

완 료 보 고 서

사 업 명	쇼핑몰 사이트의 인프라 구축 및 데이터 분석
주 관 사	Find Customers
담 당 자	김 건 호

2024. 06. 23.

내용

1.	사업 개요	4
1-1.	개요	4
1-2.	배경 및 필요성	4
1-3.	사업 범위	6
2.	추진 전략	8
2-1.	추진 목표	8
3.	적용 기술	8
3-1.	적용 기술	8
4.	기술 및 기능	14
4-1.	시스템 요구 사항	14
4-2.	기능 요구 사항	19
4-3.	보안 요구 사항	20
4-4.	데이터 요구 사항	21
4-5.	시스템 운영 요구 사항	22
4-6.	제약 사항	22
5.	성능 및 품질	23
5-1.	성능 요구 사항	23
5-2.	품질 및 인터페이스 요구사항	23
6.	프로젝트 관리	24
6-1.	프로젝트 관리 방법	24
6-2.	일정 계획	24
7.	프로젝트 지원	25
8.	구성 예상도	26
8-1.	예상도	26
9.	기대 효과	30

9-1. 기대 효과	30
10. 예상 비용	31
10-1. 컴퓨팅	31
월간 비용	31
11. 완료 보고	33
11.1. 3tier	33
11.2. CI/CD	37
11.3. 모니터링	39
11.4. 백업	42
11.5. 보안	44
11.6. 데이터 파이프라인	46

1. 사업 개요

1-1. 개요

사업명: xx 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. 북미를 중심 웹 쇼핑몰
2. 아메리카 대륙 및 유럽 등 전세계적으로 사용자 증가
3. 고객들의 다양한 데이터 수집
4. 쇼핑몰 인프라 AWS 이전 계획

필요성

1 쇼핑몰 사용자에게 지속적인 서비스 제공

- 사용자에게 지속적인 서비스를 제공함에 따라 신뢰도 향상
- 다양한 국가에서 고객이 증가하고 있으므로 트래픽이 변동이 이전보다 다양해질 것으로 예상
- 따라서 변화하는 트래픽을 유연하게 처리할 수 있는 인프라 환경 필요
- 쇼핑몰 특성 상 사진과 같은 대용량의 콘텐츠를 효율적으로 제공

2 인프라 안정성 확보

- 인터페이스와 로직이 하나의 계층에 있는 환경
- 속도 저하 및 원인 파악, 유지보수에 어려움이 있음
- 계층을 분리하여 모듈화 진행

- 데이터 손실 최소화 복구 및 손실 시 필요한 시점으로 복원 필요

3 장애 대응 필요

- 서버 상태를 확인하는 모니터링 필요
- 장애 발생시 담당자의 대처를 위해 알림 제공
- 외부 웹 공격 예방 규칙

4 배포 환경 개선 필요

- 빠른 애플리케이션 배포 프로세스 필요
- 서비스의 빠른 개선을 이루고자 함
- 빠른 배포 및 개선을 위해 서비스와 인프라의 종속성을 최소화 시키는 환경 필요

5 데이터 분석을 통해 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

5.1 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

5.2 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

5.3 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

1-3. 사업 범위

클라우드 컴퓨팅 서비스 구축 및 인프라 고도화

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

모니터링 구축

- 서버의 성능 메트릭을 실시간으로 수집
- 상태를 지속적으로 모니터링하고 경보를 설정하여 이상 감지
- 발생한 경보 알림을 관리자에게 전달하여 장애 대응

CI/CD 구축

- 코드 관리 및 버전 관리 툴 필요
- 지속적 코드 통합, 지속적 배포 파이프라인 구성 필요
- 코드의 통합, 테스트 및 배포 과정을 자동화
- CI/CD 과정 모니터링 수행 및 알림 필요

백업 시스템 구축

- 데이터베이스 백업으로 중요 데이터 보호

컨테이너 환경 구축

- 애플리케이션의 관리를 위한 오케스트레이션 툴 사용
- 컨테이너 이미지 저장 공간 필요

데이터 파이프라인 구축

- 인프라 환경에서 발생하는 데이터를 수집, 저장, 분석을 할 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

- 담당자가 설문조사를 직접 데이터 분석을 위한 저장소에 업로드
- 파싱 기능이 필요할 경우 자동화하여 파싱된 데이터 저장
- 사용자의 로그를 실시간으로 저장
- 특정한 시간 대에 저장된 로그를 분석해 시각화하는 인프라 구축

로그 데이터

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

보안 시스템 구축

- 특정 웹 트래픽을 필터링하고 웹 공격을 차단
- 웹사이트/서버 및 브라우저 간 암호화된 연결 설정

2. 추진 전략

2-1. 추진 목표

사업을 효과적으로 수행하기 위한 추진 전략(위험 요소를 고려하여 창의적이고 타당한 대안)을 제시

- AWS Well Architected 의 원칙을 최대한 적용한 인프라 설계
- 다중 AZ 배포와 복구 전략을 통해 높은 가용성 확보
- 트래픽에 따른 자원을 자동으로 조절하여 성능 효율성 인프라 설계
- 분석 결과를 기반으로 고객 맞춤형 시스템을 구축 및 타겟 마케팅과 프로모션을 최적화
- 사용자 행동 데이터를 통한 웹서비스 개선을 추진하여 고객 만족도와 사이트의 사용성을 향상시킴
- 감사 로그와 정기적인 보안 검토를 수행하여 보안 수준을 지속적으로 관리하고 개선
- 규제 준수와 데이터 보안을 위해 모든 데이터 처리 과정에서의 규정을 철저히 준수

3. 적용 기술

3-1. 적용 기술

1. 3tier

1.1. Virtual Private Cloud

- 정의한 가상 네트워크에서 리소스 사용
- 2 개의 가용영역 구성(us-west-1a, us-west-1b)

1.2. Public Subnet

- 인터넷 게이트웨이를 연결해 AWS 리소스가 인터넷과 통신할 수 있게 함

1.3. Private Subnet

- 가용영역 2 개에 Web, Was, DB 3 개의 subnet 으로 각각 분리
- 보안을 위해 인터넷에서 직접 접근할 수 없게 함

1.4. Internet Gateway

- VPC 와 인터넷 간 통신을 하게 함

1.5. NAT Gateway

- 가용영역 Public Subnet 에 각각의 NGW 구성
- Private Subnet 의 리소스가 인터넷에 접근하도록 함
- 인터넷에서는 그 리소스에 접근할 수 없게 함

1.6. Application Load Balancer

- Web, Was 2 개의 ALB 구성
- EC2 인스턴스에 자동으로 분산시켜 로드를 균등하게 관리

1.7. AutoScaling Group

- Web, Was 2 개의 ASG 구성
- 2 개의 가용 영역에 걸쳐 인스턴스를 배포하여 고가용성 보장
- 트래픽 변화에 따라 인스턴스 수를 조정하여 리소스 추가 및 제거

1.8. Relational Database Service

- 관리형 데이터베이스를 사용해 운영 작업을 간소화
- 2 개의 가용 영역에 걸쳐 배포
- active – standby 로 장애 발생시 다른 가용 영역에 있는 실패 조치를 수행

1.9. Session Manager

- 보안 관리 지점을 최소화

2. CI/CD

2.1. GitHub

- Web, Was 서버 소스 코드 저장 및 버전 관리
- 개발자가 코드 변경사항을 커밋

2.2. CodeBuild

- 소스 코드 컴파일 및 테스트 실행
- 서버리스로 관리 없이 빌드 환경을 사용

2.3. CodePipeline

- 빌드, 테스트, 배포 과정을 순차적 실행하는 워크플로우 구성
- GitHub 의 코드 변경 사항을 자동으로 감지하는 트리거로 자동화

3. 모니터링

3.1. CloudWatch

- CloudWatch Agent 로 Web, Was CPU 지표 모니터링
- 일정 수치 이상시 경보 발생

3.2. Simple Notification Service

- CloudWatch 에서 발생한 경보에 알람을 생성

3.3. Chatbot

- SNS 에서 발생한 알람을 Slack 으로 전달
- 복잡함 스크립트나 중개자 없이 사용자에게 전달

3.4. Slack

- 서버 이상 경고 알림 생성
- CodePipeline 성공 및 실패 알림 생성

4. 컨테이너

4.1. Elastic Container Service

- 컨테이너를 쉽게 배포, 관리, 확장하기 위해 오케스트레이션 툴 사용
- 효과적인 CI/CD 구성을 위해 EC2 Launch Type 사용으로 인스턴스에 대한 완전한 제어권을 가짐
- 개발 및 운영의 복잡성을 줄이고 소프트웨어 전달의 속도를 높임

4.2. Elastic Container Registry

- 컨테이너 이미지를 저장, 관리 및 배포하기 위한 완전 관리형 컨테이너 레지스트리 사용
- IAM 을 사용하여 사용자 및 서비스별 액세스 권한을 세밀하게 제어
- 안전한 이미지 관리와 액세스를 보장

5. 백업

5.1. AWS Backup

- RDS 백업 규칙 생성
- 매일 새벽 2 시에 PITR 방식으로 백업 수행
- 보관 주기는 7 일

6. 데이터 파이프라인

6.1. AWS S3

- 객체 데이터 저장 서비스로 다양한 형태의 데이터 저장 가능
- 거의 무한에 가까운 저장 용량으로 인해 빅 데이터 저장에 편리

- S3의 KEY값을 Value 형식으로 설정해 AWS Glue 크롤링 시 자동으로 파티셔닝 지정 가능

6.2. AWS Lambda

- 코드를 실행하는 서버리스 컴퓨터
- 엑셀 데이터 파싱 후 패키지 변환, Slack 알람 등 일시적인 작업을 진행할 때 사용
- 최대 15분만 진행하기 때문에, 과도한 데이터 처리 작업이 필요한 것들은 AWS Glue 등의 서비스를 사용

6.3. AWS Glue ETL Job

- 데이터를 전처리하고 변환하는 데 사용하는 작업
- RDS에서 특정 테이블의 데이터를 추출가능
- 불필요한 데이터 제거로 데이터 처리의 효율성 향상
- 서로 다른 테이블을 합쳐 분석을 위한 하나의 데이터를 만들때도 사용
- Data Quality 기능을 내장하고 있어 분석 전(혹은 전처리 후)의 데이터가 올바른지 확인 가능

6.4. AWS Glue Crawler

- 분석을 위한 데이터의 메타 데이터를 추출하기 위함

6.5. AWS Glue WorkFlow

- 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

6.6. AWS Kinesis Firehose, Data Stream

- 실시간성 데이터를 수집하기 위함

- 동적 파티셔닝을 활용해 추가적인 전처리 기능을 생략하게 함

6.7. AWS Athena

- 추출한 데이터를 기반으로 쿼리를 날려 분석 결과를 확인

6.8. AWS QuickSight

- 아테나로 쿼리를 한 결과를 시각화
- Spice 방식을 활용해 쿼리 결과를 일부분 저장해 비용을 감소

7. 보안

7.1. Route 53

- 사용자 도메인 관리 및 인터넷 트래픽을 ALB 로 라우팅
- HTTPS 를 통한 안전한 연결 위해 DNS 설정 및 SSL/TLS 인증서와 연동

