



Google Developer Group
INCHEON NATIONAL UNIVERSITY

aws 기반 고가용성 아키텍처 설계

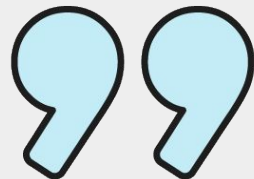
GDGoC_INU 백엔드 이윤서

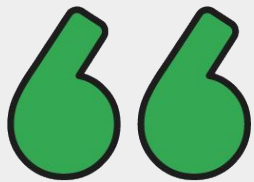


가용성(Availability)

서비스 운영 중 서비스가 정상적으로 동작하는
비율

$$A = \frac{uptime}{uptime + downtime}$$

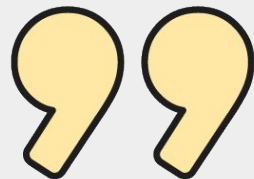




고가용성(High Availability)

서비스의 다운타임을 제거하거나 최소화하여
거의 **100%** 상시 액세스 가능하고 신뢰성을 유지하는
능력.

말그대로 높은 가용성을 보장하는 상태.



Availability %	Downtime per year ^[note 1]	Downtime per quarter	Downtime per month	Downtime per week	Downtime per day (24 hours)
90% ("one nine")	36.53 days	9.13 days	73.05 hours	16.80 hours	2.40 hours
95% ("one and a half nines")	18.26 days	4.56 days	36.53 hours	8.40 hours	1.20 hours
97% ("one and three quarters nines")	10.96 days	2.74 days	21.92 hours	5.04 hours	43.20 minutes
98% ("one and seven eights nines")	7.31 days	43.86 hours	14.61 hours	3.36 hours	28.80 minutes
99% ("two nines")	3.65 days	21.9 hours	7.31 hours	1.68 hours	14.40 minutes
99.5% ("two and a half nines")	1.83 days	10.98 hours	3.65 hours	50.40 minutes	7.20 minutes
99.8% ("two and seven eighths nines")	17.53 hours	4.38 hours	87.66 minutes	20.16 minutes	2.88 minutes
99.9% ("three nines")	8.77 hours	2.19 hours	43.83 minutes	10.08 minutes	1.44 minutes
99.95% ("three and a half nines")	4.38 hours	65.7 minutes	21.92 minutes	5.04 minutes	43.20 seconds
99.99% ("four nines")	52.60 minutes	13.15 minutes	4.38 minutes	1.01 minutes	8.64 seconds
99.995% ("four and a half nines")	26.30 minutes	6.57 minutes	2.19 minutes	30.24 seconds	4.32 seconds
99.999% ("five nines")	5.26 minutes	1.31 minutes	26.30 seconds	6.05 seconds	864.00 milliseconds
99.9999% ("six nines")	31.56 seconds	7.89 seconds	2.63 seconds	604.80 milliseconds	86.40 milliseconds
99.99999% ("seven nines")	3.16 seconds	0.79 seconds	262.98 milliseconds	60.48 milliseconds	8.64 milliseconds
99.999999% ("eight nines")	315.58 milliseconds	78.89 milliseconds	26.30 milliseconds	6.05 milliseconds	864.00 microseconds
99.9999999% ("nine nines")	31.56 milliseconds	7.89 milliseconds	2.63 milliseconds	604.80 microseconds	86.40 microseconds

고가용성 설계의 필요성

[2022 국내 10대뉴스]유례없는 '카카오 먹통 사태'

발행일 : 2022-12-29 16:00 지면 : 2022-12-30  12면

  가 

SK(주) C&C 판교 데이터센터에서 발생한 화재로 지난 10월 15일 오후 카카오톡을 비롯한 카카오 주요 서비스가 전면 중단되는 사상 초유의 사태가 발생했다. 모든 서비스가 안정화될 때까지 걸린 시간은 127시간 33분으로, 플랫폼 업계 '역대 최장 서비스 장애'라는 오명을 기록했다.

