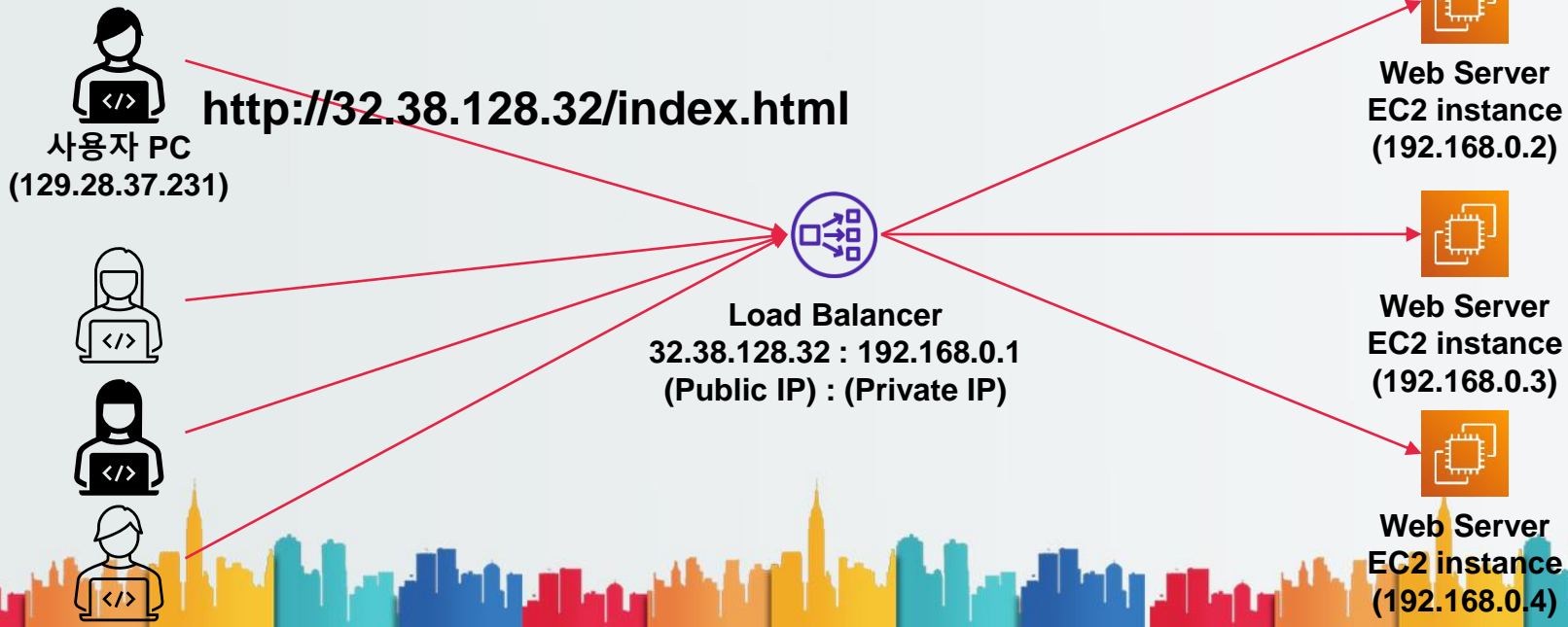


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing)

- 네트워크 트래픽을 하나 이상의 서버나 장비로 **분산**하기 위해 사용되는 기술
- S/W나 H/W를 통해 로드 밸런싱을 수행할 수 있음
- 서비스 사용자의 요청을 서버로 분산해서 전달