7.2. AWS Certificate Manager

- 도메인 이름에 대한 SSL/TLS 인증서 발급
- 사용자 도메인에 적용하여 HTTPS 를 통한 보안 연결을 활성화
- 데이터를 안전하게 암호화하여 서버 간에 전송되도록 보장

7.3. AWS WAF

- 웹 애플리케이션을 보호하기 위해 트래픽 규칙 설정
- SQL injection 공격으로부터 보호

4. 기술 및 기능

4-1. 시스템 요구 사항

요구 사항 분류		시스템 요구 사항
요구 사항 고유 번호		SYS-001
요구 사항 명칭		가용성
요구 사항 상세	정의	시스템은 고가용성을 유지해야 하며, 모든 서비스는 최소한 두 개의 가용 영역에 분산 배치되어야 한다
	세부 내용	<ul style="list-style-type: none">- 모든 리소스는 ap-northeast-2 지역의 두 개 가용 영역(ap-northeast-2a, ap-northeast-2b)에 배포- 데이터베이스 이중화로 가용성 보장
	제안 내용	<ul style="list-style-type: none">- 각 가용 영역에 최소 하나 이상의 인스턴스 배포- Auto Scaling Group을 사용하여 가용 영역 간에 인스턴스를 균등하게 분산- Load Balancer를 사용하여 트래픽 분산
요구 사항 출처		

요구 사항 분류		시스템 요구 사항
요구 사항 고유 번호		SYS-002
요구 사항 명칭		Web 서버시스템
요구 사항 상세	정의	AWS 시스템 Web 서버
	세부 내용	<p>서버 고려 사항</p> <ul style="list-style-type: none">- 컨테이너화된 애플리케이션을 호스팅할 수 있도록 ECS EC2 인스턴스 사용- 대규모 사용자를 수용하기 위한 충분한 계산 능력과 메모리 제공- 높은 동시성을 지원하기 위한 네트워크 대역폭 필요- 탄력적인 확장을 위해 Auto Scaling 설정- 관리형 데이터베이스 서비스 (RDS) 사용- 대규모 데이터를 처리하기 위한 충분한 계산 능력

		력과 메모리 제공	
		- 높은 동시성을 지원하기 위한 네트워크 대역폭 필요	
	제안 내용	WEB Server	
		구분	사양
		Type	M5.large
		vCPU	2
		Memory	8Gbps
		I/O	최대 4,750Mbps
		BW	10Gbps
		비용	시간당 0.118 USD (서울 리전)
요구 사항 출처			

요구 사항 분류		시스템 요구 사항							
요구 사항 고유 번호		SYS-003							
요구 사항 명칭		Was 서버시스템							
요구 사항 상세	정의	AWS 시스템 Was 서버							
	세부 내용	서버 고려 사항 <ul style="list-style-type: none">- 컨테이너화된 애플리케이션을 호스팅할 수 있도록 ECS EC2 인스턴스 사용- 대규모 사용자를 수용하기 위한 충분한 계산 능력과 메모리 제공- 높은 동시성을 지원하기 위한 네트워크 대역폭 필요- 탄력적인 확장을 위해 Auto Scaling 설정							
	제안 내용	<div>Was Server</div> <table><tr><th>구분</th><th>사양</th></tr><tr><td>Type</td><td>M5.large</td></tr><tr><td>vCPU</td><td>2</td></tr><tr><td>Memory</td><td>8Gbps</td></tr></table>	구분	사양	Type	M5.large	vCPU	2	Memory
구분	사양								
Type	M5.large								
vCPU	2								
Memory	8Gbps								

		I/O	최대 4,750Mbps
		BW	최대 10Gbps
		비용	시간당 0.118 USD (서울 리전)
요구 사항 출처			

요구 사항 분류		시스템 요구 사항														
요구 사항 고유 번호		SYS-004														
요구 사항 명칭		DB 서버시스템														
요구 사항 상세	정의	AWS 시스템 DB 서버														
	세부 내용	서버 고려 사항 <ul style="list-style-type: none">- 관리형 데이터베이스 서비스 (RDS) 사용- 대규모 데이터를 처리하기 위한 충분한 계산 능력과 메모리 제공- 높은 동시성을 지원하기 위한 네트워크 대역폭 필요														
	제안 내용	<div>DB Server</div> <table><tr><th>구분</th><th>사양</th></tr><tr><td>Type</td><td>db.m5.large</td></tr><tr><td>vCPU</td><td>2</td></tr><tr><td>Memory</td><td>8Gbps</td></tr><tr><td>I/O</td><td>최대 4,750Mbps</td></tr><tr><td>BW</td><td>최대 10Gbps</td></tr><tr><td>비용</td><td>시간당 0.472 USD (서울 리전)</td></tr></table>		구분	사양	Type	db.m5.large	vCPU	2	Memory	8Gbps	I/O	최대 4,750Mbps	BW	최대 10Gbps	비용
구분	사양															
Type	db.m5.large															
vCPU	2															
Memory	8Gbps															
I/O	최대 4,750Mbps															
BW	최대 10Gbps															
비용	시간당 0.472 USD (서울 리전)															
요구 사항 출처																

요구 사항 분류		시스템 요구 사항	
요구 사항 고유 번호		SYS-005	
요구 사항 명칭		컨테이너화	
요구 사항 상세	정의	애플리케이션은 컨테이너화 되어 배포, 관리 및 확장 가능해야 한다	

	세부 내용	<ul style="list-style-type: none"> - 컨테이너 오케스트레이션 ECS를 사용 - ECR을 사용하여 컨테이너 이미지를 저장 및 관리
	제안 내용	<ul style="list-style-type: none"> - ECS 클러스터 구성 및 서비스 정의 - ECR 레지스트리를 설정하여 컨테이너 이미지 저장 - ECS 서비스에서 Auto Scaling을 설정하여 태스크 수를 동적으로 조정
요구 사항 출처		

요구 사항 분류		시스템 요구 사항
요구 사항 고유 번호		SYS-006
요구 사항 명칭		CI/CD
요구 사항 상세	정의	서버 및 애플리케이션 배포는 CI/CD 파이프라인을 통해 자동화되어야 한다
	세부 내용	<ul style="list-style-type: none"> - CodePipeline을 사용하여 빌드, 테스트, 배포 과정을 자동화 - CodeBuild를 사용하여 소스 코드 컴파일 및 테스트 실행 - GitHub과 통합하여 코드 변경 사항을 자동으로 감지하고 배포
	제안 내용	<ul style="list-style-type: none"> - CodePipeline을 설정하여 각 단계별로 작업 자동화 - CodeBuild 프로젝트를 생성하여 소스 코드를 빌드 및 테스트 - GitHub 웹훅을 설정하여 코드 변경 사항이 발생할 때마다 파이프라인이 트리거되도록 구성 - Slack과 연동하여 배포 상태 알림을 실시간으로 수신 - 컨테이너 이미지를 빌드하고 ECR에 푸시하여, ECS로 배포 자동화
요구 사항 출처		

요구 사항 분류		시스템 요구 사항
요구 사항 고유 번호		SYS-007
요구 사항 명칭		데이터 파이프라인
요구 사항 상세	정의	시스템은 데이터 파이프라인을 이용하여 데이터를 효율적으로 저장, 처리 및 분석할 수 있어야 한다.
	세부 내용	<p>AWS S3</p> <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 <p>AWS Lambda</p> <ul style="list-style-type: none"> - 서버리스 환경에서 코드 실행 가능 - 엑셀 데이터 파싱, 패키지 변환, Slack 알람 등의 일시적 작업 수행에 적합 - 최대 15분까지의 실행 시간 제한으로 인해 긴 데이터 처리 작업은 AWS Glue 등의 서비스 사용 권장 <p>AWS Glue ETL Job</p> <ul style="list-style-type: none"> - 데이터 전처리 및 변환을 위한 작업 수행 - RDS에서 필요한 컬럼만 추출하여 데이터 가공 가능 - 서로 다른 테이블을 조인하여 분석을 위한 데이터 통합 가능 - 데이터 품질 관리 기능으로 데이터의 정합성 검증 가능 <p>AWS Glue Crawler</p> <ul style="list-style-type: none"> - 데이터 분석을 위한 메타 데이터 추출을 위한 서비스 - RDS에서 필요한 컬럼의 메타 데이터 추출 및 관리 가능 <p>AWS Kinesis Firehose 및 Data Stream</p>

		<ul style="list-style-type: none"> - 실시간 데이터 스트림을 수집하고 처리하는 서비스 - 동적 파티셔닝을 통해 전처리 과정 최적화 및 간소화 가능 <p>AWS Athena</p> <ul style="list-style-type: none"> - 저장된 데이터를 기반으로 SQL 쿼리를 실행하여 분석 결과 조회 가능 <p>AWS QuickSight</p> <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 효율성을 극대화 - AWS Kinesis를 통해 실시간 데이터를 수집 - AWS Athena와 QuickSight를 통해 분석 및 시각화를 지원 - 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		FUR-001
요구 사항 명칭		웹서비스 기능 요구 사항
요구 사항 상세	정의	웹서비스 기능 요구 사항
	세부 내용	1. 회원 가입

		2. 로그인 3. 판매 상품 목록 4. 판매 상품 게시 5. 주문 결제
	제안 내용	1. 회원 가입 <ul style="list-style-type: none"> • 사용자가 기본 정보를 입력하여 계정을 생성할 수 있습니다. 2. 로그인 <ul style="list-style-type: none"> • 아이디와 비밀번호를 사용하여 로그인 기능을 제공합니다. 3. 판매 상품 목록 <ul style="list-style-type: none"> • 등록된 상품을 리스트 형태로 보여주는 페이지를 구현합니다. 4. 판매 상품 게시 <ul style="list-style-type: none"> • 판매자가 상품 정보를 입력하여 게시할 수 있는 기능을 제공합니다. 5. 주문 결제 <ul style="list-style-type: none"> • 사용자가 상품을 선택하고 결제할 수 있어야 합니다.
요구 사항 출처		

4-3. 보안 요구 사항

요구 사항 분류		보안 요구 사항
요구 사항 고유 번호		SER-001
요구 사항 명칭		웹 애플리케이션 보안
요구 사항 상세	정의	웹 애플리케이션의 안전성을 확보하기 위한 보안 요구 사항을 설정

