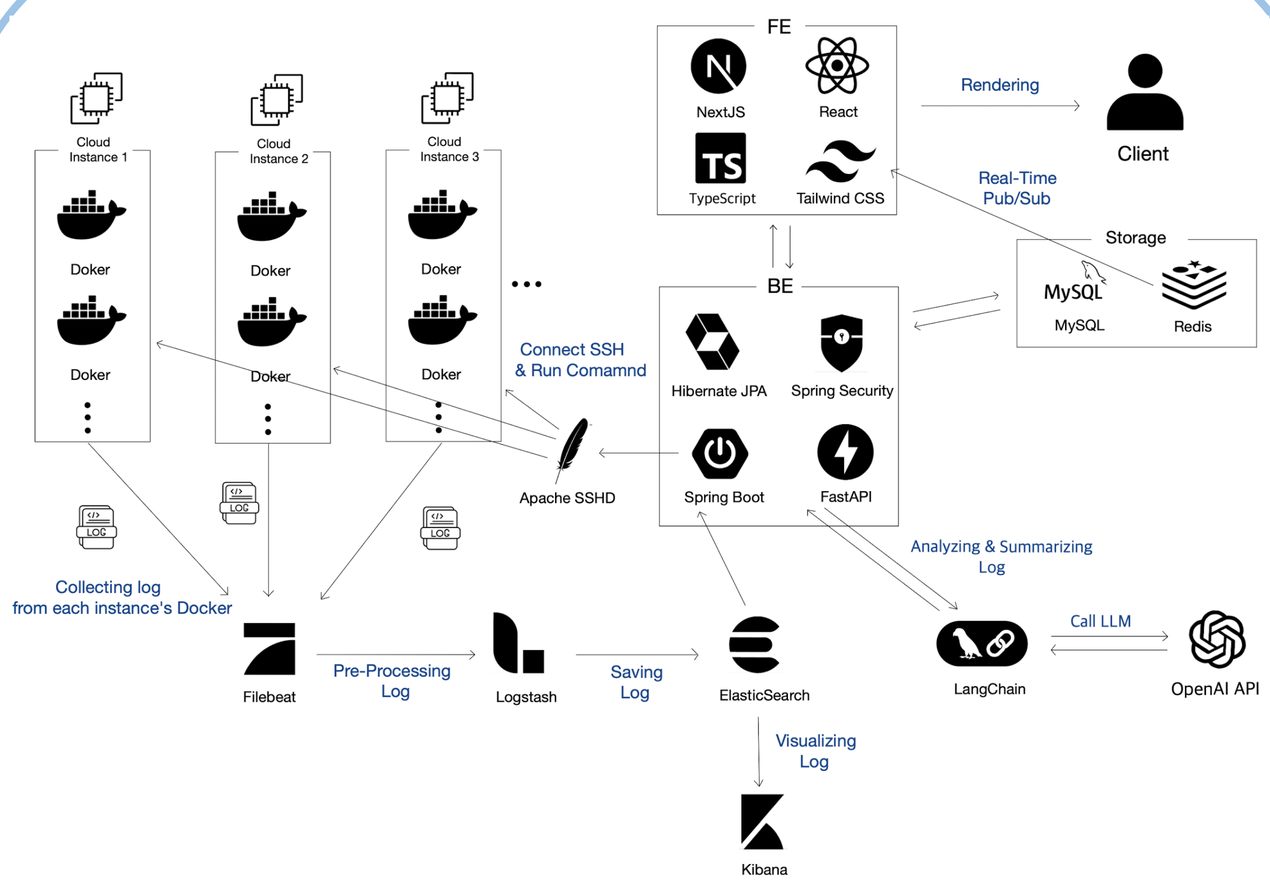
소속 정보컴퓨터공학부 분과 D 팀명 ForPaw 참여학생 이한홍, 이종일, 박재홍 지도교수 김호원

과제 배경 및 목표

- 멀티 클라우드 환경과 MSA 도입으로 애플리케이션이 여러 인프라에 **분산**됨에 따라 로그 데이터의 양이 급증했고, 이를 체계적으로 수집하고 신속히 분석할 수 있는 역량이 중요해졌다.
- 그러나 기존의 규칙 기반 탐지 방식은 **고정된 패턴**이나 **조건**에 의존하기 때문에, 새로운 문제나 예상치 못한 비정상적인 동작을 감지하는 데 **한계**가 있다.
- 이에, **LLM 기반** 로그 분석 시스템을 도입하여, 로그 데이터를 자동으로 **요약**하고, 비정상적인 패턴을 **탐지**함으로써 관리자가 핵심 정보를 빠르게 **파악**하고 **대응**할 수 있게 한다.