AWS 클라우드 장애, 기업 고객들 '타격'

아이뉴스24 원문 | 기사전송 2020-11-26 12:19

 0

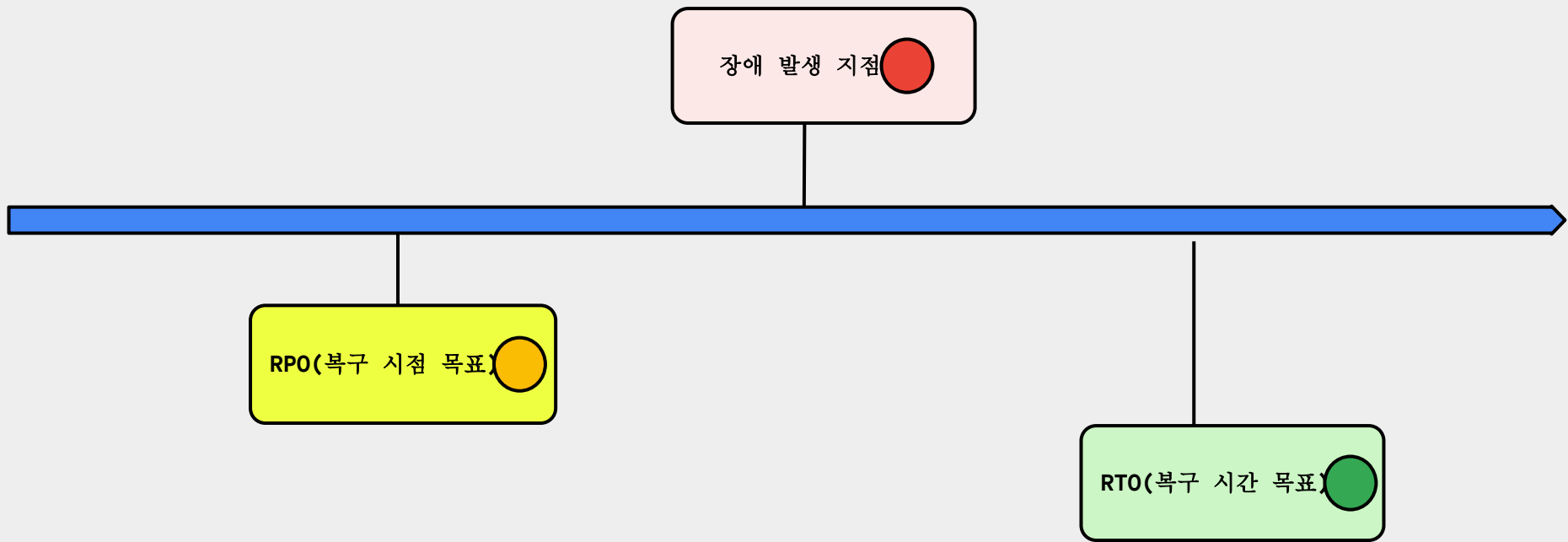
 0

AI챗으로 요약



| 동부 리전 장애...어도비 등 각종 서비스 줄줄이 먹통

RTO / RPO



고가용성 확보를 위한 전략

→ 고가용성 아키텍처 설계

Network 계층 : 멀티리전 구성 (Multi-AZ)

