AWS 아키텍처 설계

Chapter 06. AWS Network 2

설계 시나리오

AWS의 VPC 환경과 온프레미스 환경을 연결하여 하이브리드 클라우드 구성을 원하는 상황이다.

또 더 많은 사용자의 트래픽을 분산 처리하고, 가용 영역 레벨의 장애 발생시 서비스가 중단되는 것을 방지하려고 한다.

AWS 서비스 중 이러한 요구 사항을 충족할 수 있는 서비스는 무엇이 있을까?

01 VPC 연결 확장

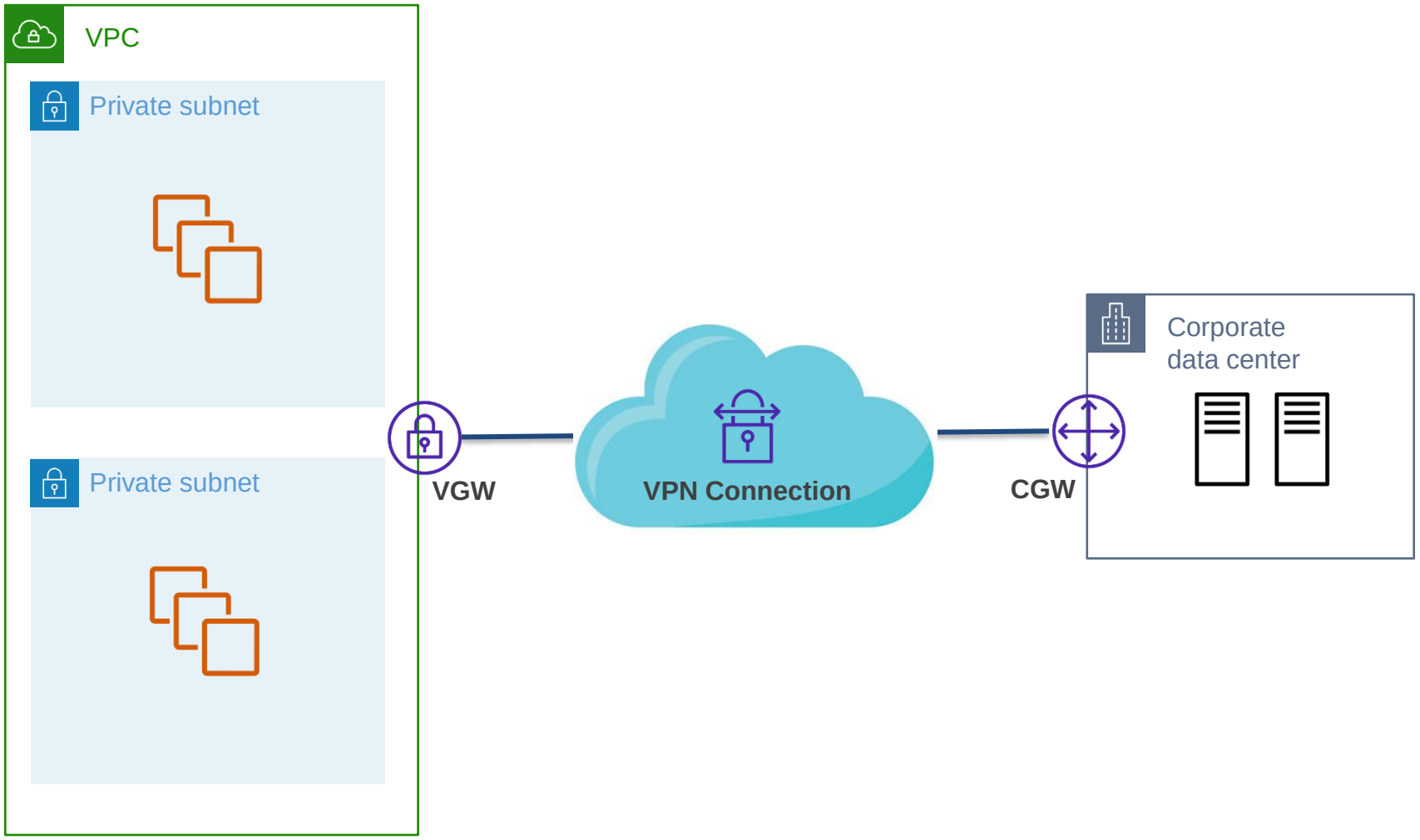
VGW (Virtual Private Gateway)

- Amazon VCP와 다른 네트워크를 VPN 혹은 Direct Connect로 연결할 수 있다.

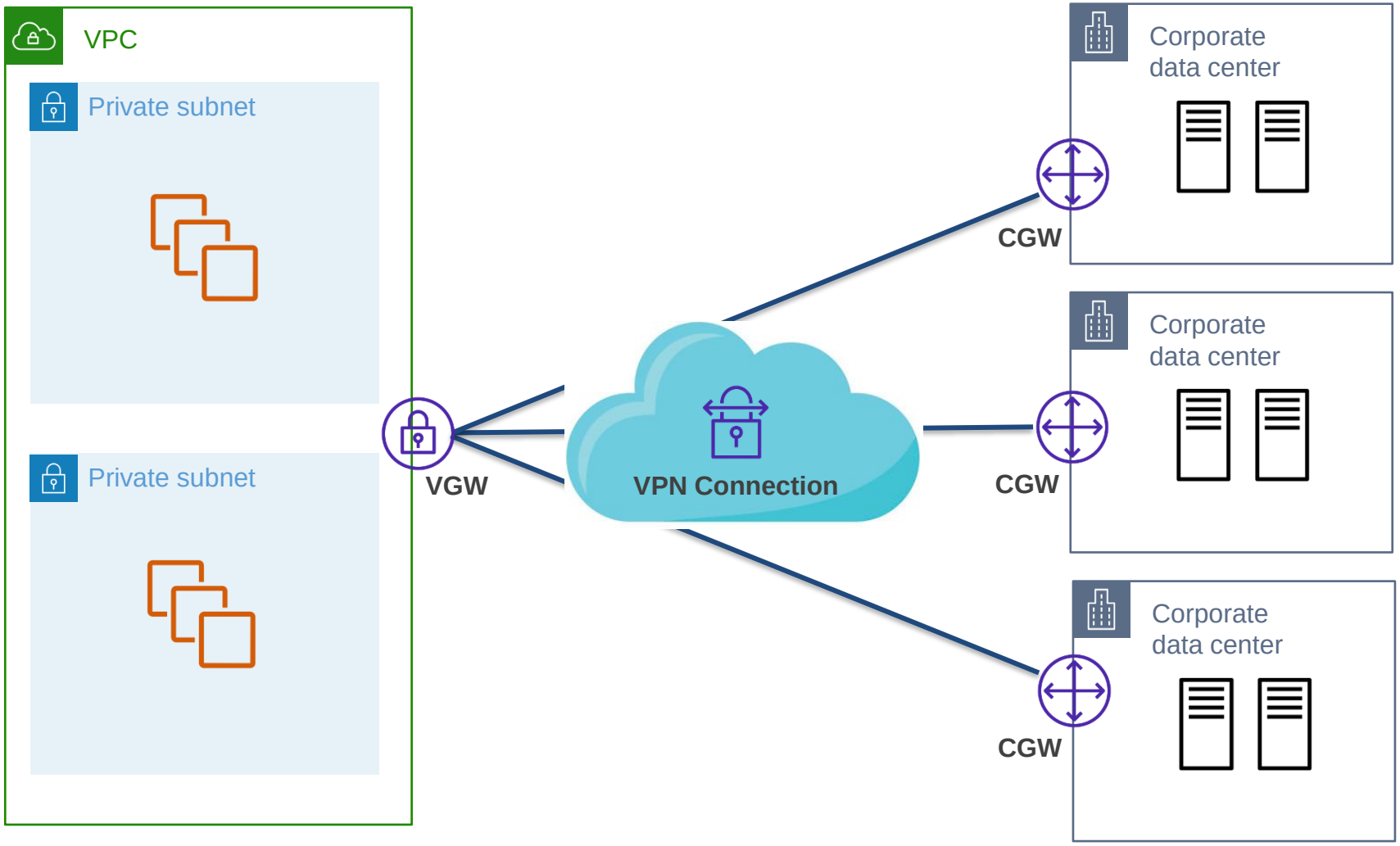
VGW

- IGW와 다르게 인터넷 이외의 외부 네트워크와 연결.

Site-to-Site VPN 연결



Site-to-Site VPN 연결

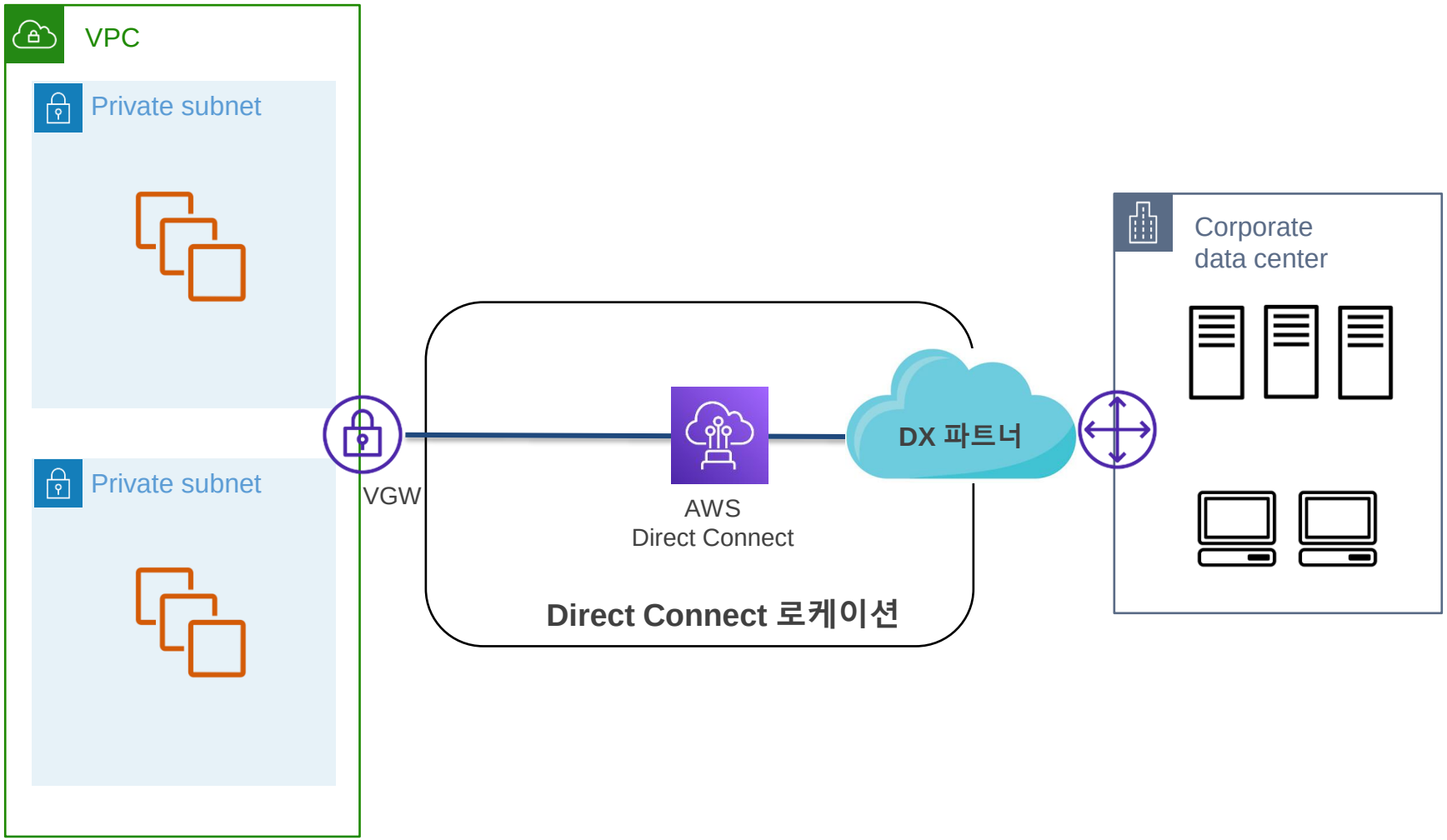


AWS Direct Connect (DX)

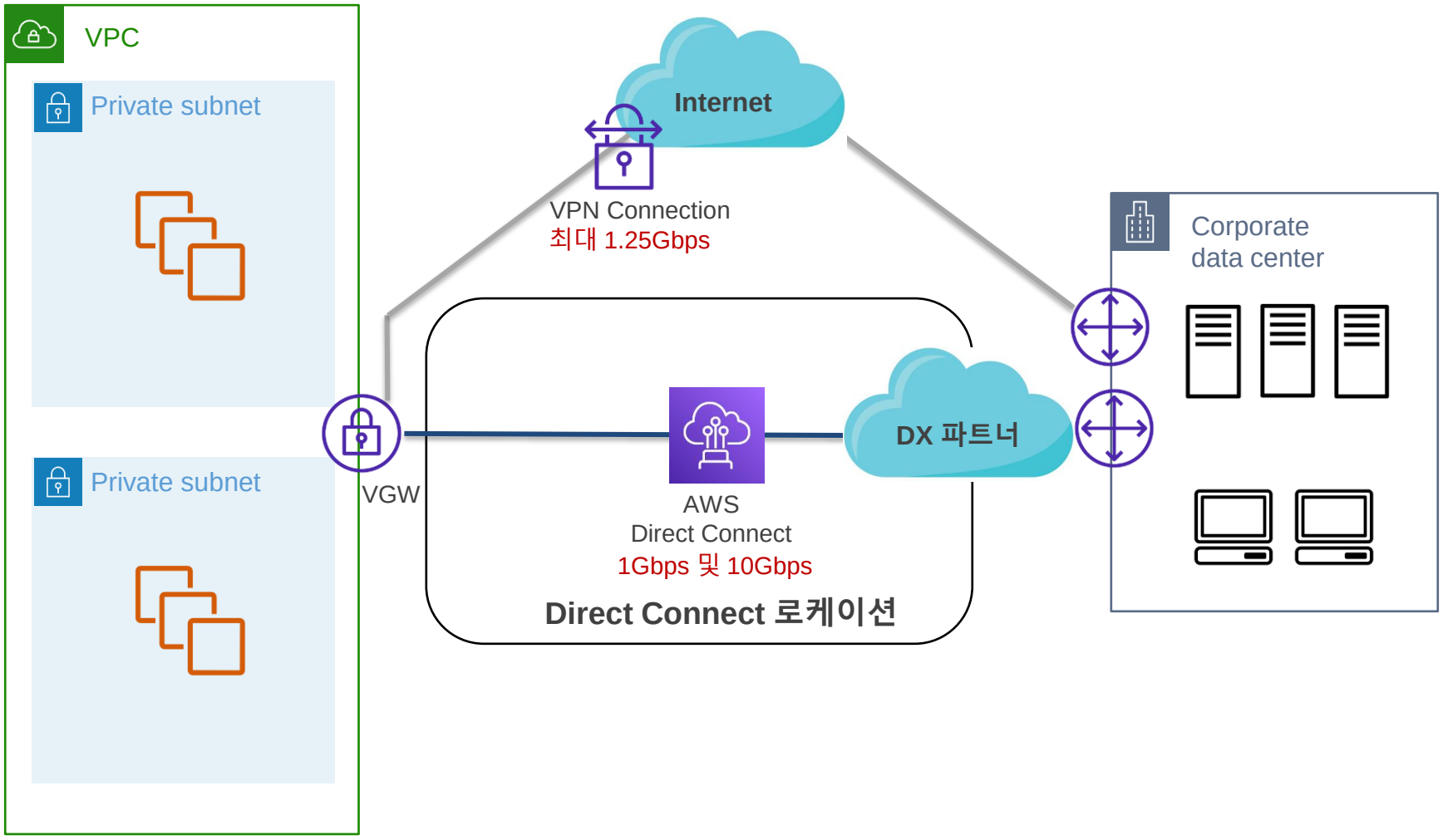
- **온프레미스에서 AWS로 전용 네트워크 연결 제공. (Private 연결)**
- **인터넷 기반 연결보다 일관된 네트워크 성능 제공.**
- **1Gbps 및 10Gbps 연결을 제공.**
- **AWS Direct Connect 적합한 사용 사례.**
  - 대용량 데이터 세트 전송
  - 실시간 데이터 피드 사용 애플리케이션
  - 하이브리드 환경
  - 보안 및 규정 준수

