

AWS 백서

AWS Outposts 고가용성 설계 및 아키텍처 고려 사항



AWS Outposts 고가용성 설계 및 아키텍처 고려 사항: AWS 백서

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon의 상표 및 브랜드 디자인은 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계와 관계없이 해당 소유자의 자산입니다.

Table of Contents

요약 및 소개	i
귀사는 Well-Architected입니까?	1
소개	1
AWS 인프라 및 서비스를 온프레미스 위치로 확장	2
업데이트된 공동 책임 모델 이해	4
장애 모드에 대해서 생각하기	7
장애 모드 1: 네트워크	7
장애 모드 2: 인스턴스	7
장애 모드 3: 컴퓨팅	8
장애 모드 4: 랙 또는 데이터 센터	8
실패 모드 5: AWS 가용 영역 또는 리전	8
AWS Outposts 랙을 이용한 고가용성 애플리케이션 및 인프라 솔루션 구축	10
네트워킹	11
네트워크 연결	11
앵커 연결	15
애플리케이션/워크로드 라우팅	18
컴퓨팅	21
용량 계획	22
용량 관리	25
인스턴스 배치	26
스토리지	29
데이터 보호	29
더 큰 장애 모드	32
결론	36
기여자	37
문서 기록	38
고지 사항	39
AWS 용어집	40
.....	xli

AWS Outposts 고가용성 설계 및 아키텍처 고려 사항

게시 날짜: 2021년 8월 12일([문서 기록](#))

이 백서에서는 IT 관리자와 시스템 아키텍트가 를 사용하여 가용성이 높은 온프레미스 애플리케이션 환경을 구축하는 데 적용할 수 있는 아키텍처 고려 사항과 권장 사례를 설명합니다 AWS Outposts.

귀사는 Well-Architected입니까?

[AWS Well-Architected Framework](#)는 클라우드에서 시스템을 구축할 때 내리는 결정의 장단점을 이해하는 데 도움이 됩니다. 이 프레임워크를 사용하여 클라우드에서 안정적이고 안전하며 효율적이고 비용 효율적인 시스템을 설계하고 운영하기 위한 아키텍처 모범 사례를 살펴볼 수 있습니다. [AWS Management Console](#)에서 무료로 제공되는 [AWS Well-Architected Tool](#)를 사용하면 각 요소에 대한 일련의 질문에 답하여, 이러한 모범 사례와 비교하여 워크로드를 검토할 수 있습니다.

참조 아키텍처 배포, 다이어그램, 백서 등 클라우드 아키텍처에 대한 더 많은 전문가 지침과 모범 사례를 보려면 [AWS 아키텍처 센터](#)를 참조하세요.

소개

이 백서는 AWS 클라우드 플랫폼을 사용하여 애플리케이션을 배포, 마이그레이션 및 운영하고 의 42U 랙 폼 팩터인 [AWS Outposts 랙](#) 를 사용하여 온프레미스에서 해당 애플리케이션을 실행하려는 IT 관리자 및 시스템 아키텍트를 위한 것입니다 [AWS Outposts](#).

랙을 포함하는 고가용성 시스템을 구축하기 위한 아키텍처 패턴, 안티 패턴 및 권장 사례를 소개합니다 AWS Outposts . AWS Outposts 랙 용량을 관리하고 네트워킹 및 데이터 센터 시설 서비스를 사용하여 고가용성 AWS Outposts 랙 인프라 솔루션을 설정하는 방법을 알아봅니다.

AWS Outposts 랙은 클라우드 컴퓨팅, 스토리지 및 네트워킹 기능의 논리적 풀을 제공하는 완전 관리형 서비스입니다. Outposts 랙을 사용하면 고객은 Amazon [Elastic Compute Cloud](#)(Amazon EC2), Amazon [Elastic Block Store](#)(Amazon EBS), Amazon [Amazon S3 on Outposts](#) , Amazon [Elastic Kubernetes Service](#)(Amazon EKS), [Amazon Elastic Container Service](#)(Amazon ECS), [Amazon Relational Database Service](#)(Amazon RDS) 및 [AWS Outposts 의 기타 서비스](#)를 포함한 온프레미스 환경에서 지원되는 AWS 관리형 서비스를 사용할 수 있습니다. Outpost의 서비스는 AWS 리전에서 사용되는 것과 동일한 [AWS Nitro 시스템](#)에서 제공됩니다.

AWS Outposts 랙을 활용하면 익숙한 AWS 클라우드 서비스 및 도구를 사용하여 고가용성 온프레미스 애플리케이션을 구축, 관리 및 확장할 수 있습니다. AWS Outposts 랙은 온프레미스 시스템에 대한 짧

은 지연 시간 액세스, 로컬 데이터 처리, 데이터 레지던시 및 로컬 시스템 상호 의존성이 있는 애플리케이션 마이그레이션이 필요한 워크로드에 적합합니다.

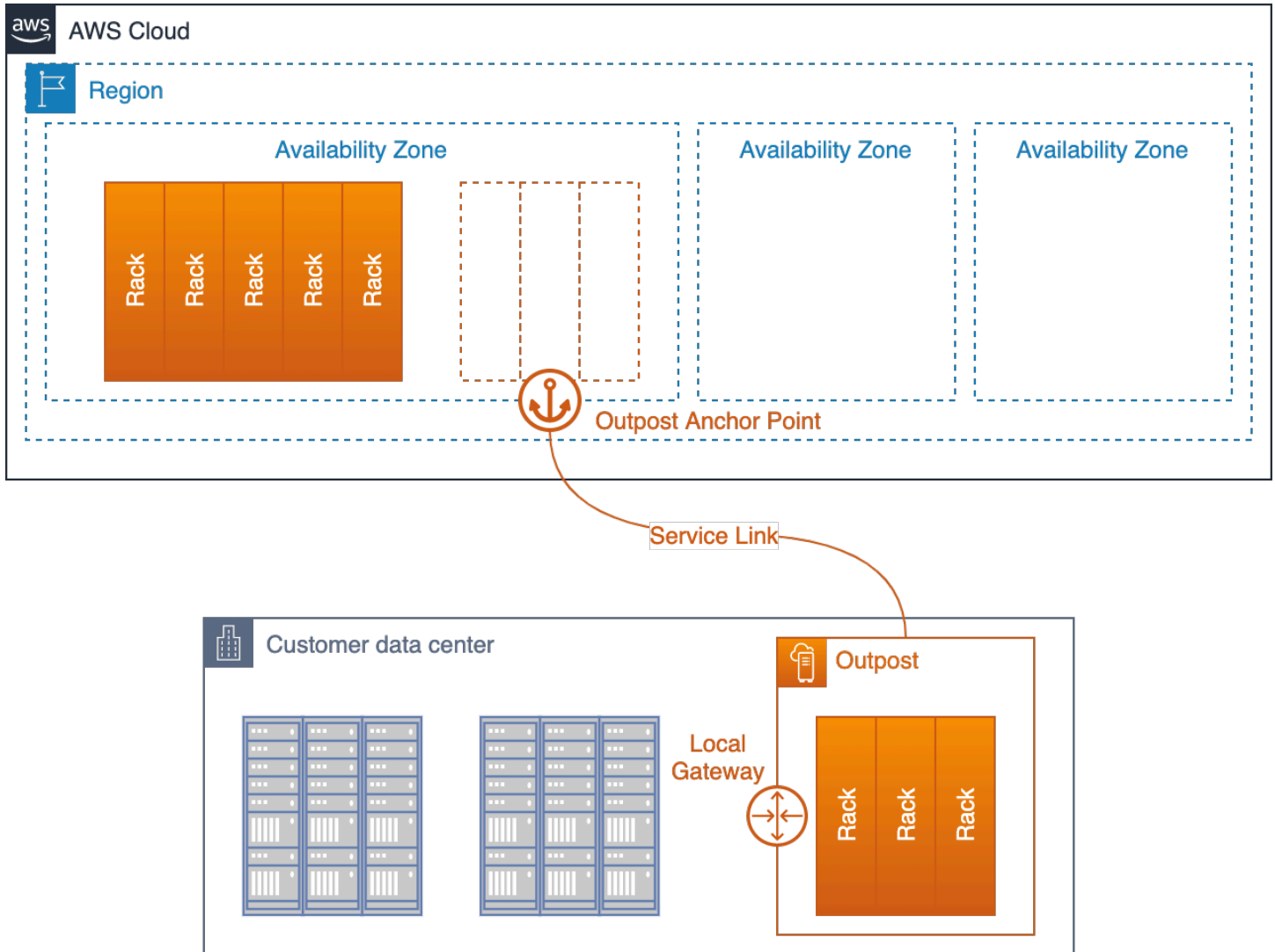
AWS 인프라 및 서비스를 온프레미스 위치로 확장

이 AWS Outposts 서비스는 [50개 이상의 국가 및 영토](#)에 있는 온프레미스 위치에 AWS 인프라와 서비스를 제공하므로 고객은 동일한 AWS 인프라, AWS 서비스APIs, 및 도구를 거의 모든 데이터 센터, 코로케이션 공간 또는 온프레미스 시설에 배포하여 진정으로 일관된 하이브리드 경험을 누릴 수 있습니다. Outposts를 사용하여 설계하는 방법을 이해하려면 AWS 클라우드를 구성하는 다양한 계층을 이해해야 합니다.

[AWS 리전](#)은 세계의 지리적 영역입니다. 각 AWS 리전은 논리적으로 [가용 영역\(\)](#)으로 그룹화된 데이터 센터 모음입니다AZs. 지연 시간이 짧고 처리량이 높으며 중복 네트워크 연결로 연결된 물리적으로 분리되고 격리된 여러(최소 2개) 가용 영역을 AWS 리전 제공합니다. 각 AZ는 하나 이상의 물리적 데이터 센터로 구성됩니다.

논리적 [Outpost](#)(이하 Outpost)는 단일 엔터티로 관리되는 물리적으로 연결된 하나 이상의 AWS Outposts 랙을 배포하는 것입니다. Outpost는 사이트 중 하나에서 AWS 컴퓨팅 및 스토리지 용량 풀에서 AZ의 프라이빗 확장으로 제공합니다 AWS 리전.

