

과제

프로필, 인증, 결제 등 강한 정합성이 필요한 정형 트랜잭션은 RDB로 처리한다. 시청기록, 좋아요, 댓글 같은 대량 비정형/고쓰기 데이터는 NoSQL과 Redis 캐시로 수평 확장한다. 클라우드 분산을 택하는 이유는 오토스케일과 글로벌 리전으로 트래픽 급증 대응과 관리형 서비스·멀티AZ로 가용성과 운영비 때문이다. OLTP와 OLAP을 분리해 성능 간섭을 차단, 분석·추천 학습을 가속한다. 데이터 흐름은 Kafka→오브젝트스토리지 (S3/GCS)→ELT(Flink/Spark)→DWH(BigQuery/Snowflake)→피처스토어로 이어지며, CDC(Debezium)로 변경을 실시간 동기화한다. 실시간 피드와 알림은 스트림 처리로 집계한다.