AWS Direct Connect

AWS Direct Connect (DX) 연결



중요 서비스를 위한 가용성 고려



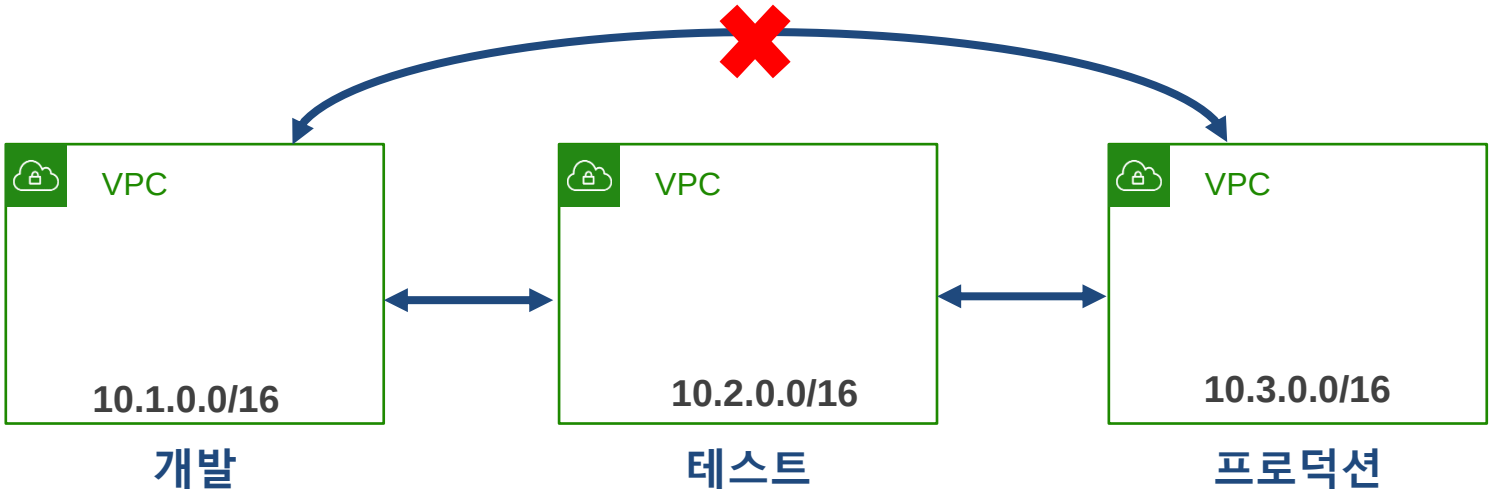
VPC 사이의 트래픽 전송

- VPC를 사용하여 각 AWS 리소스를 격리.
- 경우에 따라서 서로 다른 VPC 내부 리소스 간의 데이터 전송이 필요한 상황이 발생.

- VPC 간의 프라이빗 통신을 원하는 경우 어떻게 해야 하는가?

VPC Peering

- 사설 IP 주소를 사용하여 두 VPC 간에 트래픽을 라우팅할 수 있는 네트워킹 연결.
- 리전 내부/리전 외부의 VPC 모두 연결 가능.
- 동일 계정/서로 다른 계정 소유의 VPC 모두 연결 가능.
- 전이적 피어링은 지원되지 않음.



VPC Peering

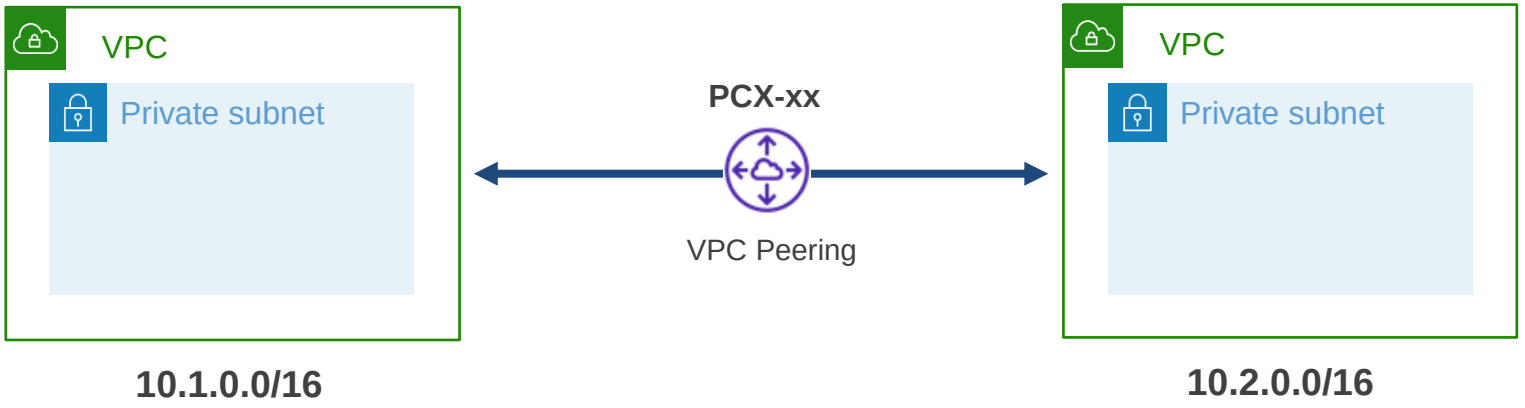
- AWS 백본 네트워크를 통해서 통신.
- Internet Gateway(IGW)와 Virtual Private Gateway(VGW) 필요 없음.
- 고가용성, 대역폭 병목에 대한 단일 지점 장애가 없음.
- 연결되는 각 VPC는 서로 다른 IP 대역을 사용해야 함.

라우팅 테이블

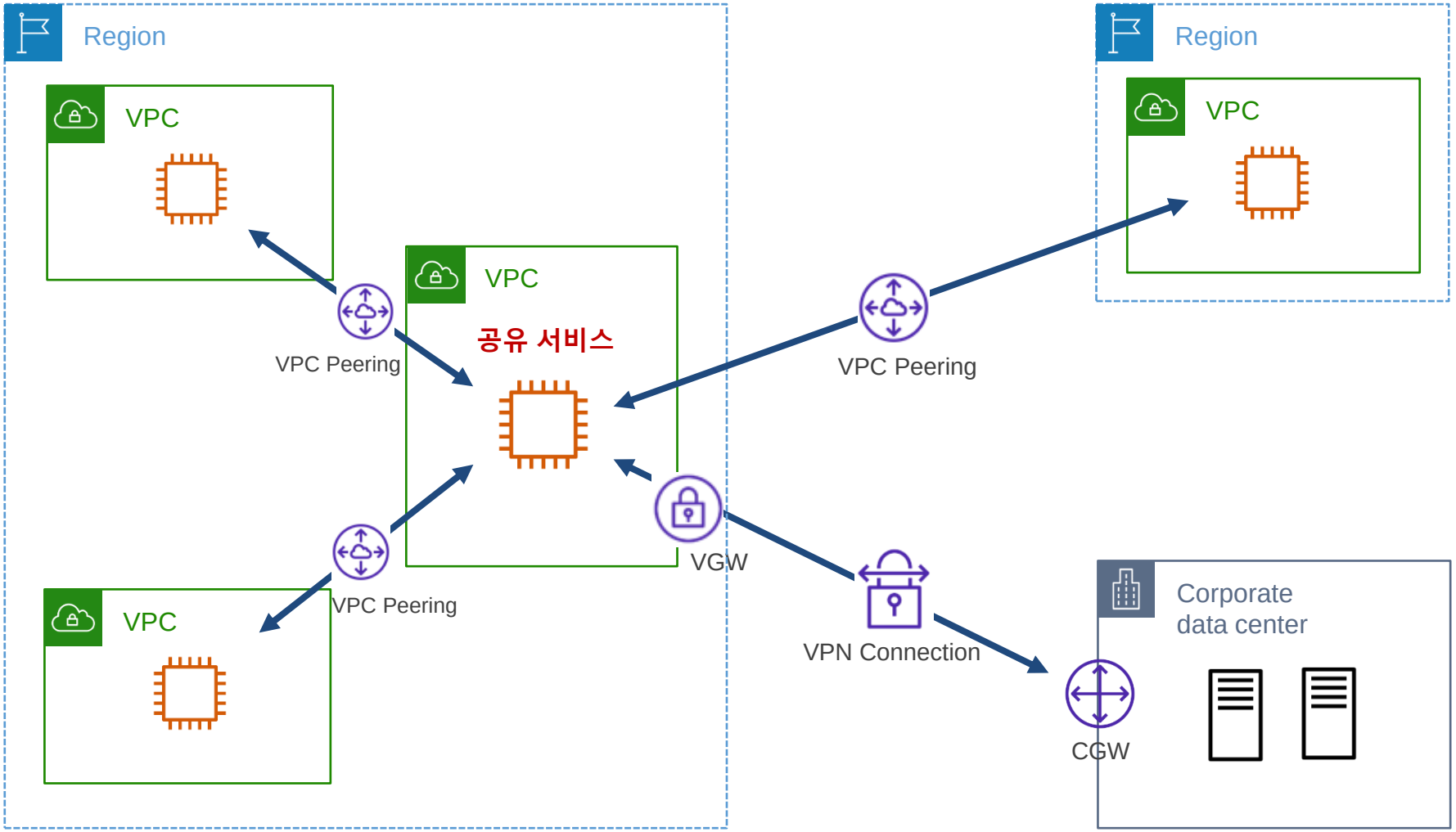
Destination	Target
10.1.0.0/16	Local
10.2.0.0/16	PCX-xx

라우팅 테이블

Destination	Target
10.2.0.0/16	Local
10.1.0.0/16	PCX-xx



VPC Peering


Region
VPC
VPC
공유 서비스
VPC Peering
VPC Peering
VPC
VPC Peering
VGW
VPN Connection
Region
VPC
Corporate
data center
CGW

Transit Gateway

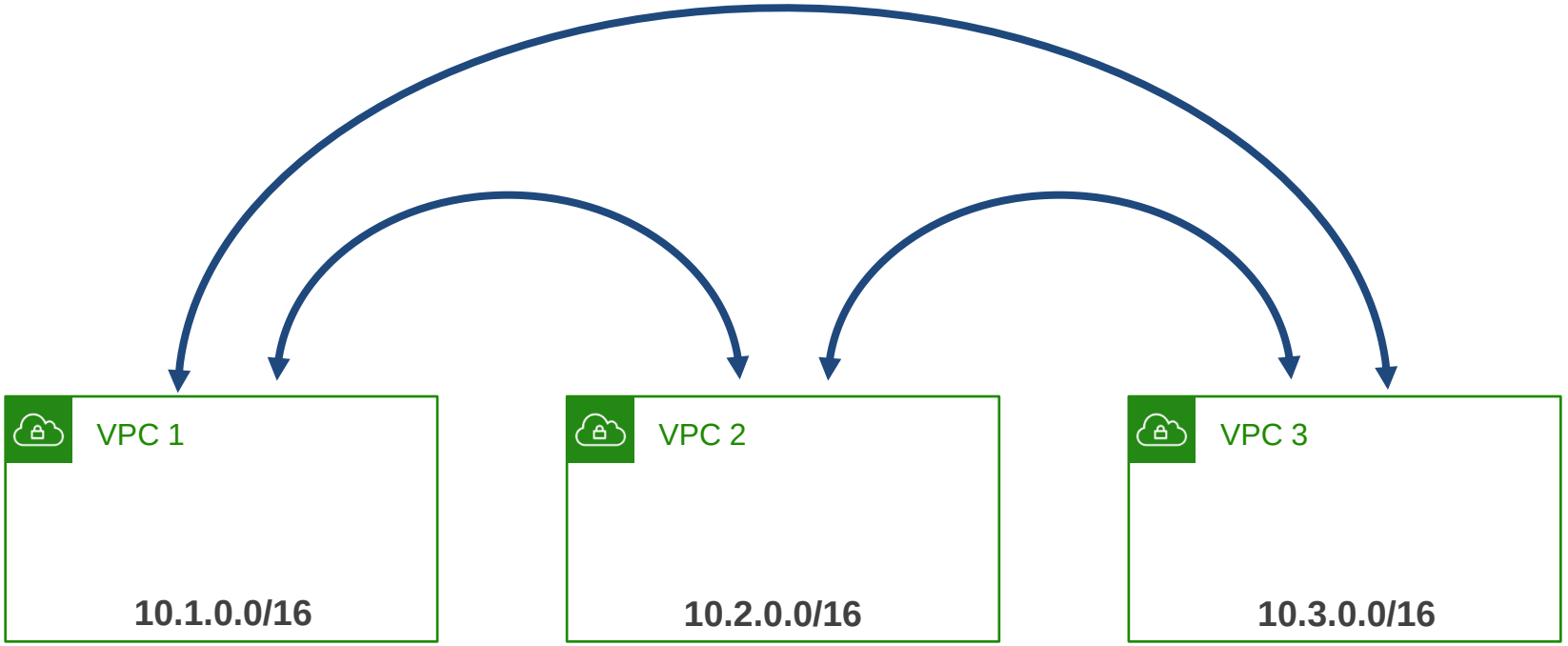
- **중앙 허브를 통해 VPC와 온프레미스 네트워크를 연결.**
- **복잡한 피어링 관계를 제거하여 네트워크를 간소화.**
- **최대 5,000개의 VPC와 온프레미스 네트워크 연결.**
- **완전 관리형 라우팅 서비스**
- **Transit Gateway 사용 이점**