에 가장 적합한 개념 모델은 의 AZ에 있는 데이터 센터에서 하나 이상의 랙을 AWS Outposts 분리하는 것입니다 AWS 리전. AZ 데이터 센터에서 귀사의 데이터 센터로 랙을 옮기면 됩니다. 그런 다음 (매우) 긴 케이블을 사용하여 랙을 AZ 데이터 센터의 앵커 포인트에 꽂아 랙이 AWS 리전의 일부로서 계속 작동하도록 합니다. 또한 로컬 네트워크에 연결하여 온프레미스 네트워크와 해당 랙에서 실행되는 워크로드 간의 연결 지연 시간을 단축할 수 있습니다.



고객 데이터 센터에 배포되고 앵커 AZ 및 상위 리전에 다시 연결된 Outpost

Outpost는 AZ가 앵커링된 AZ의 확장으로 작동합니다. 는 인프라의 일부로 인프라를 AWS 운영, 모니터링 및 관리합니다 AWS Outposts AWS 리전. Outpost는 매우 긴 물리적 케이블 대신 서비스 링크라는 암호화된 VPN 터널 세트를 통해 상위 리전에 다시 연결합니다.

서비스 링크는 Outpost의 상위 리전에 있는 가용 영역(AZ)의 앵커 포인트 집합에서 종료됩니다.

콘텐츠를 저장할 위치를 선택할 수 있습니다. AWS 리전 또는 다른 위치에 콘텐츠를 복제하고 백업할 수 있습니다. 콘텐츠는 법률 또는 정부 기관의 구속력 있는 명령을 준수하는 데 필요한 경우를 제외하고 사용자의 동의 없이 선택한 위치 외부로 이동 또는 복사되지 않습니다. 자세한 내용은 [AWS 데이터 프라이버시 섹션을 FAQ](#)참조하세요.

해당 랙에 배포한 워크로드는 로컬에서 실행됩니다. 또한 이러한 랙에서 사용할 수 있는 컴퓨팅 및 스토리지 용량은 한정되어 있고 의 클라우드 규모 서비스를 실행할 수 없지만 랙에 배포된 AWS 리전리

소스(인스턴스 및 로컬 스토리지)는 관리 영역이 에서 계속 작동하는 동안 로컬에서 실행되는 이점을 누릴 수 있습니다 AWS 리전.

Outpost에 워크로드를 배포하려면 Virtual Private Cloud(VPC) 환경에 서브넷을 추가하고 Outpost를 서브넷의 위치로 지정합니다. 그런 다음 , AWS Management Console, CLI, APIs CDK또는 인프라를 코드(IaC) 도구로 통해 지원되는 AWS 리소스를 배포할 때 원하는 서브넷을 선택합니다. Outpost 서브넷의 인스턴스는 VPC 네트워킹을 통해 Outpost 또는 리전의 다른 인스턴스와 통신합니다.

Outpost Service Link는 Outpost 관리 트래픽과 고객 VPC 트래픽(VPCOutpost의 서브넷과 리전의 서브넷 간 트래픽)을 모두 전달합니다.

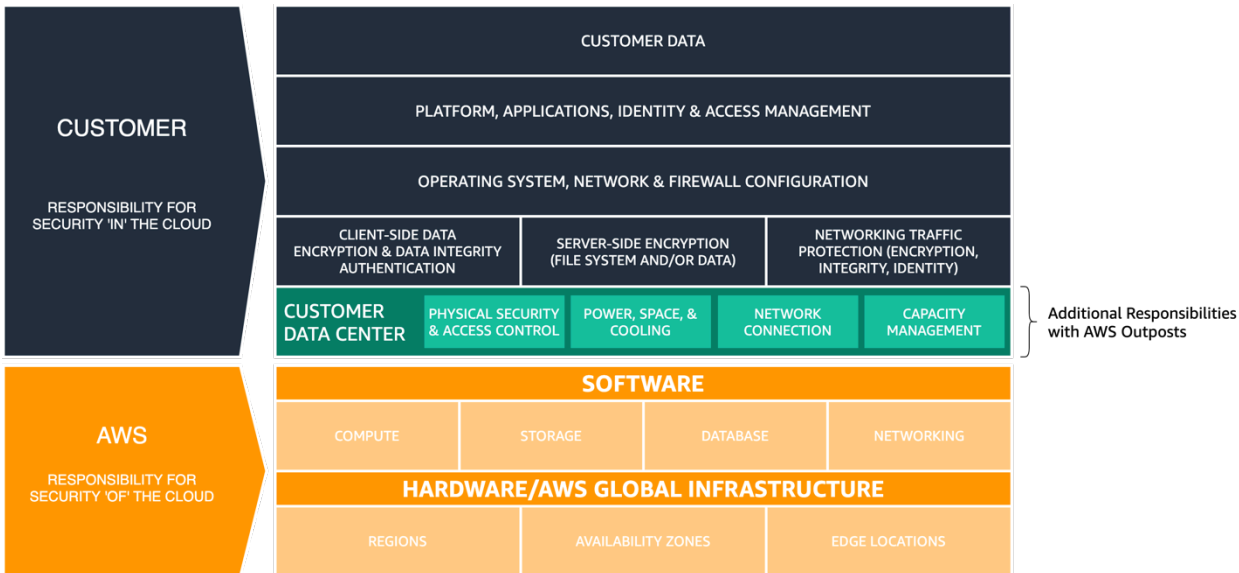
중요 용어

- AWS Outposts - 는 진정으로 일관된 하이브리드 경험을 위해 거의 모든 데이터 센터, 코로케이션 공간 또는 온프레미스 시설에 동일한 AWS 인프라, AWS 서비스, APIs 및 도구를 제공하는 완전 관리형 서비스입니다.
- Outpost - 고객 사이트에 배포된 AWS 컴퓨팅, 스토리지 및 네트워킹 풀과 단일 논리적 엔터티로 관리되는 물리적으로 연결된 하나 이상의 AWS Outposts 랙을 배포하는 것입니다.
- 상위 리전 - Outpost 배포를 위한 관리, 제어 영역 서비스 및 리전 AWS 서비스를 AWS 리전 제공하는 입니다.
- 앵커 가용 영역(앵커 AZ) - Outpost의 앵커 포인트를 호스팅하는 상위 리전의 가용 영역입니다. Outpost는 앵커 가용 영역을 확장한 역할을 합니다.
- 앵커 포인트 - 원격으로 배포된 Outpost로부터 연결을 수신하는 앵커 AZ의 엔드포인트입니다.
- 서비스 링크 - Outpost를 상위 리전의 앵커 가용 영역에 연결하는 암호화된 VPN 터널 세트입니다.
- 로컬 게이트웨이(LGW) - Outpost와 온프레미스 네트워크 간의 통신을 가능하게 하는 논리적 상호 연결 가상 라우터입니다.

업데이트된 공동 책임 모델 이해

데이터 센터 또는 코로케이션 시설에 AWS Outposts 인프라를 배포할 때 [AWS 공동 책임 모델](#)에서 추가 책임을 말합니다. 예를 들어, 리전에서 AWS 는 다양한 전원, 중복 코어 네트워킹 및 탄력적 광역 통신망(WAN) 연결을 제공하여 하나 이상의 구성 요소 장애 발생 시 서비스를 사용할 수 있도록 합니다.

Outpost를 사용하면 Outpost에서 실행되는 워크로드에 대한 가용성 요구 사항을 충족하기 위해 Outpost 랙에 탄력적인 전력 및 네트워크 연결을 제공해야 합니다.



AWS 에 대한 공동 책임 모델이 업데이트되었습니다. AWS Outposts

를 사용하면 데이터 센터 환경의 물리적 보안 및 액세스 제어를 AWS Outposts 담당할 수 있습니다. Outpost를 리전에 다시 연결하려면 Outpost를 계속 운영하고 네트워크 연결을 유지할 수 있도록 충분한 전력, 스페이스 및 냉각 장치를 제공해야 합니다.

Outpost 용량은 유한하며 사이트에 AWS 설치된 랙의 크기와 수에 따라 결정되므로 초기 워크로드를 실행하고 향후 성장을 수용하고 서버 장애 및 유지 관리 이벤트를 완화하기 위한 추가 용량을 제공하는 데 필요한 EC2EBS, 및 S3 on Outposts 용량을 결정해야 합니다.

AWS 는 AWS Outposts 랙 내의 전원 공급 장치, 서버 및 네트워킹 장비를 포함한 Outposts 인프라의 가용성을 담당합니다. 는 AWS 또한 가상화 하이퍼바이저, 스토리지 시스템 및 Outposts에서 실행되는 AWS 서비스를 관리합니다.

각 Outposts 랙의 중앙 전원 선반은 AC 전원을 DC 전원으로 변환하고 버스 바 아키텍처를 통해 랙의 서버에 전원을 공급합니다. 버스 바 아키텍처에서는 랙에 있는 전원 공급 장치의 절반에 장애가 발생해도 모든 서버가 중단 없이 계속 작동할 수 있습니다.

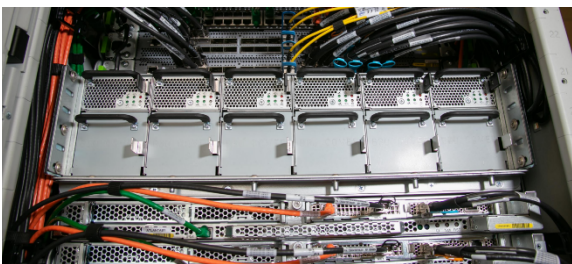


그림 3 - AWS Outposts AC-to-DC 전원 공급 장치 및 버스 바 전원 분배

Outpost 랙 내부 및 랙 사이의 네트워크 스위치와 케이블도 완전히 중복되어 있습니다. 광케이블 패치 패널은 Outpost 랙과 온프레미스 네트워크 간의 연결을 제공하며 고객 관리형 데이터 센터 환경과 관리형 AWS Outposts 환경 간의 경계점 역할을 합니다.

리전에서와 마찬가지로 AWS 는 Outposts에서 제공되는 클라우드 서비스를 담당하며 Amazon RDS on Outposts와 같은 상위 수준 관리형 서비스를 선택하고 배포할 때 추가 책임을 집니다. 고려할 때 개별 서비스에 대한 [AWS 공동 책임 모델](#) 및 자주 묻는 질문(FAQ) 페이지를 검토하고 Outposts에 배포할 서비스를 선택해야 합니다. 이러한 리소스는 사용자와 간의 책임 분할에 대한 추가 세부 정보를 제공합니다 AWS.