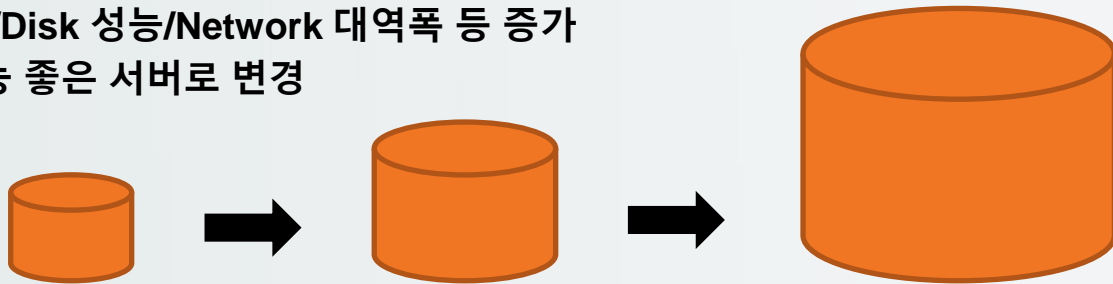


# ELB(Elastic Load Balancing)

- 웹 트래픽 증가에 따른 대응 방법

- 1. Scale Up

- CPU/RAM/Disk 성능/Network 대역폭 등 증가
    - 비싸고 성능 좋은 서버로 변경



- 2. Scale Out (로드 밸런싱과 함께 활용)