  - 간편한 연결 / 가시성 및 제어 향상 / 보안 향상 / 멀티캐스트 지원

AWS Transit Gateway

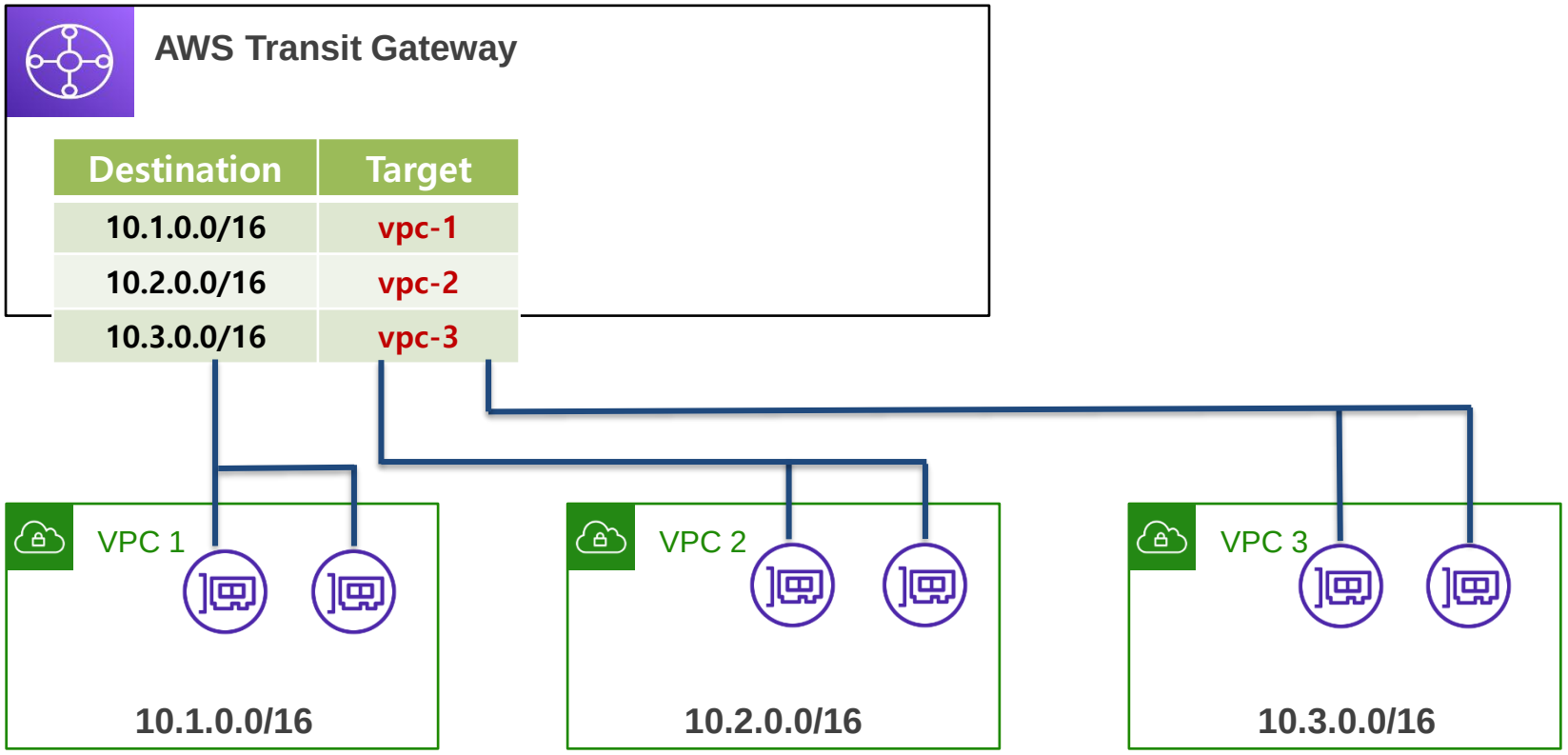
Transit Gateway 사용 예: 모든 VPC 연결

- VPC 피어링을 사용하는 경우 각 VPC를 모두 연결해야 한다.



Transit Gateway 사용 예: 모든 VPC 연결

- VPC 피어링 대신 Transit Gateway를 사용하여 모든 VPC 연결



라우팅 테이블

Destination	Target
10.1.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

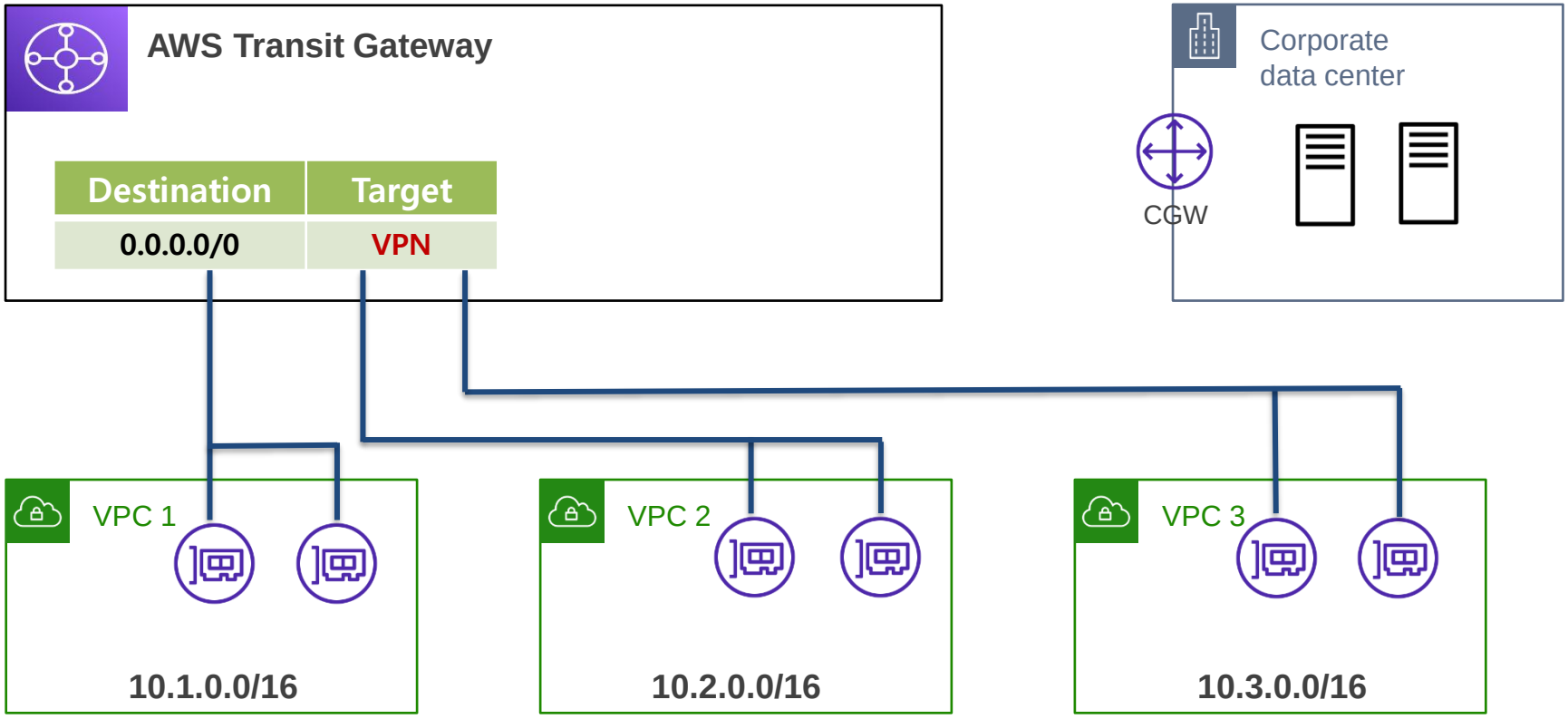
Destination	Target
10.2.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

Destination	Target
10.3.0.0/16	Local
10.0.0.0/8	tgw-xx

Transit Gateway 사용 예: VPN 연결

- VPC 간 통신 차단 후 VPN 연결



라우팅 테이블

Destination	Target
10.1.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

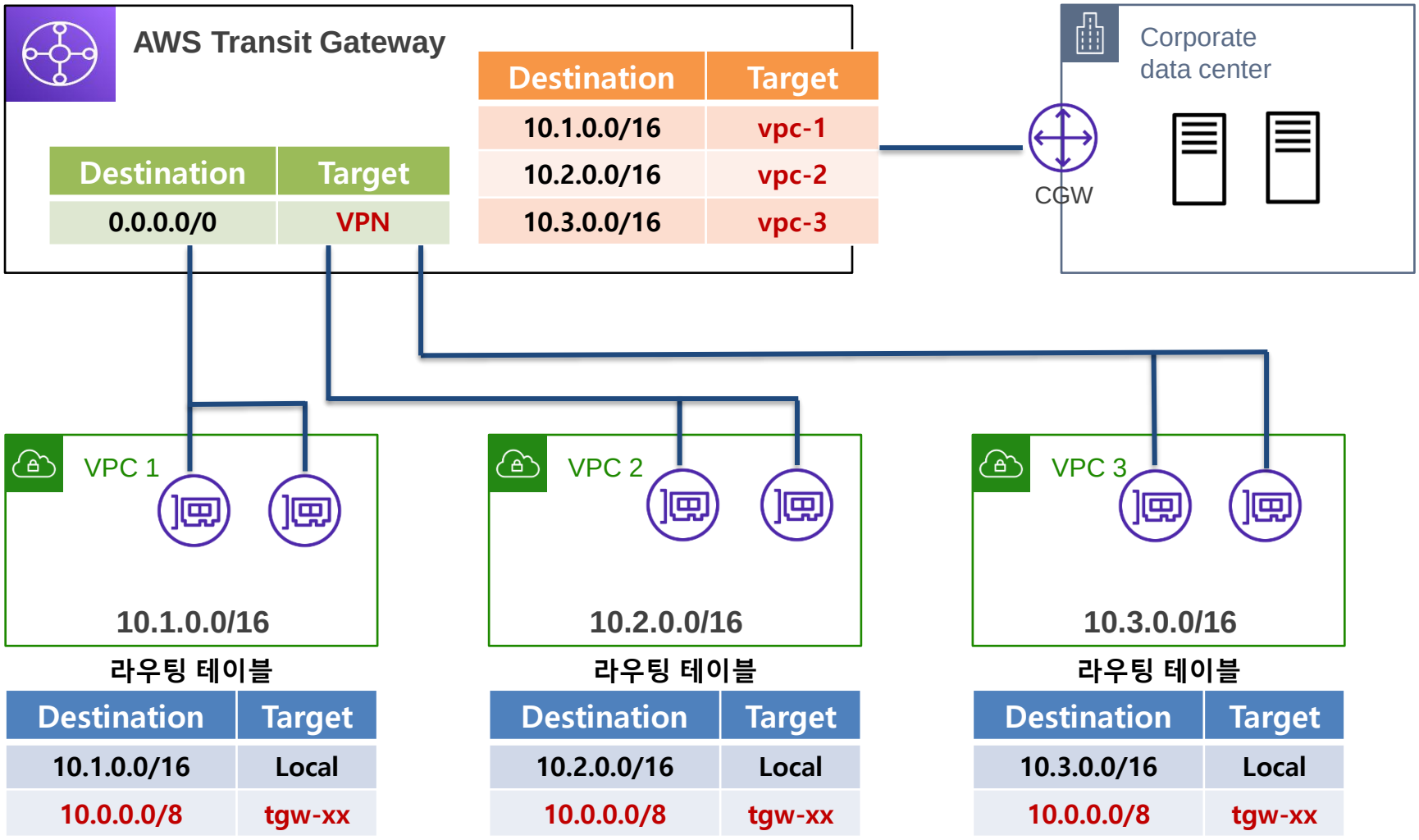
Destination	Target
10.2.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

Destination	Target
10.3.0.0/16	Local
10.0.0.0/8	tgw-xx

Transit Gateway 사용 예: VPN 연결

- VPC 간 통신 차단 후 VPN 연결



라우팅 테이블

Destination	Target
10.1.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

Destination	Target
10.2.0.0/16	Local
10.0.0.0/8	tgw-xx

라우팅 테이블

Destination	Target
10.3.0.0/16	Local
10.0.0.0/8	tgw-xx

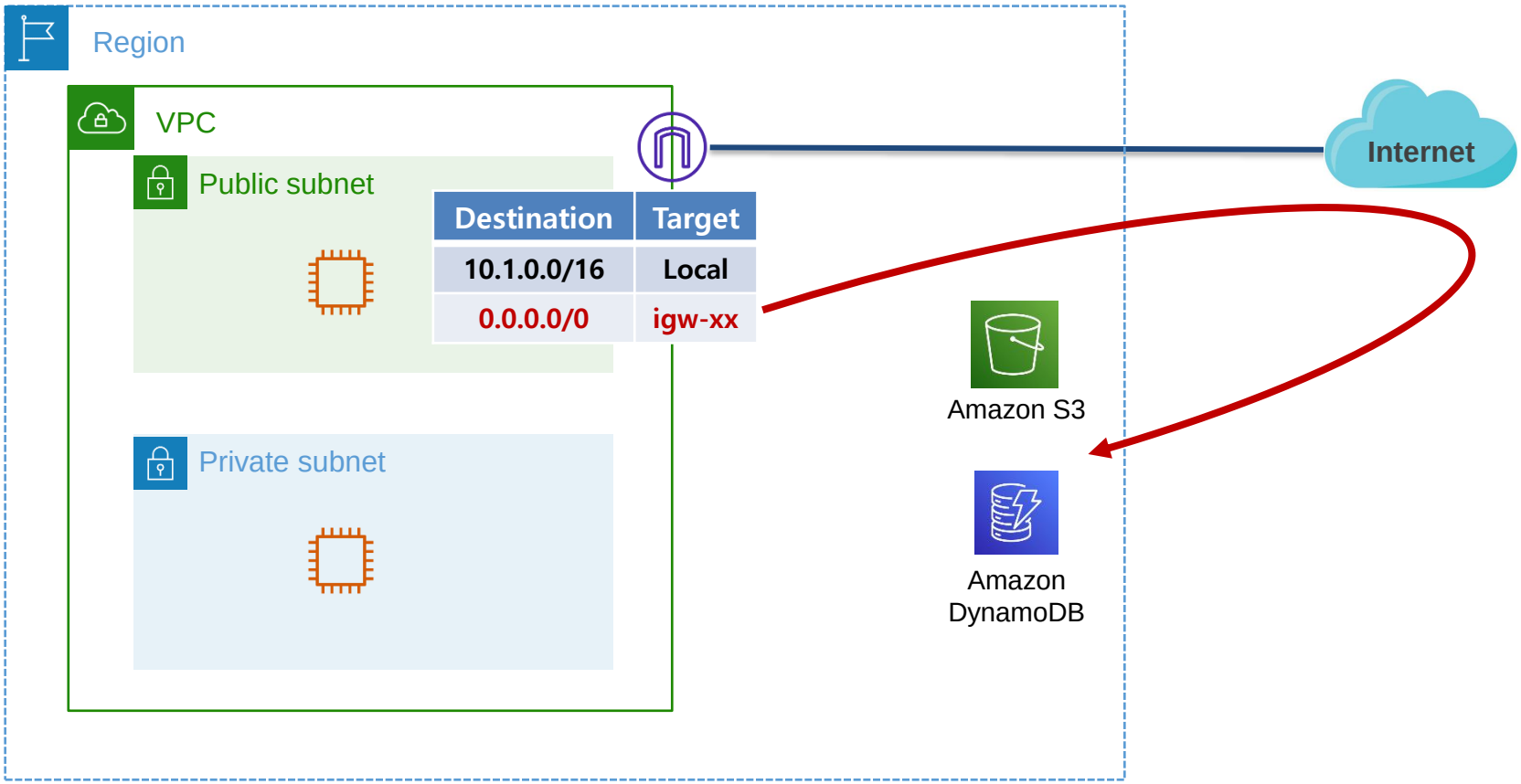
VPC Endpoint

- VPC 외부에 위치한 AWS 퍼블릭 서비스, 혹은 직접 생성한 AWS 서비스에 대한 프라이빗 접근이 가능.
- IGW, VGW, NAT Gateway 등이 필요 없음.
- 접속하는 서비스는 동일 리전에 위치해야 한다.
- 가용성, 확장성이 제공되도록 설계되어 있다.