장애 모드에 대해서 생각하기

가용성이 높은 응용 프로그램 또는 시스템을 설계할 때는 장애가 발생할 수 있는 구성 요소, 구성 요소 장애가 시스템에 미치는 영향, 구성 요소 오류의 영향을 완화 또는 제거하기 위해 구현할 수 있는 메커니즘을 고려해야 합니다. 애플리케이션이 단일 서버, 단일 랙 또는 단일 데이터 센터에서 실행되나요? 서버, 랙 또는 데이터 센터에 일시적 또는 영구적 장애가 발생하면 어떻게 되나요? 네트워킹과 같은 중요한 하위 시스템이나 애플리케이션 자체에 장애가 발생하면 어떻게 되나요? 이러한 것들이 장애 모드입니다.

Outpost 및 애플리케이션 배포를 계획할 때는 이 섹션의 장애 모드를 고려해야 합니다. 다음 섹션에서는 이러한 장애 모드를 완화하여 애플리케이션 환경에 향상된 수준의 고가용성을 제공하는 방법을 검토합니다.

장애 모드 1: 네트워크

Outpost 배포는 관리 및 모니터링을 위한 상위 리전과의 탄력적인 연결을 기반으로 합니다. 네트워크 장애는 운영자 오류, 장비 장애, 서비스 제공업체 운영 중단과 같은 다양한 장애로 인해 발생할 수 있습니다. 현장에 연결된 하나 이상의 랙으로 구성될 수 있는 Outpost는 서비스 링크를 통해 해당 리전과 통신할 수 없는 경우 연결이 끊긴 것으로 간주됩니다.

네트워크 경로를 중복시키면 연결 해제 이벤트의 위험을 완화하는 데 도움이 될 수 있습니다. 애플리케이션 종속성과 네트워크 트래픽을 매핑하여 연결 해제 이벤트가 워크로드 운영에 미치는 영향을 이해해야 합니다. 애플리케이션 가용성 요구 사항을 충족할 수 있도록 충분한 네트워크 중복을 계획하세요.

연결 해제 이벤트 중에 Outpost에서 실행되는 인스턴스는 계속 실행되며 Outpost Local Gateway()를 통해 온프레미스 네트워크에서 액세스할 수 있습니다. 로컬 워크로드 및 서비스가 해당 리전의 서비스를 사용하는 경우 로컬 워크로드 및 서비스가 손상되거나 실패할 수 있습니다. Outpost가 리전에서 연결 해제된 동안 요청(예: Outpost에서 인스턴스 시작 또는 중지), 제어 영역 작업 및 서비스 원격 측정(예: CloudWatch 지표)을 변경하면 실패합니다.

장애 모드 2: 인스턴스

EC2 실행 중인 서버에 문제가 있거나 인스턴스에 운영 체제 또는 애플리케이션 장애가 발생하는 경우 인스턴스가 손상되거나 실패할 수 있습니다. 애플리케이션이 이러한 유형의 장애를 처리하는 방법은 애플리케이션 아키텍처에 따라 다릅니다. 모놀리식 애플리케이션은 일반적으로 복구에 애플리케이션 또는 시스템 기능을 사용하는 반면, 모듈식 서비스 지향 또는 마이크로 서비스 아키텍처는 일반적으로 장애가 발생한 구성 요소를 대체하여 서비스 가용성을 유지합니다.

EC2 Auto Scaling 그룹과 같은 자동화된 메커니즘을 사용하여 실패한 인스턴스를 새 인스턴스로 교체할 수 있습니다. 인스턴스 자동 복구는 나머지 서버에 충분한 여유 용량이 있는 경우 서버 장애로 인해 장애가 발생한 인스턴스를 다시 시작할 수 있습니다.

장애 모드 3: 컴퓨팅

서버에 장애가 발생하거나 손상이 발생할 수 있으며 구성 요소 장애 및 예정된 유지 관리 작업과 같은 다양한 이유로, 일시적 또는 영구적으로 운영을 중단해야 할 수 있습니다. Outpost 랙의 서비스가 서버 장애 및 장애를 처리하는 방법은 다양하며 고객이 고가용성 옵션을 구성하는 방법에 따라 달라질 수 있습니다.

N+M 가용성 모델을 지원하려면 충분한 컴퓨팅 용량을 주문해야 합니다. 이때 N은 필요한 용량이며 M은 서버 장애를 수용할 수 있도록 할당된 예비 용량입니다.

장애가 발생한 서버에 대한 하드웨어 교체는 완전 관리형 AWS Outposts 랙 서비스의 일부로 제공됩니다. Outpost 배포에서 모든 서버 및 네트워킹 디바이스의 상태를 AWS 능동적으로 모니터링합니다. 물리적인 유지 관리가 필요한 경우 AWS 는 시간을 내서 현장을 방문하여 장애가 발생한 구성 요소를 교체해 드립니다. 예비 용량을 프로비저닝하면 장애가 발생한 서버가 서비스를 중단하고 교체되는 동안에도 워크로드를 계속 실행할 수 있습니다.

장애 모드 4: 랙 또는 데이터 센터

랙의 완전한 전원 손실이나 냉방 손실과 같은 환경적 장애로, 또는 홍수나 지진으로 인한 데이터 센터의 물리적 손상으로 인해 랙 장애가 발생할 수 있습니다. 데이터 센터 배전 아키텍처에 결함이 있거나 표준 데이터 센터 전원 유지 관리 중에 오류가 발생하면 하나 이상의 랙 또는 전체 데이터 센터의 전력이 손실될 수 있습니다.

동일한 캠퍼스 또는 대도시 지역 내에서 서로 독립된 여러 데이터 센터 층이나 위치에 인프라를 배포하면 이러한 시나리오를 완화할 수 있습니다.

AWS Outposts 랙을 사용하여 이 접근 방식을 수행하려면 애플리케이션 가용성을 유지하기 위해 여러 개의 개별 논리적 Outpost에서 애플리케이션을 설계하고 배포하는 방법을 신중하게 고려해야 합니다.

실패 모드 5: AWS 가용 영역 또는 리전

각 Outpost는 AWS 리전내의 특정 가용 영역(AZ)에 고정되어 있습니다. 앵커 AZ 또는 상위 리전 내에서 장애가 발생하면 Outpost 관리 및 변형 가능성이 손실되고 Outpost와 리전 간의 네트워크 통신이 중단될 수 있습니다.

네트워크 장애와 마찬가지로 AZ 또는 리전 장애로 인해 Outpost와 리전의 연결이 끊길 수 있습니다. Outpost에서 실행되는 인스턴스는 계속 실행되며 앞서 설명한 대로 Outpost Local Gateway(LGW)를 통해 온프레미스 네트워크에서 액세스할 수 있으며 리전의 서비스에 의존하는 경우 손상되거나 실패할 수 있습니다.

AWS AZ 및 리전 장애의 영향을 완화하기 위해 각각 다른 AZ 또는 리전에 고정된 여러 Outpost를 배포할 수 있습니다. 그런 다음 AWS 현재 설계하고 배포하는 데 사용하는 유사한 [메커니즘과 아키텍처 패턴](#)을 많이 사용하여 분산 다중 Outpost 배포 모델에서 작동하도록 워크로드를 설계할 수 있습니다.

AWS Outposts 랙을 사용하여 HA 애플리케이션 및 인프라 솔루션 구축

AWS Outposts 랙을 사용하면 익숙한 AWS 클라우드 서비스 및 도구를 사용하여 가용성이 높은 온프레미스 애플리케이션을 구축, 관리 및 확장할 수 있습니다. 클라우드 고가용성 아키텍처 및 접근 방식은 일반적으로 현재 데이터 센터에서 실행 중인 기존 온프레미스 고가용성 아키텍처와 다르다는 점을 이해하는 것이 중요합니다.

기존 온프레미스 HA 애플리케이션 배포를 사용하면 애플리케이션이 가상 머신(VM)에 배포됩니다. VMs. 복잡한 IT 시스템 및 인프라를 배포하고 유지 관리하여 가상 머신이 정상적으로 실행되도록 유지합니다. 이는 VMs 종종 특정 자격 증명이 있으며 각 VM은 전체 애플리케이션 아키텍처에서 중요한 역할을 할 수 있습니다.

아키텍처 역할은 VM 자격 증명과 밀접하게 연결되어 있습니다. 시스템 설계자는 IT 인프라 기능을 활용하여 각 VM에 컴퓨팅 용량, 스토리지 볼륨 및 네트워크 서비스에 대한 안정적인 액세스를 제공하는 고가용성 VM 런타임 환경을 제공합니다. VM에 장애가 발생하면 자동 또는 수동 복구 프로세스를 실행하여 장애가 발생한 VM을 정상 상태로 복원합니다. 대개 다른 인프라 또는 다른 데이터 센터에서 전적으로 복원됩니다.

클라우드 HA 아키텍처는 다른 접근 방식을 취합니다. AWS 클라우드 서비스는 안정적인 컴퓨팅, 스토리지 및 네트워킹 기능을 제공합니다. 애플리케이션 구성 요소는 EC2 인스턴스, 컨테이너, 서버리스 함수 또는 기타 관리형 서비스에 배포됩니다.

인스턴스는 애플리케이션 구성 요소를 인스턴스화한 것으로, 해당 역할을 수행하는 여러 구성 요소 중 하나일 수 있습니다. 애플리케이션 구성 요소는 서로 느슨하게 결합되어 있으며 전체 애플리케이션 아키텍처에서 수행하는 역할에도 영향을 미칩니다. 인스턴스의 개별 자격 증명은 일반적으로 중요하지 않습니다. 수요에 따라 스케일 업 또는 스케일 다운을 위해 추가 인스턴스를 만들거나 폐기할 수 있습니다. 장애가 발생한 인스턴스나 비정상 인스턴스는 새로운 정상 인스턴스로 간단히 교체됩니다.