- 부하를 처리할 서버 대수를 늘림
    - 저렴한 서버 여러 대를 이용해 더 많은 부하를 감당

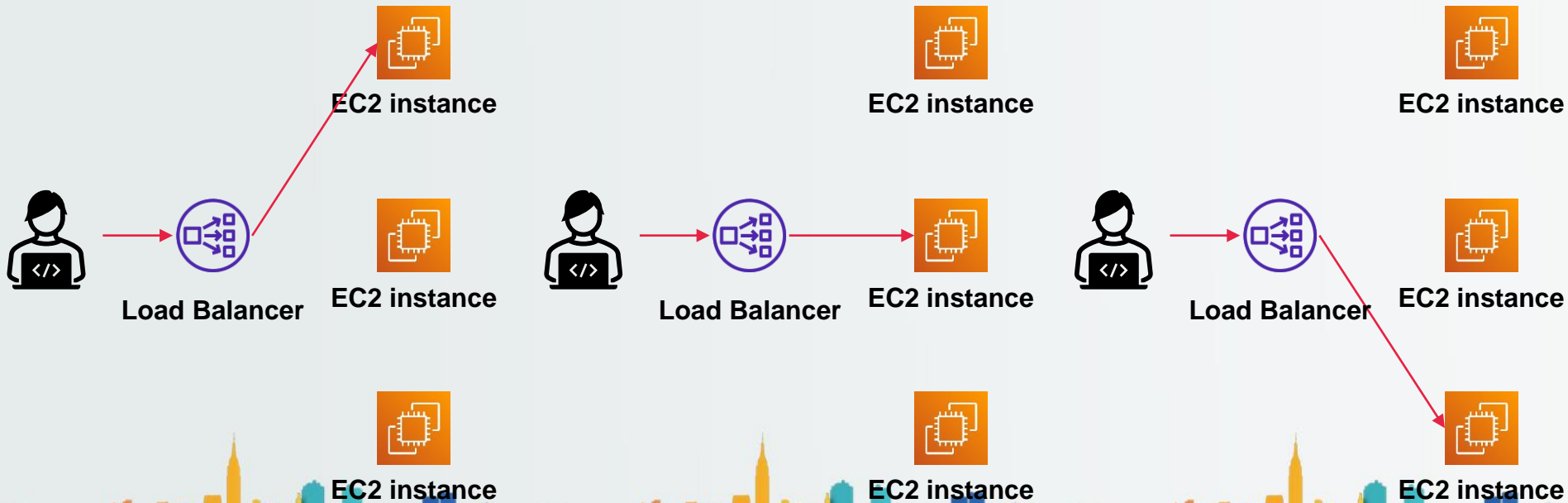


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- 1. Round Robin

- 로드 밸런서에서 서버 선택 시, 다음 순차적으로 서버를 선택하는 방법

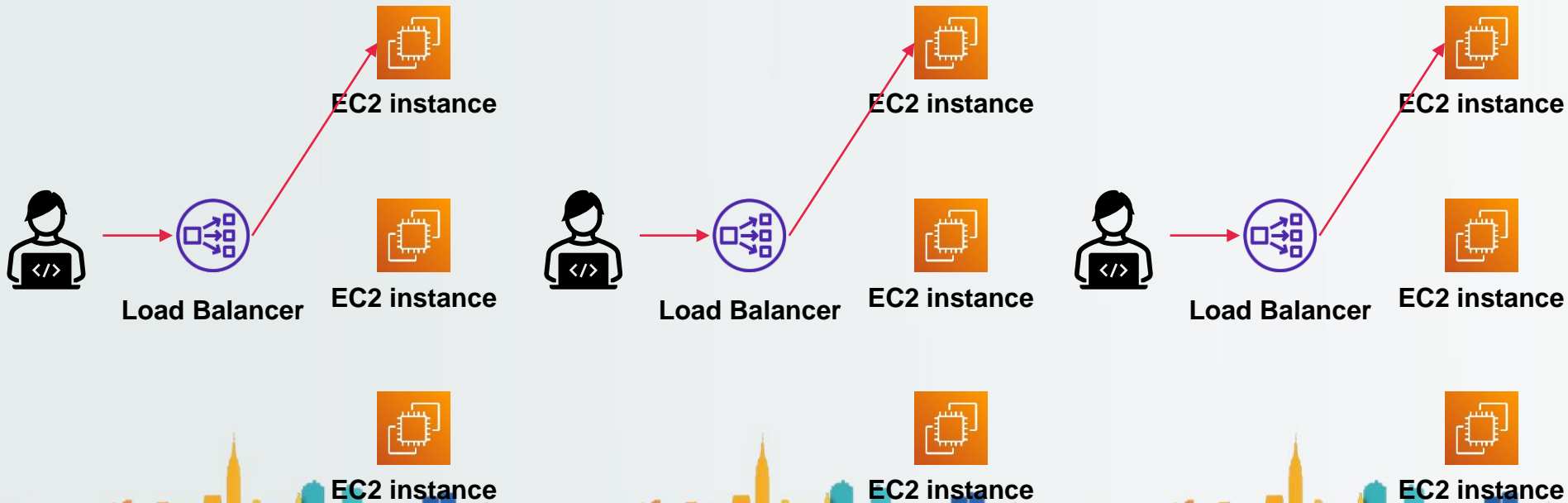


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- **2. Hash (== Sticky Session)**