	세부 내용	Route 53 <ul style="list-style-type: none"> - 사용자 도메인 관리 및 인터넷 트래픽을 ALB로 라우팅 - HTTPS를 통한 안전한 연결을 위해 DNS 설정 및 SSL/TLS 인증서와 연동 AWS Certificate Manager (ACM) <ul style="list-style-type: none"> - 도메인 이름에 대한 SSL/TLS 인증서 발급 및 관리 - 사용자 도메인에 적용하여 HTTPS를 통한 보안 연결을 활성화 - 데이터 전송 중 데이터의 안전성을 보장하기 위한 암호화 구현 AWS WAF (Web Application Firewall) <ul style="list-style-type: none"> - 웹 애플리케이션을 보호하기 위한 정책 및 규칙 설정 - 다양한 웹 애플리케이션 공격 유형 (예: SQL Injection)으로부터 보호
	제안 내용	<ul style="list-style-type: none"> - Route 53을 사용하여 도메인 관리 및 ALB로의 트래픽 라우팅 설정 - ACM을 활용하여 SSL/TLS 인증서 발급 및 HTTPS 연결을 강화 - AWS WAF를 이용하여 웹 애플리케이션에 대한 보안 정책을 설정하고 공격으로부터 보호
요구 사항 출처		

4-4. 데이터 요구 사항

요구 사항 분류		데이터 요구 사항
요구 사항 고유 번호		DAR-001
요구 사항 명칭		데이터 백업 및 복구
요구 사항 상세	정의	데이터베이스의 백업 및 복구를 위한 자동화된 시스템
	세부 내용	- AWS Backup을 사용하여 RDS 백업 규칙 생성

		- 매일 새벽 2시에 PITR 방식으로 백업 수행
	제안 내용	<ul style="list-style-type: none"> - AWS Backup 서비스를 사용하여 백업 정책 및 규칙을 구성 - RDS 인스턴스에 대한 자동 백업 규칙을 설정하여 매일 새벽 2시에 백업 수행 - 백업 데이터는 AWS Backup Vault에 안전하게 저장 - 백업 복구 시점을 유연하게 선택할 수 있는 PITR 기능 활성화
요구 사항 출처		

4-5. 시스템 운영 요구 사항

요구 사항 분류		시스템 운영 요구 사항
요구 사항 고유 번호		MAR-001
요구 사항 명칭		시스템 모니터링 및 알림
요구 사항 상세	정의	CloudWatch와 SNS를 이용해 시스템 상태를 모니터링 하고, 이상 발생 시 즉시 알림을 제공해야 한다
	세부 내용	<ul style="list-style-type: none"> - CloudWatch Agent를 설치하여 CPU, 메모리, 네트워크 등의 지표 모니터링 - 특정 지표가 임계값을 초과할 경우 경고 생성 - SNS를 통해 경보를 Slack으로 전달
	제안 내용	<ul style="list-style-type: none"> - CloudWatch 대시보드를 통해 실시간 모니터링 - 경고 설정을 통해 문제가 발생할 경우 즉시 알림 수신 - SNS와 Chatbot을 연동하여 알림 메시지 수신
요구 사항 출처		

4-6. 제약 사항

요구 사항 분류	제약 사항 요구 사항
요구 사항 고유 번호	COR-001

요구 사항 명칭		소프트웨어 제약
요구 사항 상세	정의	소프트웨어 호환성
	세부 내용	- 시스템 호환성 유지
	제안 내용	- WEB, WAS, DB 버전 및 호환성 테스트 - CloudWatch agent 호환성 테스트
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		성능 일반
요구 사항 상세	정의	성능 일반 요구
	세부 내용	1. 안정적 운영 및 사용자 지원 방안을 제시하여야 함 2. 시스템 성능 상태를 실시간으로 모니터링 하도록 하여, 성능 문제를 파악하여야 함 3. 문제 발생 시, 즉각적인 조치가 이루어져야 함
	제안 내용	
요구 사항 출처		

5-2. 품질 및 인터페이스 요구사항

요구 사항 분류		품질 및 인터페이스 요구 사항
요구 사항 고유 번호		QUR-001
요구 사항 명칭		코드 품질 및 관리
요구 사항 상세	정의	코드 품질 및 관리 개념 정의

	세부 내용	- 코드는 저장소에 보관하여 손실 위험에 대비 및 버전 관리가 되어야 함
	제안 내용	- 리포지토리를 통하여 코드를 관리하며, 변경 사항은 Git 버전을 통한 관리
요구 사항 출처		

6. 프로젝트 관리

6-1. 프로젝트 관리 방법

1. 형상 관리

- 형상 관리를 위해 Git 버전 관리 시스템을 사용

2. 문서 관리

- 모든 문서를 저장소(Naver MyBox)에 저장하여 관리

3. 문제 발생 보고 체계

- 문제를 식별하고 적절한 대응 팀 또는 관리자에게 신속하게 보고

4. 위험 관리

- 프로젝트 시작 단계에서 모든 잠재적 위험을 식별하고 평가
- 식별된 위험에 대하여 완화 전략 개발
- 정기적으로 위험을 평가하고 완화 조치의 효과를 검토 및 필요한 조정 수행

6-2. 일정 계획

7. 프로젝트 지원

7-1. 시험 운영

1. **테스트 전략 개발:** 개발된 각 기능과 모듈에 대한 체계적인 테스트 전략을 수립합니다. 이는 유닛 테스트, 통합 테스트, 시스템 테스트 및 사용자 수용 테스트(UAT)를 포함합니다.
2. **테스트 케이스 및 시나리오 작성:** 각 기능의 요구 사항을 충족하는지 확인하기 위해 상세한 테스트 케이스와 시나리오를 작성합니다.
3. **동작 확인 절차:** 실제 운영 환경에서 시뮬레이션을 통해 시스템의 동작을 확인합니다. 문제 발생 시 신속한 해결 방안을 마련하고, 재테스트를 진행합니다.

7-2. 교육 훈련

1. **기술 문서 제공:** 시스템 운영자 및 사용자가 시스템을 효과적으로 사용할 수 있도록 기술 매뉴얼 및 가이드를 제공합니다.
2. **운영 교육 세션:** 운영자 및 사용자를 대상으로 실시하는 교육 세션을 계획하고 진행합니다. 이는 이론 교육과 실습을 포함하여 구성됩니다.
3. **FAQ 및 온라인 지원:** 자주 묻는 질문(FAQ) 목록을 제공하고, 온라인 헬프 데스크를 통한 지원을 계획합니다.

7-3. 하자 보수

1. **유지보수 계획:** 정기적인 시스템 점검과 업데이트 계획을 수립합니다. 하자 발생 시 신속히 대응할 수 있는 로드맵을 제공합니다.
2. **시스템 점검 계획서 및 보고서:** 정기적인 시스템 점검 결과를 문서화하여 보고서 형태로 제공합니다. 이는 잠재적인 문제를 사전에 예방하는 데 도움이 됩니다.

7-4. 기밀 보안

1. **보안 프로토콜 및 정책 수립:** 시스템의 보안성을 강화하기 위한 보안 프로토콜 및 정책을 개발합니다. 이는 데이터 암호화, 접근 제어, 감사 트레일 등을 포함합니다.
2. **개인정보 보호 정책:** 개인정보 보호법과 관련된 최신 규정을 준수하며, 개인정보 보호 대책을 강화합니다.
3. **보안 점검 및 감사:** 정기적인 보안 점검 및 감사를 실시하여 시스템의 취약점을 식별하고 보완합니다.

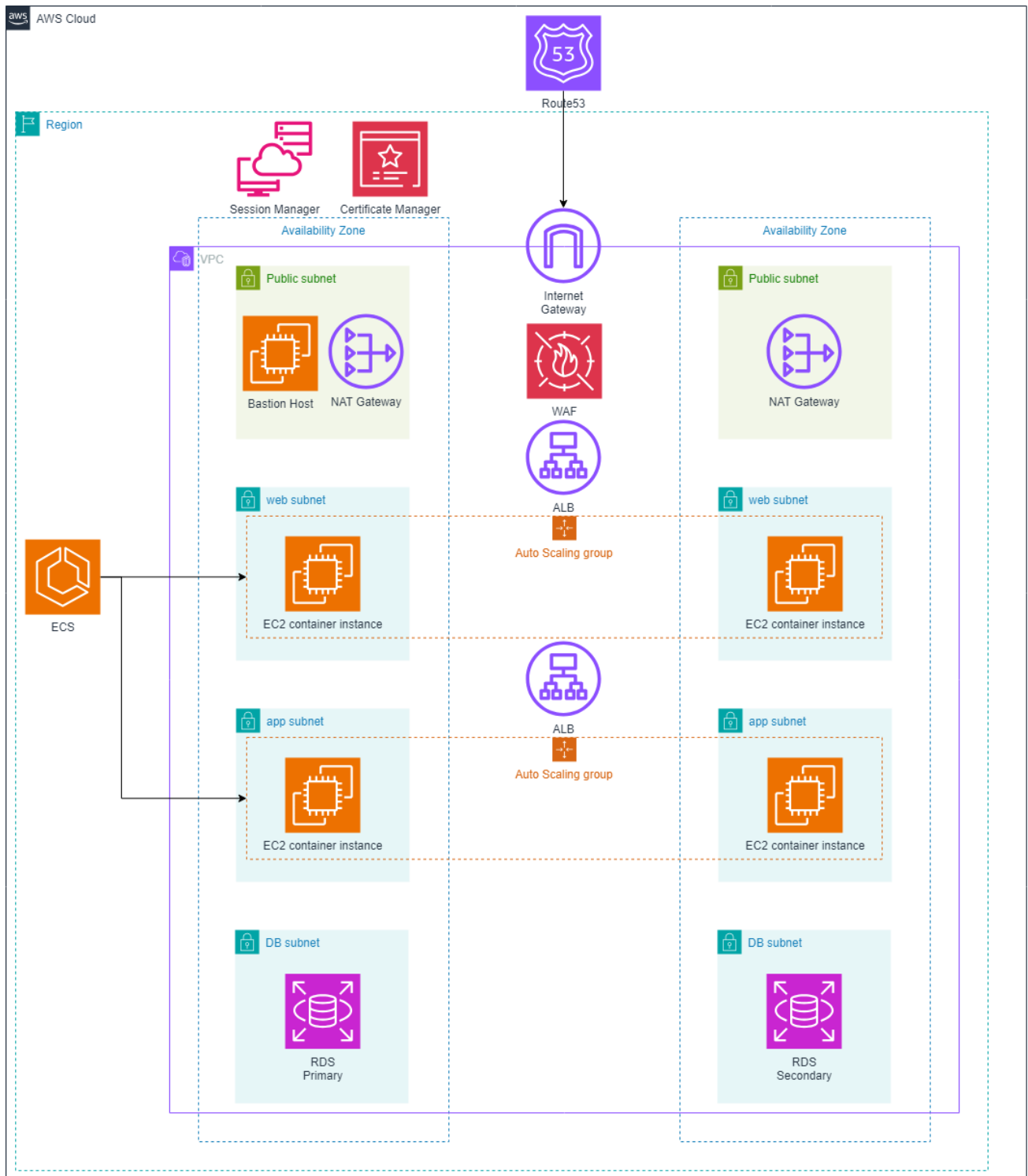
7-5. 비상 대책

1. **백업 및 복구 계획:** 데이터 손실 방지를 위한 철저한 백업 및 복구 계획을 수립합니다. 이는 데이터베이스 백업, 중요 문서, 그리고 다른 핵심적인 데이터 파일을 포함할 수 있습니다.
2. **백업 주기 설정:** 데이터의 중요도와 변경 빈도에 따라 일일, 주간, 월간 백업 주기를 설정합니다.
3. **복구 절차 시험:** 정기적으로 복구 절차를 시험하여, 실제 재해 상황에서 데이터를 신속하고 정확하게 복원할 수 있는지 확인합니다.
4. **오프사이트 백업:** 지리적으로 분산된 위치에 백업 데이터를 저장하여 자연 재해나 사이트 특정 문제로부터 보호합니다.