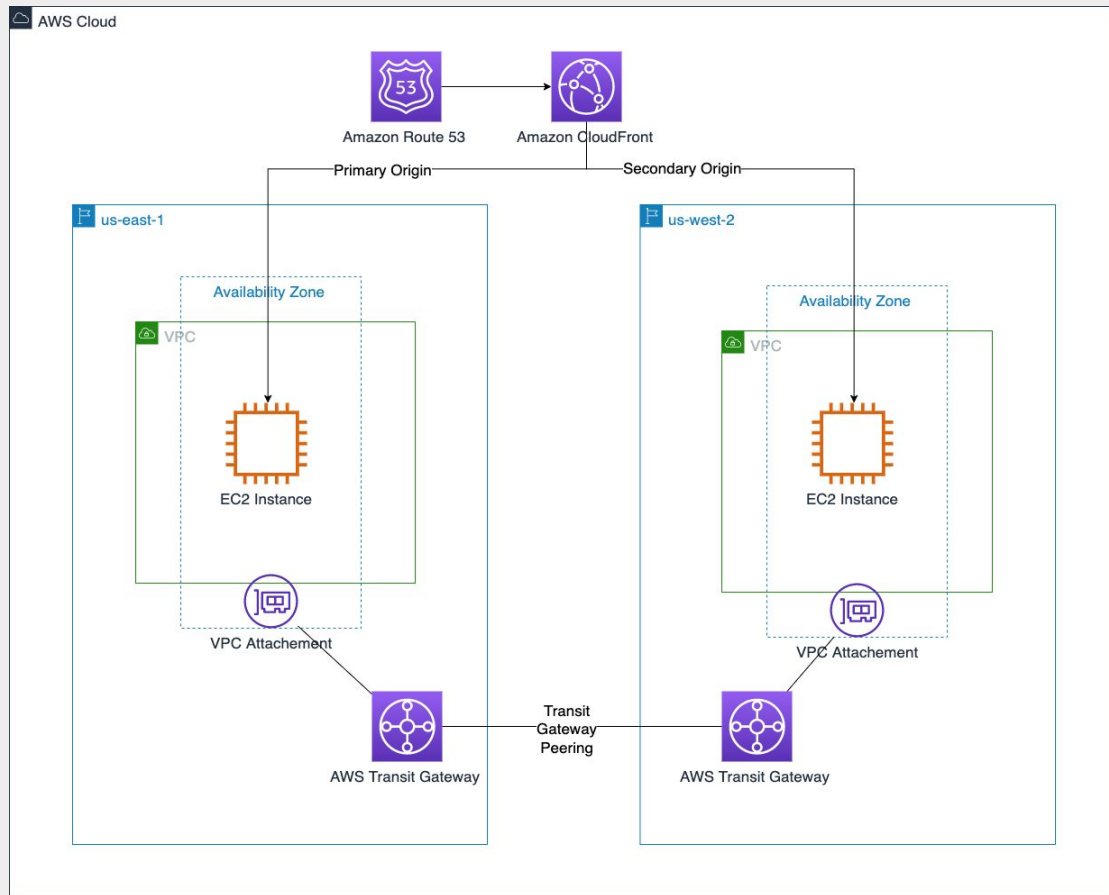
Compute 계층 : Auto Scailing + LoadBalancer