- 로드 밸런서에서 서버 선택 시, 클라이언트가 서버와 한 번 연결되면, 그 이후에는 항상 같은 서버로 연결하는 방식

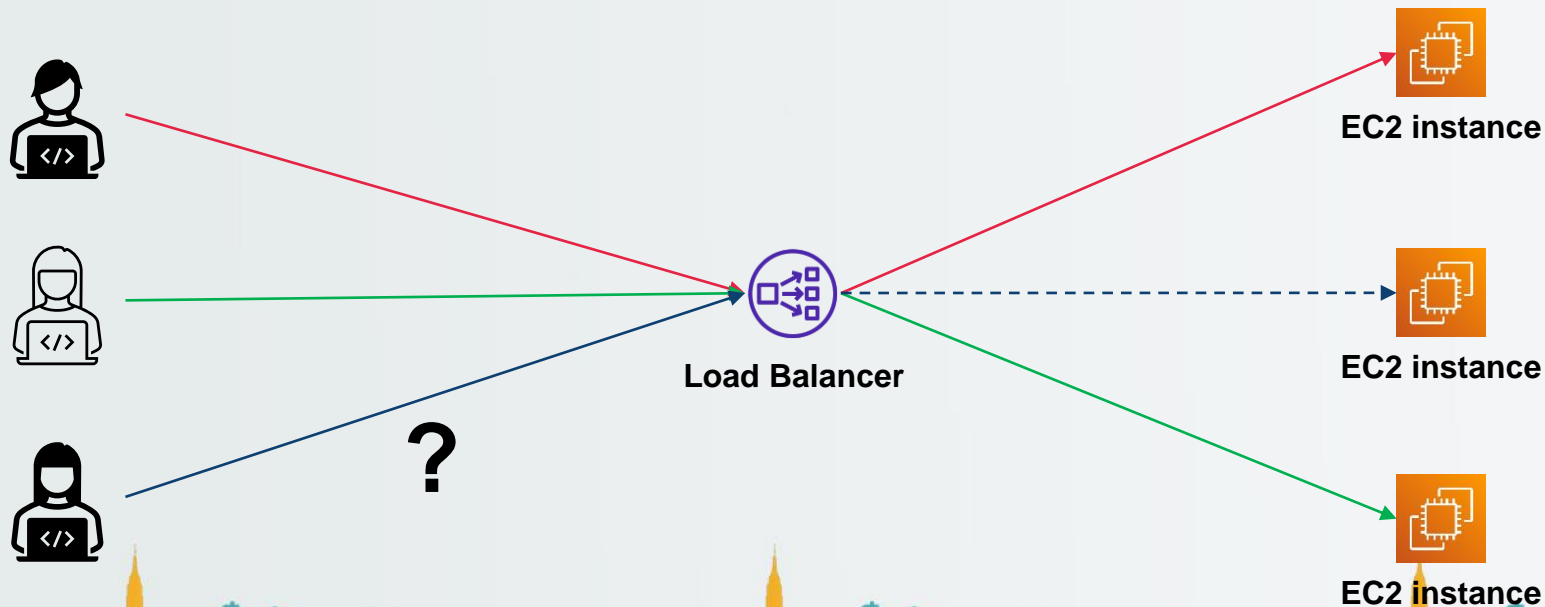


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- 3. Least Connection

- 로드 밸런서에서 서버 선택 시, 연결 수가 가장 적은 서버를 선택

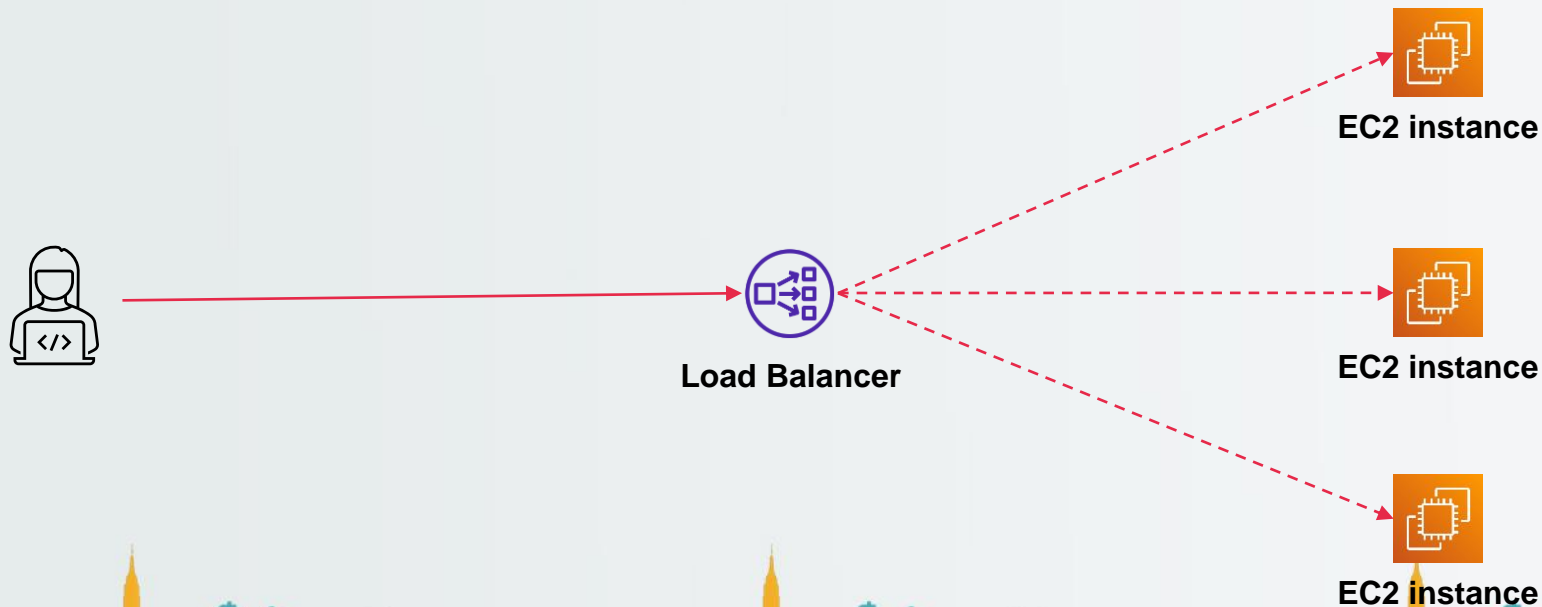


