**제 안 서**

|  |  |
| --- | --- |
| **사 업 명** | **xx 의류 쇼핑몰 마케팅 전략을 위한 데이터 인프라 구축** |
| **주 관 사** |  |

**2024. 05. 18.**

**목 차**

[**1.** **사업 개요** 3](#_Toc166832455)

[**1-1.** **개요** 3](#_Toc166832456)

[**1-2.** **배경 및 필요성** 3](#_Toc166832457)

[**1-3.** **사업 범위** 4](#_Toc166832458)

[**1-4.** **기대효과** 5](#_Toc166832459)

[**2.** **추진 전략** 6](#_Toc166832460)

[**2-1.** **추진 목표** 6](#_Toc166832461)

[**3.** **적용 기술** 6](#_Toc166832462)

[**3-1.** **적용 기술** 6](#_Toc166832463)

|  |
| --- |
| **사업 개요** |

## **개요**

사 업 명: xx 의류 쇼핑몰 마케팅 전략을 위한 데이터 인프라 구축

사 업 예 산: 150,000,000원(일억 오천만원)

사 업 기 간: 계약 체결일로부터 60일 이내

## **배경 및 필요성**

**데이터 분석을 통한 마케팅 인사이트 도출**

* 비즈니스 활동의 온라인 이동으로 인해 사용자 상호작용 데이터 증가
* 다양한 소스에서 발생하는 데이터를 효과적으로 수집하고 분석할 수 있는 솔루션 필요
* 이를 위한 데이터 파이프라인 구축으로 의사결정 효율성을 높여 기업 경쟁력 강화
* 클라우드 환경에서의 데이터 분석 환경 구축으로 On-premise 환경의 기술적 한계 극복, 관리 비용 절감

**사용자 경험 개선 및 이탈 방지를 위한 데이터 분석**

* 사용자 경험이 비즈니스 성공의 핵심 요소로 부상
* 스트림 데이터를 분석해 사용자의 요구와 행동을 이해하고 사용자 경험 최적화
* 사용자가 선호하는 페이지의 이유 분석과 이탈 방지에 대한 조치 마련 가능
* (체크아웃 과정의 단순화, 사용자 인터페이스 개선, 고객 지원 강화 등)

**효율적인 홍보 방안 도출을 위한 이벤트 설문 조사 데이터 분석**

* 웹 페이지 내에서 수집하기 어려운 마케팅 유입 경로 데이터를 이벤트 설문 조사로 수집 예정
* (TV 광고, 지면 광고, 옥외 광고 등 마케팅 유입 경로 조사)
* 때문에 설문 조사를 통해 수집된 데이터를 효율적으로 저장하고 분석해 시각화하는 환경 필요
* 이를 활용해 효과적인 마케팅 유입 경로 파악 후 추후 홍보 전략에 적용

**마케팅 최적화를 위한 사용자 정보 및 판매율, 리뷰 상관 관계 분석**

* 고객 정보와 구매 이력 데이터를 활용해 고객 특성 별 인기 상품 파악 및 신제품 출시에 활용
* 판매율 및 고객 리뷰 상관 관계를 파악해 제품 개선 사항과 품질 향상에 대한 방향성 제시

## **사업 범위**

**클라우드 컴퓨팅 서비스 구축 및 인프라 고도화**

* GCP 클라우드 컴퓨팅 서비스에서 운영 중인 인프라를 AWS 컴퓨팅 환경에서 재구축 및 고도화
* 재구축 시 기존의 2Tier 아키텍처를 3Tier로 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**

* 인프라 관리의 운영 부담을 줄임
* 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

**보안 시스템 구축**

* 특정 웹 트래픽을 필터링하고 웹 공격을 차단
* DDoS 공격으로부터 리소스를 보호
* 웹사이트/서버 및 브라우저 간 암호화된 연결 설정

## **기대효과**

* 사용자의 행동 패턴과 선호도를 파악하여 사용자 경험을 향상
* 사용자 데이터를 기반으로 한 마케팅 전략을 수립하여 효율적인 광고 및 프로모션을 진행
* 클라우드 기반의 데이터 인프라의 확장성과 유연성을 제공하여 비용을 절감하고 운영 효율을 높임

|  |
| --- |
| **추진 전략** |

## **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 타깃 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

|  |
| --- |
| **적용 기술** |

## **적용 기술**

**Virtual Private Cloud**

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

**Public Subnet**

* 인터넷 게이트웨이를 연결해 AWS 리소스가 인터넷과 통신할 수 있게 함

**Private Subnet**

* 가용영역 2개에 Web, App, DB 3개의 subnet으로 각각 분리
* 3tier
* 보안을 위해 인터넷에서 직접 접근할 수 없게 함

**Internet Gateway**

* VPC와 인터넷 간 통신을 하게 함

**NAT Gateway**

* 가용영역 Public Subnet에 각각의 NGW 구성
* Private Subnet의 리소스가 인터넷에 접근하도록 함
* 인터넷에서는 그 리소스에 접근할 수 없게 함