AWS Outposts 랙은 AWS 컴퓨팅, 스토리지, 네트워킹, 데이터베이스 및 기타 클라우드 서비스를 온프레미스 위치로 확장하여 진정으로 일관된 하이브리드 경험을 제공하는 완전 관리형 서비스입니다. Outpost 랙 서비스를 기존의 온프레미스 고가용성 메커니즘을 사용하는 IT 인프라 시스템을 즉시 대체하는 것으로 생각해서는 안 됩니다. AWS 서비스 및 Outposts를 사용하여 기존 온프레미스 HA 아키텍처를 지원하려는 시도는 패턴 방지입니다.

AWS Outposts 랙에서 실행되는 워크로드는 [Amazon EC2 Auto Scaling](#)(워크로드 수요에 맞게 수평으로 확장), [EC2 상태 확인](#)(상태가 좋지 않은 인스턴스를 감지 및 제거), [Application Load Balancer](#)(수신되는 워크로드 트래픽을 확장 또는 대체된 인스턴스로 리디렉션)와 같은 클라우드 HA 메커니즘을 사

용합니다. 애플리케이션을 클라우드로 마이그레이션할 때 AWS 리전, 또는 AWS Outposts 랙으로 마이그레이션할 때 HA 애플리케이션 아키텍처를 업데이트하여 관리형 클라우드 서비스 및 클라우드 HA 메커니즘을 활용해야 합니다.

다음 섹션에서는 온프레미스 환경에 AWS Outposts 랙을 배포하여 고가용성 요구 사항으로 워크로드를 실행하기 위한 아키텍처 패턴, 패턴 방지 및 권장 사례를 소개합니다. 이 섹션에서는 패턴과 방법을 소개하지만 구성 및 구현 세부 정보는 제공하지 않습니다. Outposts [AWS Outposts 랙FAQs](#)을 위한 환경과 서비스로 마이그레이션하기 위한 애플리케이션을 준비할 때 Outposts 랙에서 실행되는 서비스에 대한 랙 FAQs 및 [사용 설명서](#)와 및 AWS 서비스 설명서를 읽고 숙지해야 합니다.

주제

- [네트워킹](#)
- [컴퓨팅](#)
- [스토리지](#)
- [더 큰 장애 모드](#)

네트워킹

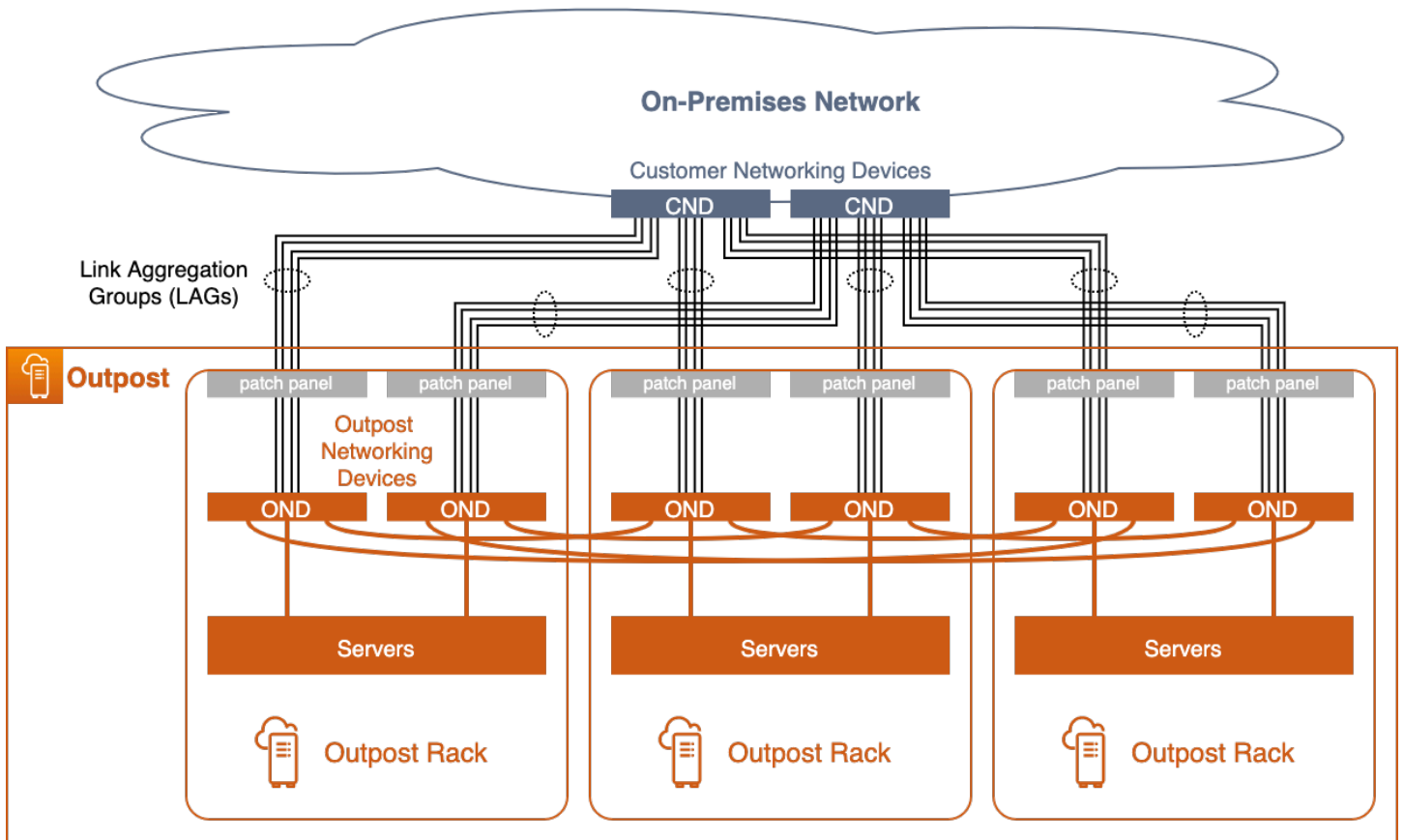
Outpost 배포는 관리, 모니터링 및 서비스 운영이 제대로 작동할 수 있도록 앵커 AZ에 대한 탄력적인 연결을 필요로 합니다. 각 Outpost 랙에 중복 네트워크 연결을 제공하고 AWS 클라우드의 앵커 포인트에 다시 안정적인 연결을 제공하도록 온프레미스 네트워크를 프로비저닝해야 합니다. 또한 Outpost에서 실행되는 애플리케이션 워크로드와 해당 워크로드가 통신하는 다른 온프레미스 및 클라우드 시스템 간의 네트워크 경로를 고려해 보세요. 네트워크에서 이 트래픽을 어떻게 라우팅하시겠습니까?

주제

- [네트워크 연결](#)
- [앵커 연결](#)
- [애플리케이션/워크로드 라우팅](#)

네트워크 연결

각 AWS Outposts 랙은 Outpost 네트워크 디바이스()라는 중복 top-of-rack 스위치로 구성됩니다. ONDs. 각 랙의 컴퓨팅 및 스토리지 서버는 두 에 모두 연결됩니다. ONDs. 각 Outpost 랙 OND에 대해 다양한 물리적 및 논리적 경로를 제공하려면 데이터 센터의 고객 네트워크 디바이스(CND)라는 별도의 스위치에 각각 연결해야 합니다. ONDs 광섬유 케이블 및 광학 트랜시버 CNDs를 사용하여 하나 이상의 물리적 연결을 사용하여 에 연결합니다. [물리적 연결](#)은 논리적 [링크 집계 그룹\(LAG\) 링크](#)로 구성됩니다.



중복 네트워크 연결이 있는 멀티랙 Outpost

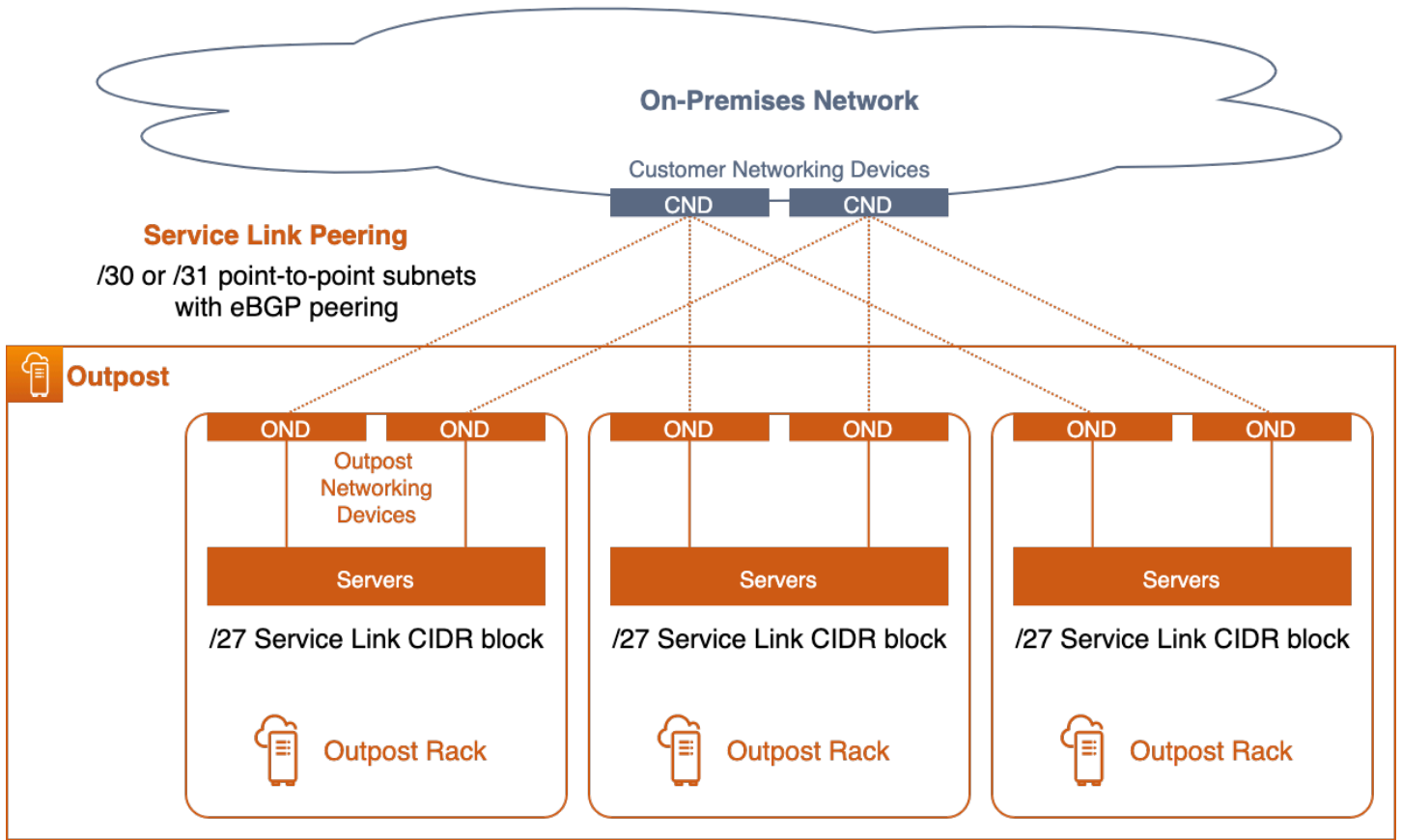
물리적 연결이 단일 광섬유 케이블 LAG인 경우에도 OND to CND 링크는 항상 에서 구성됩니다. 링크를 LAG 그룹으로 구성하면 논리적 그룹에 물리적 연결을 추가하여 링크 대역폭을 늘릴 수 있습니다. LAG 링크는 Outpost와 온프레미스 네트워크 간에 분리된 네트워킹을 활성화하기 위해 IEEE802.1q 이더넷 트렁크로 구성됩니다.

