

Auto Scaling, Load Balancing

임승훈

목차

- 스케일링이란
- 오토 스케일링
- 오토 스케일링의 장단점
- 로드밸런싱
- 로드밸런싱 종류
- 로드밸런서 방법
- 로드밸런싱 유형
- 로드밸런서 종류

스케일링이란?

인스턴스, 컴퓨팅 파워를 늘리는 것

스케일링의 방법

스케일 업(Scale up)

스케일 아웃(Scale out)

Scale Up

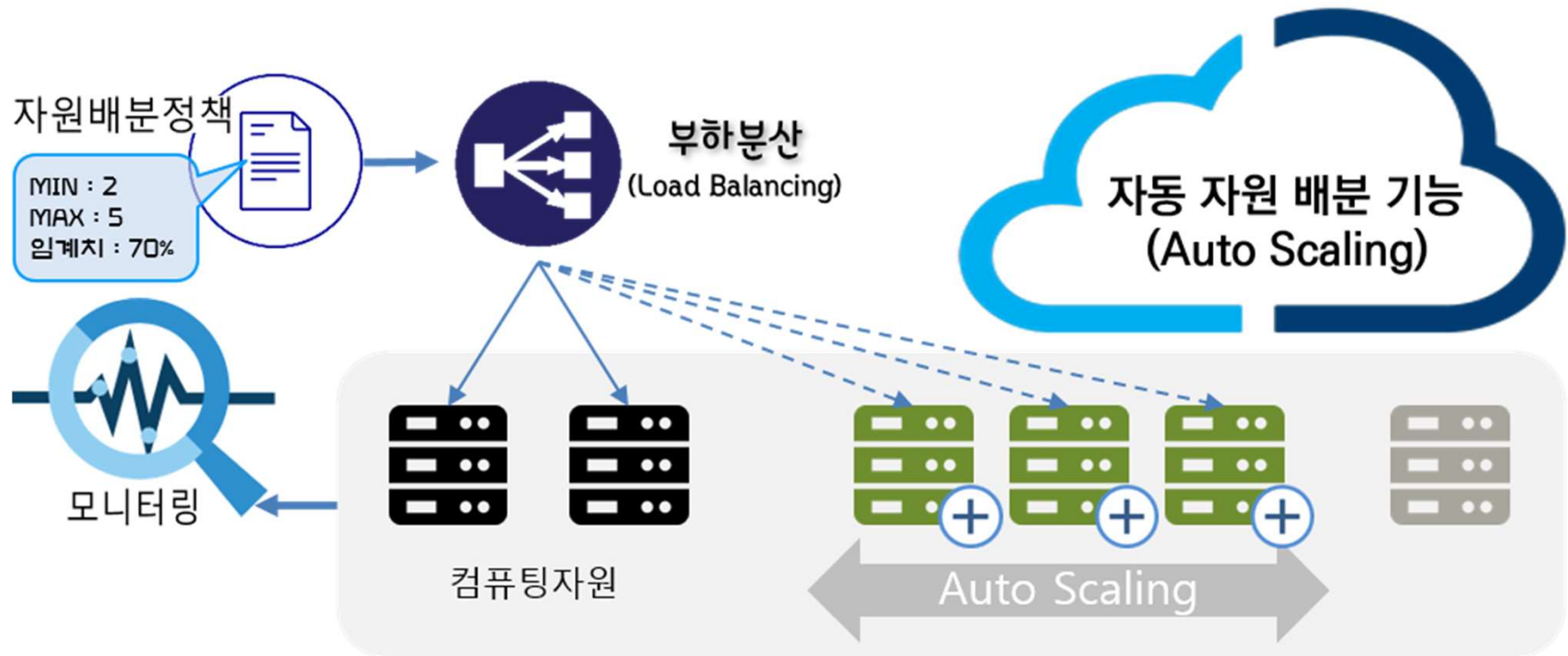


Scale Out



오토 스케일링이란?

