

제안서

| | |
|-------|---------------------------------|
| 사 업 명 | 의류 쇼핑몰 마케팅 전략을 위한 데이터 인프라 구축 |
| 주 관 사 | Find Customers |
| 담 당 자 | 김민경 |

2024. 06. 23.

목 차

| | |
|-------------------------|----|
| 1. 사업 개요 | 4 |
| 1-1. 개요 | 4 |
| 1-2. 배경 및 필요성 | 4 |
| 1-3. 사업 범위 | 6 |
| 1-4. 기대효과 | 8 |
| 2. 추진 전략 | 8 |
| 2-1. 추진 목표 | 8 |
| 3. 적용 기술 | 9 |
| 3-1. 적용 기술 | 9 |
| 4. 기술 및 기능 | 14 |
| 4-1. 시스템 요구 사항 | 14 |
| 4-2. 기능 요구 사항 | 18 |
| 4-3. 보안 요구 사항 | 18 |
| 4-4. 데이터 요구 사항 | 19 |
| 4-5. 시스템 운영 요구 사항 | 24 |
| 4-6. 제약 사항 | 25 |
| 5. 성능 및 품질 | 26 |
| 5-1. 성능 요구 사항 | 26 |
| 5-2. 품질 요구사항 | 26 |
| 5-3. 인터페이스 요구사항 | 27 |
| 6. 프로젝트 관리 | 27 |
| 6-1. 프로젝트 관리 방법 | 27 |
| 6-2. 일정 계획 | 28 |
| 7. 프로젝트 지원 | 31 |
| 7-1. 시험 운영 | 31 |

| | | |
|-------|---------------|----|
| 7-2. | 교육 훈련 | 32 |
| 7-3. | 하자 보수 | 32 |
| 7-4. | 기밀 보안 | 32 |
| 7-5. | 비상 대책 | 33 |
| 8. | 구성 예상도 | 33 |
| 8-1. | 예상도 | 33 |
| 9. | 기대 효과 | 37 |
| 9-1. | 기대 효과 | 37 |
| 10. | 예상 비용 | 38 |
| 10-1. | 네트워크 | 38 |
| 10-2. | 컴퓨팅 | 38 |
| 10-3. | CI/CD | 39 |
| 10-4. | 모니터링 | 39 |
| 10-5. | 데이터 분석 | 40 |
| 10-6. | 백업 | 41 |
| 10-7. | 보안 | 42 |
| 10-8. | 총 비용(월) | 42 |

1. 사업 개요

1-1. 개요

사업명: 의류 쇼핑몰 마케팅 전략을 위한 데이터 인프라 구축

사업예산: 150,000,000원(일억 오천만원)

사업기간: 계약 체결일로부터 60일 이내

1-2. 배경 및 필요성

배경

1. 일본 중심 웹 쇼핑몰
2. 급성장으로 사용자 및 고객 증가
3. 고객들의 다양한 데이터 수집
4. 쇼핑몰 인프라 AWS 이전 계획

필요성

1. 데이터 분석을 통한 마케팅 인사이트 도출
 - 비즈니스가 성장함에 따라 다양한 소스에서 사용자들의 데이터 수집
 - 다양한 소스에서 적절하게 데이터를 수집하고 분석할 수 있는 솔루션 필요
 - 분석 및 시각화를 통해 의사결정 효율성을 높임
 - 데이터 분석 아키텍처 구축 시 On-premise 환경의 기술적 한계 극복

1.1 사용자 경험 개선 및 이탈 방지를 위한 데이터 분석

- 사용자 웹 페이지 경험이 이커머스 분야에서 대두되고 있음
- 스트림 데이터 분석으로 사용자 요구와 행동 이해

- 해당 데이터를 분석해 사용자가 제일 많이 접속하는 시간 대 파악
- 시간 데이터를 기반으로 적절한 마케팅 진행 시간 등을 결정 가능

1.2 효율적인 홍보 방안 및 카테고리 별 고객 선호도 파악을 위한 데이터 분석

- 지면, 옥외 등 오프라인 광고는 온라인 상에서 수집하기 어려움
- 때문에, 설문조사를 활용해 오프라인 광고 유입 파악 예정
- 설문조사를 통해 수집된 데이터를 저장, 정제 및 분석해 시각화 필요
- 추후 오프라인 마케팅 전략에 사용 가능

1.3 마케팅 최적화를 위한 사용자 정보 및 판매율, 리뷰 데이터 분석

- 데이터베이스에 저장된 고객 정보와 구매 이력 데이터
- 고객 특성 별 인기 상품 파악 및 신제품 출시에 활용
- 판매율 및 고객 리뷰 데이터 활용
- 제품 개선 사항과 품질 향상에 대한 방향성 제시
-

2. AWS 환경에서의 인프라 재구축

- 기존 의류 쇼핑몰 인프라는 온프레미스에서 2-Tier로 구성
- 초기 단순한 운영의 요구사항을 충족하기에는 적합하였으나, 쇼핑몰의 성장 및 많은 고객들의 유입으로 인해 Web 서버에 너무 많은 트래픽 발생
- 따라서 Web 서버 계층을 Web, WAS 서버로 분리해 트래픽을 분산하고, 다양한 서비스 통합 및 확장에 용이한 AWS의 웹 애플리케이션(3-Tier) 필요성 제기

3. 웹 공격에 대응하는 보안 솔루션 도입

- 웹 공격 보안 솔루션 부재로 온라인 비즈니스의 안정성이 위협적
- 웹 공격 대응을 위한 보안 서비스를 도입해 외부로부터 내부 시스템 보호

4. 자동화 환경 필요

- 서버 배포 및 데이터 백업 자동화 환경 필요
- 인적 오류 감소 및 일관성 유지 필요

5. 비용 효율적인 데이터 장기 보관 필요

- 일본 법적 규정사항에 따라 백업 데이터의 장기 보관(5년) 시 비용 효율적 전략 필요

1-3. 사업 범위

클라우드 컴퓨팅 서비스 구축 및 인프라 고도화

- 온프레미스에서 운영 중인 인프라를 AWS 컴퓨팅 환경에서 재구축 및 고도화
- 재구축 시 기존의 2-Tier를 3-Tier로 전환
- 트래픽 변동에 유연하게 대응하고 시스템의 안정성을 확보
- 다중 가용 영역에 걸쳐 리소스를 배포하여 장애가 전체 시스템에 영향을 미치지 않도록 함

모니터링 구축

- Web, WAS 서버의 지표(CPU 사용량)의 모니터링 환경 구축
- 모니터링을 바탕으로 이벤트 발생(CPU 75% 이상 사용) 시 운영자에게 Slack을 통한 알람 전송

CI/CD 구축

- 지속적인 Web, WAS 서버의 업데이트가 가능한 자동화된 서버 배포 파이프라인 구축

백업 시스템 구축

- 일본 법적 규정사항(전자상거래법), 법적 분쟁 대응, 고객 문의 대응을 목적으로 구매, 고객, 리뷰 데이터의 백업 및 장기 보관 시스템 구축

보안 시스템 구축

- 웹 공격 등의 보안 위협에 대응하기 위해 AWS WAF 구축을 통한 고급 보안 솔루션 도입

데이터 파이프라인 구축

- 인프라 환경에서 발생하는 데이터를 수집, 저장, 분석할 수 있는 환경 구축
- 데이터 처리 프로세스 사용 및 ETL 작업 구축
- 데이터를 추출, 변환하는 작업을 자동화하여 운영적 비용을 절감
- 데이터 쿼리, 분석 서비스 사용
- 분석된 데이터 시각화 및 대시보드를 제공
- 데이터 파이프라인 모니터링 시스템 운영

DB 데이터

- 사용자 이용 시간이 감소하는 오전 2시에 데이터 추출
- 데이터 추출 시 분석에 필요한 데이터만 추출 및 정제
- 업무 외 시간에 데이터가 추출되는 만큼 자동화 필요

설문조사 데이터

- 담당자가 설문조사를 직접 데이터 분석을 위한 저장소에 업로드
- 파싱 기능이 필요할 경우 자동화하여 파싱된 데이터 저장

로그 데이터

- 사용자의 로그를 실시간으로 저장
- 특정한 시간 대에 저장된 로그를 분석해 시각화하는 인프라 구축

1-4. 기대효과

- 고가용성, 탄력성, 안정성 있는 서비스 제공 가능
- 서버 배포 자동화 가능
- 효율적인 데이터베이스의 데이터 백업 시스템 구축
- 반복적인 데이터 처리 작업을 줄이고 자동화하여 운영 효율 증가
- 효율적인 데이터 처리 및 분석을 통한 정확하고 신속한 의사결정 지원
- 웹 공격으로부터 서버 보호

2. 추진 전략

2-1. 추진 목표

- 갑작스럽게 증가하는 트래픽에도 유연하게 대응할 수 있도록 탄력성, 안정성, 고가용성을 가진 인프라 구축
- 지속적으로 업데이트된 서버를 배포할 수 있는 자동화된 서버 배포 환경 구축
- 일본 법적 규정사항에 따라 데이터를 장기 보관할 수 있는 비용 효율적인 데이터 보관 전략 도입
- 서버 지표의 모니터링 시스템 구축
- AWS Well Architected 원칙을 적용하여 효율적인 데이터 분석 파이프라인 시스템 구축
- 고객이 의사결정을 효율적으로 내릴 수 있도록 데이터 분석 인사이트 도출