VPC Endpoint

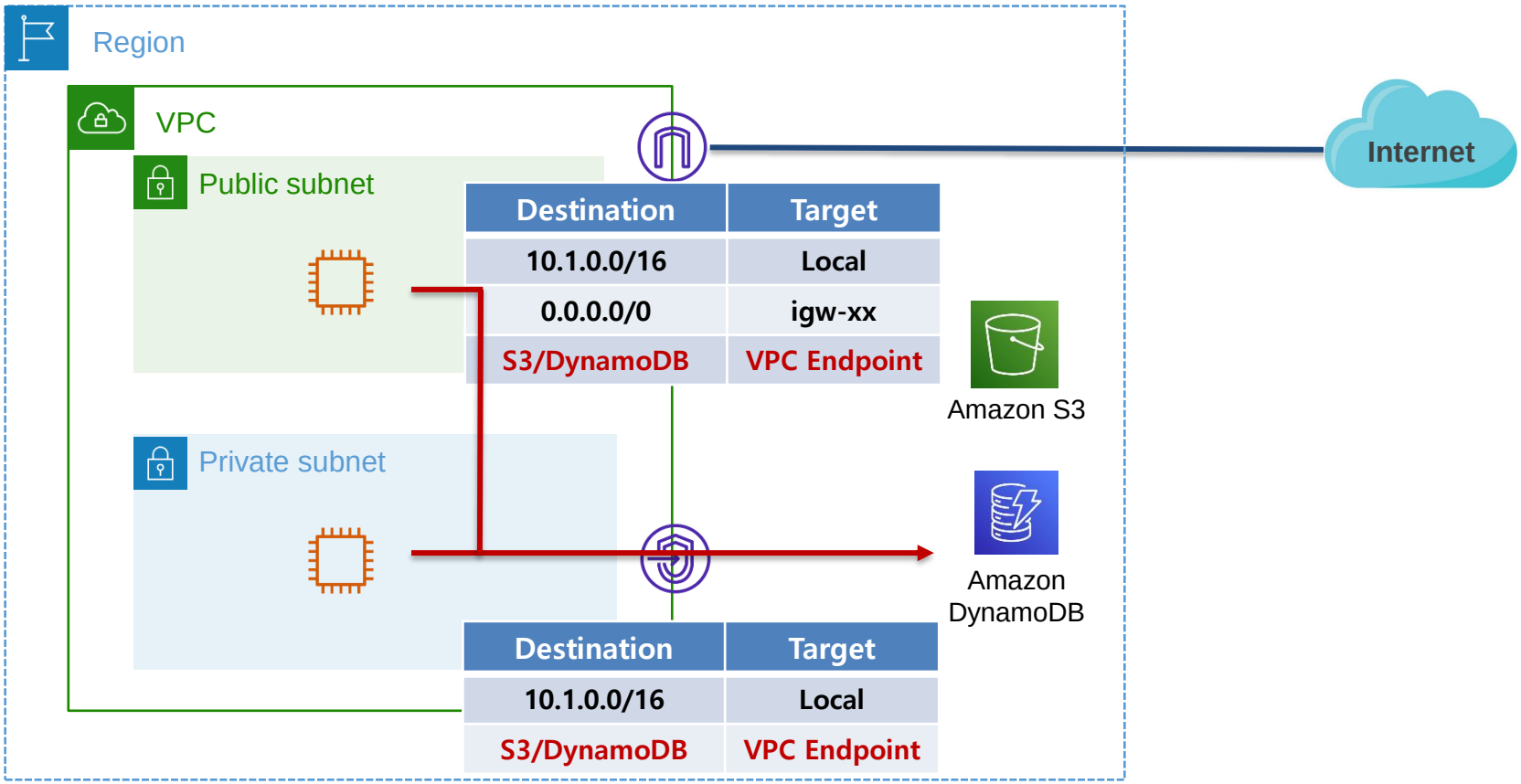
VPC Endpoint 종류

- **게이트웨이 엔드포인트**
  - Amazon S3 / Amazon DynamoDB



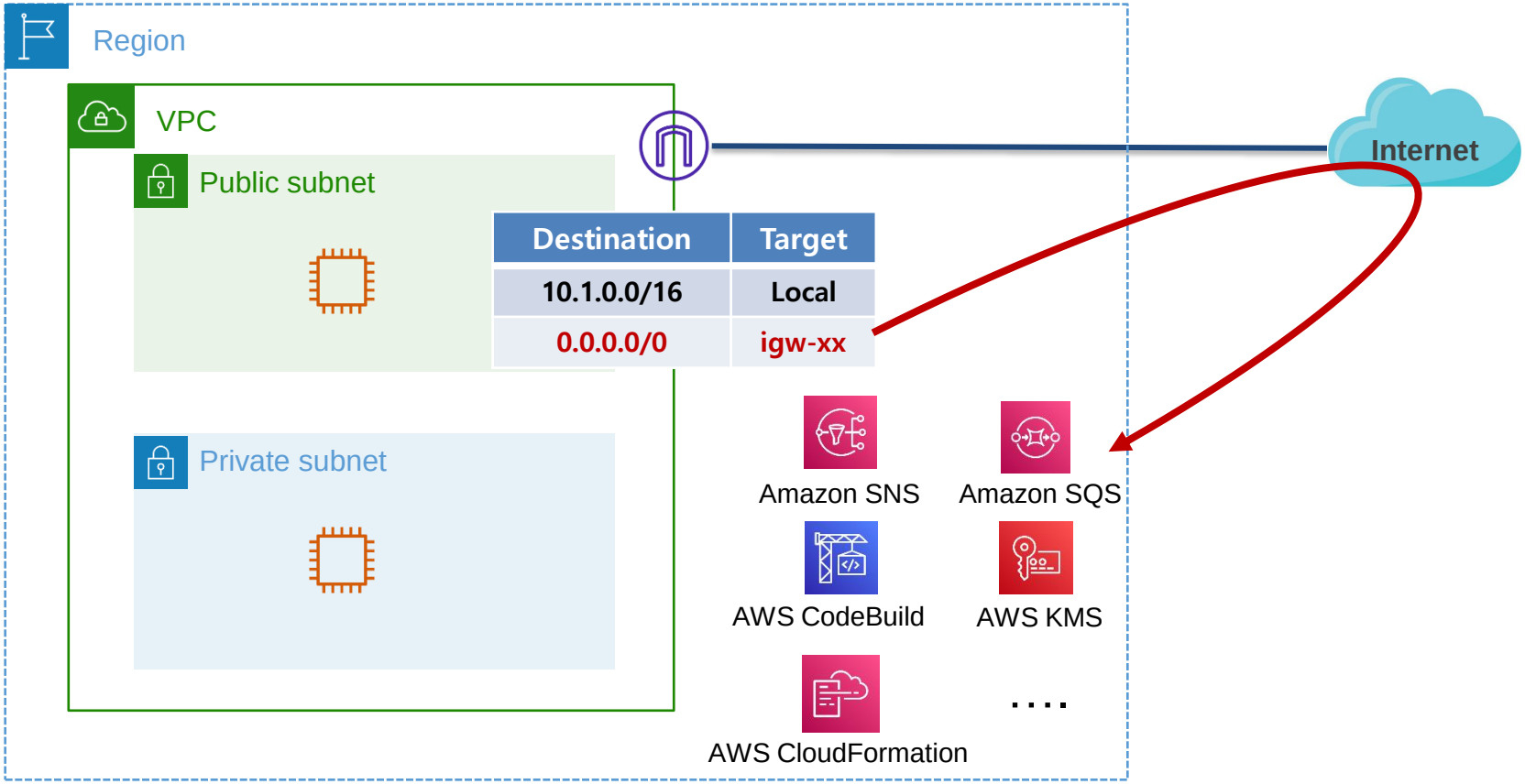
VPC Endpoint 종류

- **게이트웨이 엔드포인트**
  - Amazon S3 / Amazon DynamoDB



VPC Endpoint 종류

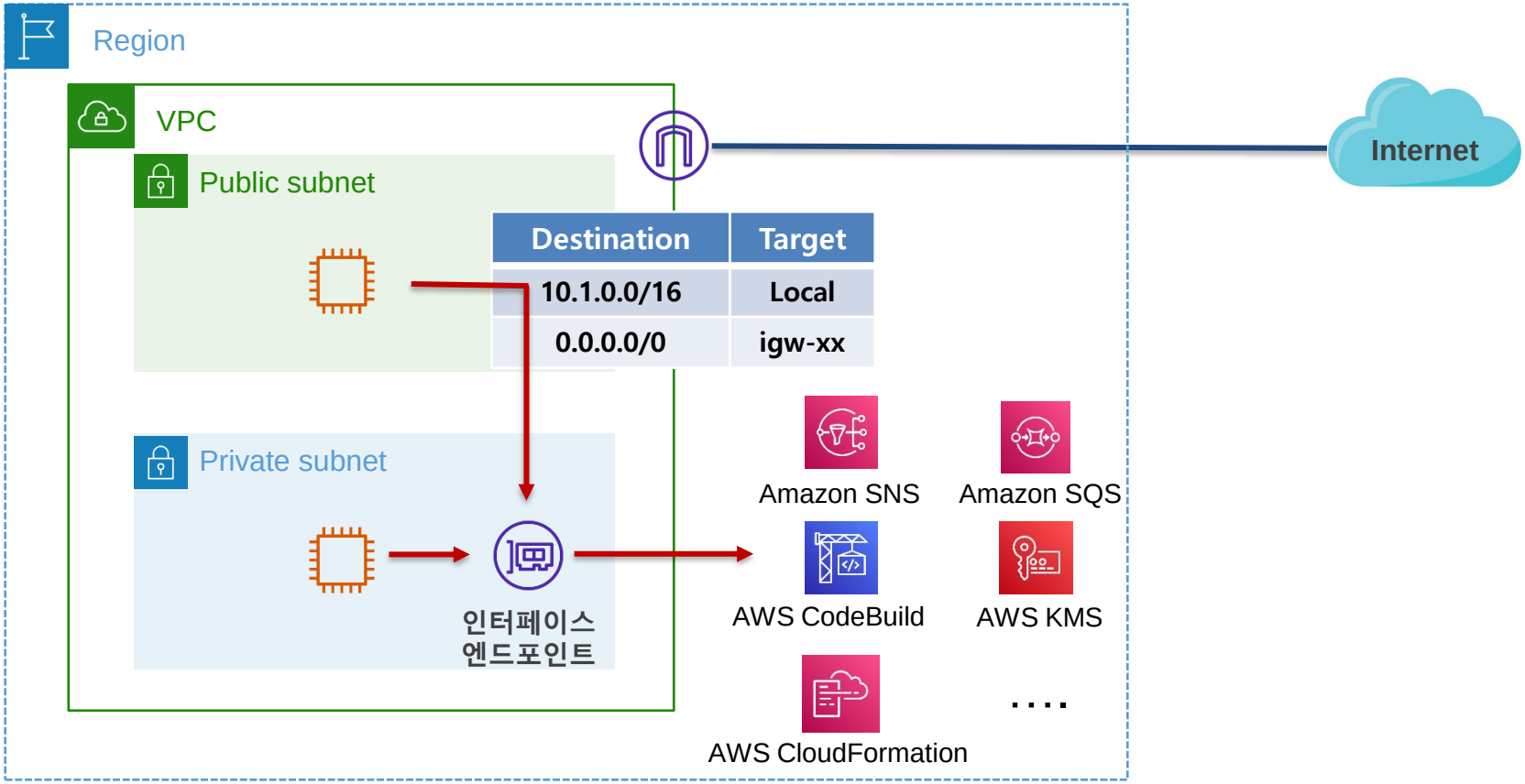
- **인터페이스 엔드포인트**
  - 다수의 서비스
    (https://docs.aws.amazon.com/ko_kr/vpc/latest/userguide/integrated-services-vpce-list.html)



VPC Endpoint 종류

- **인터페이스 엔드포인트**
  - 다수의 서비스

    (https://docs.aws.amazon.com/ko_kr/vpc/latest/userguide/integrated-services-vpce-list.html)

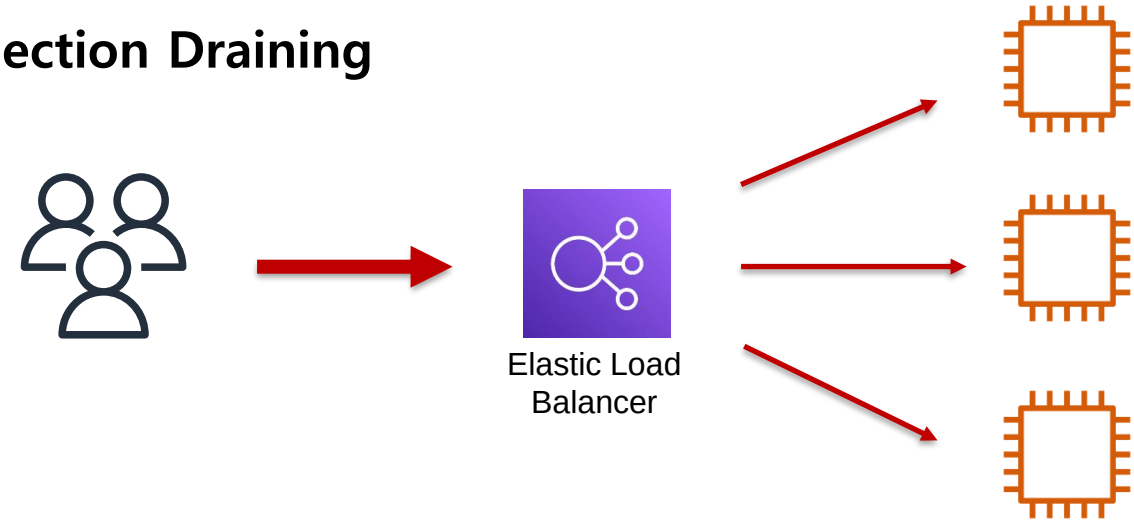


02 로드 밸런싱

Elastic Load Balancing (ELB)