Data 계층 : DB 이중화

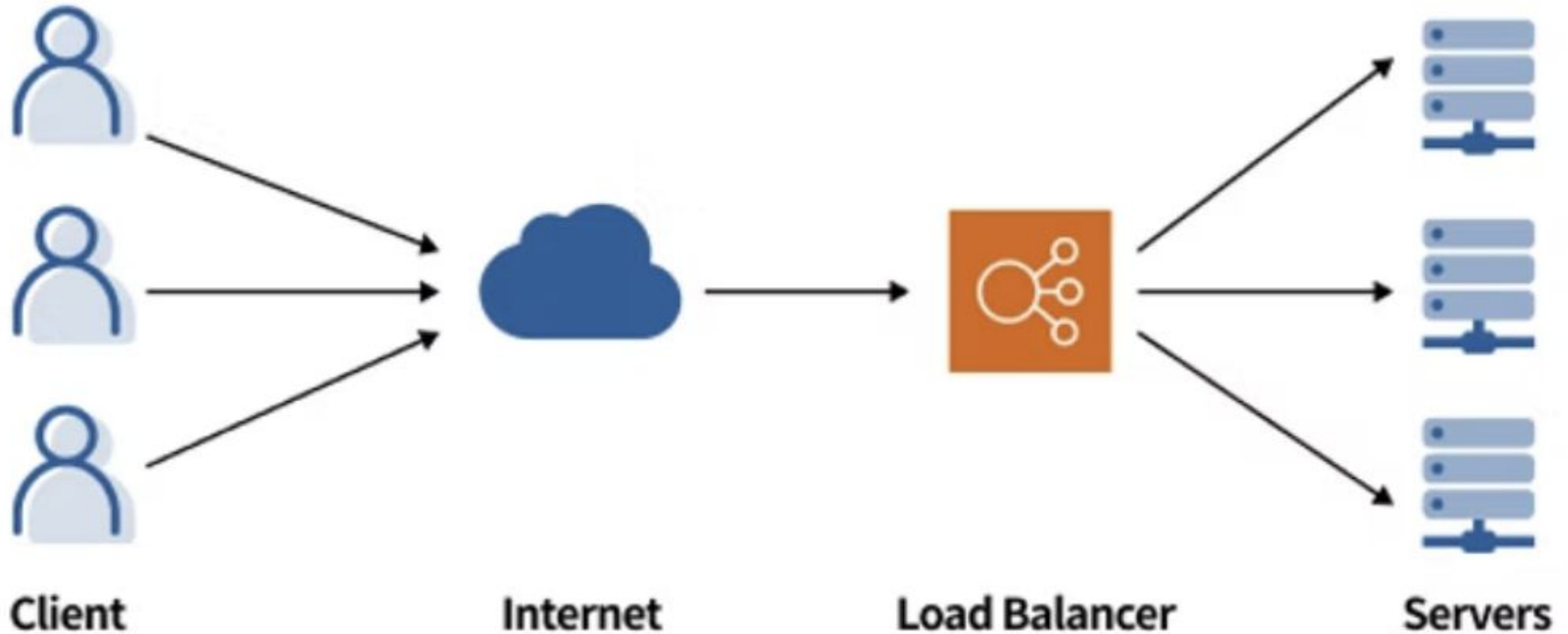
Application 계층 : Stateless 구조



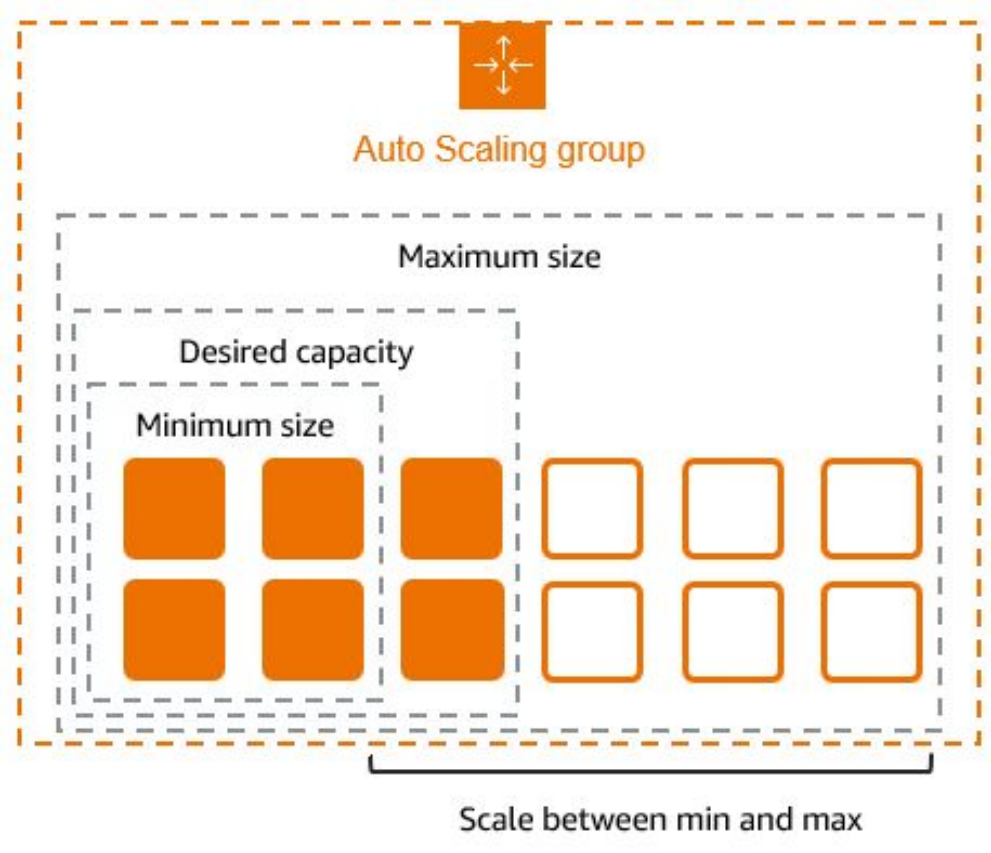
1. 멀티리전 구성(Multi-AZ)