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- 4. 응답 시간(Response Time) (1/3)

- 로드 밸런서에서 서버 선택 시, 응답속도가 가장 빠른 서버를 선택

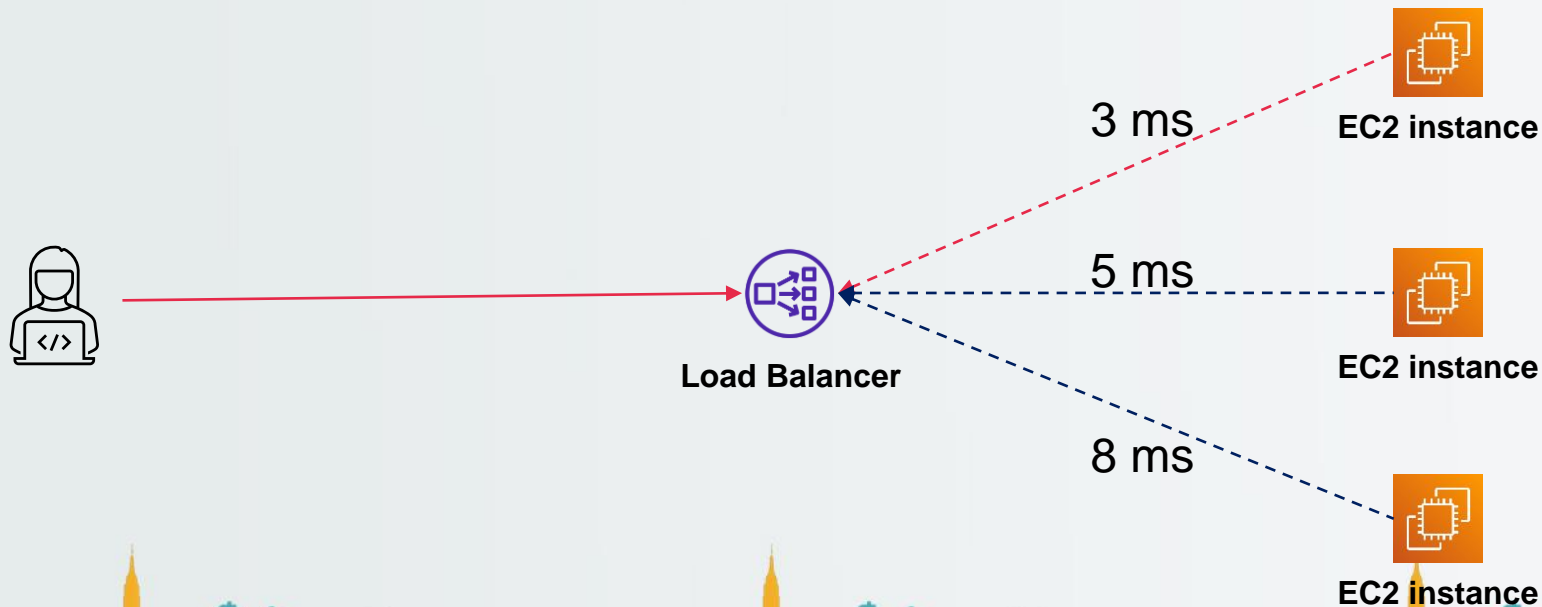


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- 4. 응답 시간(Response Time) (2/3)