- ELB는 둘 이상의 가용 영역에서 EC2 인스턴스 등 여러 대상에 걸쳐 수신되는 트래픽을 분산.
- ELB에 등록된 대상의 상태를 모니터링하여 정상 상태인 대상으로만 트래픽을 라우팅.
- AWS 완전 관리형 서비스
- Connection Draining

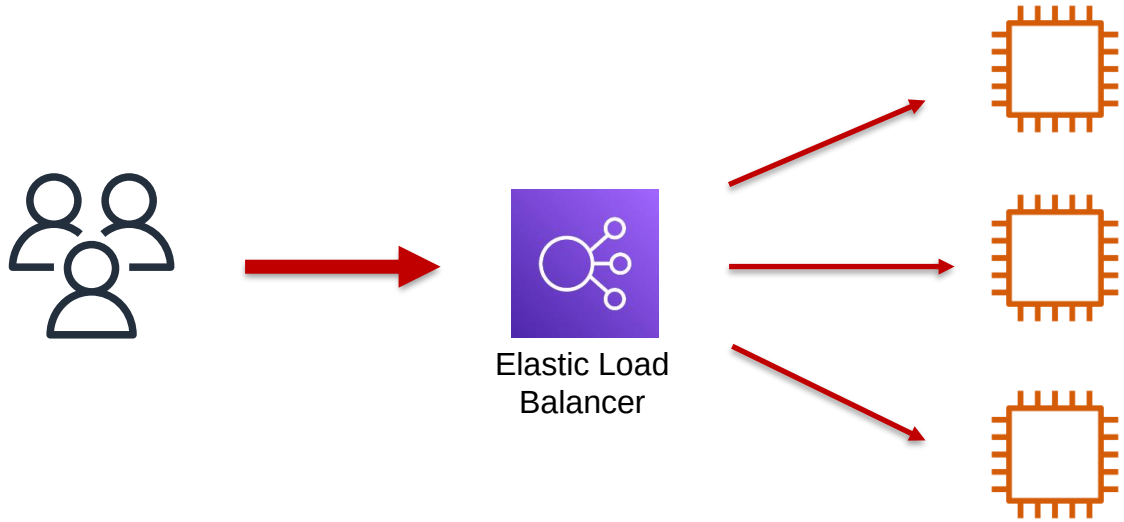
Elastic Load Balancing



Elastic Load Balancing (ELB)

- ELB의 특징들은 다음과 같다.

1. 고가용성
2. 상태 확인
3. 보안 기능

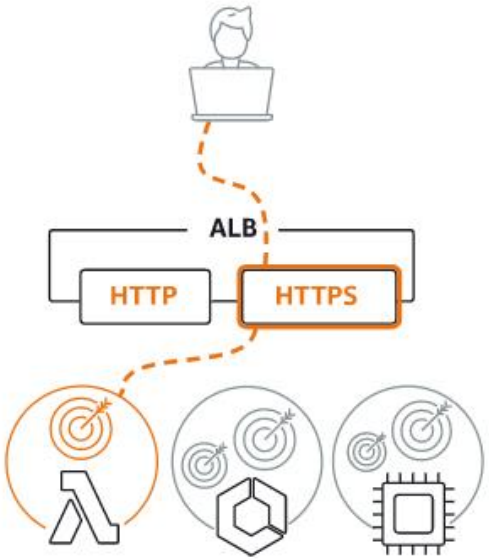


Elastic Load Balancing (ELB)

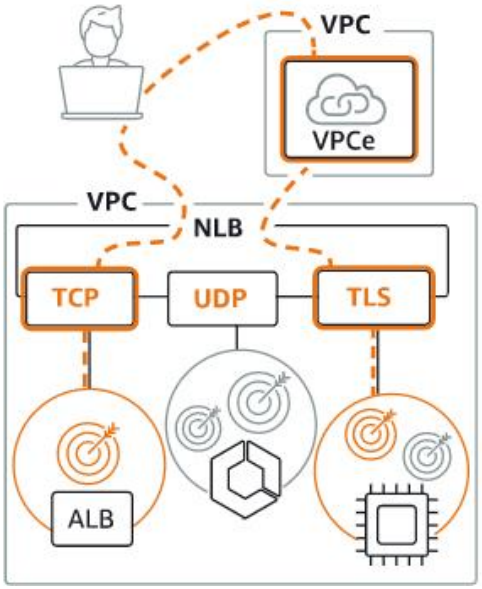
- 현재 서울 리전은 다음 네 가지 유형의 로드 밸런서를 지원

로드 밸런서 유형

Application Load Balancer 정보



Network Load Balancer 정보



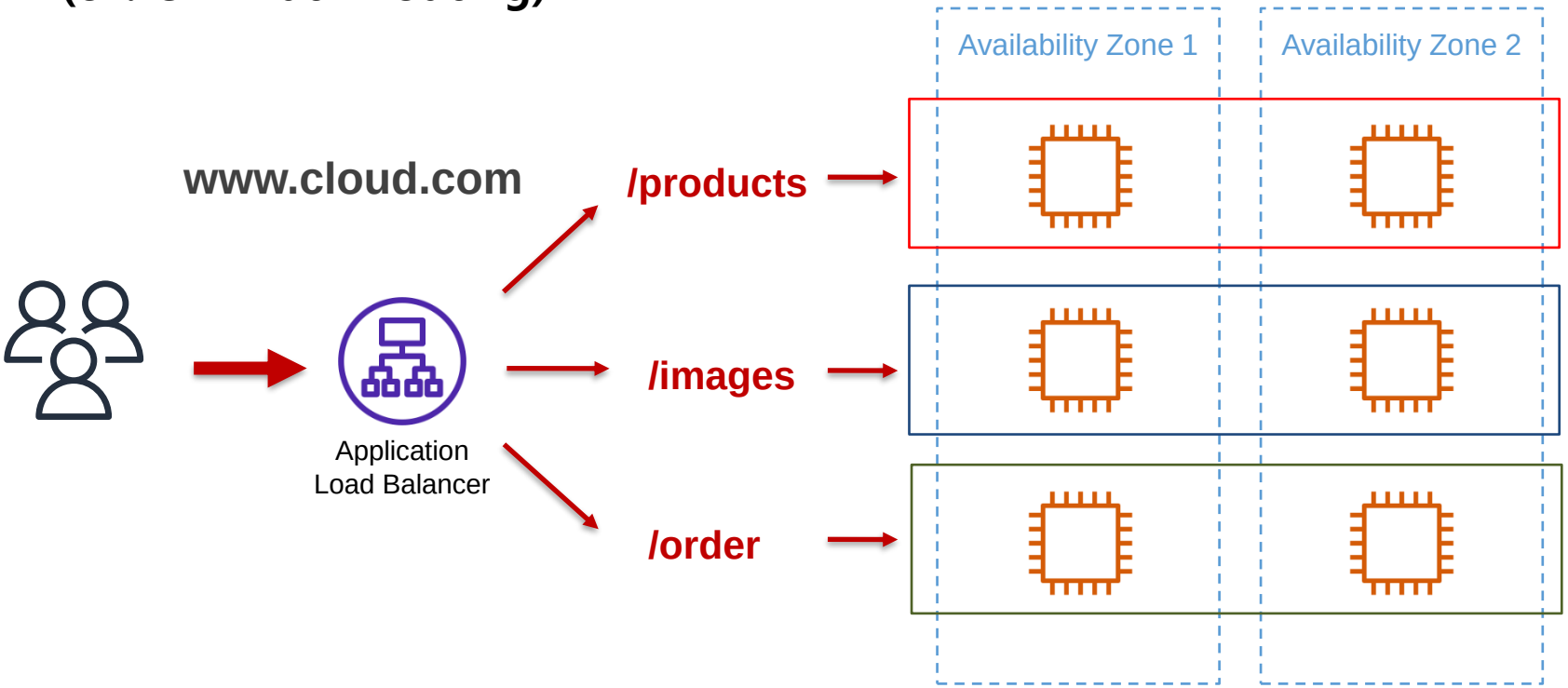
Gateway Load Balancer 정보

▼ Classic Load Balancer - *이전 세대*

Classic Load Balancer 정보

Application Load Balancer

- ALB는 웹 애플리케이션(HTTP/HTTPS)에 대한 분산 처리를 제공.
- 웹 애플리케이션에 대한 세부적이고 다양한 정책을 통해 라우팅.
  (ex. URL Path Routing)



03 고가용성을 고려한 설계

High Availability(고가용성)

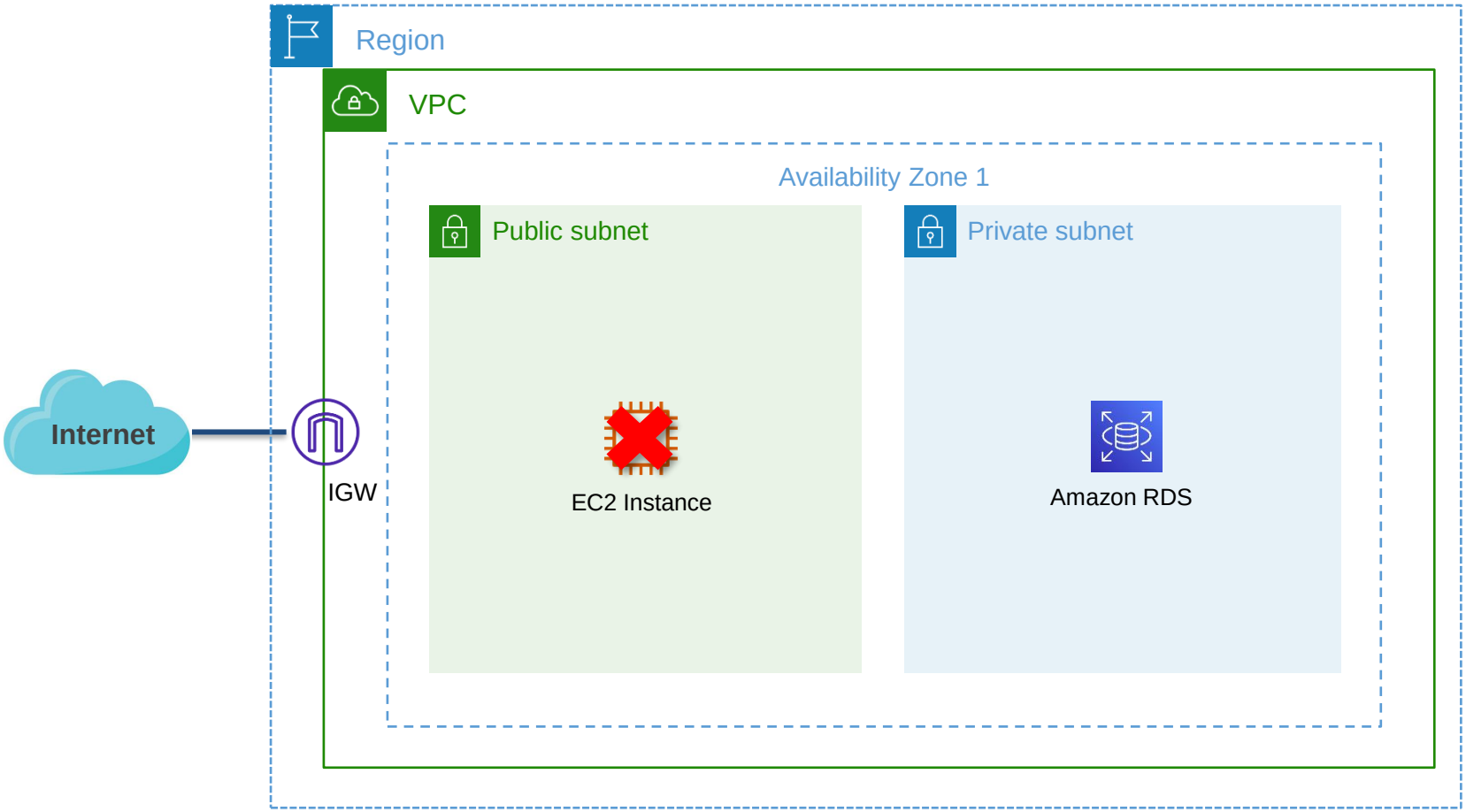
- **안정성(Reliability)은 애플리케이션에서 장애가 발생하는 것을 예방하거나 장애 발생시 신속히 복구하는 능력.**

- **가용성(Availability)은 안정성(Reliability) 기준을 측정 가능하도록 계량화 한 성능 지표.**
  - 애플리케이션이 예상한 대로 작동하는 시간을 **백분율**로 나타낸 것

가용성 비율	연간 가동 중지 시간	일일 가동 중지 시간
99%	3일 15시간 39분	14분
99.9%	8시간 45분	86초
99.99%	약 52분	8.6초
99.999%	약 5분	0.86초

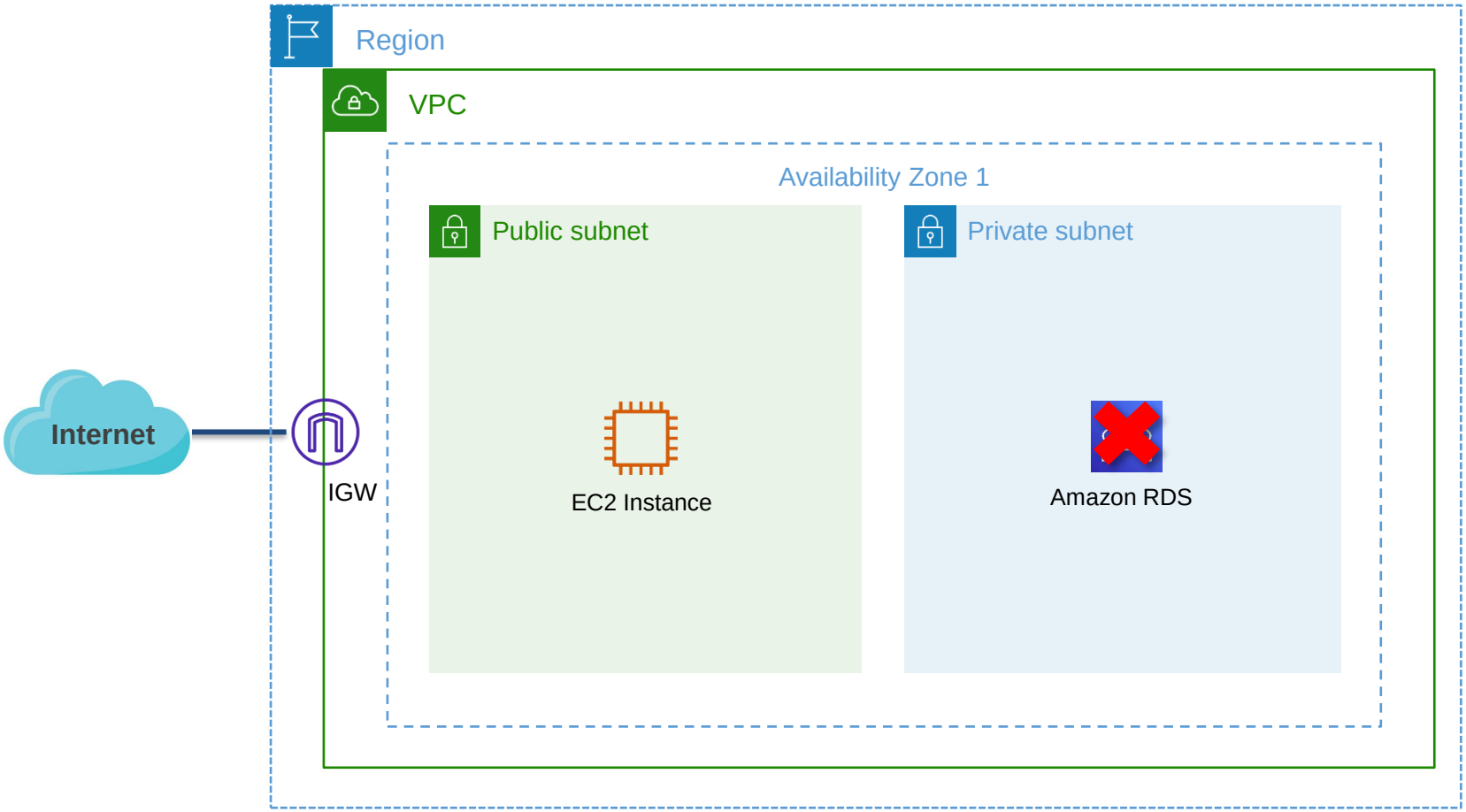
High Availability(고가용성)

- 가용성을 높이기 위해 단일 장애 지점(SPOF)을 제거.