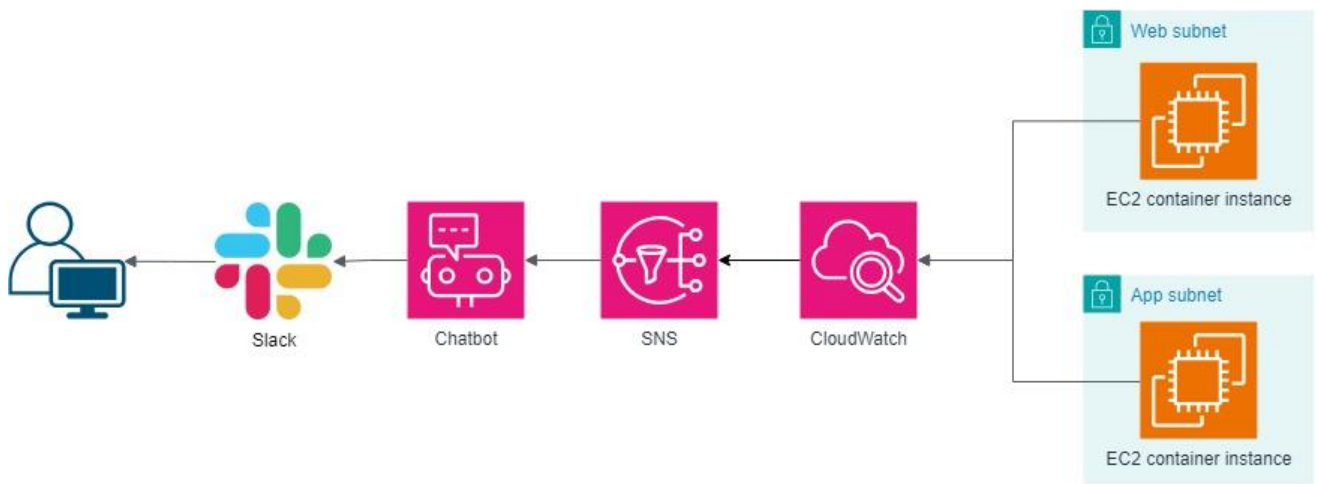
8. 구성 예상도

8-1. 예상도

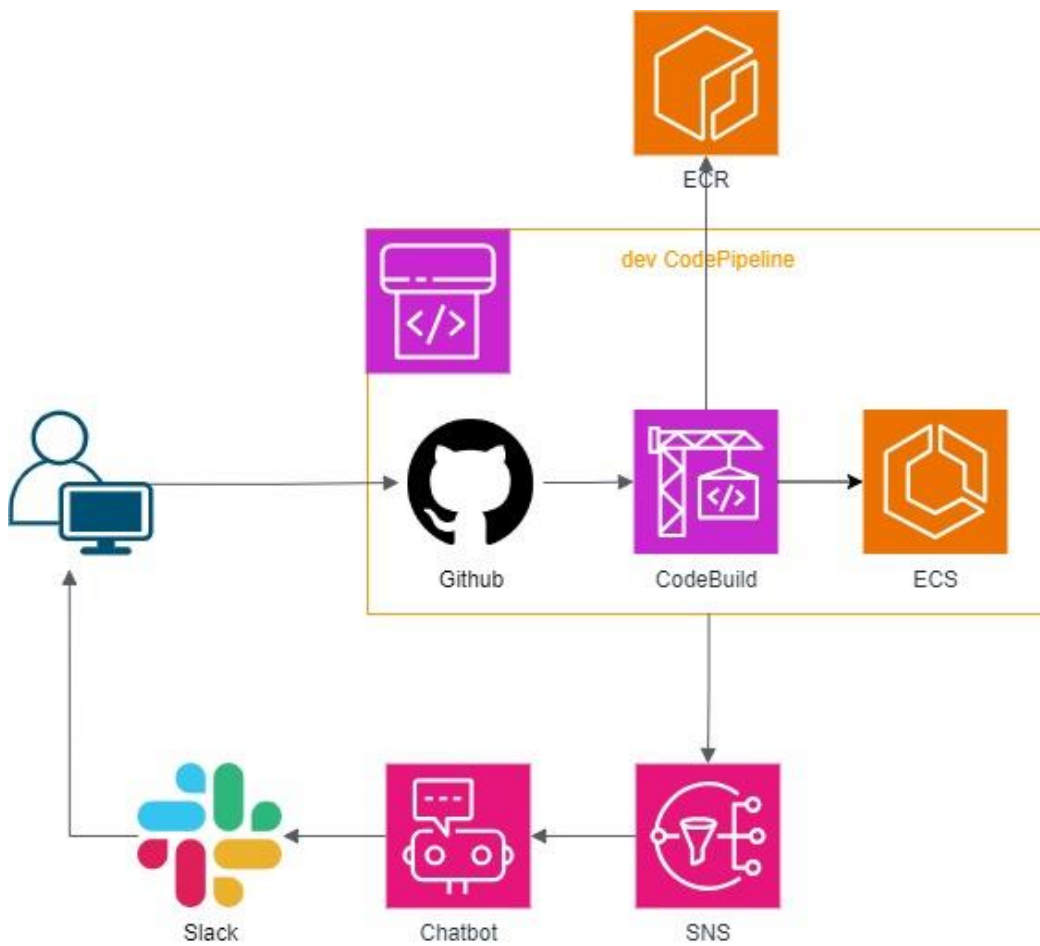
전체 예상도



모니터링



CI/.CD



백업

고가용성을 보장하고, 장애 발생 시 신속한 복구가 가능합니다.

DB 이중화 및 백업: 데이터베이스의 이중화와 정기적인 백업을 통해 데이터의 손실 위험을 크게 감소시킵니다.

2. 유효 인사이트 도출

데이터 기반 의사 결정을 강화하여 전략적 비즈니스 인사이트를 제공합니다.

데이터 분석 및 정제: 체계적인 데이터 분석을 통해 정확하고 신뢰할 수 있는 비즈니스 인사이트를 도출합니다.

인사이트 기반 의사결정 지원: 실시간 데이터 분석을 통해 의사결정 과정에서 필요한 신속하고 정확한 정보를 제공합니다.

3. 빠른 시작 대응력

시장 변화에 민첩하게 대응하고, 사용자의 요구 사항을 신속하게 충족시키기 위한 개발 프로세스를 최적화합니다.

모니터링 강화 도입: 강화된 모니터링 시스템을 통해 시스템의 성능 및 상태를 지속적으로 모니터링하고, 잠재적 문제를 예방합니다.

빠른 피드백 및 반복 개선: CI/CD 파이프라인과 자동화된 테스트를 통해 개발 및 배포 과정을 가속화하고, 빠른 피드백을 반영하여 지속적인 제품 개선을 실현합니다.

10. 예상 비용

10-1. 컴퓨팅

월간 비용

- 비용은 1달을 기준으로 함
- 고정 비용 이외에 트래픽 등에 따라 변동이 존재할 수 있음

총 월간 비용 : 856.72 USD

1. 네트워크

요소	수량	비용	사용 시간	전체 비용
NAT GW	2	\$0.059	24/7	\$ 86.14
ALB	2	\$0.0225	24/7	\$ 32.85
Route53	1	\$0.50	24/7	\$ 0.50
전체	\$ 119.49			

2. 컴퓨팅

요소	수량(용량)	비용	사용 시간	전체 비용
EC2(m5.large, WEB)	2 ~ 6	\$0.118	24/7	\$ 172.28
EC2(m5.large, WAS)	2 ~ 6	\$0.118	24/7	\$ 172.28
EC2(t2.micro, Bastion)	1	\$0.0144	24/7	\$ 10.51
RDS(db.m5.large)	100GB	0.472	24/7	\$ 372.16
전체	\$ 727.23			