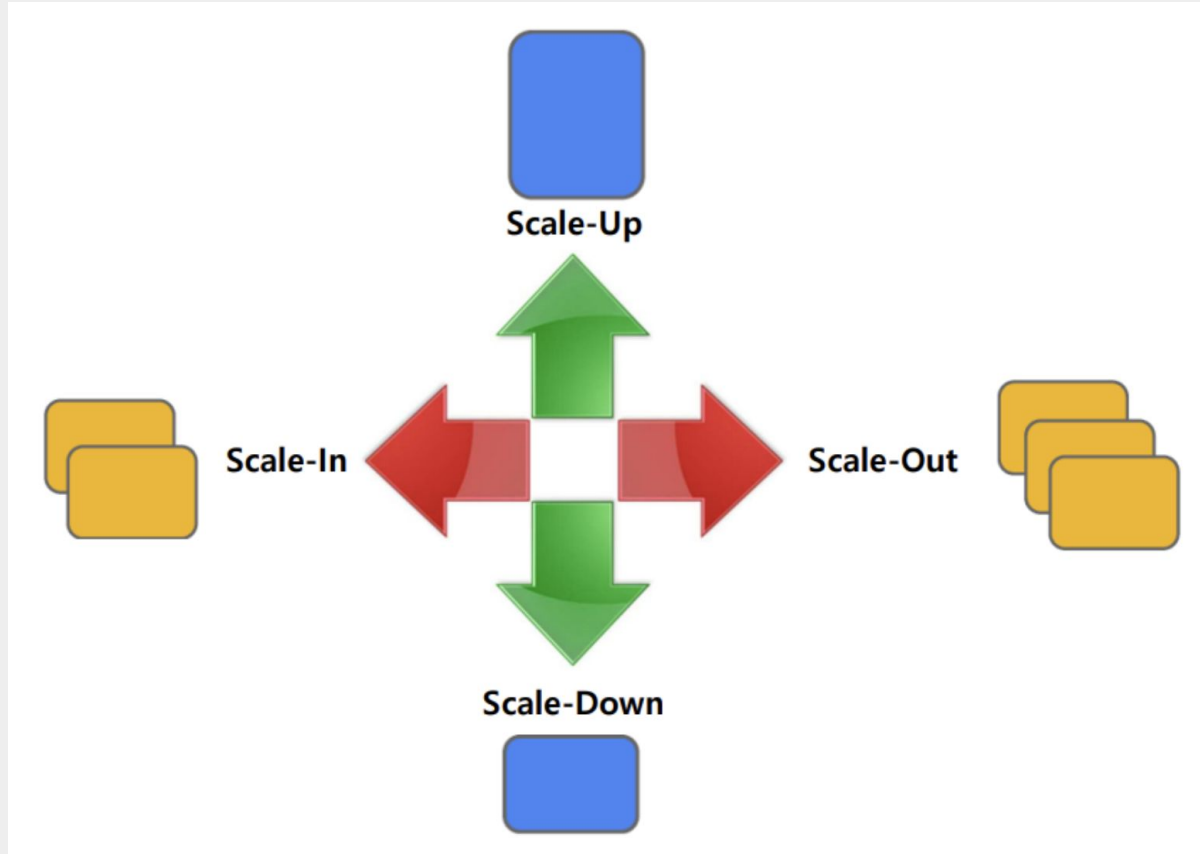
2. Load Balancer + Auto Scaling



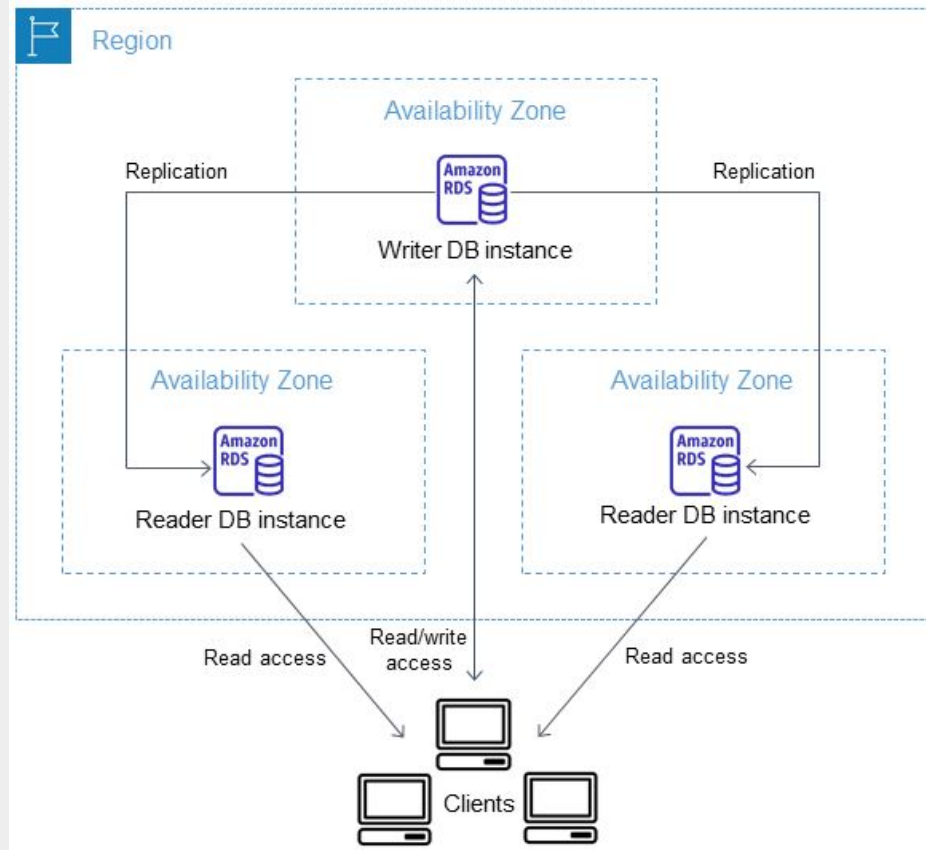
Load Balancer + Auto Scaling



Load Balancer + Auto Scaling



3. RDS Multi-AZ



4. Stateless 구성