클라우드의 유연성을 돋보이게 하는 기술



오토스케일링 장단점

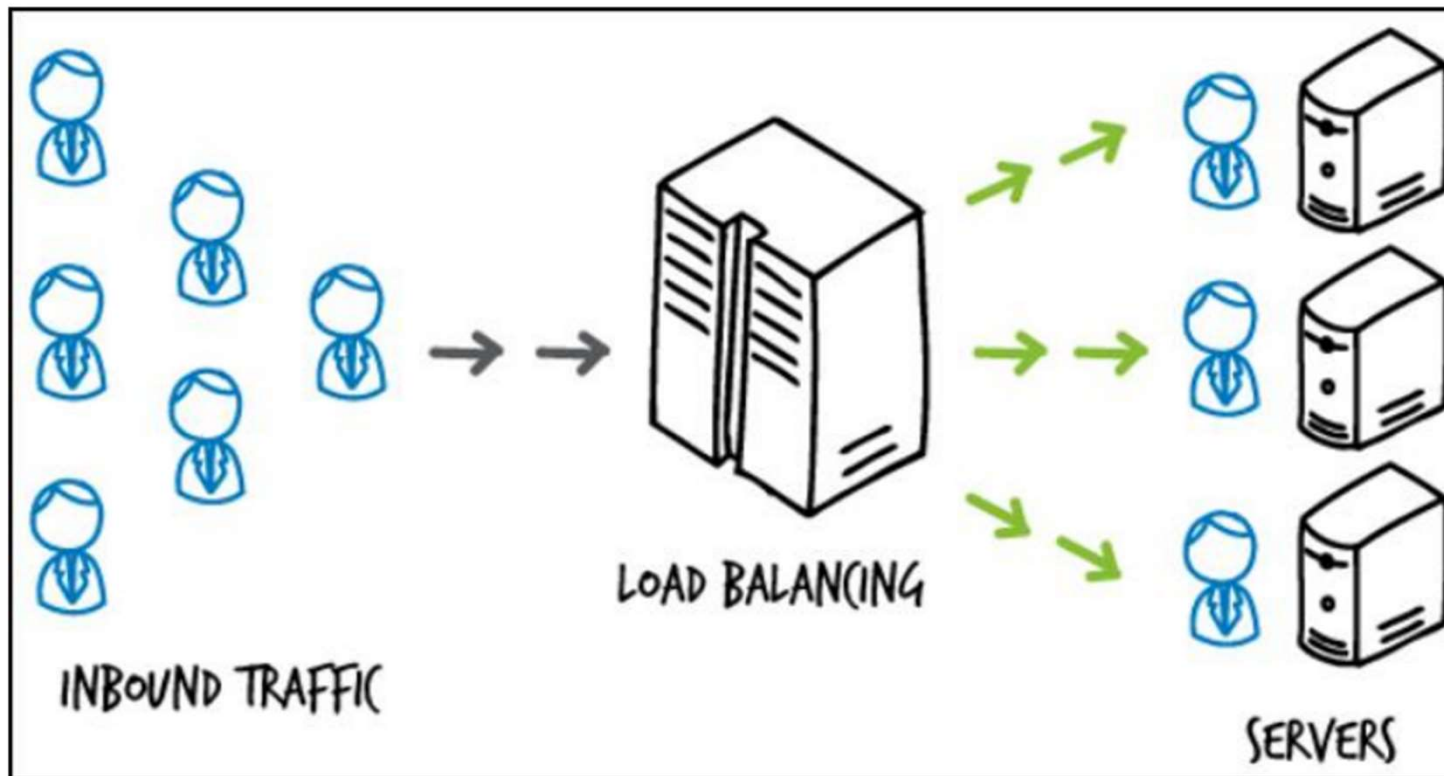
• 장점

- 1.비용 효율성
- 2.고가용성
- 3.손쉬운 관리
- 4.안전성 향상
- 5.관리 시간감소

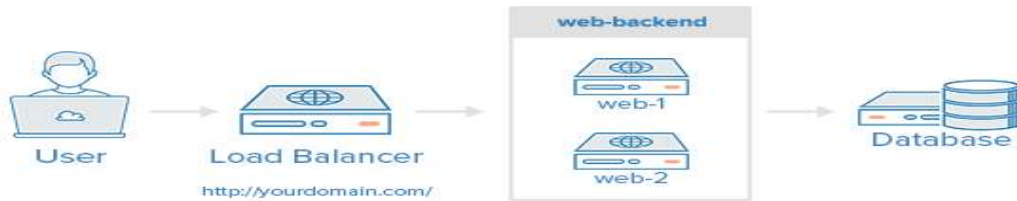
• 단점

1. 일관되지 않은 성능
2. 비용증가

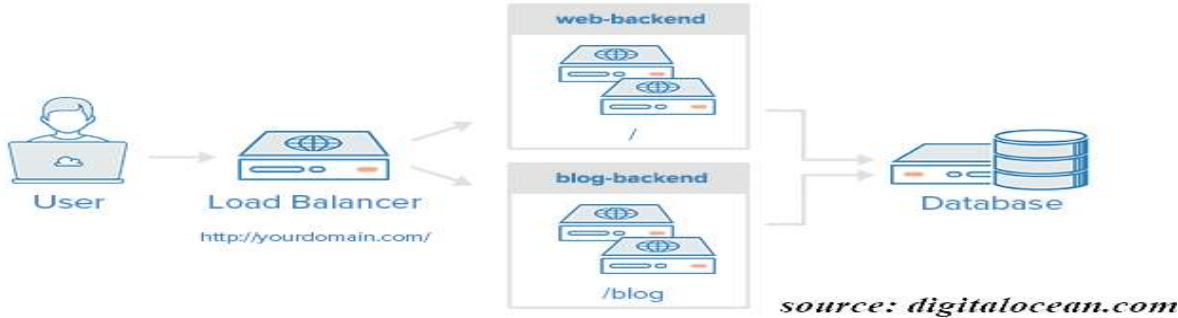
로드밸런싱



Layer 4 Load Balancing



Layer 7 Load Balancing



• L4 Load Balancer

IP, Port 를 기준으로 스케줄링 알고리즘을 통해 부하를 분산합니다, 클라이언트에서 로드밸런서로 요청을 보냈을 때 최적의 서버로 요청을 전송하고 결과를 클라이언트에게 준다.

L4 Load Balancer는 단지 부하를 분산시킴, L7 Load Balancer는 요청의 세부적인 사항을 두고 서버를 분리해서 가볍고 작은 단위로 여러 개의 서비스를 운영, 요청을 각각의 서버에 분산함

로드밸런싱 종류

로드밸런서: 로드 밸런싱 기술을 제공하는 서비스,장치

• L7 Load Balancer