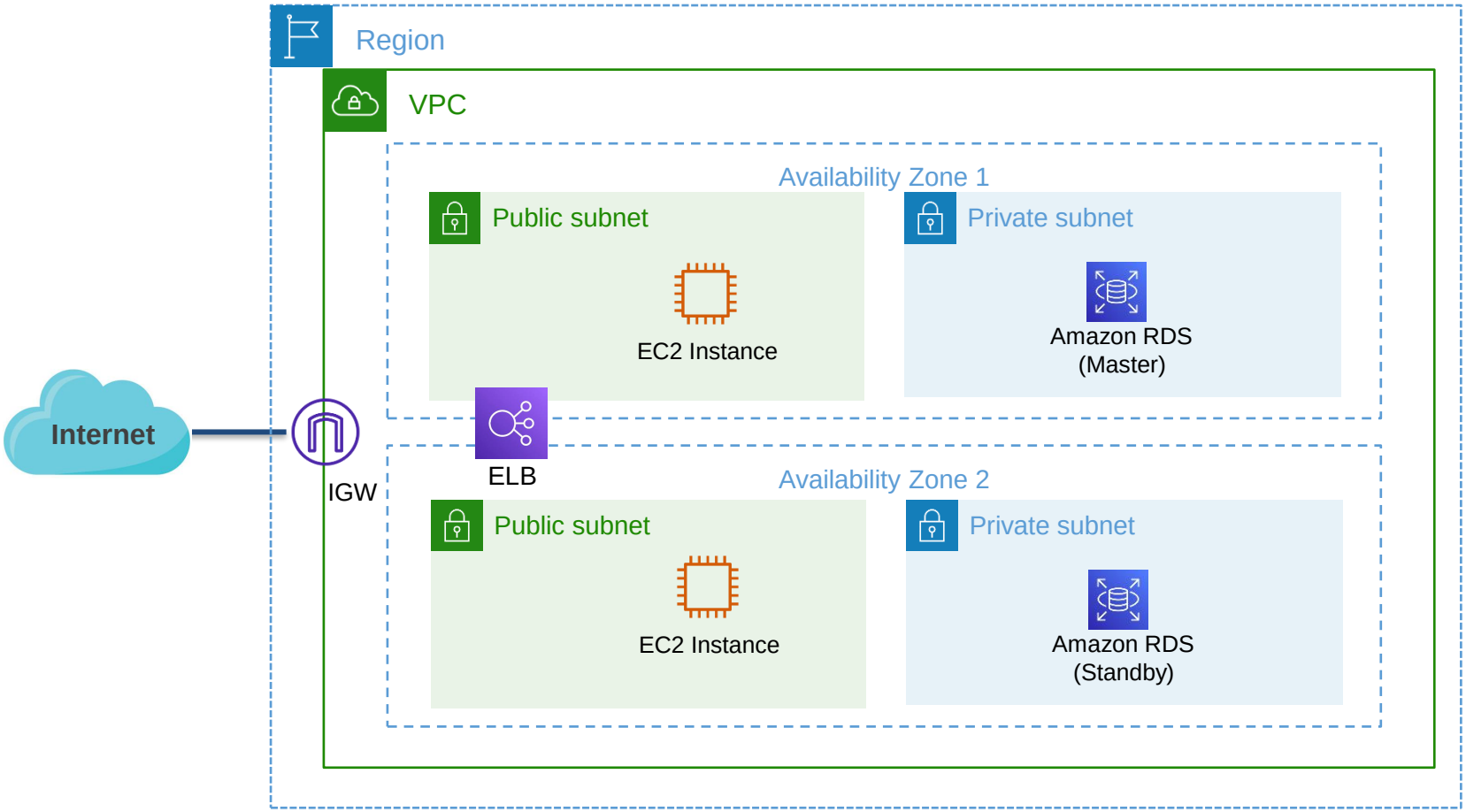
High Availability(고가용성)

- 가용성을 높이기 위해 단일 장애 지점(SPOF)을 제거.



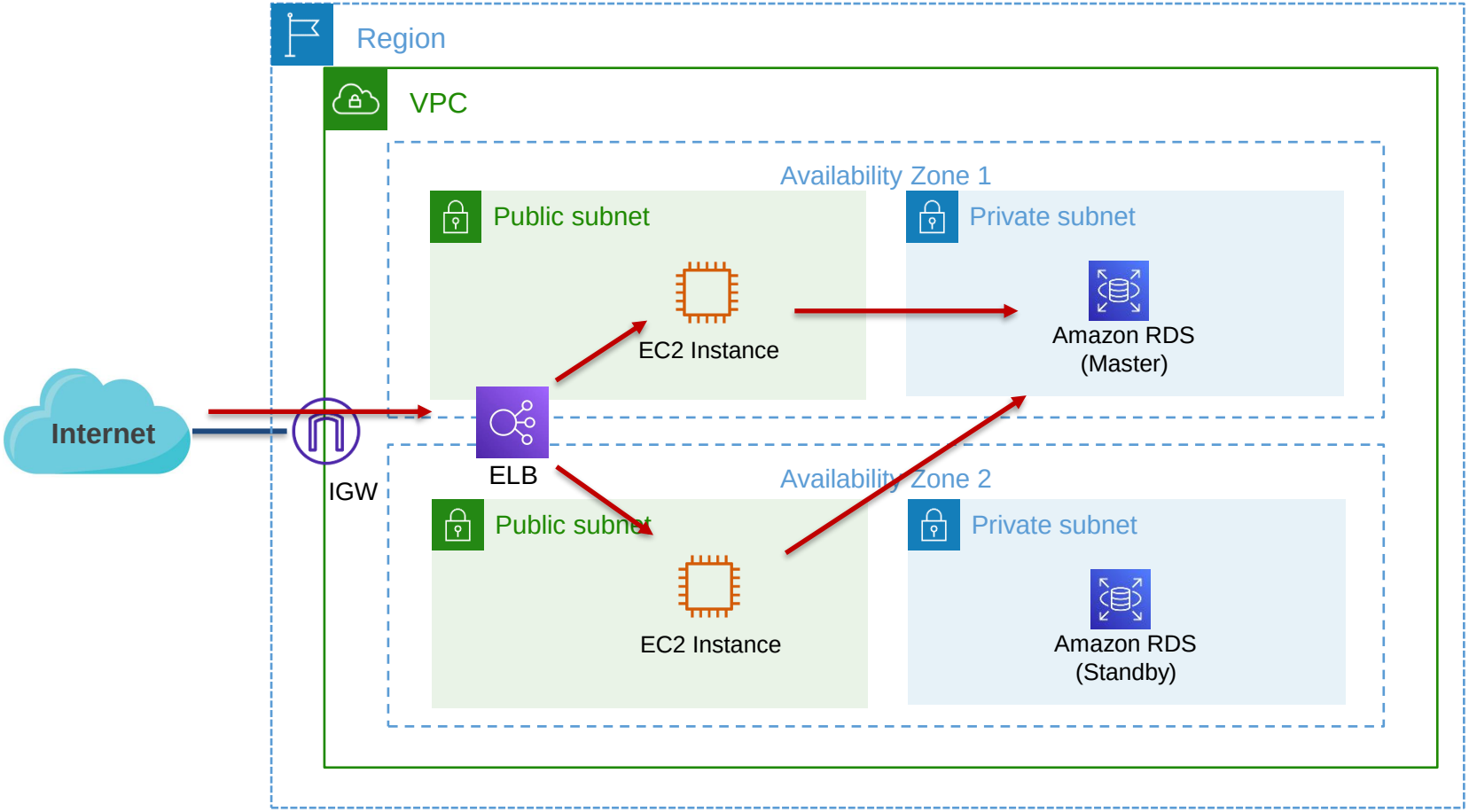
High Availability(고가용성)

- **가용성을 높이기 위해 단일 장애 지점(SPOF)을 제거.**
  - 2개의 가용 영역을 사용하면 더 높은 가용성을 얻을 수 있다.



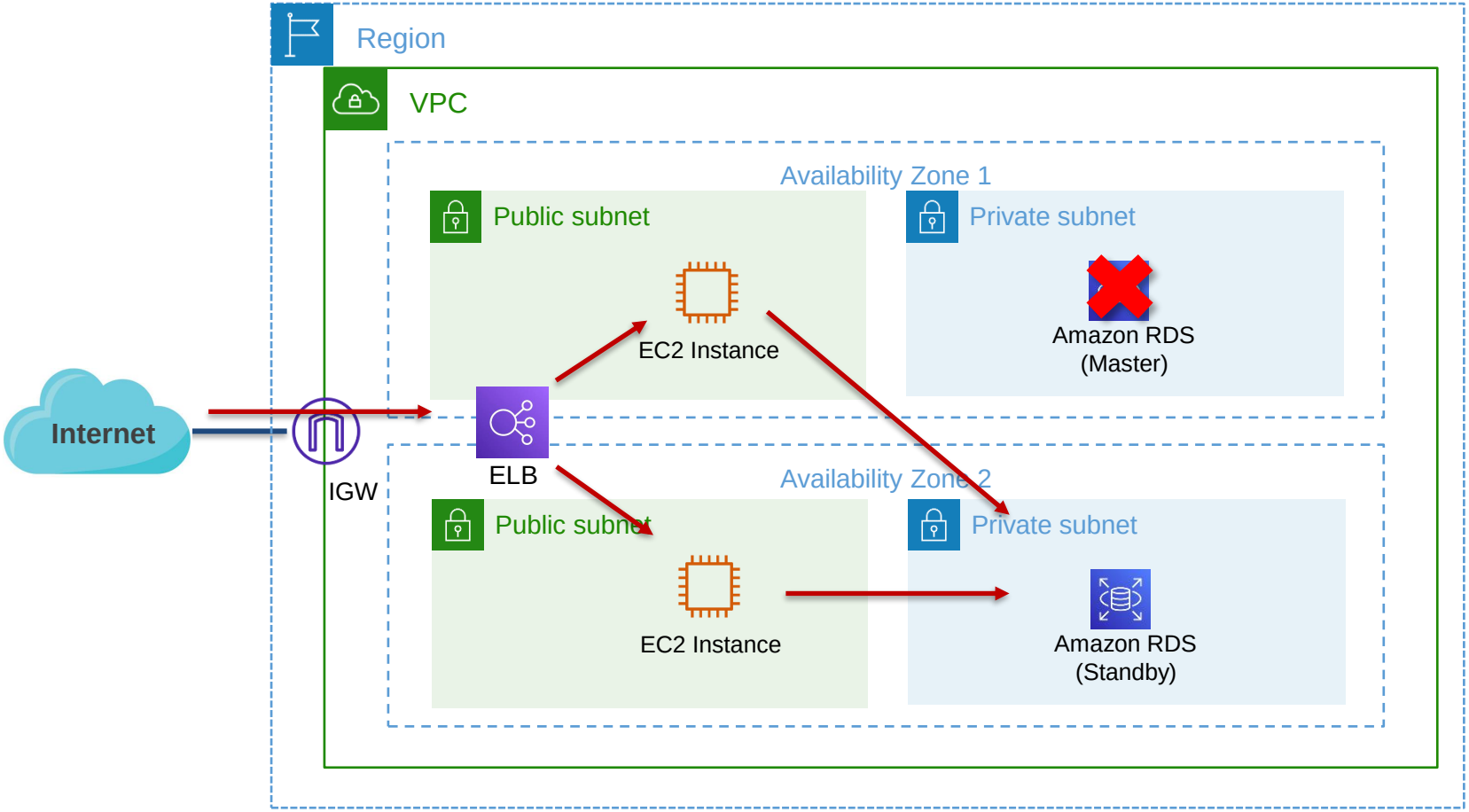
High Availability(고가용성)

- **가용성을 높이기 위해 단일 장애 지점(SPOF)을 제거.**
  - 2개의 가용 영역을 사용하면 더 높은 가용성을 얻을 수 있다.