- 웹 보안 솔루션 도입으로 안전하게 웹 공격에 대응

3. 적용 기술

3-1. 적용 기술

1) 네트워크

| 적용 기술 | 내용 |
|---------------------------|---|
| Subnet | <ul style="list-style-type: none"> • public subnet 2개, private subnet 6개 • public subnet보다 private subnet에 더 많은 ip 할당 |
| Internet Gateway | <ul style="list-style-type: none"> • 외부 트래픽이 VPC 내부에 접속하고, 내부 서비스가 외부 인터넷에 접속하도록 구축 (VPC와 인터넷 간에 통신할 수 있는 통로) |
| NAT Gateway | <ul style="list-style-type: none"> • Private 서브넷의 Web, WAS 서버가 보안 패치 및 소프트웨어 업데이트가 필요할 시 인터넷 접속이 가능하도록 구축 • NAT Gateway 이중화를 통해 하나의 AZ에 장애 발생 시에도 다른 AZ의 NAT Gateway를 사용하여 지속적인 서비스 사용 가능 |
| Application Load Balancer | <ul style="list-style-type: none"> • 외부 ALB <ul style="list-style-type: none"> - 외부에서 내부 서버로 들어오는 트래픽을 여러 Web 서버로 분산해 시스템의 안정성과 성능 향상 - 외부에서 Private 서브넷 인스턴스에 직접적인 접근을 막아 보안 강화 • 내부 ALB <ul style="list-style-type: none"> - Web 서버의 오토 스케일링 보안그룹에서부터 오는 트래픽만을 WAS 서버에 분산 접속할 수 있도록 동작 |

2) 3-Tier

| 적용 기술 | 내용 |
|------------|---|
| 멀티 티어 아키텍처 | <ul style="list-style-type: none"> • 계층 분리로 계층별 보안 강화 • 계층별 독립적인 수정, 확장 및 개발 가능 |

| | |
|-------------------------|--|
| EC2 Auto Scaling | <ul style="list-style-type: none"> 급작스러운 트래픽에도 지속적인 서비스를 제공할 수 있도록 부하에 따른 탄력성 제공 |
| Security Group | <ul style="list-style-type: none"> 인스턴스 단위에서 인/아웃바운드 접근을 제어해 보안 강화 |
| Session Manager | <ul style="list-style-type: none"> ssh 키 관리나 인스턴스에 대한 인바운드 포트를 열어줄 필요가 없어 보안 리스크를 줄이면서 서버 연결 가능 |

3) CI/CD

| 적용 기술 | 내용 |
|------------------------|---|
| AWS Code Series | <ul style="list-style-type: none"> AWS CodeCommit, AWS CodePipeline, AWS CodeBuild, AWS CodeDeploy를 활용하여 자동화된 코드 배포 파이프라인 구축 완전관리형으로 인해 관리 비용 절감 모니터링 기능 등 추후 인프라 확장 가능성이 있어 AWS와 통합성이 좋은 AWS Code Series 사용 |
| Amazon S3 | <ul style="list-style-type: none"> 빌드 후의 아티팩트 저장 |

4) 백업

| 적용 기술 | 내용 |
|------------------------------|---|
| RDS 스냅샷 | <ul style="list-style-type: none"> 일본 법적 규정사항(전자상거래법), 법적 분쟁 대응, 고객 대응 등을 위해 매일 02:00 AM에 데이터베이스의 데이터를 스냅샷 형태로 백업 |
| AWS Lambda | <ul style="list-style-type: none"> Lambda 함수를 통해 RDS 스냅샷 생성 및 S3로 내보내기 S3에 저장 시 월, 일 별 디렉토리 생성 후 생성 날짜에 맞게 저장 |
| Amazon EventBridge | <ul style="list-style-type: none"> 매일 02:00 AM에 Lambda 함수 실행하도록 자동화 |
| Amazon S3 Standard | <ul style="list-style-type: none"> 매일 생성되는 데이터의 스냅샷을 자동으로 parquet 형식으로 바꾸어 저장 데이터 생성 기간이 30일(1개월) 이하인 데이터 저장 |
| Amazon S3 Standard-IA | <ul style="list-style-type: none"> Amazon S3 Standard 데이터 중 한 달이 지난 데이터는 수명주기 정책을 통해 Amazon S3 Standard-IA로 이동 및 저장해 비용 절감 분기(3개월) 데이터 저장 |

| | |
|---|---|
| | <ul style="list-style-type: none"> • 1분기(3개월)의 데이터는 종종 접근하기 때문에 비용도 절감하고 밀리 초 단위로 액세스 가능한 Amazon S3 Standard-IA에 저장 |
| Amazon S3 Glacier Flexible Retrieval | <ul style="list-style-type: none"> • Amazon S3 Standard-IA에서 2년 이상 된 데이터는 수명주기를 통해 Amazon S3 Glacier Flexible Retrieval로 이동 • 2년간은 법적 분쟁 발생 가능성이 비교적 높음. 급할 시에는 비교적 빨리 데이터를 확인할 수 있어야 하기 때문에 검색 옵션이 1분부터 12시간까지인 S3 Glacier Flexible Retrieval로 전환 • 생성된 기간이 3개월 ~ 2년인 데이터 저장 |
| Amazon S3 Glacier Deep Archive | <ul style="list-style-type: none"> • Amazon S3 Glacier Flexible Retrieval에서 2년 이상 된 데이터는 거의 액세스할 경우가 없기 때문에 수명주기를 통해 Amazon S3 Glacier Deep Archive로 이동 • 5년간 장기 보관 (일본 법적 규정사항인 전자상거래법에 따라 5년간 구매와 고객데이터 보관 필요, 법적 분쟁 대응 시 증거 자료로 사용할 수 있도록 5년간 리뷰 데이터 보관 필요) • 5년 이상 된 데이터는 자동적으로 삭제 |

5) 모니터링 및 알람

| 적용 기술 | 내용 |
|--------------------------------|--|
| Amazon CloudWatch | <ul style="list-style-type: none"> • Web 서버와 WAS 서버의 지표(CPU 사용량) 모니터링 |
| Amazon CloudWatch Alarm | <ul style="list-style-type: none"> • CPU 사용량이 75% 초과 시 경보 활성화 |
| Amazon SNS | <ul style="list-style-type: none"> • 경보 발생 시 알림 발행 |
| AWS Chatbot | <ul style="list-style-type: none"> • Amazon SNS에서 발행한 알림을 Slack으로 간편하게 전달 |
| Slack | <ul style="list-style-type: none"> • 담당자는 경보 발생 사실에 대한 알람 수신 |

6) 데이터 파이프라인

| 기술 | 내용 |
|---|---|
| Amazon S3 | <ul style="list-style-type: none"> • 객체 데이터 저장 서비스로 다양한 형태의 데이터 저장 가능 • 거의 무한에 가까운 저장 용량으로 인해 빅 데이터 저장에 편리 • S3의 Key 값을 Value 형식으로 설정해 AWS Glue 크롤링 시 자동으로 파티셔닝 지정 가능 |
| AWS Lambda | <ul style="list-style-type: none"> • 코드를 실행하는 서버리스 컴퓨터 • 엑셀 데이터 파싱 후 parquet 변환, Slack 알람 등 일시적인 작업을 진행할 때 사용 • 최대 15분만 진행하기 때문에, 과도한 데이터 처리 작업이 필요한 것들은 AWS Glue 등의 서비스를 사용 |
| AWS Glue Crawler | <ul style="list-style-type: none"> • 분석을 위한 데이터의 메타 데이터를 추출 |
| AWS Glue ETL Job | <ul style="list-style-type: none"> • 데이터를 전처리하고 변환하는 데 사용하는 작업 • RDS에서 특정 테이블의 데이터를 추출 가능 • 불필요한 데이터 제거로 데이터 처리의 효율성 향상 • 서로 다른 테이블을 합쳐 분석을 위한 하나의 데이터를 만들도록 설정 • Data Quality 기능을 내장하고 있어 분석 전(혹은 전처리 후)의 데이터가 올바른지 확인 가능 |
| AWS Glue WorkFlow | <ul style="list-style-type: none"> • 여러 단계의 ETL 작업을 효율적으로 관리하고 실행할 수 있어, 데이터 파이프라인의 구축 및 간편한 관리가 가능 • 작업이 완료된 후 다음 작업이 자동으로 시작되도록 구성할 수 있어 Glue 작업 간 종속성 정의 가능 • 스케줄러 기능을 통해 특정 시간에 데이터 파이프라인이 동작하도록 설정할 수 있음 |
| Amazon Kinesis Data Stream, Amazon Kinesis Data Firehose | <ul style="list-style-type: none"> • 실시간성 데이터 수집 • 동적 파티셔닝을 활용해 추가적인 전처리 기능을 생략 가능 |

| | |
|-----------------------|--|
| Amazon Athena | <ul style="list-style-type: none"> 추출한 데이터를 기반으로 쿼리를 날려 분석 결과 확인 |
| AWS QuickSight | <ul style="list-style-type: none"> Amazon Athena로 쿼리한 결과 시각화 SPICE 방식을 활용해 쿼리 결과의 일부분을 메모리에 저장하여 비용 감소 |

7) 도메인 및 https 접속

| 적용 기술 | 내용 |
|--------------------------------|---|
| Route 53 | <ul style="list-style-type: none"> 사용자 편의를 위해 도메인을 통한 웹 접속 |
| AWS Certificate Manager | <ul style="list-style-type: none"> ACM을 사용하여 SSL/TLS 인증서의 손쉬운 관리 가능 ALB와 연동하여 HTTPS 접근 가능 |

8) 보안

| 적용 기술 | 내용 |
|----------------|--|
| AWS WAF | <ul style="list-style-type: none"> SQL 인젝션 등 다양한 웹 기반 공격 차단 |

9) IaC

| 적용 기술 | 내용 |
|-------------------|---|
| AWS Cloud9 | <ul style="list-style-type: none"> IaC를 활용하여 빠른 인프라 구축 |
| Terraform | |