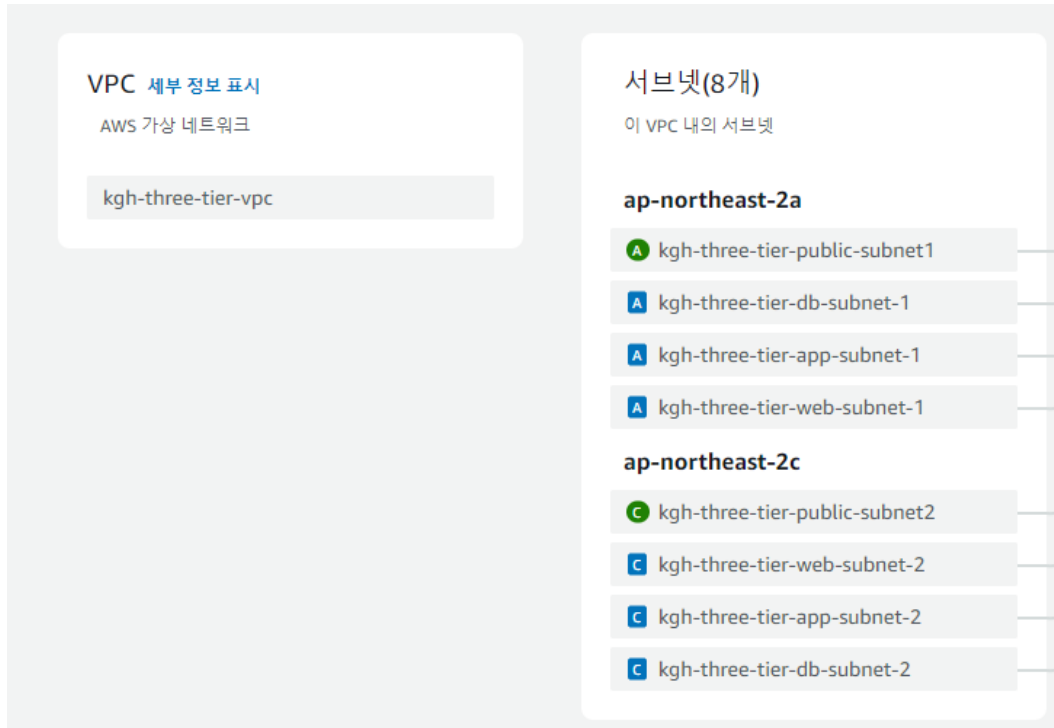
3. 보안

요소	웹 ACL	규칙	비용	전체 비용
WAF	1	3	24/7	\$ 8.00
전체	\$ 8.00			

11. 완료 보고

11.1. 3tier

VPC



Public Subnet



Private Subnet

<input type="checkbox"/>	kgh-three-tier-app-subnet-1	subnet-0462bfb8df4e2985	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.5.0/24
<input type="checkbox"/>	kgh-three-tier-app-subnet-2	subnet-0ffb7869e7ee42df8	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.6.0/24
<input type="checkbox"/>	kgh-three-tier-db-subnet-1	subnet-0c9cc494d9c8c5d41	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.7.0/24
<input type="checkbox"/>	kgh-three-tier-db-subnet-2	subnet-031f035f12f88b888	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.8.0/24
<input type="checkbox"/>	kgh-three-tier-public-subnet1	subnet-09c8519841861b4d9	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.1.0/24
<input type="checkbox"/>	kgh-three-tier-public-subnet2	subnet-0d8eca67c1b99af88	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.2.0/24
<input type="checkbox"/>	kgh-three-tier-web-subnet-1	subnet-0effddc62902bc64d	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.3.0/24
<input type="checkbox"/>	kgh-three-tier-web-subnet-2	subnet-08de36acaf10d282d	✔ Available	vpc-07c6cc47442999425 kgh-...	10.0.4.0/24


Internet Gateway

VPC > [인터넷 게이트웨이](#) > igw-088a3356e2a0ca04b

igw-088a3356e2a0ca04b / kgh-three-tier-igw

세부 정보

인터넷 게이트웨이 ID

 igw-088a3356e2a0ca04b

상태

✔ Attached

VPC ID

[vpc-07c6cc47442999425 | kgh-three-tier-vpc](#)

NAT Gateway

NAT 게이트웨이 (2) 정보



🔍 Find resources by attribute or tag

	Name ▾	NAT 게이트웨이 ID ▾	연결 유형 ▾	상태 ▾	상태 메시지 ▾	기본 퍼블릭 IPv4 ...
<input type="radio"/>	kgh-three-tier-nat-gw2	nat-0f6e5e59f85ce301a	Public	✔ Available	-	52.78.92.133
<input type="radio"/>	kgh-three-tier-nat-gw1	nat-0582ab4d952b47372	Public	✔ Available	-	13.209.104.0

EC2

인스턴스 (5) 정보

🔍 인스턴스를 속성 또는 (case-sensitive) 태그로 찾기

<input type="checkbox"/>	Name ✎ ▾	인스턴스 ID	인스턴스 상태 ▾
<input type="checkbox"/>	kgh-three-tier-web-instances	i-0b7b463db2dd9b5ea	✔ 실행 중 🔍 🔍
<input type="checkbox"/>	bastion	i-0a5ec5fb6bf877013	✔ 실행 중 🔍 🔍
<input type="checkbox"/>	kgh-three-tier-app-instances	i-0703149ee69f4b638	✔ 실행 중 🔍 🔍
<input type="checkbox"/>	kgh-three-tier-web-instances	i-053f793b9ad924e3a	✔ 실행 중 🔍 🔍
<input type="checkbox"/>	kgh-three-tier-app-instances	i-0457dd8b96e47e041	✔ 실행 중 🔍 🔍

Application Load Balancer

로드 밸런서: kgh-three-tier-alb-web

kgh-three-tier-alb-web



로드 밸런서: kgh-three-tier-alb-app

kgh-three-tier-alb-app



AutoScaling Group

Auto Scaling 그룹: kgh-three-tier-asg-web


🔍 ✕

세부 정보 | 활동 | Automatic scaling | 인스턴스 관리 | 모니터링 | 인스턴스 새로 고침

그룹 세부 정보				편집
Auto Scaling 그룹 이름 kgh-three-tier-asg-web	원하는 용량 2	원하는 용량 유형 단위(인스턴스 개수)	Amazon 리소스 이름(ARN) arn:aws:autoscaling:ap-northeast-2:381492:154999:autoScalingGroup:e50a6a57-27fa-4527-98dd-f9a9584ba11b:autoScalingGroupName/kgh-three-tier-asg-web	편집
생성된 날짜 Wed Jun 12 2024 12:01:37 GMT+0900 (한국 표준시)	최소 용량 2	상태 -		
	최대 용량 5			
로드 밸런싱				편집
로드 밸런서 대상 그룹 kgh-three-tier-tg-web	Classic Load Balancer -			

Relational Database Service

Amazon 리소스 이름(ARN)

 arn:aws:rds:ap-northeast-2:381492154999:db:my-rds-instance

리소스 ID

db-QBGCN5FHF77P6QEDN2UG7ACZTE

생성한 시간

June 12, 2024, 12:02 (UTC+09:00)

IAM DB 인증

활성화되지 **않음**

다중 AZ

예

보조 영역

ap-northeast-2a

11.2. CI/CD

GitHub

편집: Source

Source

GitHub(버전 1)

[개발자 도구](#) > [CodePipeline](#) > [파이프라인](#) > kgh-codepipeline

kgh-codepipeline

파이프라인 유형: V2 실행 모드: QUEUED

✔ Source 성공

파이프라인 실행 ID: [391e4b8f-dcea-471f-96cd-f2c84b71ff43](#)

Source

[GitHub\(버전 1\)](#)

✔ 성공 - 4일 전

[a437bbe1](#)

세부 정보 보기

[a437bbe1](#) Source: UI 개선

CodeBuild

환경

편집

이미지 aws/codebuild/amazonlinux2-x86_64-standard:4.0	환경 유형 Linux EC2	컴퓨팅 3GB 메모리, vCPU 2개	서비스 역할 arn:aws:iam::381492154999:role/service-role/kg-h-codebuild-service-role
권한이 있음 true	제한 시간 1시간 0분	대기 중인 제한 시간 8시간 0분	인증서 -
레지스트리 자격 증명 -			

▶ VPC

▶ 환경 변수

▶ 파일 시스템

Buildspec

편집

buildspec.yaml

CodePipeline

편집: Source

Source

GitHub(버전 1)

ⓘ

+ 스테이지 추가

편집: Build

Build

[AWS CodeBuild](#)

ⓘ

☐ 스테이지 장애 시 자동 롤백 구성

+ 스테이지 추가

편집: Stage

Deploy

Amazon ECS

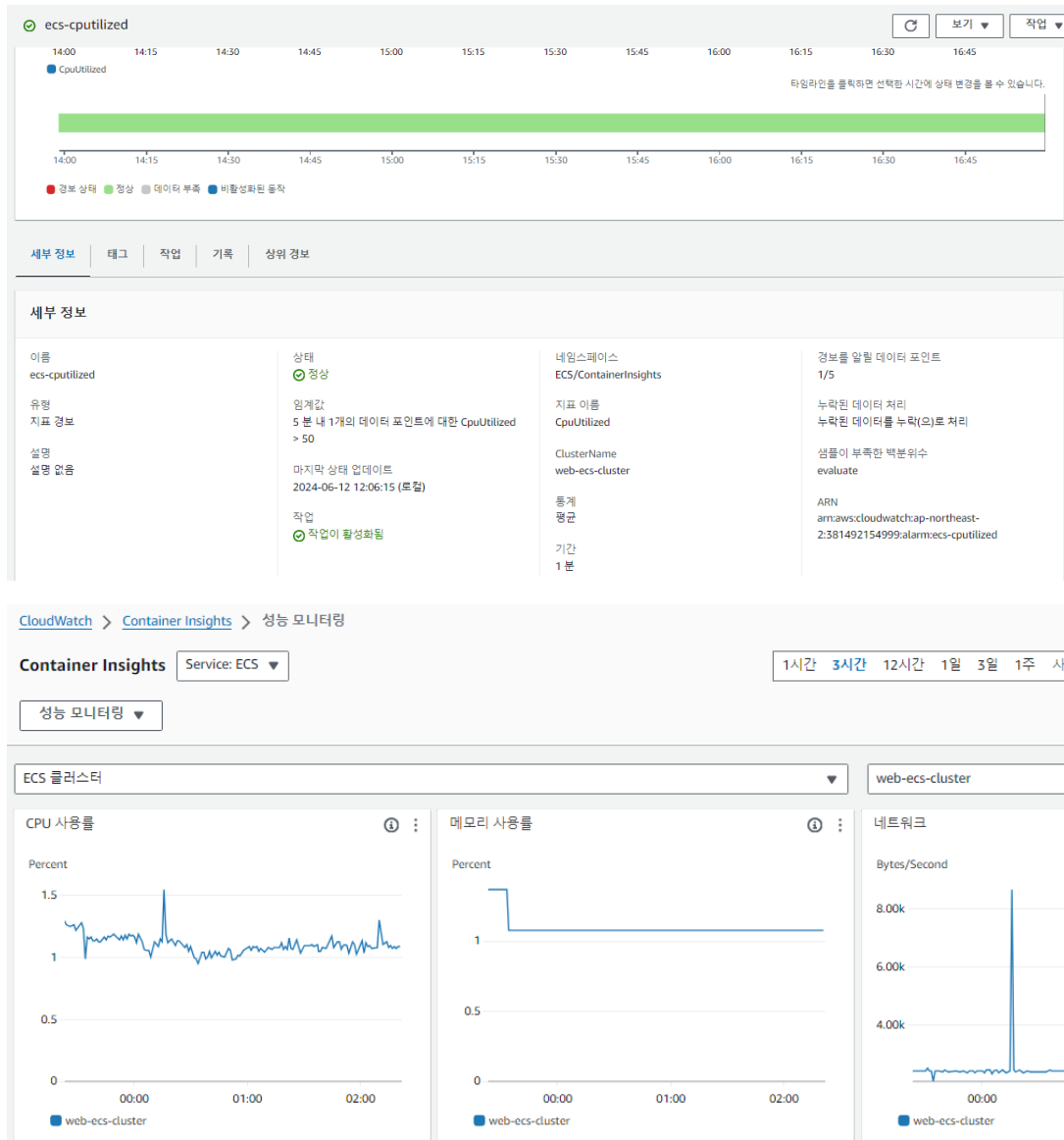
ⓘ

☐ 스테이지 장애 시 자동 롤백 구성

+ 스테이지 추가

11.3. 모니터링

CloudWatch



Simple Notification Service

Amazon SNS > 주제 > kgh-monitoring-ecs > 구독: 3fd65b36-e54c-4cac-b1ef-9a10ed6d91c5

구독: 3fd65b36-e54c-4cac-b1ef-9a10ed6d91c5

세부 정보

ARN

arn:aws:sns:ap-northeast-2:381492154999:kgh-monitoring-ecs:3fd65b36-e54c-4cac-b1ef-9a10ed6d91c5