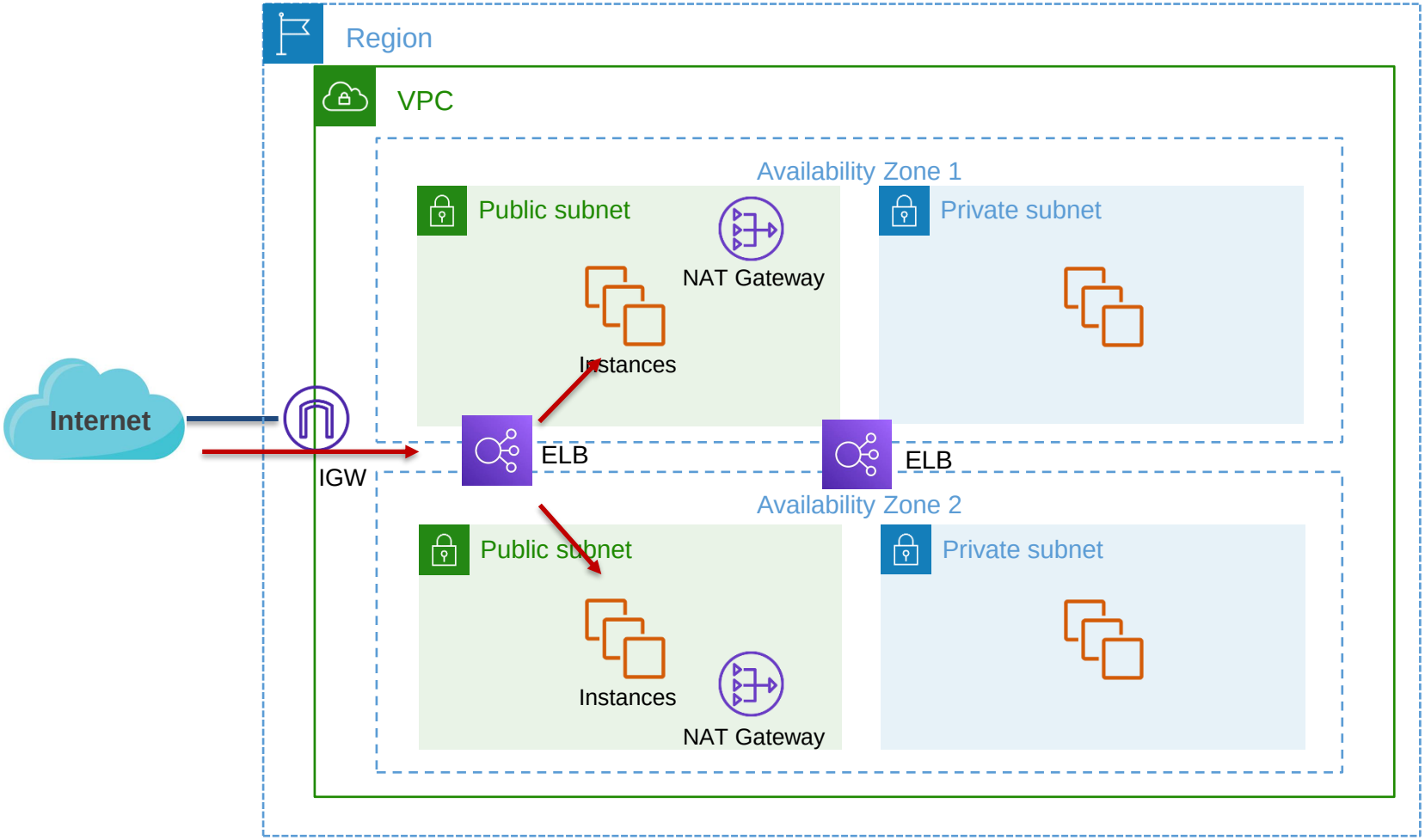
High Availability(고가용성)

- **가용성을 높이기 위해 단일 장애 지점(SPOF)을 제거.**
  - 2개의 가용 영역을 사용하면 더 높은 가용성을 얻을 수 있다.



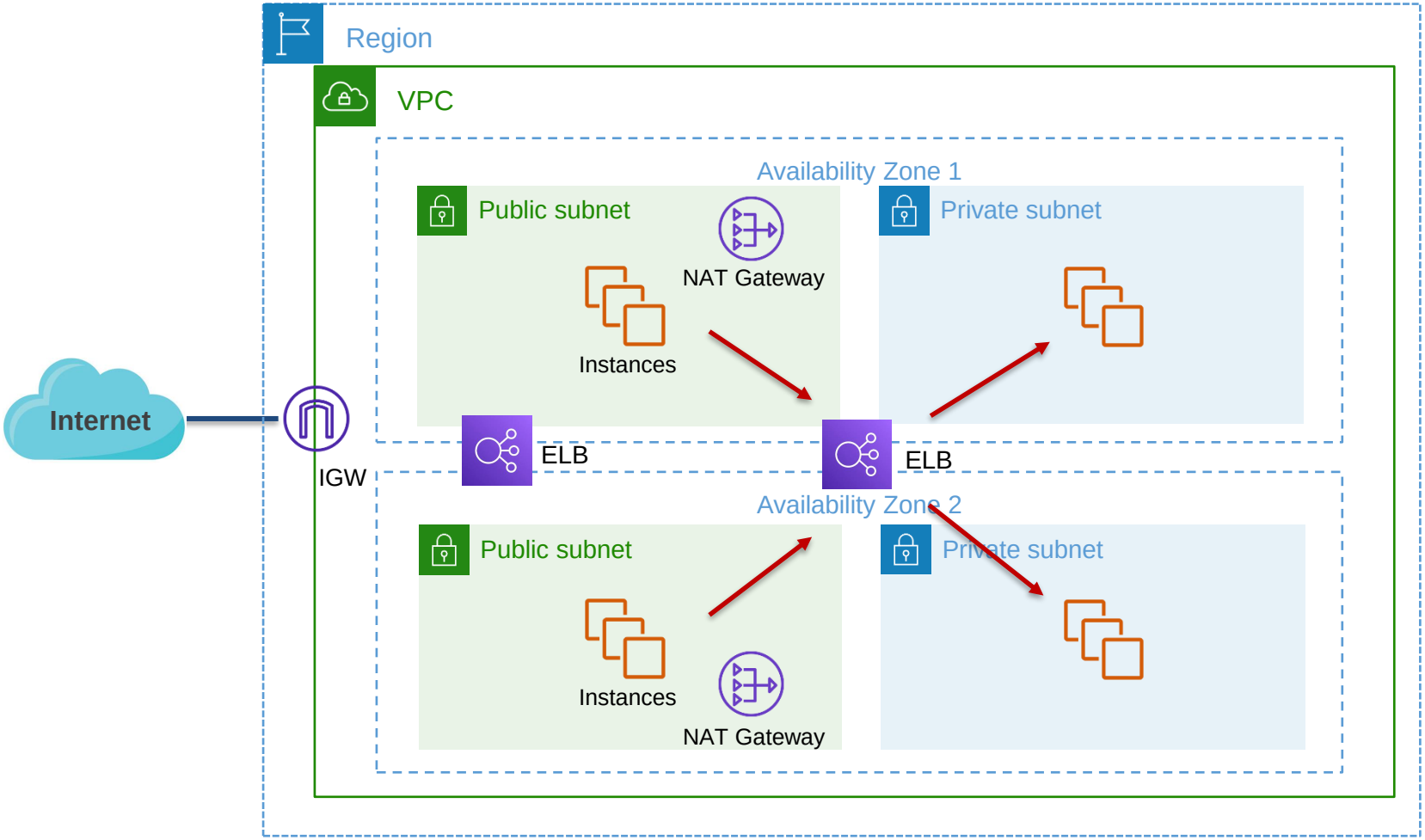
High Availability(고가용성)

- 고가용성 아키텍처 예제



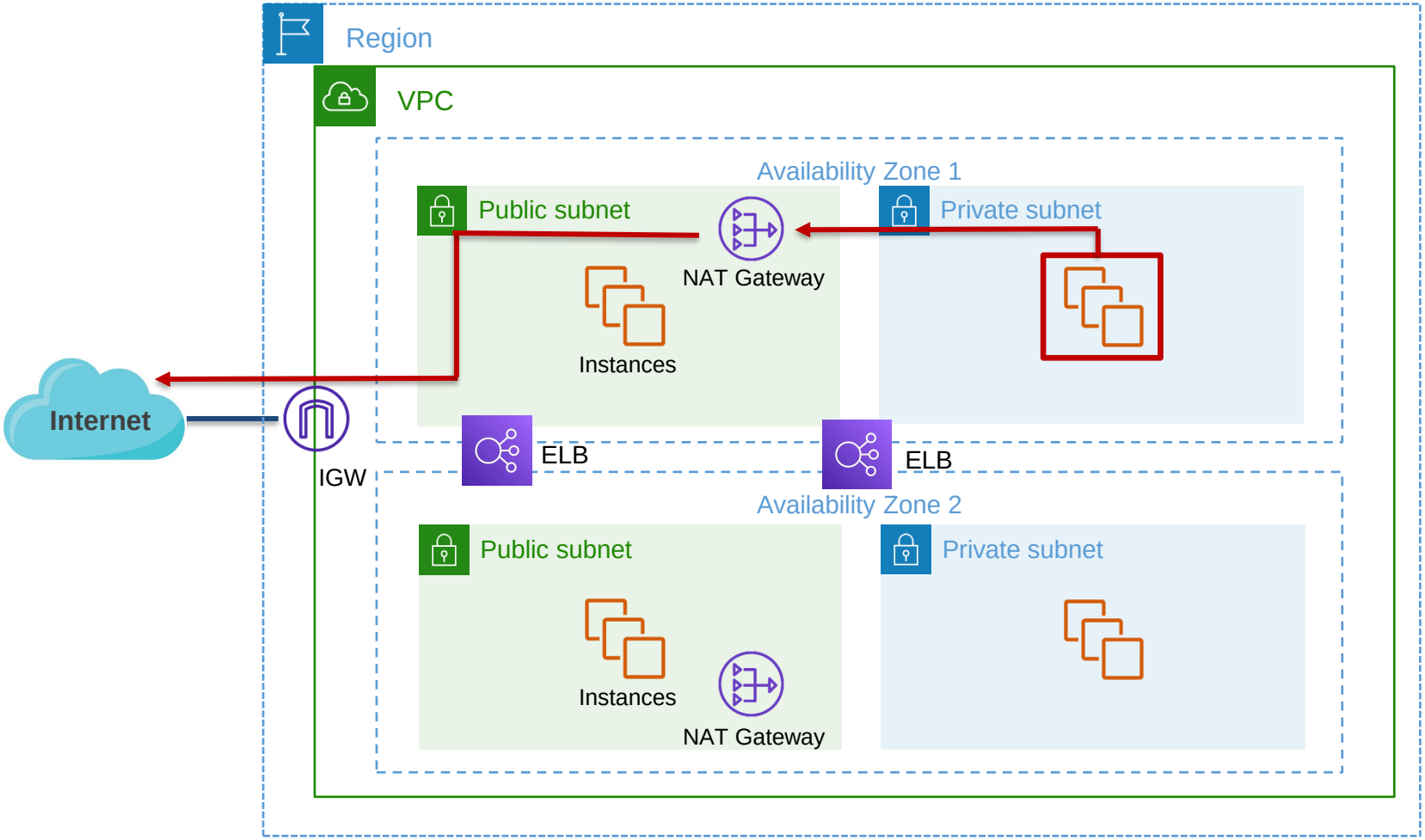
High Availability(고가용성)

- 고가용성 아키텍처 예제



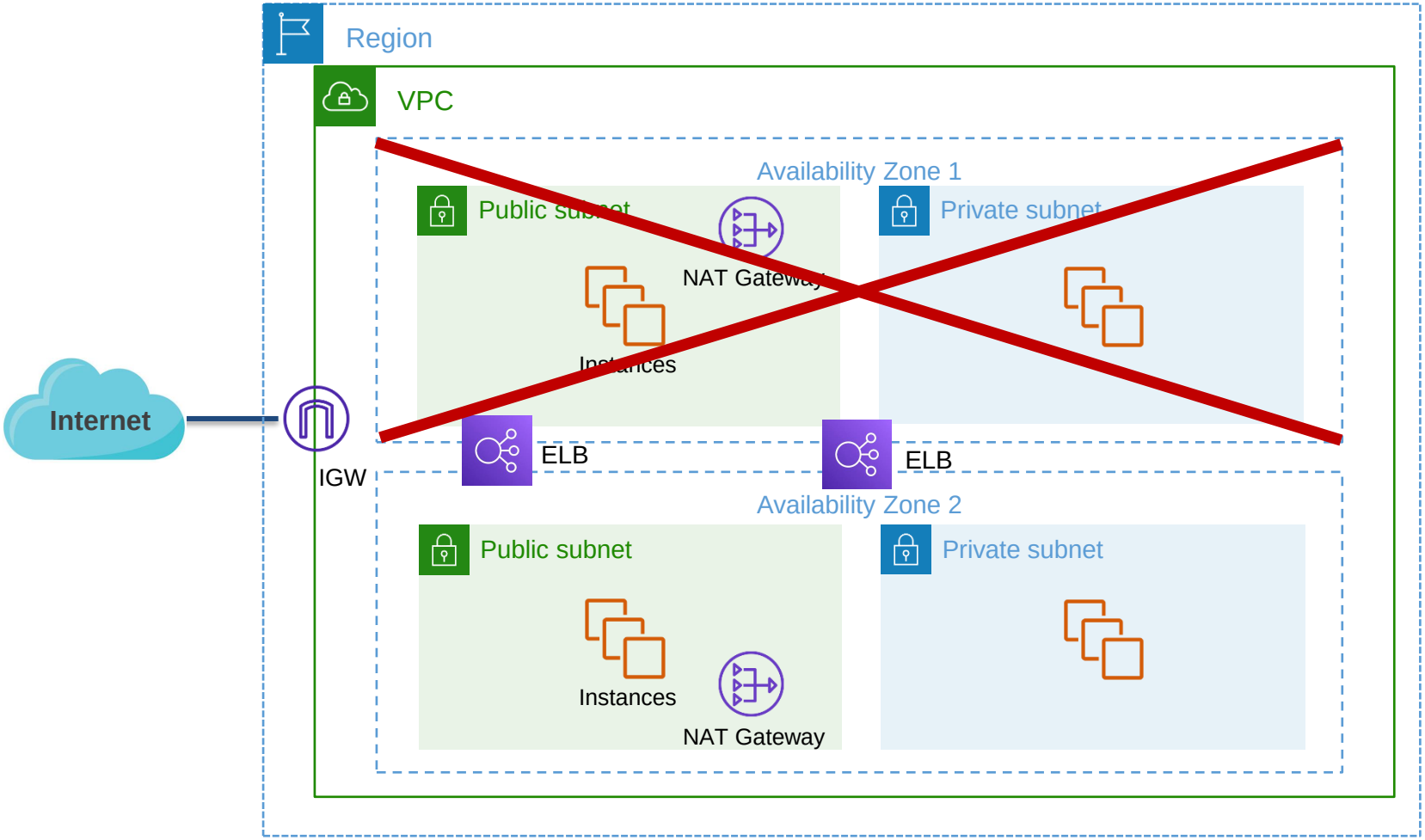
High Availability(고가용성)

- 고가용성 아키텍처 예제



High Availability(고가용성)