4. 기술 및 기능

4-1. 시스템 요구 사항

| | | | | |
|-------------|-----------------|--|---|---|
| 요구 사항 분류 | | 시스템 요구 사항 | | |
| 요구 사항 고유 번호 | | SYS-001 | | |
| 요구 사항 명칭 | | 시스템 계층 | | |
| 요구 사항 상세 | 정의 | 3계층 | | |
| | 세부 내용 | <ul style="list-style-type: none">3계층 구성으로 서버 별 독립적인 수정 및 확장 가능한 환경 구축멀티 리전으로 계층별 2개의 가용영역 사용 | | |
| | 제안 내용 | <ul style="list-style-type: none">계층별 서버 구축 | | |
| | | 1계층 | Web 서버 구축 | |
| | | 2계층 | WAS 서버 구축 | |
| | | 3계층 | DB 서버 구축 | |
| | | <ul style="list-style-type: none">각 계층별 총 1024개의 ip 할당 | | |
| | | 계층 | 가용영역 | IP 대역 |
| | | 1계층 | ap-northeast-1a | 10.0.4.0/22 (10.0.4.0 ~ 10.0.7.255) |
| | | | ap-northeast-1c | 10.0.8.0/22 (10.0.8.0 ~ 10.0.11.255) |
| 2계층 | | ap-northeast-1a | 10.0.12.0/22 (10.0.12.0 ~ 10.0.15.255) | |
| | ap-northeast-1c | 10.0.16.0/22 (10.0.16.0 ~ 10.0.19.255) | | |
| 3계층 | ap-northeast-1a | 10.0.20.0/22 (10.0.20.0 ~ 10.0.23.255) | | |
| | ap-northeast-1c | 10.0.24.0/22 (10.0.24.0 ~ 10.0.27.255) | | |
| 요구 사항 출처 | | | | |

| 요구 사항 분류 | | 시스템 요구 사항 | | | | | | | | | | | | | |
|-------------|----------------------|---|----|----|------|-----------|------|---|--------|-------|-----------|--------|----|-----------|----|
| 요구 사항 고유 번호 | | SYS-002 | | | | | | | | | | | | | |
| 요구 사항 명칭 | | 서버 시스템 | | | | | | | | | | | | | |
| 요구 사항 상세 | 정의 | Web, WAS 서버 | | | | | | | | | | | | | |
| | 세부 내용 | <ul style="list-style-type: none">• CPU와 메모리<ul style="list-style-type: none">- 대규모 사용자를 수용할 수 있는 충분한 계산 능력과 메모리 필요• 네트워크 성능<ul style="list-style-type: none">- 높은 동시성을 지원하기 위한 대역폭 및 높은 네트워크 성능 요구 | | | | | | | | | | | | | |
| | 제안 내용 | <table><tr><th>구분</th><th>사양</th></tr><tr><td>Type</td><td>m5a.large</td></tr><tr><td>vCPU</td><td>2</td></tr><tr><td>Memory</td><td>8 GiB</td></tr><tr><td>인스턴스 스토리지</td><td>EBS 전용</td></tr><tr><td>BW</td><td>최대 10Gbps</td></tr><tr><td>비용</td><td>시간당 0.112USD (도쿄 리전)</td></tr></table> | 구분 | 사양 | Type | m5a.large | vCPU | 2 | Memory | 8 GiB | 인스턴스 스토리지 | EBS 전용 | BW | 최대 10Gbps | 비용 |
| 구분 | 사양 | | | | | | | | | | | | | | |
| Type | m5a.large | | | | | | | | | | | | | | |
| vCPU | 2 | | | | | | | | | | | | | | |
| Memory | 8 GiB | | | | | | | | | | | | | | |
| 인스턴스 스토리지 | EBS 전용 | | | | | | | | | | | | | | |
| BW | 최대 10Gbps | | | | | | | | | | | | | | |
| 비용 | 시간당 0.112USD (도쿄 리전) | | | | | | | | | | | | | | |
| 요구 사항 출처 | | https://aws.amazon.com/ko/ec2/pricing/on-demand/ https://aws.amazon.com/ko/ec2/instance-types/ | | | | | | | | | | | | | |

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 시스템 요구 사항 |
| 요구 사항 고유 번호 | | SYS-003 |
| 요구 사항 명칭 | | 서버 시스템 |
| 요구 사항 상세 | 정의 | 데이터베이스 서버 |
| | 세부 내용 | <ul style="list-style-type: none"> • 기존 데이터베이스 엔진 유지 <ul style="list-style-type: none"> - 기존 데이터베이스 엔진인 MySQL 유지 • 안정적, 고성능 스토리지와 메모리 필요 <ul style="list-style-type: none"> - 데이터베이스 서버의 성능은 애플리케이션 전체 성능에 큰 영향을 미침 |

| | | | |
|----------|--|------------------------------------|-------------------|
| | | - 따라서 높은 성능의 인스턴스 유형 선택 필요 | |
| | 제안 내용 | | |
| | | 구분 | 사양 |
| | | 엔진 | AWS RDS for MySQL |
| | | Type | db.m7g.large |
| | | vCPU | 2 |
| | | Memory | 8 GiB |
| | | 인스턴스 스토리지 | EBS 전용 |
| | | BW | 최대 10Gbps |
| | 비용 | 시간당 0.234 USD (도쿄 리전, 다중 AZ 배포) | |
| 요구 사항 출처 | https://aws.amazon.com/ko/rds/instance-types/ https://aws.amazon.com/ko/rds/mysql/pricing/ | | |

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 시스템 요구 사항 |
| 요구 사항 고유 번호 | | SYS-004 |
| 요구 사항 명칭 | | 데이터 파이프라인 |
| 요구 사항 상세 | 정의 | 데이터의 효율적인 저장, 처리 및 분석 |
| | 세부 내용 | <ul style="list-style-type: none"> • Amazon S3 <ul style="list-style-type: none"> - 객체 데이터 저장 서비스로 다양한 형태의 데이터 저장 가능 - 거의 무한한 저장 용량으로 인해 빅 데이터 저장에 적합 - AWS Glue 크롤링 시 자동으로 파티셔닝 지정 가능 • AWS Lambda <ul style="list-style-type: none"> - 서버리스 환경에서 코드 실행 가능 - 엑셀 데이터 파싱, parquet 변환, Slack 알람 등 일시적 작업 수행에 적합 - 최대 15분까지의 실행 시간 제한으로 인해 긴 |

| | | |
|--|-------|---|
| | | <p>데이터 처리 작업은 AWS Glue 등의 서비스 사용 권장</p> <ul style="list-style-type: none"> • AWS Glue ETL Job <ul style="list-style-type: none"> - 데이터 전처리 및 변환을 위한 작업 수행 - RDS에서 필요한 컬럼만 추출하여 데이터 가공 가능 - 서로 다른 테이블을 조인해 분석을 위한 데이터 통합 가능 - 데이터 품질 관리 기능으로 데이터의 정합성 검증 가능 • AWS Glue Crawler <ul style="list-style-type: none"> - 데이터 분석을 위해 메타 데이터 추출 - RDS에서 필요한 컬럼의 메타 데이터 추출 및 관리 가능 • AWS Kinesis Data Streams 및 AWS Kinesis Firehose <ul style="list-style-type: none"> - 실시간 데이터 스트림을 수집하고 처리하는 서비스 - 동적 파티셔닝을 통해 전처리 과정 최적화 및 간소화 가능 • Amazon Athena <ul style="list-style-type: none"> - 저장된 데이터를 기반으로 SQL 쿼리를 실행하여 분석 결과 조회 가능 • Amazon QuickSight <ul style="list-style-type: none"> - AWS Athena를 통해 실행한 쿼리 결과를 시각화하여 데이터 분석 결과를 제공 - Spice 기능을 활용하여 쿼리 결과의 일부분을 지속적으로 저장 및 비용 절감 |
| | 제안 내용 | <ul style="list-style-type: none"> • AWS Lambda를 활용하여 데이터 파싱 및 변환 작업을 자동화 • AWS Glue를 이용하여 ETL 작업을 최적화하여 데이터 처리 효율성 극대화 • AWS Kinesis를 통해 실시간 데이터를 수집 • AWS Athena와 QuickSight를 통해 분석 및 시각화 지 |

| | | |
|----------|--|---|
| | | 원 <ul style="list-style-type: none"> 데이터 품질 검사를 위해 AWS Glue의 내장된 Data Quality 기능을 활용하여 전처리 및 분석 전 데이터의 정확성과 완전성을 확인하고, 필요 시 정정 작업을 자동화 |
| 요구 사항 출처 | | |

4-2. 기능 요구 사항

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 기능 요구 사항 |
| 요구 사항 고유 번호 | | SFR-001 |
| 요구 사항 명칭 | | 서버 접속 기능 |
| 요구 사항 상세 | 정의 | 쇼핑몰 홈페이지 접속 |
| | 세부 내용 | <ul style="list-style-type: none"> 쇼핑몰 접속 시 Web, WAS 서버로 접속 가능 |
| | 제안 내용 | <ul style="list-style-type: none"> 도메인으로 접속 시 웹 페이지로 접속하도록 구축 WAS 서버는 데이터베이스의 데이터를 가져와 사용자에게 보여지도록 구축 |
| 요구 사항 출처 | | |

4-3. 보안 요구 사항

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 보안 요구 사항 |
| 요구 사항 고유 번호 | | SER-001 |
| 요구 사항 명칭 | | 서버의 보안 접속 |
| 요구 사항 상세 | 정의 | 웹 서버 보안 접속 |
| | 세부 내용 | <ul style="list-style-type: none"> 웹 서버 보안 접속 |
| | 제안 내용 | <ul style="list-style-type: none"> 쇼핑몰 웹 홈페이지에 접속 시 SSL/TLS 암호화 기능을 적용 AWS Certificate Manager 서비스를 활용하여 ALB와 연동 후 HTTPS 연결 환경 구축 |
| 요구 사항 출처 | | |

| | | |
|-------------|-------|--|
| 요구 사항 분류 | | 보안 요구 사항 |
| 요구 사항 고유 번호 | | SER-002 |
| 요구 사항 명칭 | | 웹 서버 공격 보안 |
| 요구 사항 상세 | 정의 | 웹 서버 공격 대응 |
| | 세부 내용 | • 웹 기반 공격 차단 |
| | 제안 내용 | <ul style="list-style-type: none"> • 기본적인 웹 공격 차단을 위해 AWS가 기본적으로 제공하는 AWS Shield Standard 활용 • SQL 인젝션 등 다양한 웹 기반 공격 차단을 위해 AWS WAF 구축 |
| 요구 사항 출처 | | |