모든 Outpost에는 고객 네트워크와 통신하거나 고객 네트워크를 통해 통신해야 하는 논리적으로 분리된 네트워크가 두 개 이상 있습니다.

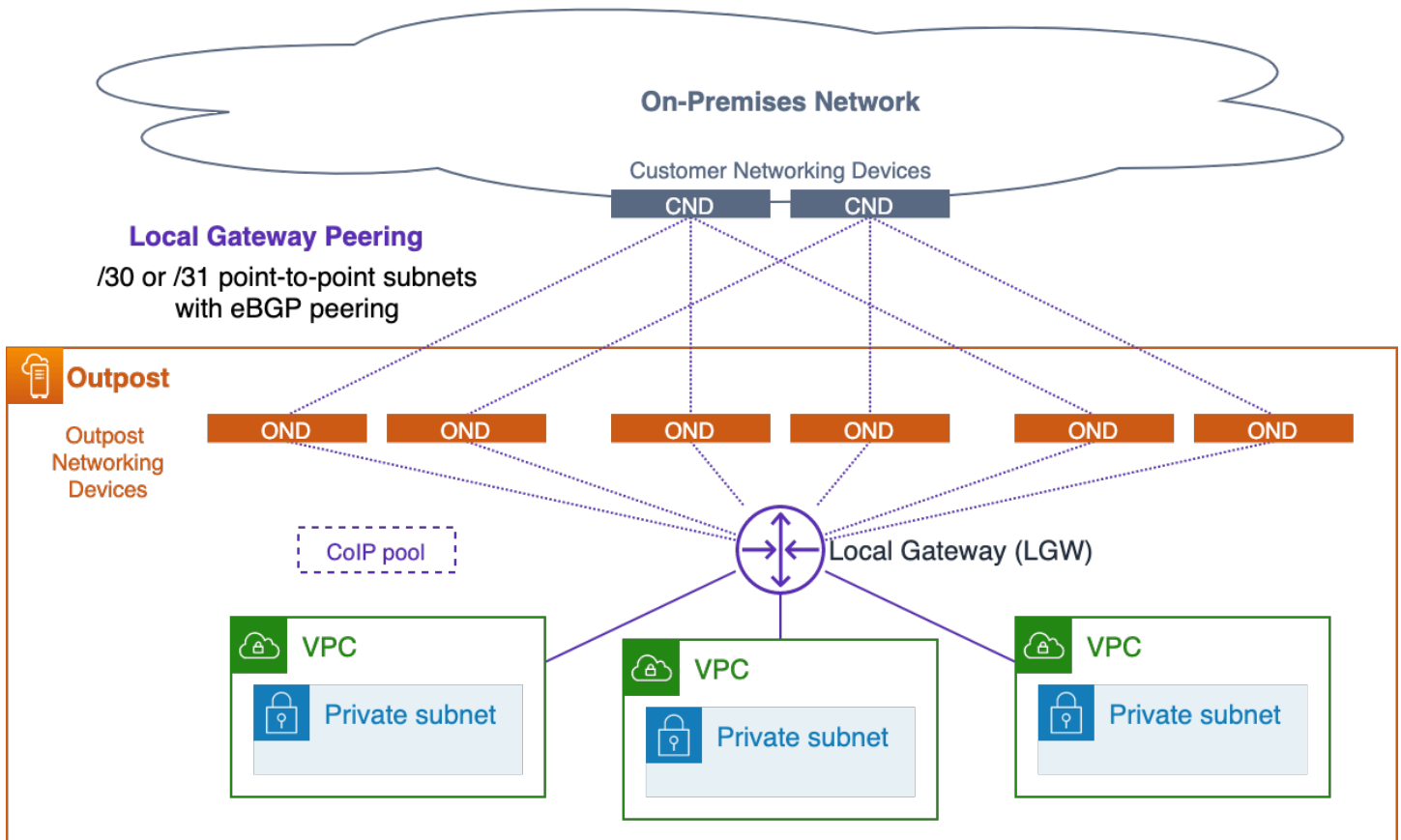
- 서비스 링크 네트워크 - 서비스 링크 IP 주소를 Outpost 서버에 할당하고 온프레미스 네트워크와의 통신을 용이하게 하여 서버가 리전의 Outpost 앵커 포인트에 다시 연결할 수 있도록 합니다.
- Local Gateway 네트워크 - Outpost Local Gateway()를 통해 Outpost의 VPC 서브넷과 온프레미스 네트워크 간의 통신을 활성화합니다 LGW.

이러한 분리된 네트워크는 LAG 링크를 통한 일련의 [point-to-point IP 연결](#)을 통해 온프레미스 네트워크에 연결됩니다. 각 OND CND LAG 연결은 각 분리된 네트워크(서비스 링크 및)에 대해 VLAN IDs, point-to-point (/30 또는 /31) IP 서브넷 및 eBGP 피어링으로 구성됩니다 LGW. point-to-point VLANs 및

서브넷이 있는 LAG 링크를 계층 2 분할, 라우팅된 계층 3 연결로 고려해야 합니다. 라우팅된 IP 연결은 Outpost의 분리된 네트워크와 온프레미스 네트워크 간의 통신을 용이하게 하는 중복 논리적 경로를 제공합니다.



서비스 링크 피어링



로컬 게이트웨이 피어링

직접 연결된 CND 스위치에서 계층 2 LAG 링크(및 해당 VLANs)를 종료하고 CND 스위치에서 IP 인터페이스 및 BGP 피어링을 구성해야 합니다. 데이터 센터 스위치 LAG VLANs 간에 브리지해서는 안 됩니다. 자세한 내용은 AWS Outposts 사용 설명서의 [네트워크 계층 연결](#)을 참조하세요.

논리적 다중 랙 Outpost 내에서 ONDs는 중복으로 상호 연결되어 랙과 서버에서 실행되는 워크로드 간에 가용성이 높은 네트워크 연결을 제공합니다. AWS 는 Outpost 내에서 네트워크 가용성을 담당합니다.

고가용성 네트워크 연결을 위한 권장 방법

- Outpost 랙의 각 Outpost 네트워킹 디바이스(OND)를 데이터 센터의 별도의 고객 네트워킹 디바이스(CND)에 연결합니다.
- 직접 연결된 Customer Networking Device(CND) 스위치에서 계층 2 링크 VLANs, , 계층 3 IP 서브넷 및 BGP 피어링을 종료합니다. 온프레미스 네트워크 OND CND VLANs 간에 CNDs 또는 온프레미스 네트워크를 통해 브리지하지 마세요.

- 링크 집계 그룹(LAGs)에 대한 링크를 추가하여 Outpost와 데이터 센터 간에 사용 가능한 대역폭을 늘립니다. 두 를 모두 통해 다양한 경로의 집계 대역폭에 의존하지 마세요 ONDs.
- 중복을 통해 다양한 경로를 사용하여 Outpost 네트워크와 온프레미스 네트워크 간에 탄력적인 연결을 ONDs 제공합니다.
- 최적의 중복성을 달성하고 중단 없는 OND 유지 관리를 허용하려면 고객이 다음과 같이 BGP 광고 및 정책을 구성하는 것이 좋습니다.
 - 고객 네트워크 장비는 BGP 속성을 변경하지 않고 BGP 다중 경로/로드 밸런싱을 활성화하여 최적의 인바운드 트래픽 흐름(고객에서 Outpost로)을 달성할 수 있도록 Outpost로부터 BGP 광고를 수신해야 합니다. AS-Path 접두사는 유지 관리가 필요한 경우 Outpost BGP 접두사가 트래픽을 특정 OND/uplink에서 멀리 이동시키는 데 사용됩니다. 고객 네트워크는 AS-Path 길이가 1인 Outpost의 경로를 AS-Path 길이가 4인 경로, 즉 AS-Path 프리펜딩에 반응하는 경로보다 선호해야 합니다.
 - 고객 네트워크는 Outpost ONDs의 모든 에 동일한 속성을 가진 동일한 BGP 접두사를 광고해야 합니다. 기본적으로 Outpost 네트워크는 모든 업링크 간에 아웃바운드 트래픽 (고객 대상) 의 부하를 분산합니다. 라우팅 정책은 Outpost 측에서 유지 관리가 필요한 OND 경우 트래픽을 특정 에서 멀리 이동시키는 데 사용됩니다. 이 트래픽 전환을 수행하고 중단 없는 방식으로 유지 관리를 수행하려면 고객 측의 BGP 접두사가 모두 같아 ONDs야 합니다. 고객 네트워크에 유지 관리가 필요한 경우 AS-Path 프리펜딩을 사용하여 특정 업링크 또는 장치에서 트래픽을 일시적으로 전환시키는 것이 좋습니다.