- 고가용성 아키텍처 예제



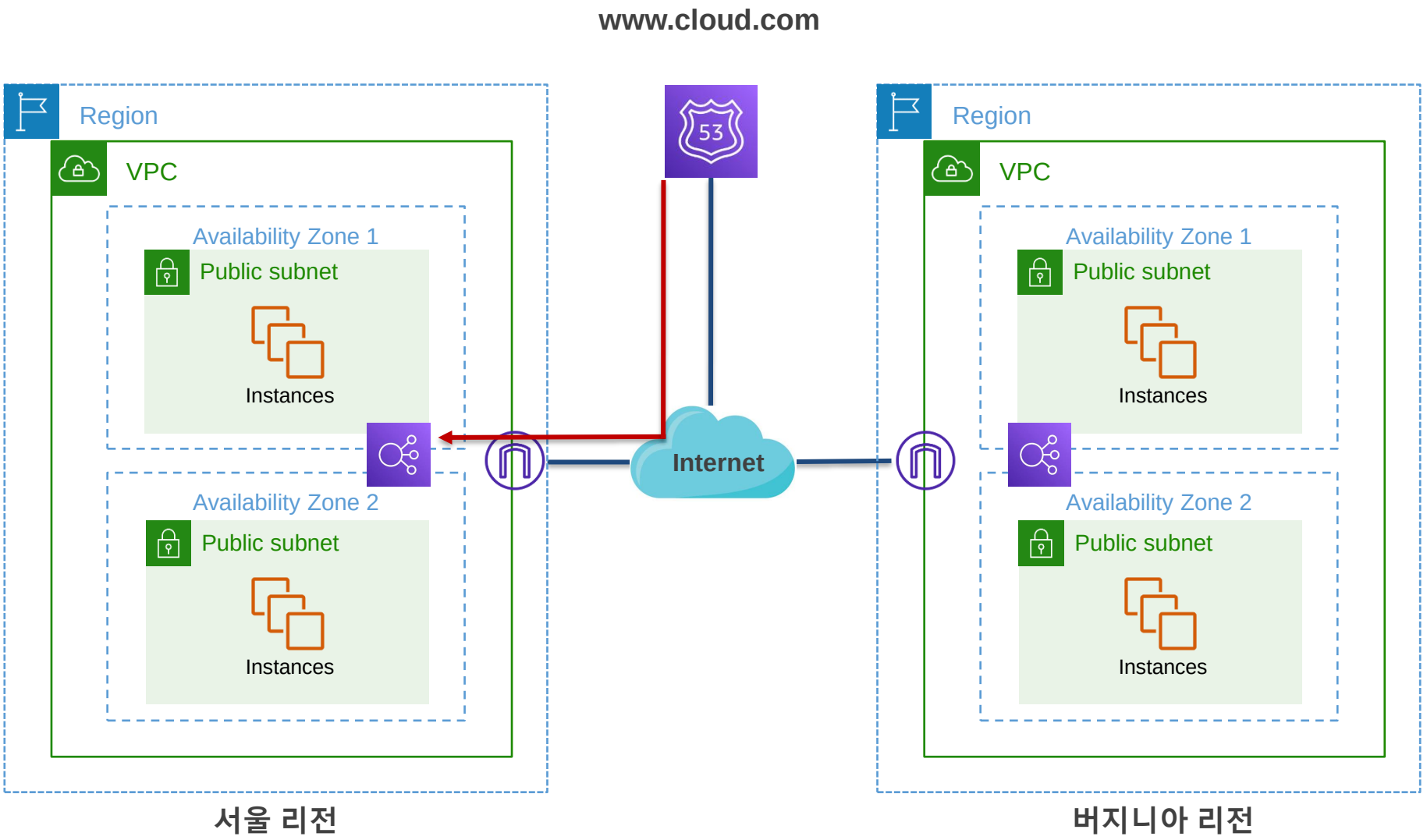
04 Amazon Route 53와 리전 단위 고가용성

Amazon Route 53

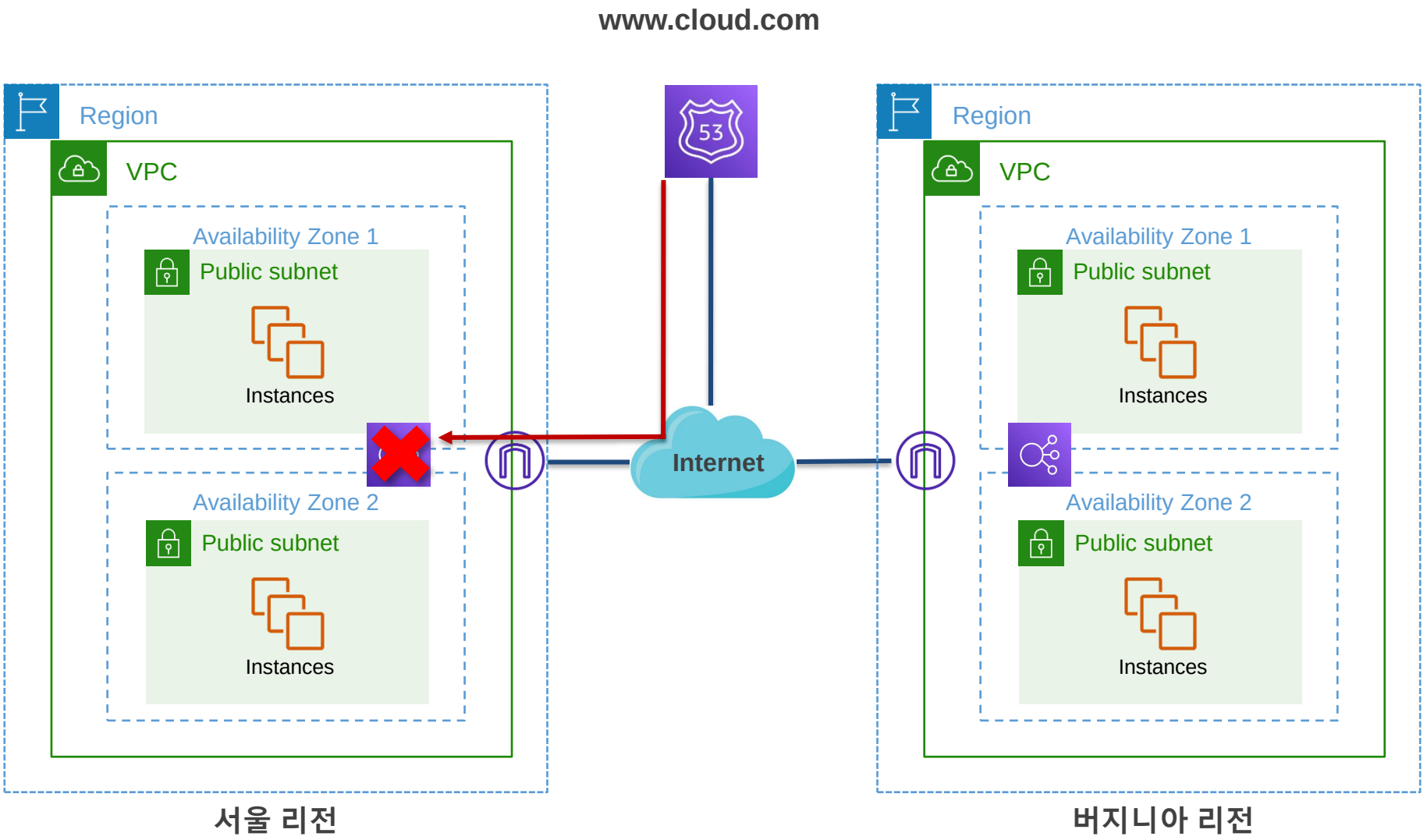
- 가용성과 확장성이 뛰어난 관리형 DNS 서비스.
- 도메인이름 등록 기능.
- 다양한 방식의 라우팅 정책 지원.
- 장애 조치를 통한 고가용성 지원.

Amazon Route 53

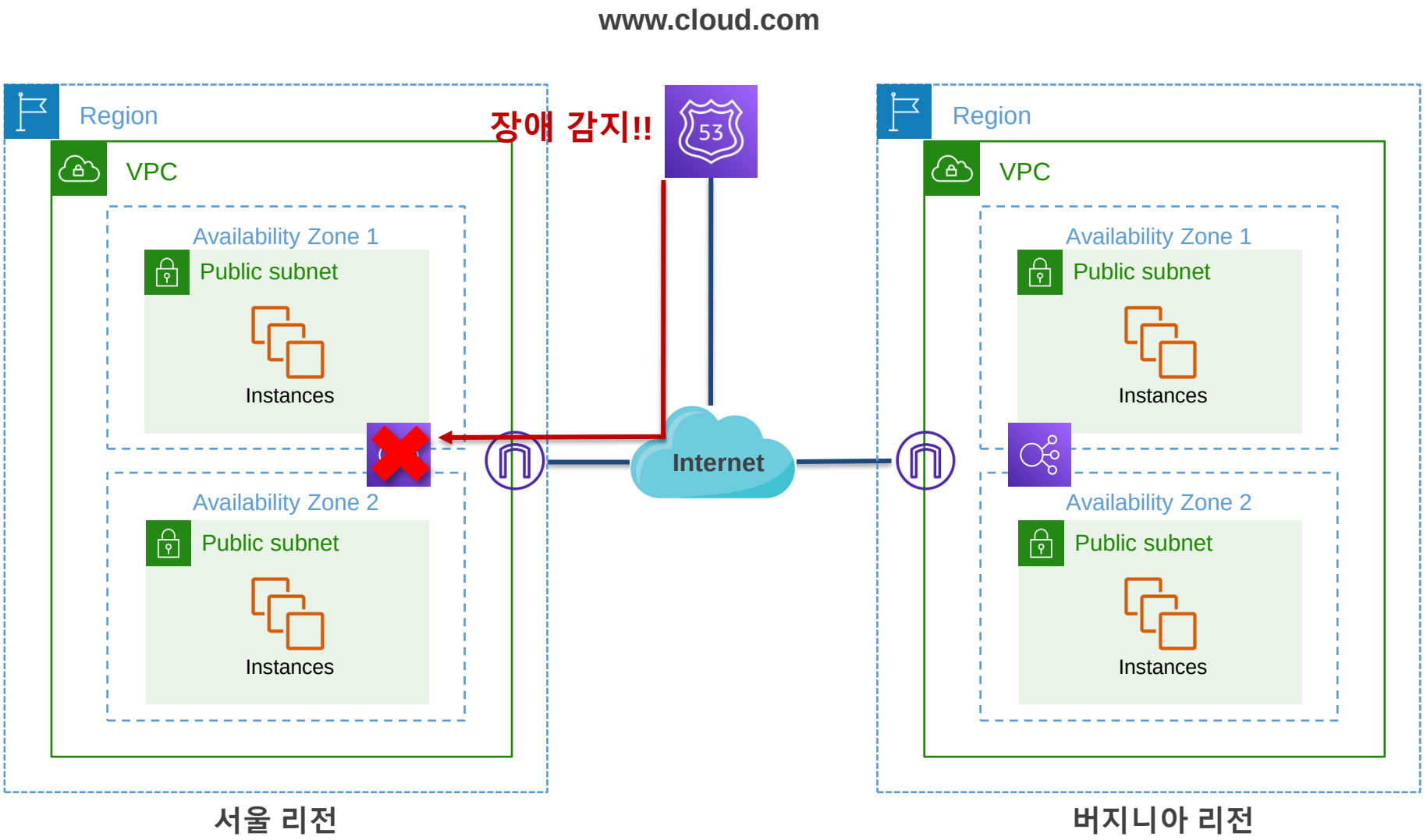
Route 53 장애 조치를 사용한 리전 고가용성 구성



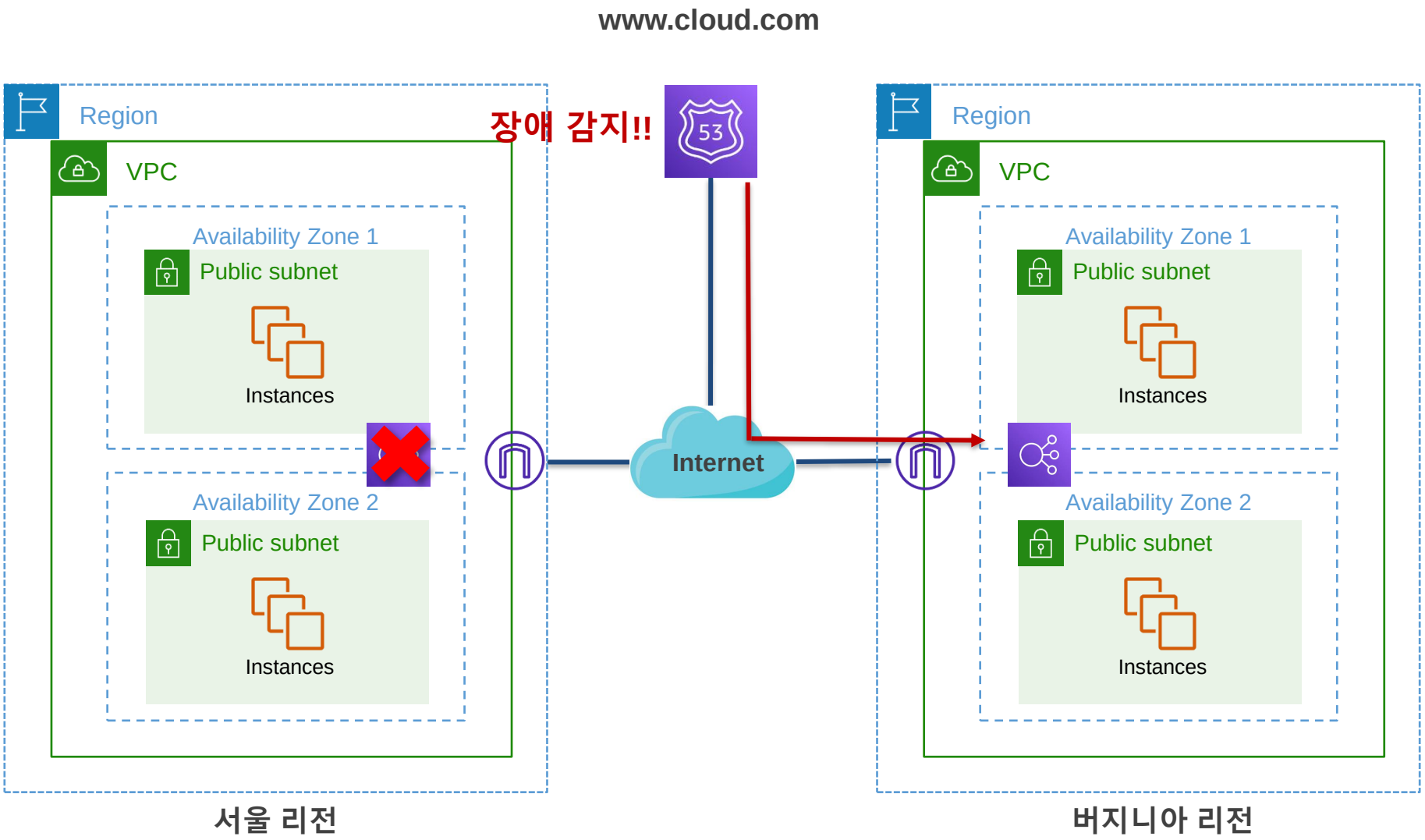
Route 53 장애 조치를 사용한 리전 고가용성 구성



Route 53 장애 조치를 사용한 리전 고가용성 구성



Route 53 장애 조치를 사용한 리전 고가용성 구성



Route 53 라우팅 정책

- 단순 라우팅
- 가중치 기반
- 지리적 위치
- 지연 시간
- 장애 조치
- 다중 값 응답