4-4. 데이터 요구 사항

| | | | | |
|-------------|-------|------------------------------|---|--|
| 요구 사항 분류 | | 데이터 요구 사항 | | |
| 요구 사항 고유 번호 | | DAR-001 | | |
| 요구 사항 명칭 | | 데이터 수집 | | |
| 요구 사항 상세 | 정의 | 데이터별 수집 개별화 | | |
| | 세부 내용 | • 데이터 소스 별 데이터 수집 시스템 자동화 | | |
| | 제안 내용 | • 로그 데이터 | | |
| | | Amazon Kinesis Data Streams | • AWS Kinesis Data Streams 서비스에서 로그 데이터의 일시적인 저장 • 데이터 급증 시 샤드 증가 후 데이터 분산 저장 • 큐 방식을 통해 AWS Kinesis Firehose로 순차적인 데이터 이동 및 처리 | |
| | | Amazon Kinesis Data Firehose | • Amazon S3 버킷으로 로그 스트림 데이터의 순차적 이동 | |

| | | | | |
|----------|--|--|---------------------------|---|
| | | | Amazon S3 Standard | <ul style="list-style-type: none">• Amazon S3 kmk-clean-log 버킷에 로그 스트림 데이터 저장 및 수집• key-value 형식 사용하여 key를 부여해 추후에 crawling 가능한 환경 구축<ul style="list-style-type: none">- year=[연도]- month=[월]- day=[일]- hour=[시간] |
| | | | • 데이터베이스 데이터 | |
| | | | AWS Glue ETL Job | <ul style="list-style-type: none">• 데이터베이스에 저장되어 있는 데이터 중 필요한 칼럼만 추출• 분석을 위해 여러 테이블을 합칠 때 사용 |
| | | | Amazon S3 Standard | <ul style="list-style-type: none">• Amazon S3 kmk-clean-rds 버킷에 정제된 데이터베이스 데이터 저장 및 수집 |
| | | | • 설문조사 데이터 | |
| | | | Amazon S3 Standard | <ul style="list-style-type: none">• Amazon S3 kmk-raw-survey 버킷에 정제되지 않은 설문조사 raw 데이터 저장• Amazon S3 kmk-clean-survey 버킷에 정제된 설문조사 데이터 저장 및 수집 |
| | | | AWS Lambda | <ul style="list-style-type: none">• 다음 기능을 파이썬 코드로 람다 함수를 생성해 실행<ul style="list-style-type: none">- 설문조사 raw 데이터의 파싱- 파싱 후의 parquet 형식으로 변- Amazon S3 kmk-clean-survey 버킷에 변환된 데이터 이동 |
| | | | | |
| | | | | |
| | | | | |
| 요구 사항 출처 | | | | |

| | | | |
|--------------------|--|--|---|
| 요구 사항 분류 | | 데이터 요구 사항 | |
| 요구 사항 고유 번호 | | DAR-002 | |
| 요구 사항 명칭 | | 데이터 분석 | |
| 요구 사항 상세 | 정의 | 데이터 품질 검사 및 분석 | |
| | 세부 내용 | • 데이터베이스 데이터와 설문조사 데이터의 품질 검사 환경 구축 • 데이터 정제 후 분석 환경 구축 | |
| | 제안 내용 | • 품질 검사 | |
| | | AWS Glue Data Quality | • 분석 전(혹은 전처리 후)의 데이터가 올바른 지 확인 |
| | | AWS Lambda | • Lambda 함수 작동시켜 설문조사 데이터의 품질 검사 진행 시 알람 발생 |
| | | Slack | • 품질 검사 진행 시 담당자에게 알람을 전송해 해당 작업 실행 확인 |
| • 데이터 정제 | | | |
| AWS ETL Job | | • 수집된 로그 스트림 데이터와 품질 검사를 마친 데이터베이스 데이터의 정제 작업 진행 | |
| Amazon S3 Standard | • Amazon S3 kmk-result-log 버킷에 정제된 로그 스트림 데이터 저장 | | |
| | • Amazon S3 kmk-result-rds 버킷에 정제된 데이터베이스 데이터 저장 | | |
| • 데이터 분석 | | | |
| crawler | • 분석을 위해 메타데이터 추출 | | |
| Amazon Athena | • 파티셔닝된 데이터를 기반으로 데이터 분석 | | |
| Amazon S3 Standard | • Amazon S3 2team-query-result 버킷에 분석된 데이터 결과 저장 | | |
| 요구 사항 출처 | | | |

| | | | |
|-------------|-------|--|---|
| 요구 사항 분류 | | 데이터 요구 사항 | |
| 요구 사항 고유 번호 | | DAR-003 | |
| 요구 사항 명칭 | | 데이터 시각화 | |
| 요구 사항 상세 | 정의 | 데이터 시각화 및 인사이트 구축 | |
| | 세부 내용 | • 분석된 데이터를 기반으로 고객이 원하는 인사이트 도출 및 대시보드 구축 | |
| | 제안 내용 | • 다음과 같은 인사이트 도출 | |
| | | 데이터 베이스 데이터 | • 카테고리별 판매수 • 카테고리별 판매율과 리뷰의 상관관계 • 연령대별 카테고리 판매량 |
| | | 클릭 데이터 | • 시간대별 클릭 스트림 수 |
| 설문 조사 데이터 | | • 연령대별 유입경로 | |
| | | • QuickSight를 통한 시각화 진행 <ul style="list-style-type: none">- spice 방식을 통해 자주 쿼리하는 데이터는 원본 데이터베이스의 부하를 줄여 비용 절감- 데이터 소스별로 섹션을 나누어 데이터 분석 결과 대시보드 구성 | |
| 요구 사항 출처 | | | |

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 데이터 요구 사항 |
| 요구 사항 고유 번호 | | DAR-004 |
| 요구 사항 명칭 | | 백업 요구사항 |
| 요구 사항 상세 | 정의 | 데이터베이스의 백업 |
| | 세부 내용 | <ul style="list-style-type: none"> 일본 법적 규정사항(전자상거래법)으로 5년간 데이터 베이스의 구매, 고객 데이터장기 보관 |

| | | | |
|--|-------|---|--|
| | | 데이터 종류 | 5년 보관 목적 |
| | | 구매 데이터 | <ul style="list-style-type: none"> 전자상거래법 준수 법적 분쟁 대응 고객 문의 대응 |
| | | 고객 데이터 | <ul style="list-style-type: none"> 전자상거래법 준수 법적 분쟁 대응 |
| | | 리뷰 데이터 | <ul style="list-style-type: none"> 법적 분쟁 발생 시 중요한 증거로 사용될 가능성 존재 ex) 특히 제품 결함이나 서비스 불만에 대한 분쟁이 발생할 경우, 과거 리뷰 데이터는 중요한 참고 자료가 됨. |
| | | | • 비용 절감을 위해 기간별 데이터 보관 방법 차별화 |
| | 제안 내용 | <ul style="list-style-type: none"> RDS 스냅샷 <ul style="list-style-type: none"> 매일 RDS 스냅샷을 통해 Amazon S3 Standard에 저장 기간별 데이터 보관 방법 차별화 | |
| | | Amazon S3 Standard | • RDS 스냅샷을 1달 동안 저장 |
| | | Amazon S3 Standard-IA | <ul style="list-style-type: none"> Amazon S3 Standard에 1달 이상 된 데이터는 수명주기를 통해 자동으로 Amazon S3 Standard-IA로 이동 분기별(3개월) 데이터 저장 |
| | | Amazon S3 Glacier Flexible Retrieval | <ul style="list-style-type: none"> Amazon S3 Standard-IA에서 2년 이상 된 데이터는 수명주기를 통해 Amazon S3 Glacier Flexible Retrieval로 이동 검색 옵션이 1분부터 12시간까지로 다양하기 때문에 급할 시 비교적 빨리 데이터 확인 가능 생성된 기간이 3개월 ~ 2년인 데이터 저장 |
| | | | |
| | | | |

| | | | | |
|----------|--|--|---------------------------------------|---|
| | | | Amazon S3 Glacier Deep Archive | <ul style="list-style-type: none"> Amazon S3 Glacier Flexible Retrieval에서 1년 이상 된 데이터는 수명주기를 통해 자동으로 Amazon S3 Standard Glacier Deep Archive로 이동 5년 이상 된 데이터는 자동 삭제 |
| 요구 사항 출처 | | | | |

4-5. 시스템 운영 요구 사항

| | | |
|-------------|-------|--|
| 요구 사항 분류 | | 시스템 운영 요구 사항 |
| 요구 사항 고유 번호 | | OPS-001 |
| 요구 사항 명칭 | | 모니터링 및 알람 |
| 요구 사항 상세 | 정의 | 서버의 모니터링 및 알람 |
| | 세부 내용 | <ul style="list-style-type: none"> Web 서버와 WAS 서버의 지표 모니터링 <ul style="list-style-type: none"> 지표: CPU 사용량 Web 서버와 WAS 서버에 이벤트 발생 시 담당자에게 알람 전송 <ul style="list-style-type: none"> 이벤트: CPU 75%이상 |
| | 제안 내용 | <ul style="list-style-type: none"> Amazon CloudWatch를 통해 Web 서버와 WAS 서버의 지표 모니터링 이벤트 발생 시 Alarm 생성 Alarm 생성시 Amazon SNS를 통해 알람 발행 AWS Chatbot이 발행된 알람을 스크립트 등의 기타 작업 없이 Slack으로 간편히 전달 이벤트 발생 시 담당자는 Slack을 통해 알람 확인 가능 |
| 요구 사항 출처 | | |

| | | |
|-------------|-------|--|
| 요구 사항 분류 | | 시스템 운영 요구 사항 |
| 요구 사항 고유 번호 | | OPS-002 |
| 요구 사항 명칭 | | 시스템 탄력성 |
| 요구 사항 상세 | 정의 | 트래픽에 유연한 서버의 탄력성 |
| | 세부 내용 | <ul style="list-style-type: none"> • 급작스러운 트래픽에도 서버의 유연한 대응으로 고객에게 지속적인 서비스 제공 • 부하도에 따른 서버 확장성 및 탄력성 환경 구축 |
| | 제안 내용 | <ul style="list-style-type: none"> • Web 서버와 WAS 서버의 Auto Scaling 기능 도입으로 트래픽에 따라 탄력성 있는 환경 도입 • ALB를 통해 동작 중인 서버에 적절히 부하 분산해 접속 |
| 요구 사항 출처 | | |