엔드포인트

https://global.sns-api.chatbot.amazonaws.com

주제

kgh-monitoring-ecs

구독 보안 주제

arn:aws:iam::381492154999:role/aws-service-role/management.chatbot.amazonaws.com/AWSServiceRoleForAWSChatbot

상태

🟢 확인됨

프로토콜

HTTPS

원시 메시지 전송

비활성화됨

Chatbot

AWS Chatbot > 구성된 클라이언트 > Slack Workspace: kgh > kgh-monitoring-ecs

세부 정보

사용자 역할

태그

kgh-monitoring-ecs

Slack 채널 세부 정보

가드 레일 설정

테스트 메시지 전송

편집

삭제

이름

kgh-monitoring-ecs

채널

monitoring-ecs [🔗](#)

역할 설정

채널 역할

로그 수준

해제

채널 역할

kgh-Chatbot-role [🔗](#)

가드 레일 정책

1

ARN

arn:aws:chatbot:381492154999:chat-configuration/slack-channel/kgh-monitoring-ecs

지표

CloudWatch에서 지표 보기 [🔗](#)

Slack

kgh-monitoring-ecs 편집

구성 세부 정보

로깅 - 선택 사항

AWS Chatbot은(는) 사용자가 시작한 명령에 대한 감사 이벤트를 Amazon CloudWatch logs에 자동으로 로깅하지만, 구성에 대한 추가 로깅을 활성화할 수도 있습니다. Amazon CloudWatch Logs에 로깅하는 데는 요금이 부과됩니다. [자세히 알아보기](#)

☐ Amazon CloudWatch Logs에 로그 게시

Slack 채널

채널 유형

목록에서 퍼블릭 채널을 선택합니다. 프라이빗 채널을 선택하려면 채널 ID를 입력합니다.

☒ 퍼블릭

Workspace의 누구나 퍼블릭 채널을 보고 참여할 수 있습니다.

☐ 프라이빗

초대를 통해서만 프라이빗 채널에 참여하거나 프라이빗 채널을 볼 수 있습니다.

퍼블릭 채널 이름

🔍 monitoring-ecs



11.4. 백업

백업 규칙

[AWS Backup](#) > [백업 계획](#) > [kgh-backup-rds](#) > kgh-backup-rule

kgh-backup-rule

삭제

편집

요약

백업 규칙 이름 kgh-backup-rule	빈도 매일 에서 오전 02:00 Asia/Seoul (UTC+09:00)	다음 시간 내에 시작 1시간	다음 시간 내에 완료 3시간
클라우드 스토리지로 전환 활성화되지 않음	아카이브 Amazon EBS 스냅샷 활성화되지 않음	총 보존 기간 1개월	백업 볼트 kgh-backup-vault
연속 백업 활성화됨	복구 시점에 추가되는 태그 - 선택 사항 key: latest		

백업 리소스

[AWS Backup](#) > [백업 계획](#) > [kgh-backup-rds](#) > kgh-backup-resource

kgh-backup-resource

요약

리소스 할당 이름
kgh-backup-resource

IAM 역할
arn:aws:iam::381492154999:role/service-role/AWSBackupDefaultServiceRole

리소스 ID
arn:aws:rds:ap-northeast-2:381492154999:db:my-rds-instance

백업 볼트

kgh-backup-vault 정보

볼트 잠금 생성 볼트 삭제 액세스 정책 편집

요약

볼트 이름 kgh-backup-vault	KMS 암호화 키 ID 43001eae-81d9-445c-bdae-81140279cd40 링크	볼트 잠금 -
볼트 ARN arn:aws:backup:ap-northeast-2:381492154999:backup-vault:kgh-backup-vault	생성 날짜 2024년 6월 8일, 17:03 (UTC+09:00)	볼트 잠금 보존 기간 최소 보존 기간: - 최대 보존 기간: -

복구 시점 (5) 정보

모두 선택 취소 작업 ▼

리소스 유형, 복구 시점 ID, 상태, 리소스 ID 또는 소스 계정 ID로 필터링

< 1 >

<input type="checkbox"/>	복구 시점 ID ▼	상태 ▼	리소스 이름 ▼	리소스 ID ▼	리소스 유형 ▼	백업 유형 ▼	생성 시간 ▼	소스 계정 ID ▼	보존 기간	월드 스트로리지로 전환
<input type="checkbox"/>	continuous:db-svuyxo7zzgr4umty2gz2waunm-6b5a0d9c	사용 가능	my-rds-instance	my-rds-instance	RDS	지속적	2024년 6월 19일, 02:00 (UTC+09:00)	-	7일	안 함
<input type="checkbox"/>	continuous:db-mhnqbm73gqjduylvzkpestsey-f56c8f86	사용 가능	my-rds-instance	my-rds-instance	RDS	지속적	2024년 6월 17일, 02:00 (UTC+09:00)	-	7일	안 함
<input type="checkbox"/>	continuous:db-qbgcn5fhf77p6qedn2ug7aczte-dde66da5	만료됨	my-rds-instance	my-rds-instance	RDS	지속적	2024년 6월 13일, 02:00 (UTC+09:00)	-	7일	안 함
<input type="checkbox"/>	continuous:db-zdonyuf2l5xtatxqd73bhhrpmu-c1860018	만료됨	my-rds-instance	my-rds-instance	RDS	지속적	2024년 6월 9일, 17:45 (UTC+09:00)	-	7일	안 함
<input type="checkbox"/>	continuous:db-wgijvqaj6szaldqjsfr4toi2u-c1262440	만료됨	my-rds-instance	my-rds-instance	RDS	지속적	2024년 6월 8일, 17:45 (UTC+09:00)	-	7일	안 함

11.5. 보안

Route 53

Route 53 > 호스팅 영역 > mkcloud.site

미분류 mkcloud.site 정보

영역 삭제레코드 테스트쿼리 로깅 구성

▶ 호스팅 영역 세부 정보

호스팅 영역 편집

레코드(8)

DNSSEC 서명

호스팅 영역 태그(1)

레코드 (8) 정보

레코드 삭제영역 파일 가져오기레코드 생성

Automatic 모드는 최상의 필터 결과에 최적화된 현재 검색 동작입니다. 모드를 변경하려면 설정(settings)으로 이동합니다.

Q 속성 또는 값을 기준으로 레코드 필터링

유형라우팅 정책별칭

< 1 > ⓘ

<input type="checkbox"/>	레코드 이름	유형	라우팅 ...	차별...	별칭	값/트래픽 라우팅 대상	TTL(초)	상태 확...	대상 ...	레...
<input type="checkbox"/>	mkcloud.site	NS	단순	-	아니요	ns-600.awsdns-11.net. ns-1169.awsdns-18.org. ns-1704.awsdns-21.co.uk. ns-490.awsdns-61.com.	172800	-	-	-
<input type="checkbox"/>	mkcloud.site	SOA	단순	-	아니요	ns-600.awsdns-11.net. awsd...	900	-	-	-
<input type="checkbox"/>	jw.mkcloud.site	A	단순	-	예	dualstack.alb-web-1534330...	-	-	아니요	-
<input type="checkbox"/>	_b300893b21a02a814ff556654a096e09.jw.mkcloud.site	CNAME	단순	-	아니요	_60556335774b30c63cef21...	300	-	-	-
<input type="checkbox"/>	kgh.mkcloud.site	CNAME	단순	-	아니요	kgh-three-tier-alb-web-2896...	86400	-	-	-
<input type="checkbox"/>	_2c645b92600daca34278268847954040.kgh.mkcloud.site	CNAME	단순	-	아니요	_8b1614aec4b29783654105...	60	-	-	-
<input type="checkbox"/>	web.mkcloud.site	A	단순	-	예	d1ousguq0s40hy.cloudfront....	-	-	아니요	-
<input type="checkbox"/>	_e1747f5dc238cb6365fe39ae7b133b45.web.mkcloud.site	CNAME	단순	-	아니요	_b3c5b6ff46f68f285dc2e79...	60	-	-	-

Certificate Manager

레코드 (1/10) 정보

레코드 삭제영역 파일 가져오기레코드 생성

Automatic 모드는 최상의 필터 결과에 최적화된 현재 검색 동작입니다. 모드를 변경하려면 설정(settings)으로 이동합니다.

Q 속성 또는 값을 기준으로 레코드 필터링

유형라우팅 정책별칭

< 1 > ⓘ

<input type="checkbox"/>	레코드 ...	유형	라우팅 ...	차별...	별칭	값/트래픽 라우팅 대상
<input type="checkbox"/>	mkcloud.s...	A	단순	-	예	mk-alb-web-742562245.ap-...
<input type="checkbox"/>	mkcloud.s...	NS	단순	-	아니요	ns-600.awsdns-11.net. ns-1169.awsdns-18.org. ns-1704.awsdns-21.co.uk. ns-490.awsdns-61.com.
<input type="checkbox"/>	mkcloud.s...	SOA	단순	-	아니요	ns-600.awsdns-11.net. awsd...
<input type="checkbox"/>	_d7c6b30...	CNAME	단순	-	아니요	_2b4900d67c276ebdf775...
<input type="checkbox"/>	jw.mkclou...	A	단순	-	예	dualstack.alb-web-1534330...
<input type="checkbox"/>	_b300893...	CNAME	단순	-	아니요	_60556335774b30c63cef21...
<input checked="" type="checkbox"/>	kgh.mkclo...	CNAME	단순	-	아니요	kgh-three-tier-alb-web-1127...
<input type="checkbox"/>	_2c645b9...	CNAME	단순	-	아니요	_8b1614aec4b29783654105...
<input type="checkbox"/>	web.mkcl...	A	단순	-	예	d3nm1bv3xjdtq.cloudfront....