**Application Load Balancer**

* Web, App 2개의 ALB 구성
* EC2 인스턴스에 자동으로 분산시켜 로드를 균등하게 관리

**AutoScaling Group**

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

**Relational Database Service**

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

**GitHub**

* Web, App 서버 소스 코드 저장 및 버전 관리
* 개발자가 코드 변경사항을 커밋

**CodeBuild**

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

**CodeDeploy**

* 컴파일 된 소스를 컴퓨팅 서비스에 배포
* 배포 실패시 롤백 구성

**CodePipeline**

* 빌드, 테스트, 배포 과정을 순차적 실행하는 워크플로우 구성
* GitHub의 코드 변경 사항을 자동으로 감지하는 트리거로 자동화

**CloudWatch**

* CloudWatch Agent로 Web, App CPU 지표 모니터링
* 일정 수치 이상시 경보 발생

**Simple Notification Service**

* CloudWatch에서 발생한 경보에 알람을 생성

**Chatbot**

* SNS에서 발생한 알람을 Slack으로 전달
* 복잡합 스크립트나 중개자 없이 사용자에게 전달

**Slack**

* 서버 이상 경보 알림 생성
* CodePipeline 성공 및 실패 알림 생성

**Elastic Container Service**

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

**Elastic Container Registry**

* 컨테이너 이미지를 저장, 관리 및 배포하기 위한 완전 관리형 컨테이너 레지스트리 사용
* ECS와 원활하게 통합되어 컨테이너화된 애플리케이션의 배포를 간소
* IAM을 사용하여 사용자 및 서비스별 액세스 권한을 세밀하게 제어
* 안전한 이미지 관리와 액세스를 보장

**Kinesis DataStream**

* 클릭 스트림 데이터 처리를 위해 대량의 데이터 레코드를 실시간으로 수집하고 처리
* Kinesis Agent로부터 수신한 데이터를 임시 저장 후 Firehose로 전송
* 자동으로 데이터를 처리하고 실패한 데이터의 재처리를 보장

**Kinesis Firehose**

* 스트리밍 데이터를 데이터 레이크, 데이터 저장소 및 분석 서비스에 로드
* 데이터 최적화 및 집합 생성을 위해 ETL 작업을 수행하여 데이터를 변환, 준비 및 로드

**Lambda**

* 온디맨드로 구성한 Glue Crawler를 설정한 작업에 따라 실행시킴
* Post 요청으로 온 http body 부분을 추출해 S3에 저장

**AWS Glue**

* 완전 관리형으로 데이터 추출, 변환, 로드 구성
* 데이터 분석 파이프라인의 각 단계에서 데이터 전처리 및 정제 작업을 자동화
* 데이터 카탈로그를 생성하고 데이터 소스의 메타데이터 관리
* 데이터 품질 검사 기능을 통해 데이터의 일관성과 정확성을 검사하여 이상 데이터를 식별
* 온디맨드로 Crawler 구성하여 lambda로 실행시켜 작동 스케줄을 설정

**AWS Athena**

* 서버리스 쿼리 서비스로, S3에 저장된 데이터를 SQL을 사용해 분석
* S3에 저장된 데이터를 간편하게 분석
* 데이터 파이프라인의 각 단계에서 생성된 데이터를 Athena를 통해 쿼리하여 필요한 정보를 추출

**AWS QuickSight**

* 비즈니스 인텔리전스 서비스로, 데이터 시각화 및 대시보드 작성을 제공
* 분석된 데이터를 시각화하여 비즈니스 인텔리전스를 실시간으로 제공

**AWS S3**

* 객체 스토리지 서비스로, 수집 및 분석 데이터를 안전하게 저장하고 검색할 수 있는 서비스 제공
* 다양한 데이터 형식을 지원하므로 다양한 데이터 소스로부터 데이터를 수집하고 통합하기 위한 유연성을 충족시킴

**Route 53**

* 사용자 도메인 관리 및 인터넷 트래픽을 ALB로 라우팅
* HTTPS를 통한 안전한 연결 위해 DNS 설정 및 SSL/TLS 인증서와 연동

**AWS Certificate Manager**

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

**AWS WAF**

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

**AWS Shield**

* DDOS 공격으로부터 AWS 리소스를 보호