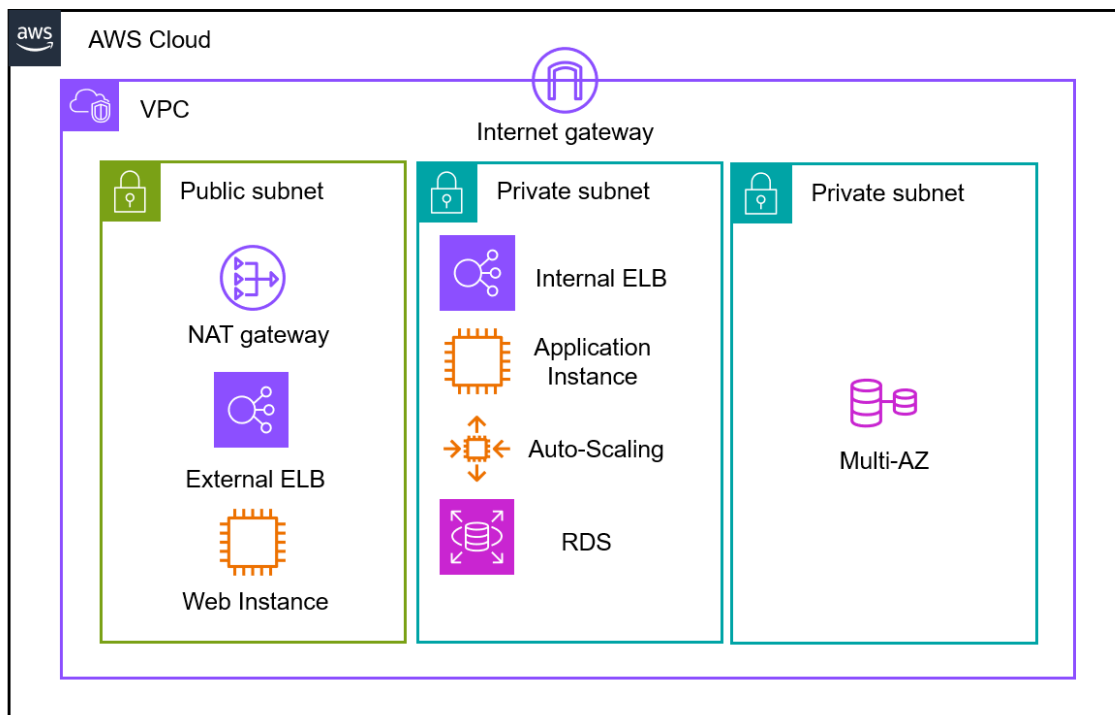


# 로드 밸런서와 오토 스케일링을 적용해 확장 가능한 애플리케이션 환경 만들기

확장성과 안정성을 높이기 위해 분리된 아키텍처에 로드 분산과 자동 확장을 적용해 보자

로드 분산 및 오토 스케일링을 통한 안정된 비즈니스 애플리케이션을 서비스하기 위해 제안하는 보고서입니다. 이렇게 개선된 구조는 단순한 성능 향상을 넘어 운영 효율성과 신뢰성을 모두 높이는 방향으로 나아가게 됩니다.



## 1. 트래픽 처리 효율 향상

- External ELB(Public-facing ALB)를 통해 클라이언트 요청을 여러 Web EC2에 분산 처리하여 고부하 상황에서도 안정적인 응답을 제공합니다.
- Web EC2는 내부 요청을 Internal ELB를 통해 Application EC2에 전달함으로써 티어 간 트래픽 분산이 이루어집니다.
- Path 기반 라우팅을 통해 /api/, /static/\* 등의 URI를 분석하고 요청을 각각 Web Tier 또는 App Tier로 정확히 전달합니다.
- SSL 종료를 External ELB에서 처리하여 EC2 인스턴스의 부하를 줄이고 전체 보안성을 강화합니다.

## 2. 탄력적인 확장성 확보

- Web 및 Application EC2는 각자 Auto Scaling Group에 의해 관리되어 트래픽에 따라 자동으로 인스턴스가 증감합니다.
- 갑작스러운 트래픽 급증에도 Application Tier의 유연한 수평 확장이 가능하며 부하가 적을 경우 인스턴스를 축소하여 비용 효율성을 극대화할 수 있습니다.
- Internal ELB가 Auto Scaling 대상 인스턴스로 동적으로 트래픽을 분배해 부하 균형을 유지합니다.

### 3. 서비스 장애 대응력 강화

- External 및 Internal ELB 모두 헬스 체크를 통해 비정상 인스턴스를 자동으로 제외합니다.
- Auto Scaling은 장애가 발생한 EC2를 감지하여 신규 인스턴스를 자동으로 생성함으로써 셀프 복구 구조를 갖추고 있습니다.
- ELB를 통한 로드 밸런싱과 Auto Scaling 조합은 특정 인스턴스 장애에도 전체 서비스의 고가용성을 확보합니다.

### 4. 운영 자동화 및 CI/CD 적용 용이

- ELB 및 Auto Scaling은 블루/그린 배포, 롤링 업데이트 등 다양한 배포 전략과 연동 가능하며 CI/CD 환경에 최적화되어 있습니다.
- EC2 인스턴스를 태그 기반 또는 이미지 기반으로 관리하며 업데이트 및 배포 시 운영자 개입 없이 자동 처리할 수 있어 운영 효율성이 높습니다.
- 배포 중에도 ELB가 정상 인스턴스로만 트래픽을 전달하므로 무중단 배포가 가능합니다.

### 5. 데이터 베이스 고가용성 확보

- RDS는 Primary 인스턴스가 Private Subnet 1, Standby 인스턴스가 Private Subnet 2에 위치하여 Multi-AZ 고가용성 구조를 갖추니다.
- 서로 다른 AZ 간 동기식 복제를 통해 실시간 데이터 일관성을 유지하고, 장애 발생 시 자동 Failover를 수행해 서비스 중단을 방지합니다.
- 백업 작업은 Standby 인스턴스에서 수행되므로 Primary 인스턴스의 처리 성능에 영향을 주지 않습니다.
- 전체적으로 Multi-AZ RDS 구성을 통해 데이터 내구성과 안정성을 확보하며 운영 중단 없이 안정적인 서비스를 제공합니다.