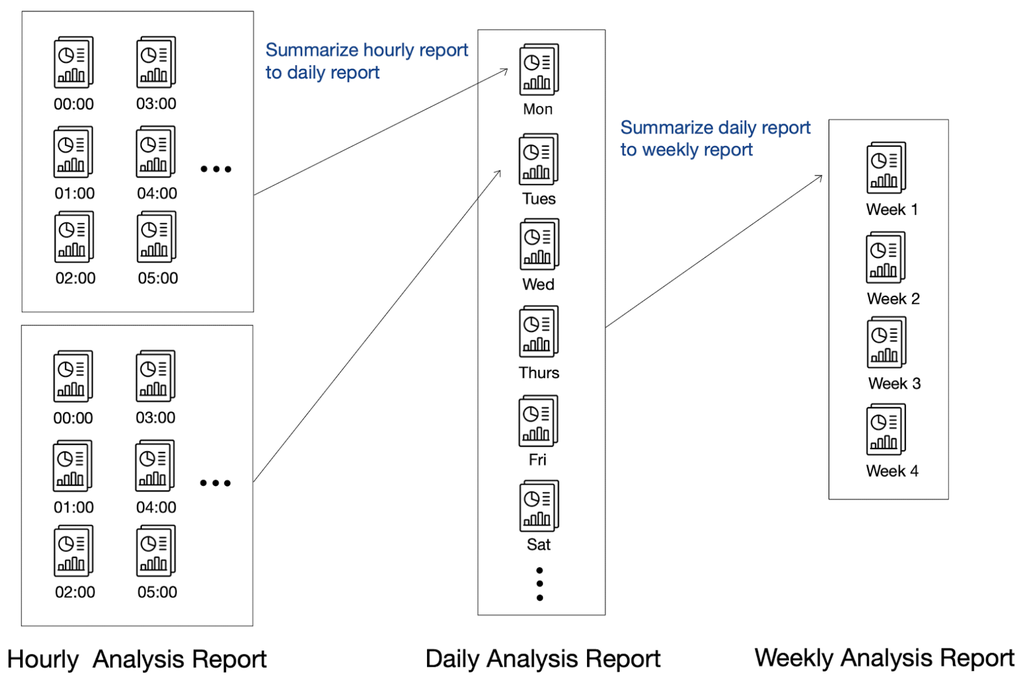
상세 내용

전체 구성도



- ✓ 각 클라우드 서버의 도커 컨테이너에서 생성된 로그는 Filebeat로 **수집**되어 Logstash로 전송된 뒤, **전처리**를 거쳐 ElasticSearch에 저장되고, **LLM**을 통해 요약과 분석이 이루어진다.
- ✓ 로그 데이터 분석 시, LangChain이 로그 데이터를 바탕으로 프롬프트를 생성하고, OpenAI API의 결과를 FastAPI가 처리하여 DB에 저장한다.
- ✓ 문제가 발생하면, 관리자는 Apache SSHD를 통해 각 클라우드 서버에 **원격**으로 접속하여, 실시간으로 명령어를 실행하고 상태를 확인한다.

로그의 계층적 관리

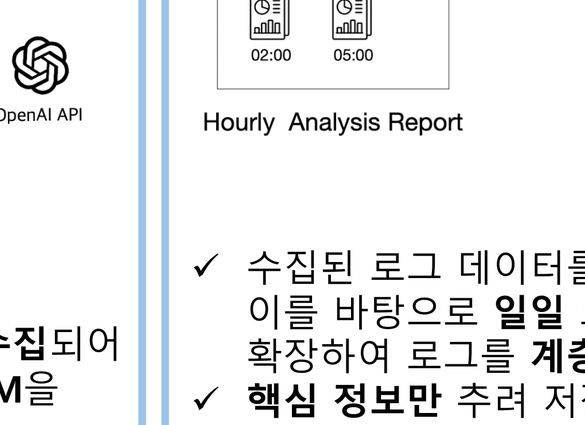


- ✓ 수집된 로그 데이터를 LLM을 통해 시간별로 요약한 후, 이를 바탕으로 **일일** 요약을 생성하고, **주간** 요약으로 확장하여 로그를 **계층적**으로 관리한다.
- ✓ **핵심 정보만** 추려 저장함으로써, 분석 과정에서 불필요한 리소스 낭비를 줄이고, 처리 효율을 높였다.
- ✓ 시간별 로그를 단계적으로 **확장**하는 다단계 분석을 통해, 관리자는 **단기적** 이벤트부터 **장기적** 추세까지 모두 파악하여 정확한 판단과 조치를 내릴 수 있다.

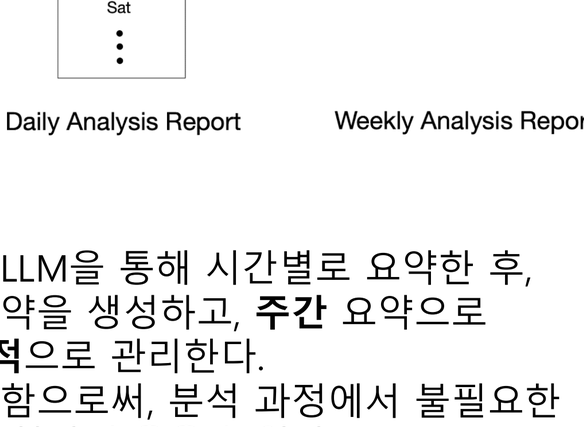
결과 화면

클라우드 서버의 상태를 실시간으로 모니터링하는 대시보드 화면

LLM을 통해 분석된 결과를 다양한 종류의 리포트로 제공하는 화면



질문할 로그를 선택한 후, LLM을 통해 궁금한 내용을 질문하고 답변을 받는 화면



클라우드 서버에 원격으로 접속해 명령어를 실행하고, 그 결과를 실시간으로 확인하는 화면