레코드 세부 정보 ⓘ ✕

레코드 편집

레코드 이름

kgh.mkcloud.site

레코드 유형

CNAME

값

kgh-three-tier-alb-web-1127373520.ap-northeast-2.elb.amazonaws.com

별칭

아니요

TTL(초)

86400

라우팅 정책

Simple

AWS Certificate Manager > 인증서 > 9f16e886-a358-418f-b262-e05caf190b85

9f16e886-a358-418f-b262-e05caf190b85 삭제

인증서 상태

식별자: 9f16e886-a358-418f-b262-e05caf190b85 상태: 보급됨

ARN: arn:aws:acm:ap-northeast-2:381492154999:certificate/9f16e886-a358-418f-b262-e05caf190b85

유형: Amazon 발급

도메인 (2) Route 53에서 레코드 생성 CSV로 내보내기

도메인	상태	경신 상태	유형	CNAME 이름	CNAME 값
mikcloud.site	상관	-	CNAME	_d7c6b30b80355e57200f5b5ded1c45a4.mikcloud.site.	_2b4900a67c276ebedf6775f629879b80.sdgtdhzh.acm-validations.aws.
kgh.mikcloud.site	상관	-	CNAME	_2c645b52600daca34278268847954040.kgh.mikcloud.site.	_8b1614aec4b29783654105bd35608042.sdgtdhzh.acm-validations.aws.

WAF

로드 밸런서 (1/2) 로드 밸런서 생성

Elastic Load Balancing은 수신 트래픽의 변화에 따라 자동으로 로드 밸런서 용량을 확장합니다.

로드 밸런서 필터링

이름	DNS 이름	상태	VPC ID	가용 영역	유형	생성
<input checked="" type="checkbox"/> alb-web	alb-web-1534330926.us-e...	활성	vpc-070d54095f5f5a5f1	2 가용 영역	application	202
<input type="checkbox"/> alb-app	internal-alb-app-3951801...	활성	vpc-070d54095f5f5a5f1	2 가용 영역	application	202

로드 밸런서: alb-web

AWS Config 정보 통합이 탐지되지 않음

최적화: 모니터링, 규정 준수

AWS 웹 애플리케이션 방화벽(WAF) 정보 통합 탐지됨

최적화: 보안

통합 상태 세부 정보
이 로드 밸런서는 AWS WAF와 통합됩니다.

AWS WAF 웹 ACL
[CreatedByALB-935cf3505975acb2](#)

[AWS WAF](#) > [Web ACLs](#) > CreatedByALB-935cf3505975acb2

CreatedByALB-935cf3505975acb2

[Traffic overview](#) | [Rules](#) | [Associated AWS resources](#) | [Custom response bodies](#) | [Logging and metrics](#)

Rules (3)

Find rules

<input type="checkbox"/>	Name	Action	Prior
<input type="checkbox"/>	AWS-AWSManagedRulesAmazonIpReputationList	Use rule actions	0
<input type="checkbox"/>	AWS-AWSManagedRulesCommonRuleSet	Use rule actions	1
<input type="checkbox"/>	AWS-AWSManagedRulesKnownBadInputsRuleSet	Use rule actions	2

11.6. 데이터 파이프라인

Glue Connections

Connections (1) [Info](#)

Actions ▼ Create connection Create job

You can manage your connections or use a connection in a job.

Filter connections by property < 1 >

Name	Status	Type	Last modified
Mysql connection	Ready	JDBC	Jun 12, 2024

Lambda

Lambda > 함수 > kgh-clean-survey-quality

kgh-clean-survey-quality

조절 ARN 복사 작업 ▼

▼ 함수 개요 정보 Application Composer로 내보내기 다운로드 ▼

다이어그램 템플릿

+ 대상 추가

설명
-
마지막 수정
11월 전
함수 ARN
arn:aws:lambda:ap-northeast-2:381492154999:function:kgh-clean-survey-quality
함수 URL [정보](#)
-

코드 테스트 모니터링 구성 별칭 버전

코드 소스 정보

File Edit Find View Go Tools Window Test Deploy

Environment

Go to Anything (Ctrl-P)

lambda_function x Environment Var x

1 import boto3
2 import sys
3 import json
4
5 def lambda_handler(event, context):
6 glue_client = boto3.client('glue')
7
8 for record in event['Records']:
9 bucket_name = record['s3']['bucket']['name']
10 object_key = record['s3']['object']['key']
11
12 # S3 경로
13 s3_path = f"s3://{bucket_name}/{object_key}"
14
15 # AWS Glue 실행
16 response = glue_client.start_job_run(
17 JobName='kgh-data-quality-survey',
18 Arguments={
19 '--S3_INPUT_PATH': s3_path
20 },
21)
22
23 return {
24 'statusCode': 200,
25 'body': json.dumps('Glue job started successfully')
26 }
27

S3 버킷 put event

이벤트 알림 (1) 편집 삭제 이벤트 알림 생성

버킷에서 특정 이벤트가 발생하면 알림을 보냅니다. [자세히 알아보기](#)

이름	이벤트 유형	필터	대상 유형	대상
kgh-raw-survey-put	전송	-	Lambda 함수	kgh-excel-parsing-survey

Glue ETL job - RAW

Script	Job details	Runs	Data quality	Schedules	Version Control
--------	-------------	------	--------------	-----------	-----------------

Script [Info](#)

```
1 import sys
2 from awsglue.transforms import *
3 from awsglue.utils import getResolvedOptions
4 from pyspark.context import SparkContext
5 from awsglue.context import GlueContext
6 from awsglue.job import Job
7 from pyspark.sql.functions import current_date, year, month, dayofmonth, col, lit
8 from datetime import datetime
9
10 # AWS Glue Context 설정
11 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
12 sc = SparkContext()
13 glueContext = GlueContext(sc)
14 spark = glueContext.spark_session
15 job = Job(glueContext)
16 job.init(args['JOB_NAME'], args)
17
18 # 현재 날짜 가져오기
19 current_date_str = datetime.now().strftime("%Y-%m-%d")
20 year_value = datetime.now().year
21 month_value = datetime.now().month
22 day_value = datetime.now().day
23
24 # S3 경로
25 input_paths = {
26     "customer": f"s3://lhs-raw-rds/customer/year={year_value}/month={month_value}/day={day_value}/",
27     "purchase": f"s3://lhs-raw-rds/purchase/year={year_value}/month={month_value}/day={day_value}/",
28     "review": f"s3://lhs-raw-rds/review/year={year_value}/month={month_value}/day={day_value}/"
29 }
30
```

Glue ETL job – Clean

Script	Job details	Runs	Data quality	Schedules	Version Control
--------	-------------	------	--------------	-----------	-----------------

Script Info

```

1 import sys
2 from awsglue.transforms import *
3 from awsglue.utils import getResolvedOptions
4 from pyspark.context import SparkContext
5 from awsglue.context import GlueContext
6 from awsglue.job import Job
7 from pyspark.sql.functions import current_date, year, month, dayofmonth, col, lit
8 from datetime import datetime
9
10 # AWS Glue Context 설정
11 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
12 sc = SparkContext()
13 glueContext = GlueContext(sc)
14 spark = glueContext.spark_session
15 job = Job(glueContext)
16 job.init(args['JOB_NAME'], args)
17
18 # 현재 날짜 가져오기
19 current_date_str = datetime.now().strftime("%Y-%m-%d")
20 year_value = datetime.now().year
21 month_value = datetime.now().month
22 day_value = datetime.now().day
23
24 # S3 경로
25 input_paths = {
26     "customer": f"s3://lhs-raw-rds/customer/year={year_value}/month={month_value}/day={day_value}/",
27     "purchase": f"s3://lhs-raw-rds/purchase/year={year_value}/month={month_value}/day={day_value}/",
28     "review": f"s3://lhs-raw-rds/review/year={year_value}/month={month_value}/day={day_value}/"
29 }
30

```

Glue ETL job – RDS 정제

Script	Job details	Runs	Data quality	Schedules	Version Control
--------	-------------	------	--------------	-----------	-----------------

Script [Info](#)

```
1 import sys
2 from awsglue.transforms import *
3 from awsglue.utils import getResolvedOptions
4 from pyspark.context import SparkContext
5 from awsglue.context import GlueContext
6 from awsglue.job import Job
7 from awsglue.dynamicframe import DynamicFrame
8 from pyspark.sql.functions import col
9 from datetime import datetime
10
11 # @params: [JOB_NAME]
12 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
13
14 # Initialize Spark and Glue contexts
15 sc = SparkContext()
16 glueContext = GlueContext(sc)
17 spark = glueContext.spark_session
18 job = Job(glueContext)
19 job.init(args['JOB_NAME'], args)
20
21 # Get the current date
22 current_date_str = datetime.now().strftime("%Y-%m-%d")
23 year_value = datetime.now().year
24 month_value = datetime.now().month
25 day_value = datetime.now().day
26
27
28 # Read data from S3 in Parquet format
29 data_frame_purchase = glueContext.create_dynamic_frame.from_options(
30     connection_type="s3",
```

Glue ETL job – Log 정제

Script	Job details	Runs	Data quality	Schedules	Version Control
--------	-------------	------	--------------	-----------	-----------------

Script [Info](#)

```

1 import sys
2 from awsglue.transforms import *
3 from awsglue.utils import getResolvedOptions
4 from pyspark.context import SparkContext
5 from awsglue.context import GlueContext
6 from awsglue.job import Job
7 from awsglue.dynamicframe import DynamicFrame, DynamicFrameCollection
8 from pyspark.sql.functions import hour, countDistinct
9 from datetime import datetime
10
11 # Get job arguments
12 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
13
14 # Initialize Spark and Glue contexts
15 sc = SparkContext()
16 glueContext = GlueContext(sc)
17 spark = glueContext.spark_session
18 job = Job(glueContext)
19 job.init(args['JOB_NAME'], args)
20
21 # Get the current date
22 current_date = datetime.now()
23 year_value = current_date.year
24 month_value = current_date.month
25 day_value = datetime.now().day # Fixed as per your path
26
27 # Script generated for node Amazon S3
28 AmazonS3_node1718213052888 = glueContext.create_dynamic_frame.from_options(
29     format_options={"multiline": False},
30     connection_type="s3",

```

Glue Data Quality

Script	Job details	Runs	Data quality - <i>updated</i>	Schedules	Version Control
Script Info					
<pre> 1 import sys 2 from awsglue.transforms import * 3 from awsglue.utils import getResolvedOptions 4 from pyspark.context import SparkContext 5 from awsglue.context import GlueContext 6 from awsglue.job import Job 7 from awsgluedq.transforms import EvaluateDataQuality 8 from datetime import datetime 9 10 args = getResolvedOptions(sys.argv, ['JOB_NAME']) 11 sc = SparkContext() 12 glueContext = GlueContext(sc) 13 spark = glueContext.spark_session 14 job = Job(glueContext) 15 job.init(args['JOB_NAME'], args) 16 17 current_date_str = datetime.now().strftime("%Y-%m-%d") 18 year_value = datetime.now().year 19 month_value = datetime.now().month 20 day_value = datetime.now().day 21 22 root_path="kgh-clean-rds" 23 24 # Read data from S3 in Parquet format 25 data_frame_purchase = glueContext.create_dynamic_frame.from_options(26 connection_type="s3", 27 format="parquet", 28 connection_options={ 29 "paths": [f"s3://{root_path}/purchase/year={year_value}/month={month_value}/day={day_value}/"], 30 "recurse": True 31 }, 32 transformation_ctx="data_frame_purchase" 33) 34 </pre>					