4-6. 제약 사항

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 제약 사항 요구 사항 |
| 요구 사항 고유 번호 | | COR-001 |
| 요구 사항 명칭 | | 배포 제약사항 |
| 요구 사항 상세 | 정의 | 서버 배포 제약사항 |
| | 세부 내용 | <ul style="list-style-type: none"> • 허가된 사용자만 업데이트된 서버 배포 가능 |
| | 제안 내용 | <ul style="list-style-type: none"> • 허가된 사용자만 코드 커밋, 빌드, 배포 가능하도록 IAM 권한 세부 제어 |
| 요구 사항 출처 | | |

5. 성능 및 품질

5-1. 성능 요구 사항

| | | |
|-------------|-------|--|
| 요구 사항 분류 | | 성능 요구 사항 |
| 요구 사항 고유 번호 | | PER-001 |
| 요구 사항 명칭 | | 배포 중 가용성 |
| 요구 사항 상세 | 정의 | 배포 중 안정적인 성능 유지 |
| | 세부 내용 | • 서버 업데이트 시에도 지속적으로 서비스를 제공할 수 있도록 배포 자동화 환경 구축 |
| | 제안 내용 | • 롤링 업데이트 방식을 통해 서버를 하나씩 업데이트 하여 시스템을 무중단으로 업데이트 |
| 요구 사항 출처 | | |

5-2. 품질 요구사항

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 품질 요구 사항 |
| 요구 사항 고유 번호 | | QUR-001 |
| 요구 사항 명칭 | | 고가용성 |
| 요구 사항 상세 | 정의 | 서비스 고가용성 |
| | 세부 내용 | • 장애가 발생할 시라도 재빠른 대안 투입으로 고가용성 있는 인프라 환경 구축 |
| | 제안 내용 | • 이중화 환경 구축으로 장애 발생 시 재빠른 대응 <ul style="list-style-type: none">- NAT 이중화- 멀티 AZ- 멀티 Web, WAS 서버- DB의 Active-Standby 환경 |
| 요구 사항 출처 | | |

5-3. 인터페이스 요구사항

| | | |
|-------------|-------|---|
| 요구 사항 분류 | | 인터페이스 요구 사항 |
| 요구 사항 고유 번호 | | INR-001 |
| 요구 사항 명칭 | | 접속 프로토콜 변환 |
| 요구 사항 상세 | 정의 | 프로토콜 변환 |
| | 세부 내용 | • HTTP 프로토콜을 HTTPS로 변환하여 접속 |
| | 제안 내용 | • IaC를 통해 HTTP 프로토콜을 HTTPS로 변환하여 Web 서버에 접속하도록 구축 |
| 요구 사항 출처 | | |

6. 프로젝트 관리

6-1. 프로젝트 관리 방법

1) 일정 관리

- WBS 및 간트 차트 구성 후 해당 일정 관리
- 팀 공통과 개인 구축 부분은 별도로 분리하여 일정 관리

| | |
|---------|---|
| 팀 공통 구축 | • 데이터 분석 파이프라인 |
| 개인 구축 | • 웹 애플리케이션(3-Tier) • CI/CD • 모니터링 • 백업 |

2) 투입 인력에 대한 일반 요건

- 본 사업에 투입되는 인력은 시스템의 안정적 지원을 위해 계약기간 동안 근무 가능한 인력 투입
- 투입 인력 수에 변동이 생길 시 WBS 및 간트 차트 재수정 후 계약 기간까지 모든 인 프라가 구축할 수 있도록 역할 재분배

3) 수시 보고

- 긴급 변경사항 및 특이사항 발생 시 해당 담당자에게 빠른 수시 보고 필요

4) 사업 수행 성과물

- 계약 기간 안에 팀 공통 관련 부분은 모든 팀원들이 결과물로 도출

5) 문서 관리 방법

- 팀 전체가 문서를 공유 및 관리할 수 있도록 체계적인 공유, 관리 환경 구축

6-2. 일정 계획

1) 팀 공통 일정

| 진행 단계 | 수행내용 | 상세수행내용 | 담당 | 기간 |
|-------------------|--------------------|--|------|-------------|
| 1. 사전작업 | 프로젝트 투입 | 데이터 분석 대시보드 구성 수집 데이터 종류 파악 기존의 데이터 확인 R&R 설정 | 팀 공통 | 05.13~05.27 |
| | 사전 테스트 | DB) Glue를 사용한 RDS 데이터 추출 | 이혜성 | 05.26~05.27 |
| | | 설문) 엑셀 데이터 임시 수집 | 나지원 | 05.17~05.17 |
| | | 설문) 엑셀 데이터 전처리 | | |
| | | 설문) 전처리된 데이터 분석 | | |
| | | 클릭) 클릭 스트림 데이터 수집 | 김건호 | 05.20~05.21 |
| 2. 데이터 수집 | 데이터 추출 | DB) Glue를 사용한 RDS 데이터 추출 | 이혜성 | 05.26~05.27 |
| | | 설문) 기존의 설문 엑셀 데이터 수집 | 나지원 | 05.21~05.21 |
| | | 클릭) 데이터 수집 | 김건호 | 05.21~05.21 |
| | 데이터 적제 | Data Lake 구축 | 김건호 | 05.21~05.21 |
| 3. 데이터 품질검사 | 데이터 품질 검사 규칙 설정 | DB) 추출된 DB 데이터 품질 검사를 위 한 규칙 설정 설문) 추출된 설문조사 데이터 품질 검 사를 위한 규칙 설정 | 나지원 | 05.28~05.29 |

| | | | | |
|---------------------------|--------------------|---|------------|-------------|
| | 데이터 품질 검사 정제 구축 | DB) 추출된 DB 데이터 품질 검사 구축 설문) 추출된 설문조사 데이터 품질 검사 구축 | 나지원 | 05.30~05.31 |
| 4. 데이터 정제 | 데이터 정제 | DB) 데이터 전처리 및 parquet 형태로 결과 데이터 파티셔닝 DB) 데이터 품질 검사 규칙 설정 | 이혜성 | 05.28~05.30 |
| | | 설문) 엑셀 데이터 파싱 후 CSV 변환 설문) 엑셀 데이터 분기별 파티셔닝 및 전처리 | 나지원 이혜성 | 05.21~06.01 |
| | | 클릭) 데이터 파티셔닝 및 전처리 | 이혜성 | 06.03~06.04 |
| 5. 데이터분석 및 시각화 | 데이터 분석 | SQL 작성 Athena에서 SQL 테스트 SQL 결과 저장 | 김건호 | 05.22~05.31 |
| | 데이터 시각화 | SQL 결과 기반 데이터 시각화 시각화된 결과 대시보드 구축 | 김민경 | 05.23~05.29 |
| 6. 데이터 파이프라인 자동화 | 데이터 파이프라인 자동화 | Step Function 혹은 Glue WorkFlow를 사 용한 자동화 | 나지원 이혜성 | 06.03~06.07 |
| 7. 데이터 파이프라인 알람 | 알람 지표 구성 | 데이터 파이프라인 알람 구성 데이터 파이프라인 알람 구축 | 나지원 | 05.22~05.24 |

2) 개인 일정

| 진행 단계 | 수행내용 | 상세수행내용 | 담당 | 기간 |
|----------|---------|---|-----|-------------|
| 1. 분석 | 프로젝트 투입 | 구축요구조건식별 백업 요구조건식별 CI/CD 요구조건 식별 데이터 분석 요구조건 식별 모니터링 및 알람 | 김민경 | 04.24~05.15 |

| | | | | | |
|-------|-----------------|-----------|--|------|-------------|
| 2. 설계 | 아키텍처 | | AWS Resource 구성안 CI/CD 구성안 데이터 분석 파이프라인 구성안 모니터링 및 알람 구성안 백업 구성안 | 김민경 | 04.26~05.15 |
| 3. 구축 | 네트워크 | | VPC 구축 서브넷 구축 인터넷 게이트웨이 구축 EIP 생성 Nat 게이트웨이 구축 보안 그룹 구축 | 김민경 | 05.02~05.06 |
| | 서버 | | Web 서버 구축 Web ASG 구축 Web ALB 구축 Web TG 구축 App 서버 구축 App ASG 구축 App ALB 구축 App TG 구축 | 김민경 | 05.02~05.10 |
| | CI/CD | | AWS CodeCommit 연동 AWS CodeBuild AWS Bucket 연동 AWS Code Deploy 구축 AWS Code Pipeline 구축 | 김민경 | 05.23~06.07 |
| | 도메인 및 HTTPS | | Route53을 통한 도메인 접속 HTTPS 보안 접속 | 김민경 | 05.02~05.10 |
| | 데이터 분석 파이프라인 | 데이터 추출 | f를 사용한 RDS 데이터 추출 기존의 설문 엑셀 데이터 수집 클릭 스트림 데이터 수집 | 팀 공통 | 05.27~06.07 |
| | | 데이터 품질 검사 | DB 품질 검사를 위한 규칙 설정 설문조사 품질 검사를 위한 규칙 설정 DB 품질 검사 구축 설문조사 품질 검사 구축 | 팀 공통 | 05.29~06.04 |
| | | 데이터 정제 | RDS 데이터 정제 엑셀 데이터 파싱 후 CSV 변환 로그 스트리밍 파티셔닝 | 팀 공통 | 05.15~06.11 |
| | 백업 | | 데이터베이스 백업 구축 | 김민경 | 06.10~06.14 |
| | 모니터링 | | Amazon CloudWatch 연동 | 김민경 | 06.05~06.11 |

| | | | | |
|----------|--------|--|-----|-------------|
| | | Amazon SNS 연동 Slack 연동 | | |
| 4. 모니터링 | 수행 | 3-Tier 전체 모니터링 데이터 파이프라인 모니터링 정기점검 가이드 작성 | 김민경 | 06.10~06.14 |
| 5. 운영 계획 | 보고서 작성 | 정기점검 가이드 작성 비정기 점검 가이드 작성 변경 관리 가이드 작성 백업 관리 가이드 작성 보안 관리 가이드 작성 장애 관리 가이드 작성 문제 관리 가이드 작성 최종 결과물 제출 및 발표 | 김민경 | 06.10~06.14 |