- 로드 밸런서에서 서버 선택 시, 응답속도가 가장 빠른 서버를 선택

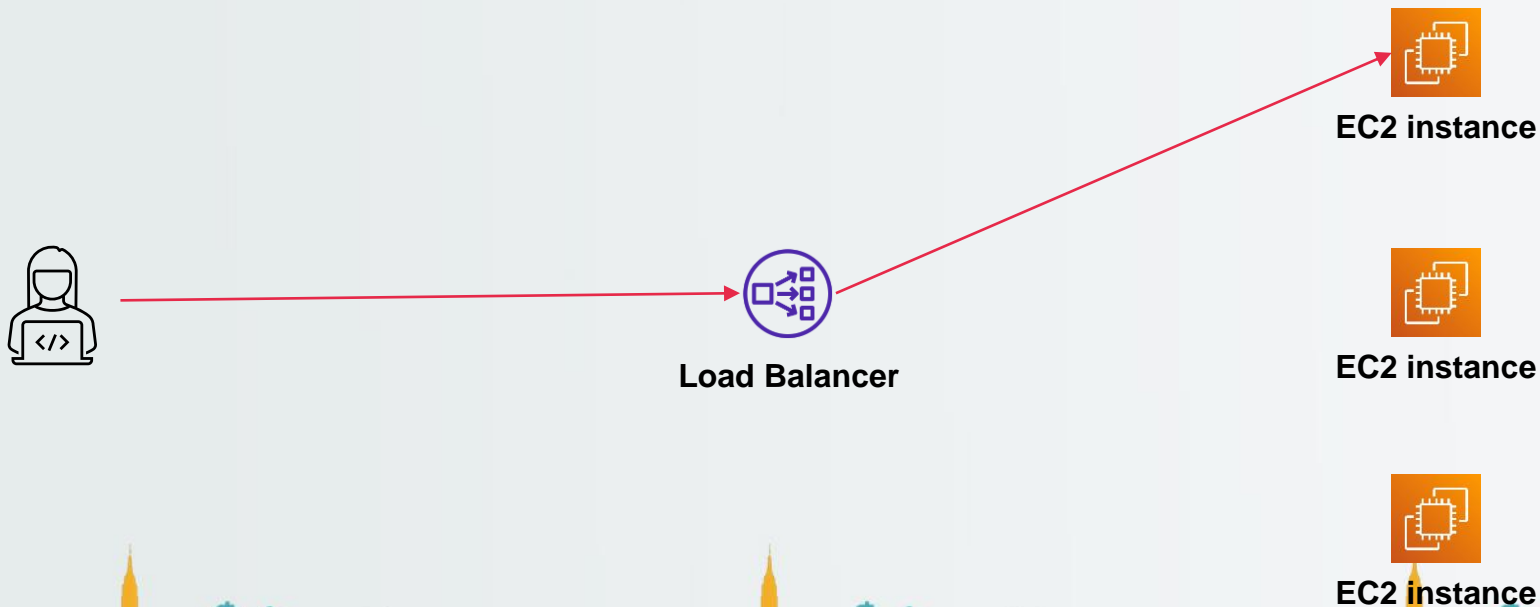


# ELB(Elastic Load Balancing)

- 로드 밸런싱(Load Balancing) 방식

- 4. 응답 시간(Response Time) (3/3)

- 로드 밸런서에서 서버 선택 시, 응답속도가 가장 빠른 서버를 선택





# ELB(Elastic Load Balancing)

- **ELB 유형**

- **1. ALB(Application Load Balancer)**

- OSI 7 Layer 중 Application Layer에 속하는 패킷을 처리할 수 있음
    - HTTP/HTTPS 프로토콜

- **2. NLB(Network Load Balancer)**

- OSI 7 Layer 중 Transport Layer에 속하는 패킷을 처리할 수 있음
    - TCP 프로토콜

- **3. CLB(Classic Load Balancer)**

- OSI 7 Layer 중 Network 및 Transport Layer에 속하는 패킷을 처리할 수 있음
    - 과거에 사용되던 방식으로 호환성을 위해 유지되는 상태