앵커 연결

[Outpost 서비스 링크](#)는 Outpost의 상위 리전에 있는 특정 가용 영역(AZ)의 퍼블릭 또는 프라이빗 앵커(둘 다 아님)에 연결됩니다. Outpost 서버는 서비스 링크 IP 주소에서 앵커 AZ의 앵커 포인트로의 아웃바운드 서비스 링크 VPN 연결을 시작합니다. 이러한 연결은 UDP 및 TCP 포트 443을 사용합니다. AWS 는 리전의 앵커 포인트 가용성을 담당합니다.

Outpost 서비스 링크 IP 주소가 네트워크를 통해 앵커 AZ의 앵커 포인트에 연결될 수 있는지 확인해야 합니다. 서비스 링크 IP 주소는 온프레미스 네트워크의 다른 호스트와 통신할 필요가 없습니다.

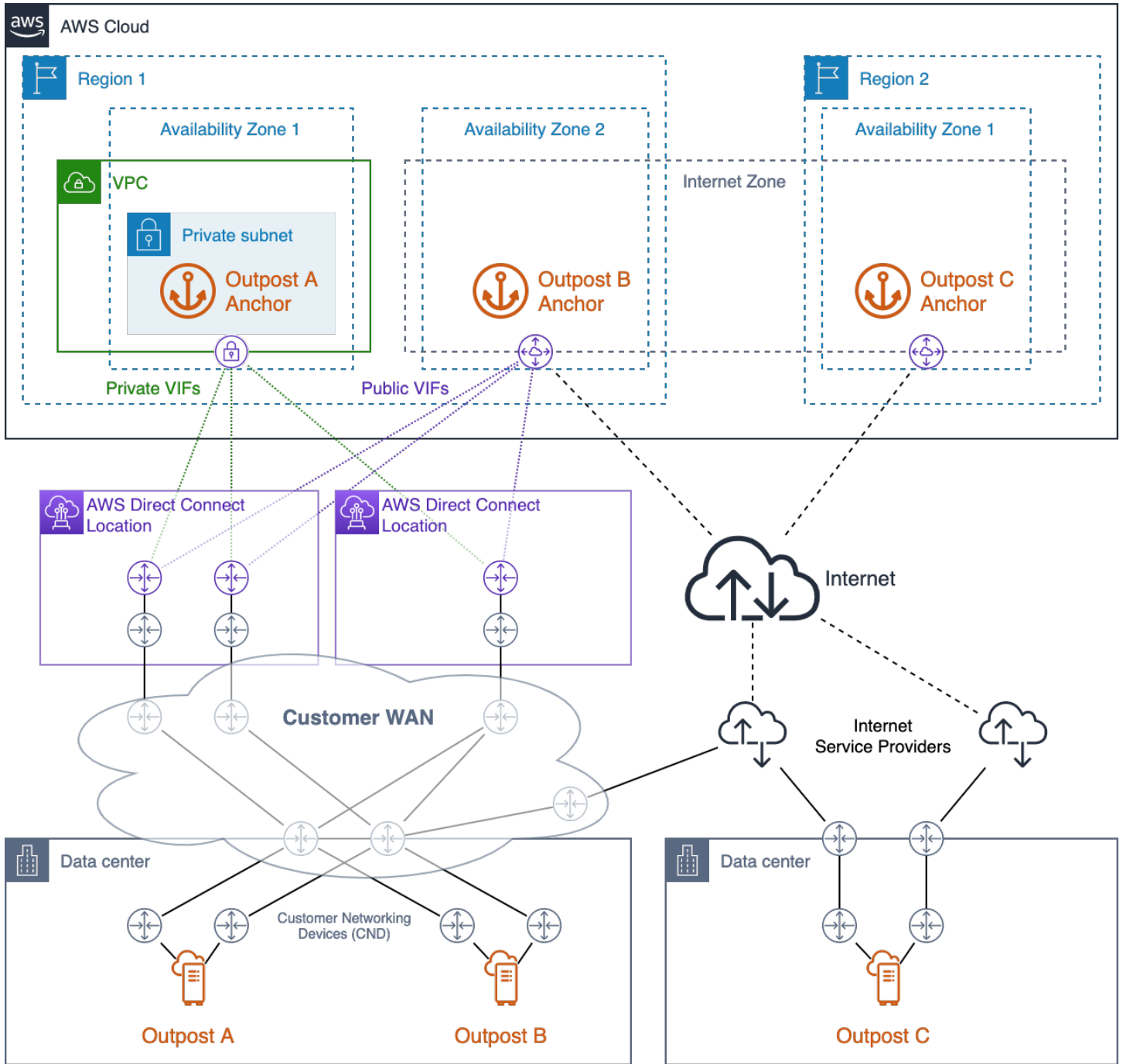
퍼블릭 앵커 포인트는 리전의 [퍼블릭 IP 범위](#)(EC2서비스 CIDR 블록)에 있으며 인터넷 또는 [AWS Direct Connect](#) (DX) 퍼블릭 가상 인터페이스()를 통해 액세스할 수 있습니다 VIFs. 퍼블릭 앵커 포인트를 사용하면 서비스 링크 트래픽이 퍼블릭 인터넷의 앵커 포인트에 성공적으로 도달할 수 있는 사용 가능한 경로를 통해 라우팅될 수 있으므로 보다 유연한 경로 선택이 가능합니다.

프라이빗 앵커 포인트를 사용하면 IP 주소 범위를 앵커 연결에 사용할 수 있습니다. 프라이빗 앵커 포인트는 고객이 할당한 IP 주소를 사용하여 [전용 내의 프라이빗 서브넷VPC](#)에서 생성됩니다. VPC 는 Outpost 리소스를 소유 AWS 계정 하는 에 생성되며 사용자는 VPC 가 사용 가능하고 제대로 구성되어

있는지 확인할 책임이 있습니다(삭제하지 마세요!). 프라이빗 앵커 포인트는 [Direct Connect 프라이빗 VIFs](#)를 사용하여 액세스해야 합니다.

Outpost와 리전 내 앵커 포인트 사이에 중복 네트워크 경로를 제공해야 하며, 두 곳 이상의 위치에 있는 개별 장치에서 연결이 종료됩니다. 연결 또는 네트워킹 장치에 장애가 발생할 경우 트래픽을 대체 경로로 자동으로 다시 라우팅하도록 동적 라우팅을 구성해야 합니다. 한 WAN 경로의 실패가 나머지 경로를 압도하지 않도록 충분한 네트워크 용량을 프로비저닝해야 합니다.

다음 다이어그램은 3개 AZs 사용하는 앵커에 대한 중복 네트워크 경로 AWS Direct Connect 와 퍼블릭 인터넷 연결이 있는 3개의 Outpost를 보여줍니다. Outpost A와 Outpost B는 같은 리전의 서로 다른 가용 영역에 고정되어 있습니다. Outpost A는 리전 1의 AZ 1에 있는 프라이빗 앵커 포인트와 연결됩니다. Outpost B는 리전 1의 AZ 2에 있는 퍼블릭 앵커 포인트와 연결됩니다. Outpost C는 리전 2의 AZ 1에 있는 퍼블릭 앵커와 연결됩니다.



AWS Direct Connect 및 퍼블릭 인터넷 액세스를 통한 고가용성 앵커 연결

Outpost A에는 프라이빗 앵커 포인트에 도달하기 위한 세 개의 중복 네트워크 경로가 있습니다. 단일 Direct Connect 위치의 중복 Direct Connect 회로를 통해 두 개의 경로를 사용할 수 있습니다. 세 번째 경로는 두 번째 Direct Connect 위치의 Direct Connect 회로를 통해 사용할 수 있습니다. 이 설계는 프라이빗 네트워크에서 Outpost A의 서비스 링크 트래픽을 유지하고 Direct Connect 회로 중 하나에 장애가 발생하거나 전체 Direct Connect 위치에 장애가 발생할 수 있는 경로 중복성을 제공합니다.

Outpost B에는 퍼블릭 앵커 포인트에 도달하기 위한 네 개의 중복 네트워크 경로가 있습니다. Outpost A에서 사용하는 Direct Connect 회로 및 위치에 VIFs 프로비저닝된 퍼블릭을 통해 세 가지 경로를 사용할 수 있습니다. 네 번째 경로는 고객과 퍼블릭 인터넷을 통해 사용할 WAN 수 있습니다. Outpost B의 서비스 링크 트래픽은 퍼블릭 인터넷의 앵커 포인트에 성공적으로 도달할 수 있는 사용 가능한 경로를 통해 라우팅될 수 있습니다. Direct Connect 경로를 사용하면 보다 일관된 대기 시간과 더 높은 대역폭 가용성을 제공할 수 있는 반면 퍼블릭 인터넷 경로는 재해 복구(DR) 또는 대역폭 확대 시나리오에 사용될 수 있습니다.

Outpost C에는 퍼블릭 앵커 포인트에 도달하기 위한 두 개의 중복 네트워크 경로가 있습니다. Outpost C는 Outposts A 및 B와 다른 데이터 센터에 배포됩니다. Outpost C의 데이터 센터에는 고객에 연결되는 전용 회로가 없습니다. 대신 데이터 센터에는 서로 다른 두 인터넷 서비스 공급자(ISPs)가 제공하는 중복 인터넷 연결이 있습니다. Outpost C의 서비스 링크 트래픽은 퍼블릭 인터넷의 앵커 포인트에 도달하기 위해 ISP 네트워크 중 하나를 통해 라우팅될 수 있습니다. 이 설계를 사용하면 가능한 모든 퍼블릭 인터넷 연결을 통해 서비스 링크 트래픽을 라우팅할 수 있습니다. 그러나 end-to-end 경로는 대역폭 가용성과 네트워크 지연 시간이 변동하는 퍼블릭 타사 네트워크에 따라 달라집니다.

Outpost와 서비스 링크 앵커 포인트 간의 네트워크 경로는 다음 대역폭 사양을 충족해야 합니다.