7. 프로젝트 지원

7-1. 시험 운영

1) ALB 테스트

- Web, Was 서버의 ALB 테스트 진행
- 서버 별로 다른 index.html 작성 후 서버 접속 시 여러 가용영역에 부하 분산되어 접속되는지 테스트

2) CI/CD 테스트

- 서버 배포 파이프라인이 제대로 구축되었는지 테스트
- 시각적으로 한 눈에 파악할 수 있게끔 Web, WAS 서버의 코드 업데이트 후 배포하는 방식으로 배포 테스트 진행
- 배포 작업 후 업데이트된 서버가 작동되는지 확인

3) 모니터링 및 알람 테스트

- Web, WAS 서버에 직접 접속 후 stress 기능을 사용하여 부하도 증가
- CPU 사용률이 75% 이상일 시 경보 발생 및 알람 전달 여부 확인

4) 백업 테스트

- RDS 스냅샷이 자동으로 생성되어 S3 Standard에 저장되는지 확인
- 실제로는 새벽 2시에 백업이 진행되어야 하지만, 테스트 시에는 테스트하는 시간에 맞게 Cron을 설정하여 스냅샷이 자동으로 생성되는지 테스트

7-2. 교육 훈련

- 모든 인프라의 각 세부사항에 대한 내용은 문서화하여 고객사에게 전달
- 인프라 환경 별 운영 과정은 고객사 미팅을 통해 문서화한 내용을 토대로 담당자에게 전달 및 교육

7-3. 하자 보수

- 프로젝트 계약 기간 만료 후 2년 내에 문제 발생 시 협의 후 프로젝트 연장 및 하자 보수 가능
- 프로젝트 계약 기간 만료 후에도 원활히 서비스가 작동하고 있는지 확인할 수 있도록 시스템 점검 보고서 제공

7-4. 기밀 보안

- 프로젝트 진행 시 쇼핑몰 데이터베이스의 데이터 유출 금지

- 허가된 사용자만 IAM 역할을 통해 서버 배포와 코드 수정 및 관리 가능
- 허가된 사용자는 MFA를 생성하여 이중 로그인 기능 사용

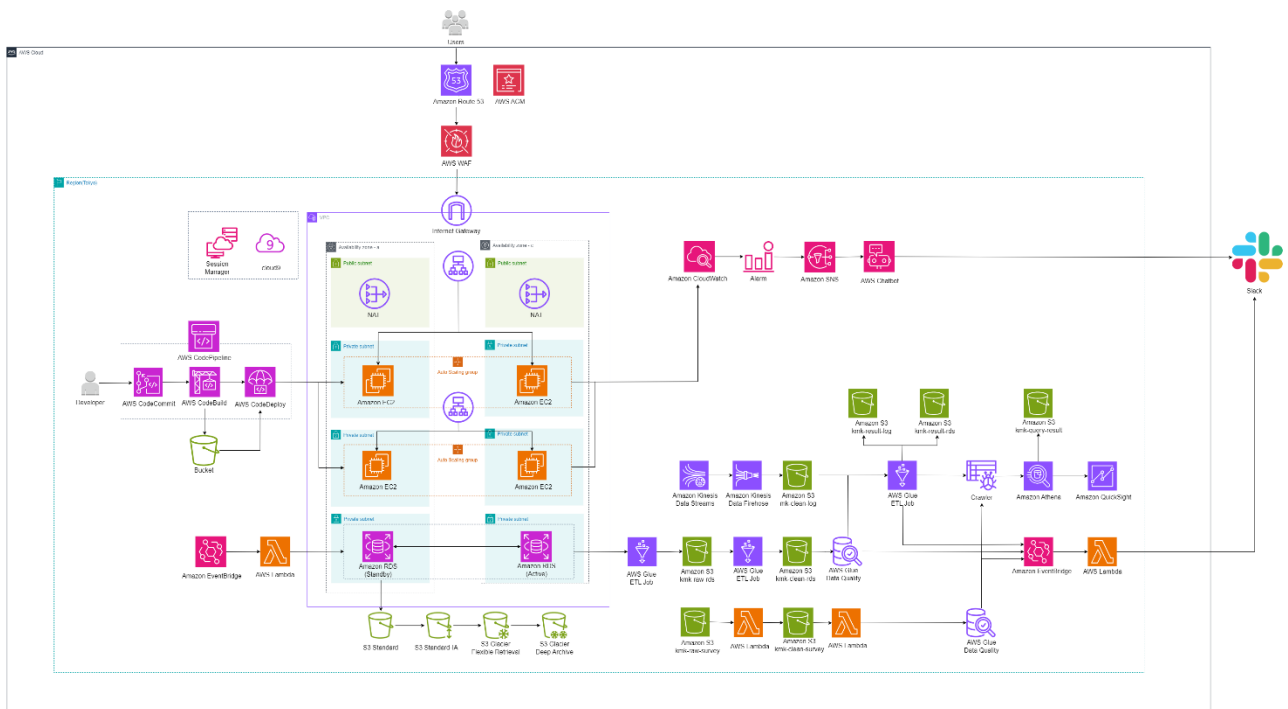
7-5. 비상 대책

- 직원의 실수로 데이터베이스 삭제 시 전날 생성된 데이터베이스의 스냅샷으로 복구
- 장애 발생에 대비하여 고객에게 지속적인 서비스를 제공할 수 있는 환경 구축하기(이중화, 멀티 리전, 백업 등)