Script	Job details	Runs	Data quality - <i>updated</i>	Schedules	Version Control
Script Info					
<pre> 1 import sys 2 from urllib.parse import unquote 3 from awsglue.transforms import * 4 from awsglue.utils import getResolvedOptions 5 from pyspark.context import SparkContext 6 from awsglue.context import GlueContext 7 from awsglue.job import Job 8 from awsgluedq.transforms import EvaluateDataQuality 9 import boto3 10 import logging 11 12 # 로그 설정 13 logging.basicConfig(level=logging.INFO) 14 logger = logging.getLogger(__name__) 15 16 # Lambda에서 전달된 파라미터를 받아옵니다. 17 args = getResolvedOptions(sys.argv, ['JOB_NAME', 'S3_INPUT_PATH']) 18 sc = SparkContext() 19 glueContext = GlueContext(sc) 20 spark = glueContext.spark_session 21 job = Job(glueContext) 22 job.init(args['JOB_NAME'], args) 23 24 # 전달된 S3 경로를 로그에 출력 25 s3_input_path = args['S3_INPUT_PATH'] 26 logger.info(f"Received S3 input path: {s3_input_path}") 27 28 # URL 인코딩된 S3 경로를 디코딩 29 decoded_s3_input_path = unquote(s3_input_path) 30 logger.info(f"Decoded S3 input path: {decoded_s3_input_path}") 31 32 # S3 경로에서 파일이 존재하는지 확인 33 try: 34 s3 = boto3.client('s3') </pre>					

Data Streams

Amazon Kinesis > 데이터 스트림 > kgh-kinesis-datastreams		
kgh-kinesis-datastreams 정보		
데이터 스트림 요약		
상태 🟢 활성	용량 모드 온디맨드 데이터 보존 기간 1일	ARN 📄 <code>arn:aws:kinesis:ap-northeast-2:381492154999:stream/kgh-kinesis-datastreams</code>

Data firehose

Amazon Data Firehose > Firehose 스트림 > kgh-kinesis-firehose		
kgh-kinesis-firehose 정보		
Firehose 스트림 세부 정보		
상태 🟢 활성	대상 Amazon S3	데이터 변환 활성화되지 않음
소스 Amazon Kinesis Data Streams	ARN 📄 <code>arn:aws:firehose:ap-northeast-2:381492154999:deliverystream/kgh-kinesis-firehose</code>	동적 파티셔닝 활성화됨

Glue Crawler

[AWS Glue](#) > Crawlers

Crawlers

A crawler connects to a data store, progresses through a prioritized list of classifiers to determine the schema for your data, and then creates meta

Crawlers (10) [Info](#)

View and manage all available crawlers.

Q

Filter crawlers

5 matches

kgh

X

Clear filters

<input type="checkbox"/>	Name	State	Schedule	Last run
<input type="checkbox"/>	kgh-age-category-sales-crawler	✔ Ready		✔ Succeeded
<input type="checkbox"/>	kgh-category-sales-crawler	✔ Ready		✔ Succeeded
<input type="checkbox"/>	kgh-log-crawler	✔ Ready		✔ Succeeded
<input type="checkbox"/>	kgh-review-purchase-crawler	✔ Ready		✔ Succeeded
<input type="checkbox"/>	kgh-survey-crawler	✔ Ready		✔ Succeeded

Glue Workflow

[AWS Glue](#) > [Workflows](#) > kgh-RDS-Workflow

kgh-RDS-Workflow

Last updated (UTC)
June 23, 2024 at 10:28:02

Run workflow

Edit

Delete

Workflow details

Advanced properties

Name
kgh-RDS-Workflow

Last run
June 22, 2024 at 14:30:25

Description
-

Last modified
June 13, 2024 at 03:15:38

Max concurrency
-

Blueprint name
-

Last run status
✔ Completed

Blueprint run id
-

Graph

History

Tags

Legend: ● Start ◆ Trigger 📄 Job 📄 Crawler ✔ Incomplete ✖ Error ⏸ Deleting

Remove

Action ▼

Trigger details

X

Name
start RDS data cleaning

Description
-

Trigger type
CONDITIONAL

Trigger state
ACTIVATED

Starts
kgh-s3-to-clean-rds

Trigger logic
Start after ANY watched event

Watching
kgh-rds-data-save - SUCCEEDED

Athena

Amazon Athena > 쿼리 편집기

편집기 | 최근 쿼리 | 저장된 쿼리 | 설정

데이터

데이터 원본
AwsDataCatalog

데이터베이스
2team_data_analytics_db

테이블 및 보기
생성 ▼

테이블 (5)
 2team_clean_survey 파티션됨
 age_category_sales 파티션됨
 category_sales 파티션됨
 log-hourly_session_counts 파티션됨
 review_purchase 파티션됨
 보기 (0)

쿼리 15

```
1 SELECT * FROM "2team_data_analytics_db"."log-hourly_session_counts"
```

SQL Ln 1, Col 1

다시 실행 | 설명 | 취소 | 지우기 | 생성 ▼

쿼리 결과 | 쿼리 통계

완료됨

EventBridge

Amazon EventBridge > 규칙 > kgh-glue-slack-event-state

kgh-glue-slack-event-state

규칙 세부 정보

<p>규칙 이름 kgh-glue-slack-event-state</p> <p>설명</p>	<p>상태 활성화됨</p> <p>규칙 ARN arn:aws:events:ap-northeast-2:381492154999:rule/kgh-glue-slack-event-state</p>	<p>이벤트 버스 이름 default</p> <p>이벤트 버스 ARN arn:aws:events:ap-northeast-2:381492154999:event-bus/default</p>
---	---	---

이벤트 패턴

대상 | 모니터링 | 태그

이벤트 패턴 정보

```
1 {
2   "source": ["aws.glue"],
3   "detail-type": ["Glue Job State Change"]
4 }
```

복사

[Amazon EventBridge](#) > [규칙](#) > kgh-job-change-state

kgh-job-change-state

규칙 세부 정보

규칙 이름

kgh-job-change-state

설명

상태

활성화됨

규칙 ARN

arn:aws:events:ap-northeast-2:381492154999:rule/kgh-job-change-state

이벤트 버스 이름

default

이벤트 버스 ARN

arn:aws:events:ap-northeast-2:381492154999:event-bus/default

이벤트 패턴

대상

모니터링

태그

이벤트 패턴 정보

```
1 {
2   "source": ["aws.glue"],
3   "detail-type": ["Glue Job State Change"],
4   "detail": {
5     "jobName": ["kgh-rds-data-quality", "kgh-data-quality-survey", "kgh-hourly_session_counts", "kgh-review-purchase", "kgh-purchase-country", "kgh-category-sales", "kgh-category-sales", "kgh-age-category-sales"]
6   }
7 }
```

복사

SPICE

데이터 세트

이름		
	연령대별 카테고리 판매량	SPICE
	연령대별 유입경로	SPICE
	시간대별 클릭 스트림 수	SPICE
	카테고리별 판매율 & 리뷰	SPICE
	카테고리별 판매수	SPICE

QuickSight 대시보드

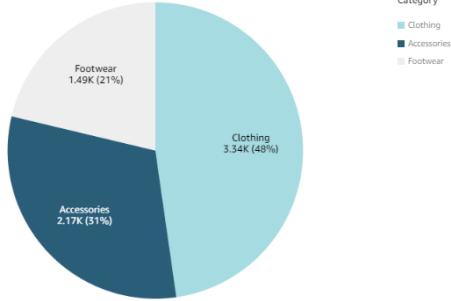
Shopping Mall Data Insights

의류 쇼핑몰의 데이터 분석 인사이트

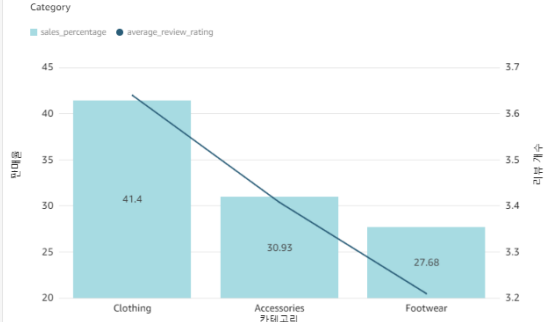
Database Insights

1, 2분기 (2024/01/01 ~ 2024/06/17)

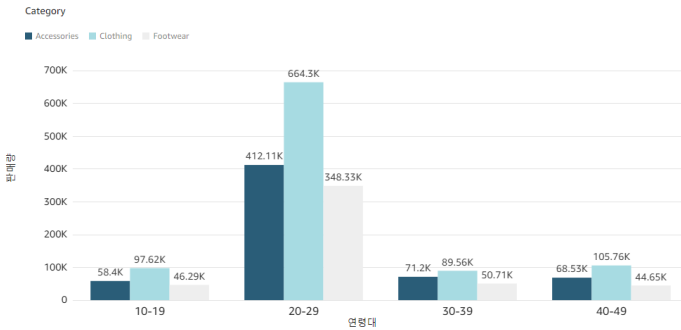
카테고리별 판매수



카테고리별 판매율 & 리뷰



연령대별 카테고리 판매량



연령대

전체

카테고리

전체

Log Insights

1, 2분기 (2024/01/01 ~ 2024/06/17)

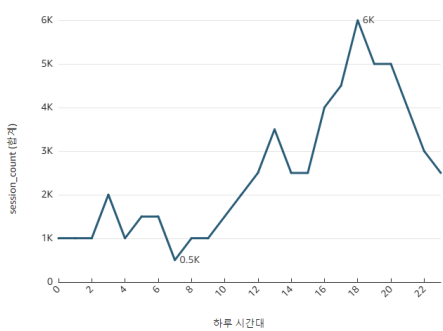
session

- 모두 선택
- 500
 - 1000
 - 1500
 - 2000
 - 2500
 - 3000
 - 3500
 - 4000
 - 4500
 - 5000
 - 6000

시간대

전체

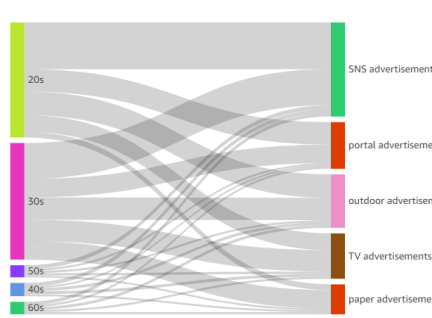
시간대별 클릭 스트림 수



Survey Insights

2분기 (2024/04/01 ~ 2024/06/17)

연령대별 유입경로



연령대

전체

유입 경로

전체