:고가용성 구조를 쉽게 만들 수 있도록 도와주는 설계 방식

서버가 요청 간에 아무런 사용자 상태를 저장하지 않음

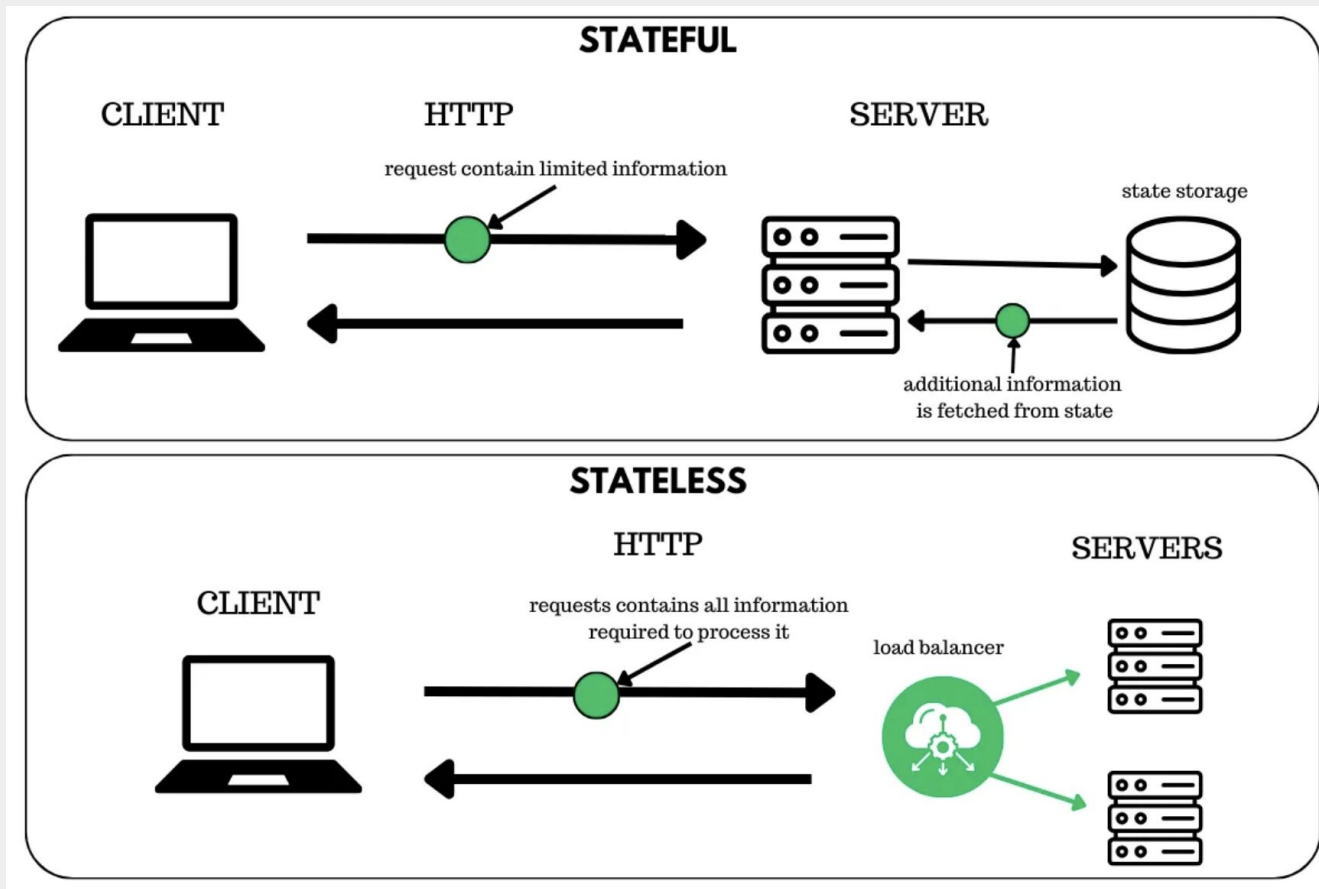


그렇기 때문에 요청을 어떤 서버가 처리하든 결과가
일관됨

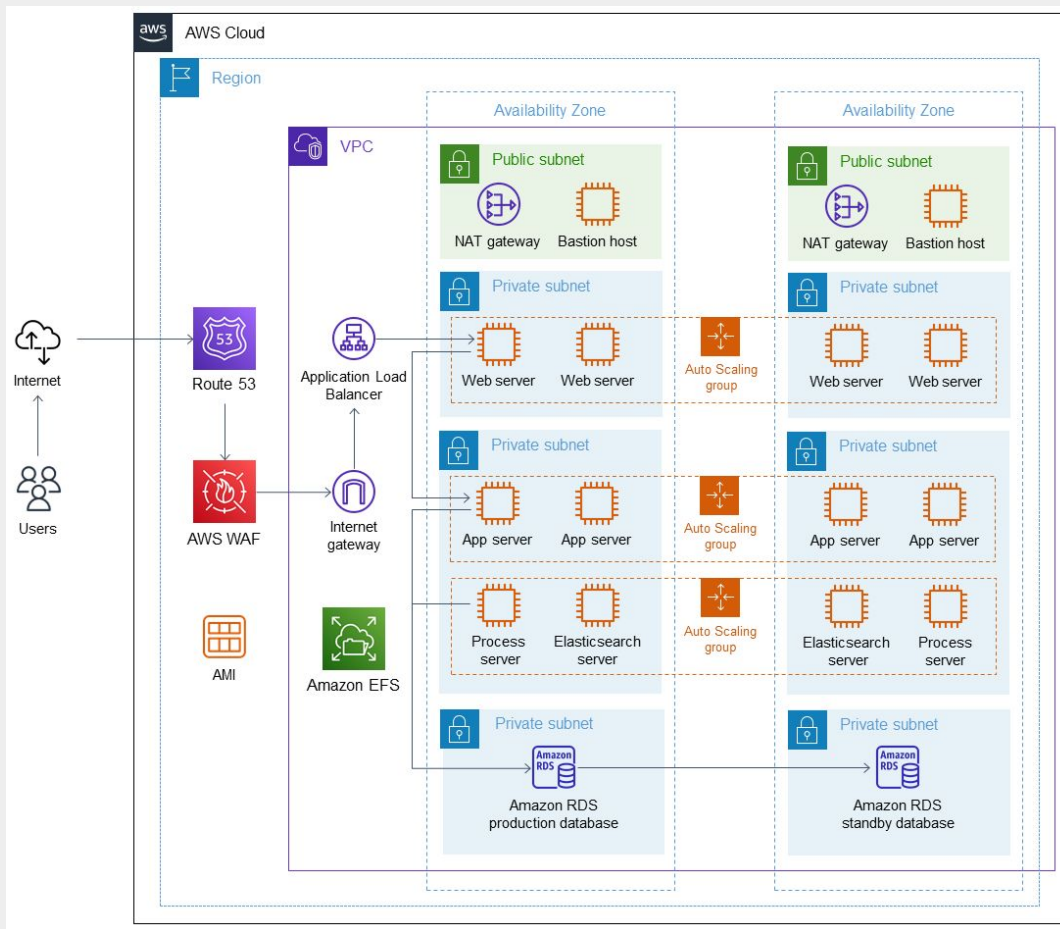


즉, 특정 서버가 죽어도, 다른 서버가 아무 문제 없이
대체 가능

4. Stateless 구성 - Stateless vs Stateful



고가용성 아키텍처



7년 걸쳐 AWS로 이전'... 넷플릭스 엔지니어가 설명하는 클라우드 여정

기획

2018.09.12 · 8분

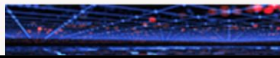
클라우드 컴퓨팅

데이터센터

IT 전략

넷플릭스는 1억 3,000만 명에서 비디오 스트리밍을 제공하기까지 인프라를 어떻게 구축해왔을까?
넷플릭스의 선임 엔지니어 데이브 한이 설명한다.

AD



쿠팡, IT인프라 전체를 AWS 클라우드로 이전했다

2017.08.10 16:00

가 가

쿠팡이 자사의 IT인프라를 100% 클라우드 기반으로 전환했다. 쿠팡은 이커머스 서비스 전체의 클라우드 이전을 완료했다고 10일 밝혔다. 쿠팡이 선택한 클라우드는 아마존웹서비스(AWS)다.