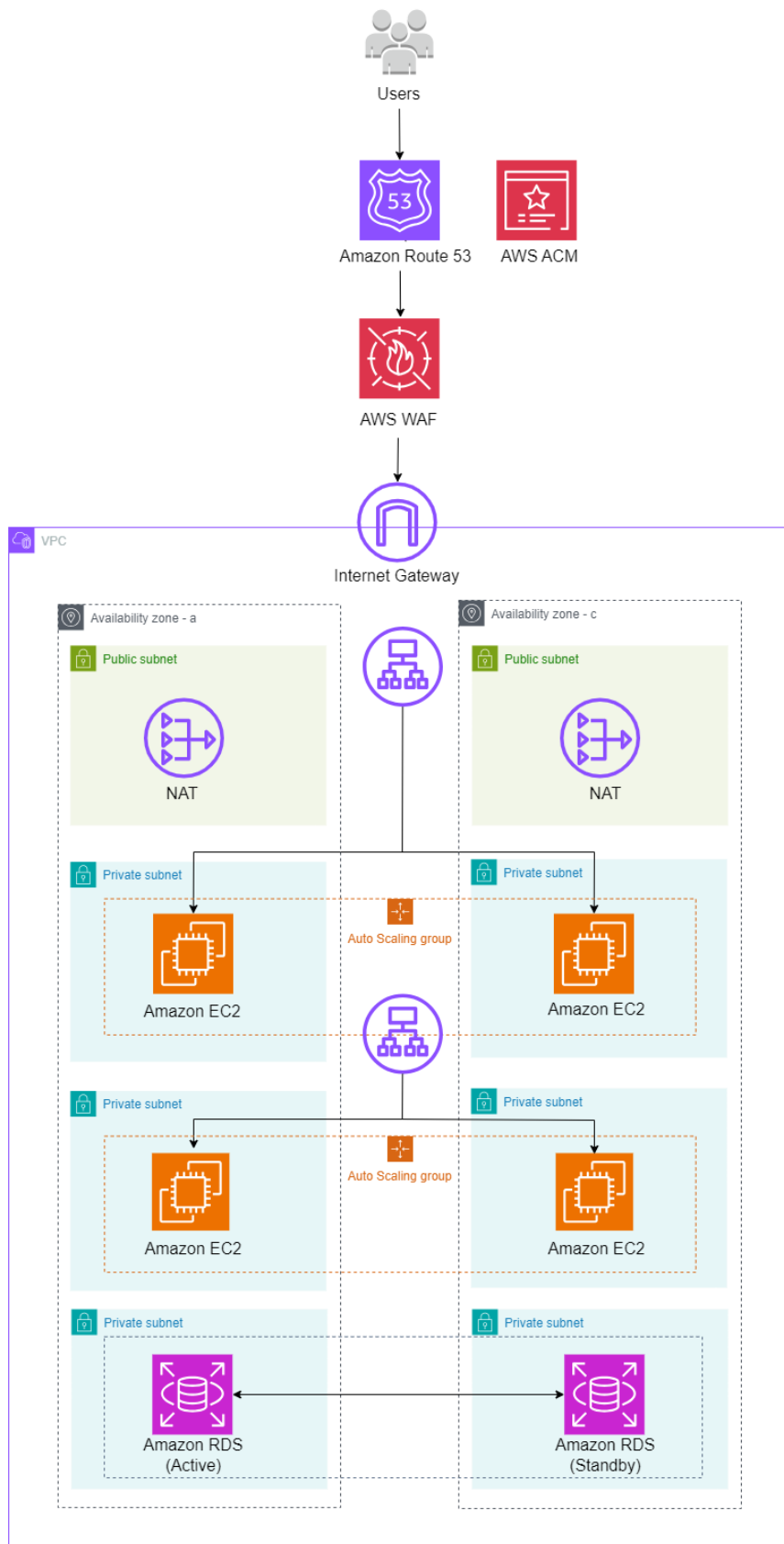
8. 구성 예상도

8-1. 예상도

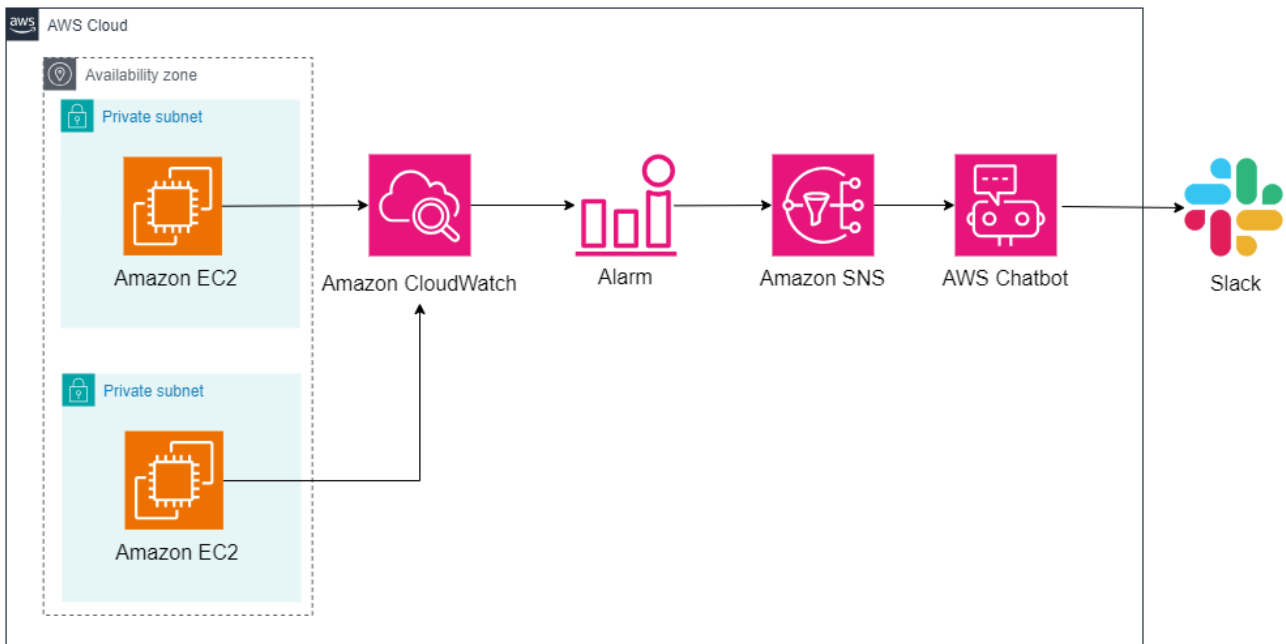
1) 전체 예상도



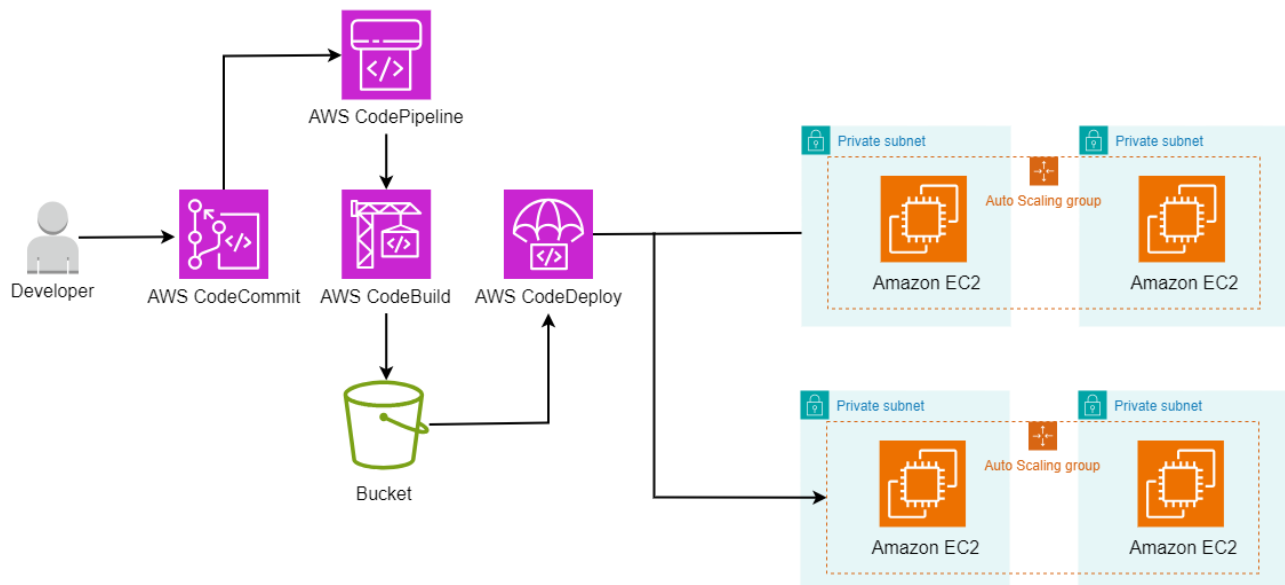
2) 웹 애플리케이션(3-Tier)



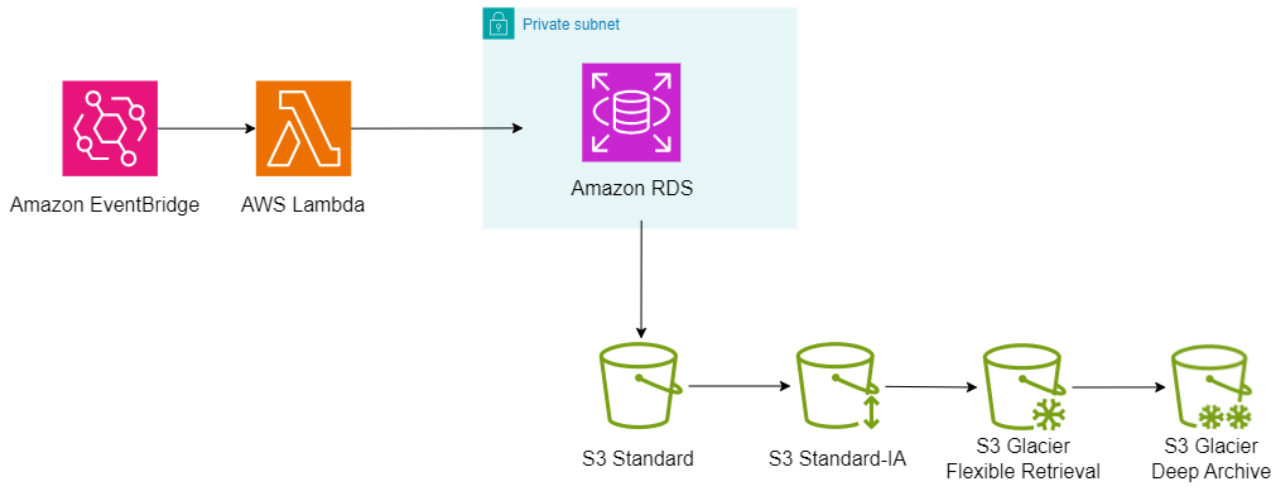
3) 모니터링



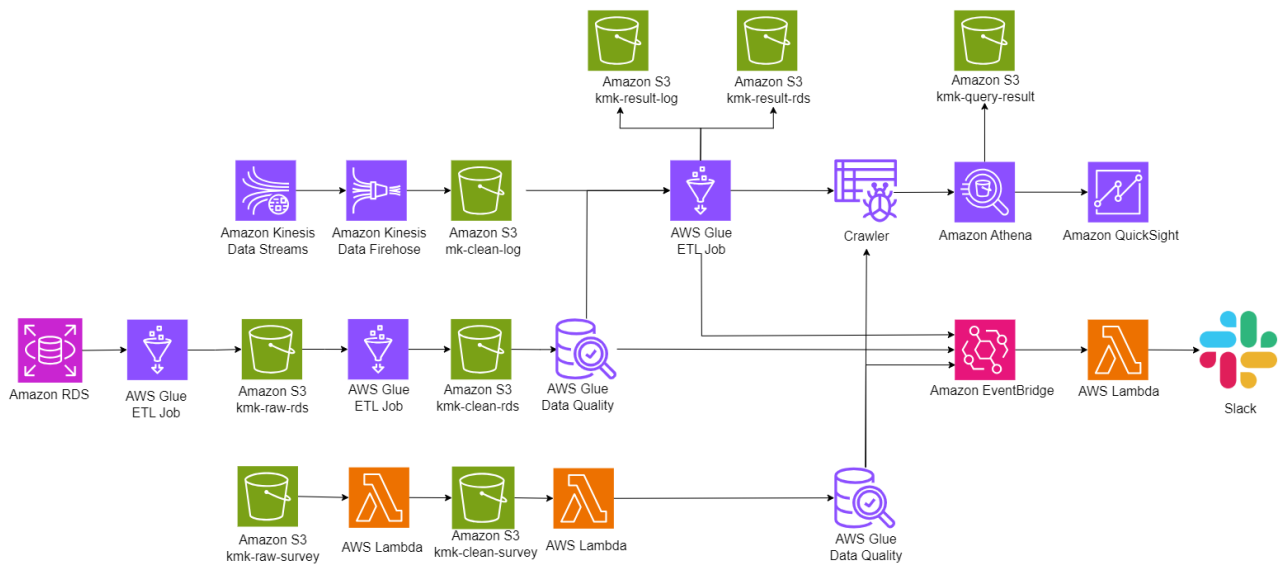
4) CI/CD



5) 백업



6) 데이터 분석 파이프라인



9. 기대 효과

9-1. 기대 효과

1. 안정적 서비스 제공

- 트래픽 분산으로 안정적, 지속적 서비스를 제공하여 고가용성 확보

2. 자동적인 서버 배포

- 업데이트 및 수정된 Web, WAS 서버의 빠른 적용 가능
- 서버 배포 자동화 가능

3. 효율적 마케팅 전략 수립

- 체계적인 데이터 수집, 정제, 분석 및 시각화로 생성된 결과를 통해 효율적인 마케팅 전략 설계 가능

4. 자동 백업 및 비용 효율적인 데이터 장기 보관

- 매일 새벽 2시에 자동적인 데이터베이스 증분 백업 가능
- 일본 법적 규정사항(전자상거래법), 법적 분쟁 대응, 고객 대응을 목적으로 비용 효율적으로 데이터 장기 보관 가능
- 일정 기간이 지나면 자동적으로 수명주기 정책을 통해 S3 스토리지 클래스 변환으로 비용 절감 가능
- 5년이 지나면 수명주기 정책을 통해 자동적으로 데이터가 삭제되어 효율적으로 비용 및 스토리지 용량 확보가능