- Outpost 랙당 500Mbps - 1Gbps의 가용 대역폭(예: 랙 3개: 1.5~3Gbps의 가용 대역폭)

고가용성 앵커 연결을 위한 권장 사례:

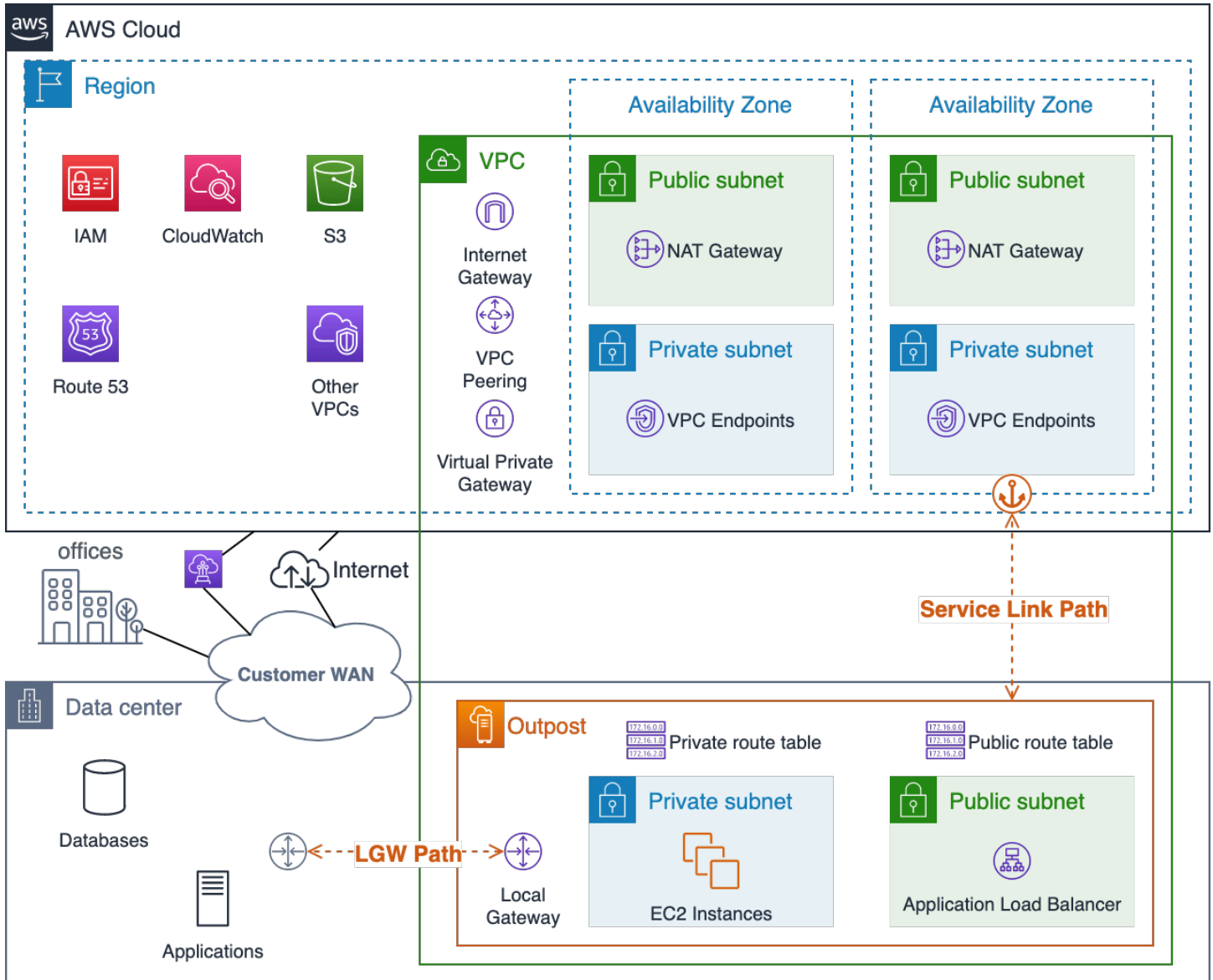
- 각 Outpost와 해당 리전의 앵커 포인트 사이에 중복 네트워크 경로를 제공합니다.
- Direct Connect(DX) 경로를 사용하여 지연 시간과 대역폭 가용성을 제어할 수 있습니다.
- Outpost Service Link CIDR 블록에서 상위 리전의 [EC2 IP 주소 범위](#)로 TCP 및 UDP 포트 443이 열려 있는지(아웃바운드) 확인합니다. 모든 네트워크 경로에서 포트가 열려 있는지 확인합니다.
- 각 경로가 대역폭 가용성 및 지연 시간 요구 사항을 충족하는지 확인합니다.
- 동적 라우팅을 사용하여 네트워크 장애에 대한 트래픽 리디렉션을 자동화합니다.
- 경로가 예상대로 작동하는지 확인하기 위해 계획된 각 네트워크 경로를 통해 서비스 링크 트래픽 라우팅을 테스트합니다.

애플리케이션/워크로드 라우팅

Outpost에서 애플리케이션 워크로드를 처리하는 데는 두 가지 경로가 있습니다.

- 서비스 링크 경로
- 로컬 게이트웨이(LGW) 경로

대상 네트워크에 도달하기 위한 경로 선택을 제어하도록 Outpost 서브넷 라우팅 테이블을 구성합니다. 를 가리키는 경로LGW는 트래픽을 로컬 게이트웨이와 온프레미스 네트워크로 보냅니다. Internet Gateway, Gateway, NAT Virtual Private Gateway 및 와 같이 리전의 서비스 및 리소스를 가리키는 경로TGW는 [서비스 링크](#)를 사용하여 이러한 대상에 도달합니다. 동일한 OutpostVPCs에 여러 가 있는 VPC 피어링 연결이 있는 경우 간의 트래픽은 Outpost에 VPCs 남아 있으며 서비스 링크를 리전으로 다시 사용하지 않습니다. VPC 피어링에 대한 자세한 내용은 Amazon VPC 사용 설명서의 [VPC 피어링을 VPCs 사용하여 연결](#)을 참조하세요.



Outpost 서비스 링크 및 LGW 네트워크 경로 시각화

애플리케이션 라우팅을 계획할 때는 네트워크 장애 시 정상 작동과 제한된 라우팅 및 서비스 가용성을 모두 고려하도록 주의를 기울여야 합니다. Outpost가 리전과 연결 해제된 경우 서비스 링크 경로를 사용할 수 없습니다.

다양한 경로를 프로비저닝하고 OutpostLGW와 중요한 온프레미스 애플리케이션, 시스템 및 사용자 간에 동적 라우팅을 구성해야 합니다. 중복 네트워크 경로를 통해 네트워크에서 장애가 발생한 경우 트래픽을 라우팅하고 부분적인 네트워크 장애 발생 시 온프레미스 리소스가 Outpost에서 실행되는 워크로드와 통신할 수 있습니다.

Outpost VPC 라우팅 구성은 정적입니다. AWS Management Console, CLI APIs, 및 기타 IaC(Infrastructure as Code) 도구를 통해 서브넷 라우팅 테이블을 구성합니다. 그러나 연결 해제 이벤트 중에는 서브넷 라우팅 테이블을 수정할 수 없습니다. 라우팅 테이블을 업데이트하려면 Outpost와 리전 간의 연결을 다시 설정해야 합니다. 연결 해제 이벤트 중에 사용할 계획과 동일한 경로를 일반 운영에도 사용하세요.

Outpost의 리소스는 서비스 링크와 리전의 Internet Gateway(IGW) 또는 Local Gateway(LGW) 경로를 통해 인터넷에 연결할 수 있습니다. LGW 경로 및 온프레미스 네트워크를 통해 인터넷 트래픽을 라우팅하면 기존 온프레미스 인터넷 수신/송신 지점을 사용할 수 있으며, 리전IGW의 에 대한 서비스 링크 경로를 사용하는 경우와 비교하여 지연 시간이 짧고, 높고 MTUs, 데이터 송신 요금이 절감 AWS 될 수 있습니다.

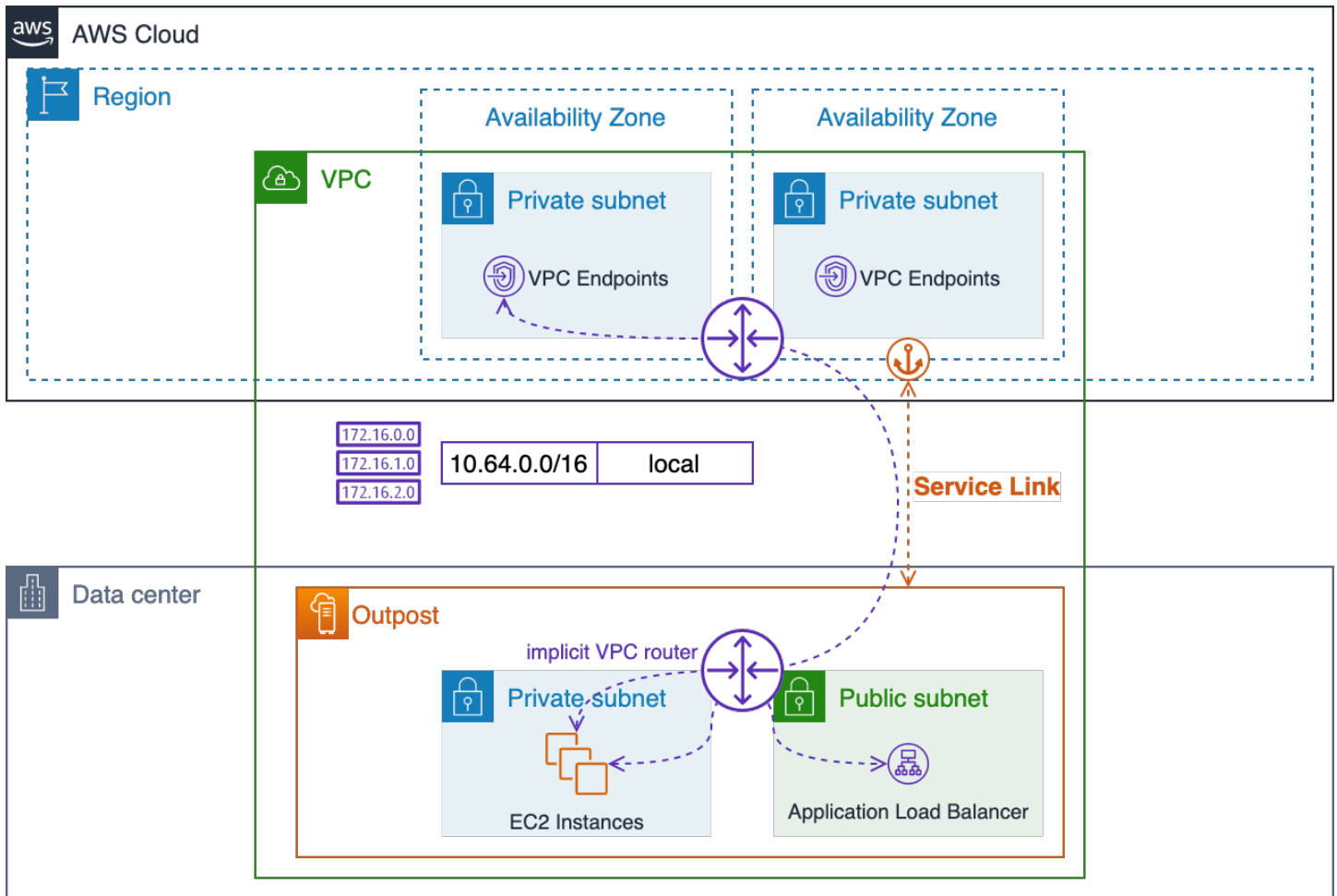
애플리케이션이 온프레미스에서 실행되어야 하고 퍼블릭 인터넷에서 액세스할 수 있어야 하는 경우, Outpost의 리소스에 도달하기 위해 온프레미스 인터넷 연결(들)을 통해 애플리케이션 트래픽을 LGW로 라우팅해야 합니다.