L7 위에서 동작하기 때문에 IP, Port 이외에도 URI, Payload, Http Header 등의 내용을 기준으로 부하를 분산합니다.

로드밸런서 방법

- L4

라운드로빈 방식: 순서대로 배정

가중 라운드로빈 방식: 가중치를 매기고 배정

최소 연결 방식: 가장 연결이 적은 서버에 트래픽을 배정

- L7

URL 스위칭 방식: 특정 하위 URL들은 특정 서버로 처리

컨텍스트 스위칭 방식: 클라이언트가 요청한 특정 리소스에 따라 특정서버로 이동

로드밸런싱 유형

- 하드웨어 로드밸런싱: 하드웨어 장비를 이용
- 소프트웨어 로드밸런싱: 관리자가 모니터링후 분배
- DNS 기반 로드밸런싱: 도메인 이름을 IP주소로 변환하는 기술

로드밸런서 종류

- ALB(Application Load Balancer)

HTTP,HTTPS 트래픽 처리에 적합한 로드 밸런서

scale-out을 할때 유용하게 사용

- CLB(Classic Load Balancer)

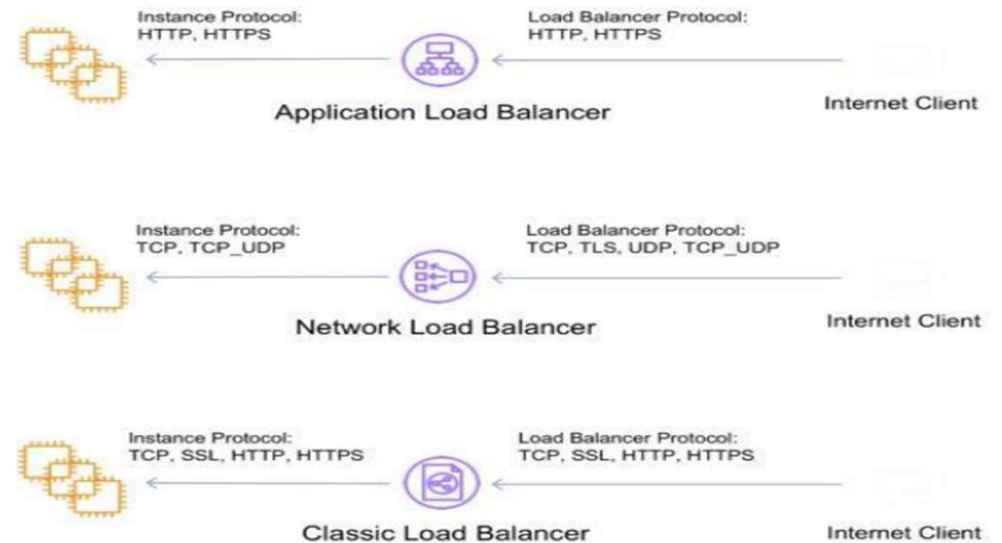
HTTP, HTTPS, TCP, TLS 트래픽을 처리가 가능하다.

가장 오래된 로드 밸런서

- NLB(Network Load Balancer)

TCP,UDP 및 TLS 트래픽 처리에 적합한 로드 밸런서

다른 로드 밸런서보다 속도적인 면에서 가장 빠르다.



감사합니다