10. 예상 비용

10-1. 네트워크

- 총 비용: 약 \$ 113.88 +a

| 서비스 | 월 비용(약) | 내용 |
|---------------------------------|-------------|--|
| Amazon Route 53 | \$ 0.90 +a | <ul style="list-style-type: none">• 호스팅 영역 1개: 약 \$0.50• 표준 쿼리 100만개: \$ 0.40 |
| NAT Gateway | \$ 95.48 +a | <ul style="list-style-type: none">• NAT 게이트웨이당 요금: \$ 0.062• 24시간 구동• 처리된 데이터 GB당 요금: \$ 0.062• 처리된 데이터 \$ 0.31 (100GB 시) |
| Application Load Balancer (ALB) | \$ 17.496 | <ul style="list-style-type: none">• 실행 시간: 24시간• 시간당 \$ 0.0243 |

10-2. 컴퓨팅

- 총 비용: 약 \$ 659.52

| 서비스 | 월 비용(약) | 내용 |
|--------|-----------|--|
| Web 서버 | \$ 161.28 | <ul style="list-style-type: none">• 인스턴스 유형: m5a.large• 인스턴스 개수: 기본 2개• 시간당 \$ 0.112 |
| WAS 서버 | \$ 161.28 | <ul style="list-style-type: none">• 인스턴스 유형: m5a.large• 인스턴스 개수: 기본 2개• 시간당 \$ 0.112 |
| RDS | \$ 336.96 | <ul style="list-style-type: none">• db.m5.large• 다중 AZ 배포• 시간당 \$ 0.47 |

10-3. CI/CD

- 총 비용: 약 \$ 5.25 + a

| 서비스 | 월 비용(약) | 내용 |
|------------------|---------|---|
| AWS CodeCommit | \$ 0 | <ul style="list-style-type: none">• 5명까지 무료<ul style="list-style-type: none">- 현재 AWS CodeCommit 사용하는 개발자 5명 |
| AWS CodeBuild | \$ 5 | <ul style="list-style-type: none">• 빌드 개수 x 빌드 기간(분 단위)<ul style="list-style-type: none">- 100개 빌드 x 5분 시)• 컴퓨팅 유형: general1.medium (분당 \$ 0.01) |
| AWS CodeDeploy | \$ 0 | <ul style="list-style-type: none">• 추가 비용이 부과되지 않음 |
| AWS CodePipeline | \$ 0.25 | <ul style="list-style-type: none">• V2 유형 파이프라인<ul style="list-style-type: none">: 작업 실행 분당 0.002 USD가 부과• 파이프라인 실행 수 x 파이프라인 작업 수 x 각 작업 실행 기간<ul style="list-style-type: none">- 파이프라인 실행 수 5번- 파이프라인 작업 수 5번- 각 작업 실행 기간 5분 시 |

10-4. 모니터링

- 총 비용: 약 \$ 16 + a

| 서비스 | 월 비용(약) | 내용 |
|-------------------------|---------|---|
| Amazon CloudWatch | \$ 15 | <ul style="list-style-type: none">• 지표: 월 \$0.30• 데이터 수집: GB 당 \$ 0.50<ul style="list-style-type: none">- 수집 데이터 100GB 시 |
| Amazon CloudWatch Alarm | \$ 0.5 | <ul style="list-style-type: none">• 경고 지표당 \$ 0.10<ul style="list-style-type: none">- 경고 5개 시 |
| Amazon SNS | \$ 0.5 | <ul style="list-style-type: none">• 모바일 푸시 알림• 알림 1백 만개당 \$ 0.50 |

| | | |
|-------------|------|--------------|
| | | - 알림 1백 만개 시 |
| AWS Chatbot | \$ 0 | • 추가 비용 없음 |

10-5. 데이터 분석

• 총 비용: 약 \$ 564.02 + a

| 서비스 | 월 비용(약) | 내용 |
|------------------------------|-----------|--|
| Amazon Kinesis Data Streams | \$ 121.68 | <ul style="list-style-type: none"> 스트림당 시간당 요금: \$0.052 <ul style="list-style-type: none"> 매일 10GB의 데이터 처리 하루 동안 2개의 shard 사용 GB당 수집된 데이터 요금: \$0.104 <ul style="list-style-type: none"> 일일 데이터 처리량: 10GB GB당 데이터 검색 요금: \$0.052 <ul style="list-style-type: none"> 일일 데이터 처리량: 10GB |
| Amazon Kinesis Data Firehose | \$ 27.57 | <ul style="list-style-type: none"> 처리된 데이터(GB당): \$ 0.0919 <ul style="list-style-type: none"> KDS를 소스로 사용할 시 일별 10GB의 데이터 S3에 적재 |
| Amazon S3 | \$ 12.5 | <ul style="list-style-type: none"> 스토리지 요금(GB당): \$ 0.025 <ul style="list-style-type: none"> S3 Standard 사용 월별 저장 데이터량: 500GB |
| AWS Glue ETL Job | \$ 39.6 | <ul style="list-style-type: none"> DPU 시간당 \$0.44의 비용이 발생 <ul style="list-style-type: none"> 일일 1시간 동안 3 DPU 사용 시 |
| AWS Glue Data Quality | \$ 39.6 | <ul style="list-style-type: none"> DPU 시간당 \$0.44의 비용이 발생 <ul style="list-style-type: none"> 일일 1시간 동안 3 DPU 사용 시 |
| AWS Lambda | \$ 1.67 | <ul style="list-style-type: none"> 실행 시간과 메모리 할당에 따른 비용 발생 1백만 요청당 \$0.20 <ul style="list-style-type: none"> 요청 1백만 건 시 컴퓨팅 요금: GB-초당 \$0.0000166667 <ul style="list-style-type: none"> 평균 실행시간: 100ms 시 |
| AWS Glue Crawler | \$ 4.40 | <ul style="list-style-type: none"> DPU 시간당 \$0.44의 비용이 발생 <ul style="list-style-type: none"> 월간 10 시간 시 |

| | | |
|---------------------------|--------|--|
| Amazon Athena | \$ 250 | <ul style="list-style-type: none"> 스캔한 데이터의 TB당 5.00 USD 데이터 압축 및 파티셔닝 - 월간 데이터 100GB 스캔 시 |
| Amazon Quicksight | \$ 67 | <ul style="list-style-type: none"> Author 당(월): \$ 24 - 2명 시 SPICE 사용 시(GB/월): \$ 0.38 - 월간 50GB 시 |
| Amazon EventBridge | \$0 | <ul style="list-style-type: none"> 500,000 이벤트 (첫 100만 이벤트 무료) |

10-6. 백업

- 총 비용: 약 \$ 20.10 + a

| 서비스 | 월 비용(약) | 내용 |
|---------------------------|---------|---|
| Amazon EventBridge | \$ 0 | <ul style="list-style-type: none"> 매월 1,400만 건의 호출 무료 1,400만 건 초과 시 예약 기반 호출 100만건 당 월 \$1.25 |
| AWS Lambda | \$ 9 | <ul style="list-style-type: none"> 매일 30분, 30일 동안 실행(54,000초) 매일 10GB 메모리로 람다 실행 GB-초당 \$0.0000166667 |
| Amazon RDS | \$ 3.6 | <ul style="list-style-type: none"> 스냅샷 생성: 비용 x 스냅샷 내보내기 : 스냅샷 크기 GB 당 \$ 0.012 - 매일 10GB로 내보낼 시 (실제로는 증분 백업이기 때문에 매일 메모리 양에 따라 변경됨) |
| Amazon S3 | \$ 7.5 | <ul style="list-style-type: none"> S3 Standard 스토리지: 약 \$7.5 (GB당 \$ 0.025 (50TB 이하 시)) 300GB 데이터 저장 예정 |

10-7. 보안

• 총 비용: 약 \$ 6.06

| 서비스 | 월 비용(약) | 내용 |
|---------|---------|---|
| AWS WAF | \$ 6.06 | <ul style="list-style-type: none">• 웹 ACL 요금(월): \$ 5.00• 규칙(월): \$ 1.00<ul style="list-style-type: none">- 규칙 1개 시• 요청: 1백만건당 \$ 0.060 |

10-8. 총 비용(월)

| | |
|--------|---------------|
| 네트워크 | \$ 113.876 |
| 컴퓨팅 | \$ 659.52 |
| CI/CD | \$ 5.25 |
| 모니터링 | \$ 16 |
| 데이터 분석 | \$ 564.02 |
| 백업 | \$ 20.10 |
| 보안 | \$ 6.06 |
| 총비용 | 약 \$ 1,377.63 |