# ELB(Elastic Load Balancing)

- **ELB 옵션**

- 인터넷 트래픽용 vs 내부 네트워크 트래픽용
  - 1. 인터넷 트래픽용 (Internet-Facing)
    - 인터넷을 통해 ELB에 접근하는 경우
      - » 즉, Public IP를 가진 클라이언트에서 접근하는 경우
    - 단, ELB가 Private IP:Public IP 쌍을 가지므로, Private IP 클라이언트에서도 사용 가능
  - 2. 내부 네트워크 트래픽용
    - 내부 네트워크에서 ELB에 접근하는 경우
      - » 즉, Private IP를 가진 클라이언트에서 접근하는 경우
    - 이 경우는 주로 전용선을 통해 인터넷을 거치지 않고 VPC에 연결되는 경우나 혹은 VPN을 통해 Private IP를 이용해서 통신하는 경우에 활용됨

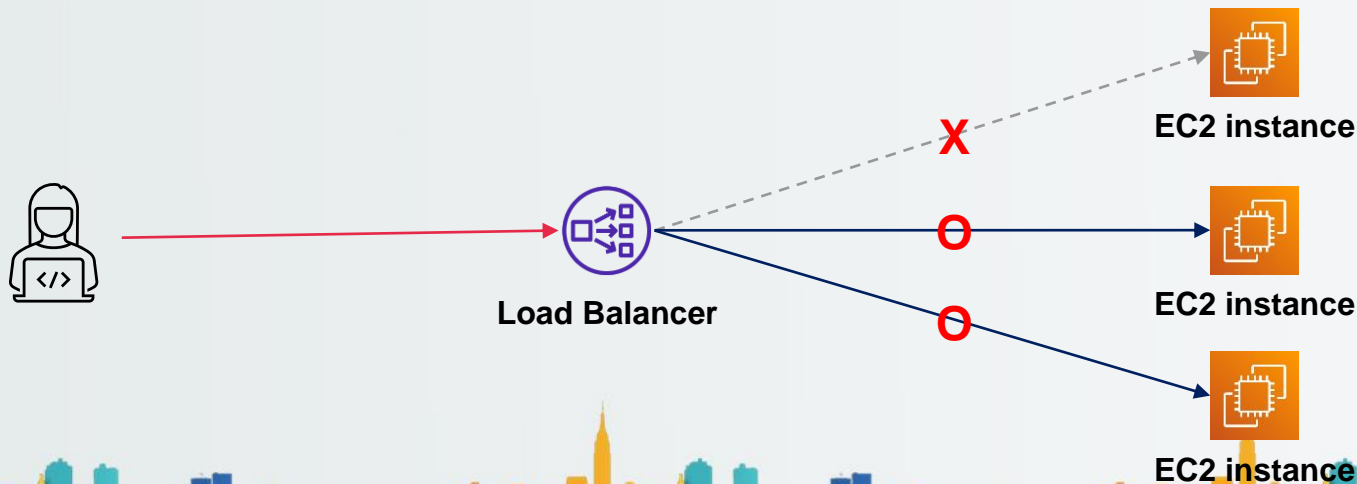


# ELB(Elastic Load Balancing)

- ELB 동작 특징

- 상태 확인(Health Check) 서비스

- ELB에서 연결할 서버들의 상태를 상시 체크
    - 비정상 상태의 서버로 트래픽을 전달하지 않음
    - ELB에서 서버 측에 요청을 전달해서 응답을 받음으로써 처리되므로, 서버의 보안그룹에서 해당 요청을 받을 수 있도록 지정된 포트에 대해 접근을 허용해야 함



# ELB(Elastic Load Balancing)

- ELB 동작 특징

- 고가용성(High Availability) 지원

- ELB를 서로 다른 가용 영역에 해당하는 서브넷에 생성
    - 특정 가용 영역에 문제가 발생하더라도 다른 가용 영역의 ELB가 트래픽을 처리