Outpost에서 서브넷을 리전의 퍼블릭 서브넷처럼 구성할 수 있지만 대부분의 사용 사례에서는 바람직하지 않을 수 있습니다. 인바운드 인터넷 트래픽은 를 통해 들어오 AWS 리전 고 서비스 링크를 통해 Outpost에서 실행되는 리소스로 라우팅됩니다.

응답 트래픽은 서비스 링크를 통해 라우팅되고 AWS 리전의 인터넷 연결을 통해 다시 라우팅됩니다. 이 트래픽 패턴은 지연 시간을 가중시킬 수 있으며, 트래픽이 Outpost로 가는 도중에 리전을 떠나 돌아오고, 돌아오는 트래픽이 리전을 통해 다시 들어와 인터넷으로 나가는 경우 데이터 송신 요금이 부과됩니다. 해당 리전에서 애플리케이션을 실행할 수 있는 경우 해당 리전에서 애플리케이션을 실행하는 것이 가장 좋습니다.

VPC 리소스 간 트래픽(동일한 의VPC)은 항상 로컬 VPC CIDR 경로를 따르며 암시적 VPC 라우터에 의해 서브넷 간에 라우팅됩니다.

예를 들어 Outpost에서 실행되는 EC2 인스턴스와 리전의 VPC 엔드포인트 간의 트래픽은 항상 서비스 링크를 통해 라우팅됩니다.



암시적 라우터를 통한 로컬 VPC 라우팅

애플리케이션/워크로드 라우팅 권장 사례:

- 가능한 경우 서비스 링크 경로 대신 로컬 게이트웨이(LGW) 경로를 사용합니다.
- LGW 경로를 통해 인터넷 트래픽을 라우팅합니다.
- Outpost 서브넷 라우팅 테이블을 표준 라우팅 세트 구성합니다. 표준 라우팅 테이블은 일반 운영 및 연결 해제 이벤트 중에 모두 사용됩니다.
- Outpost LGW와 중요한 온프레미스 애플리케이션 리소스 간에 중복 네트워크 경로를 프로비저닝합니다. 동적 라우팅을 사용하여 온프레미스 네트워크 장애에 대한 트래픽 리디렉션을 자동화하세요.

컴퓨팅

의 Amazon EC2 용량은 무한해 AWS 리전 보이지만 Outposts의 용량은 유한합니다. Outposts 배포의 컴퓨팅 용량을 계획하고 관리할 책임은 귀하에게 있습니다.

주제

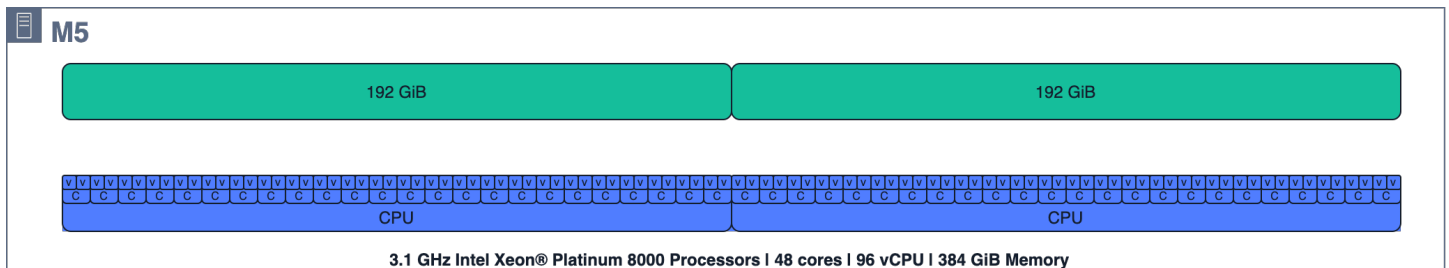
- [용량 계획](#)
- [용량 관리](#)
- [인스턴스 배치](#)

용량 계획

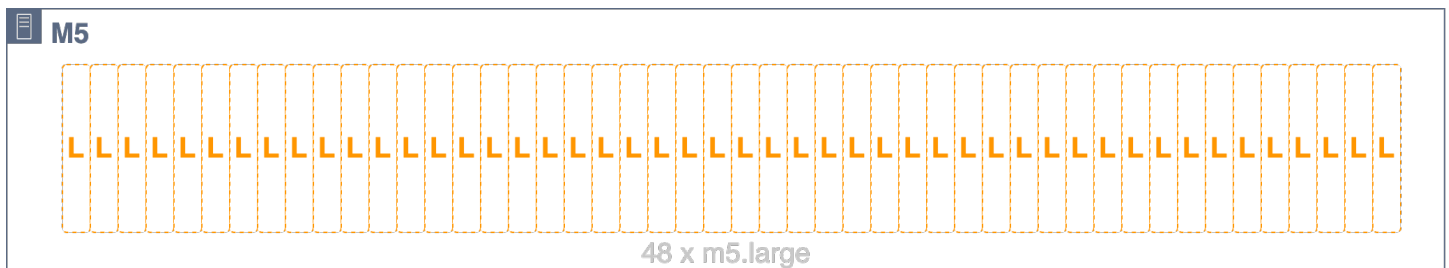
의 Amazon EC2 용량은 무한해 AWS 리전 보이지만 Outposts의 용량은 유한합니다. 이는 주문된 컴퓨팅 용량의 총 볼륨에 의해 제한됩니다. Outposts 배포의 컴퓨팅 용량을 계획하고 관리할 책임은 귀하에게 있습니다. N+M 가용성 모델을 지원하려면 충분한 컴퓨팅 용량을 주문해야 합니다. 여기서 N은 필요한 서버 수이고 M은 서버 장애를 수용하기 위해 프로비저닝된 예비 서버 수입니다. N+1 및 N+2가 가장 일반적인 가용성 수준입니다.

각 서버(C5, M5R5, 등)는 단일 EC2 인스턴스 패밀리를 지원합니다. EC2 컴퓨팅 서버에서 인스턴스를 시작하기 전에 각 서버가 제공할 [EC2 인스턴스 크기](#)를 지정하는 슬롯 레이아웃을 제공해야 합니다. 는 요청된 슬롯 레이아웃으로 각 서버를 AWS 구성합니다.

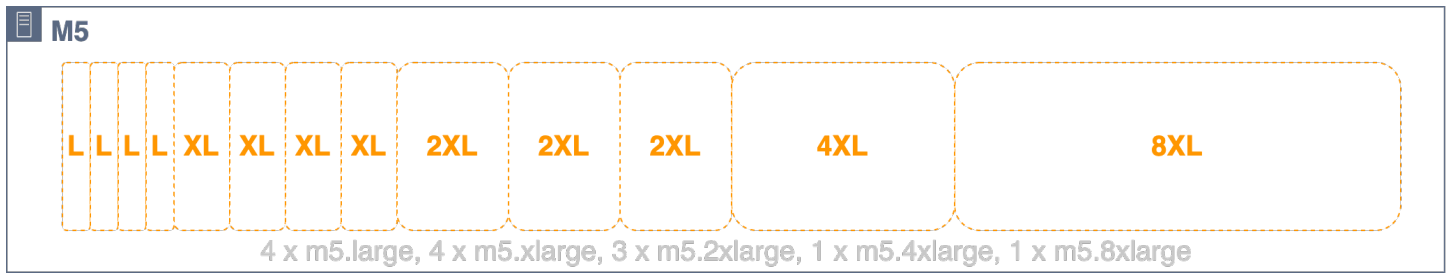
모든 슬롯이 동일한 인스턴스 크기(예: 슬롯 48m5.large개)이거나 인스턴스 유형(예: 4, 4, m5.xlarge3, m5.large, m5.2xlarge, 1m5.4xlarge, 1)이 혼합된 이기종 슬롯이 있는 서버는 균일 슬롯이 될 수 있습니다. 이러한 슬롯 구성의 시각화는 다음 세 가지 그림을 m5.8xlarge참조하세요.



m5.24xlarge 서버 컴퓨팅 리소스



m5.24xlarge 48개의 슬롯에 균등하게 m5.large 슬롯된 서버



m5.24xlarge 4, *4m5.large*, *3m5.xlarge*, *m5.2xlarge* 1 *m5.4xlarge* 및 1 슬롯에 이기종 *m5.8xlarge*으로 슬롯된 서버

전체 서버 용량을 슬롯에 넣을 필요는 없습니다. 할당되지 않은 가용 용량이 있는 서버에 슬롯을 추가할 수 있습니다. 지원 티켓을 열어 슬롯팅 레이아웃을 수정합니다. 실행 중인 인스턴스가 특정 슬롯을 차지하고 있는 상태에서 새 슬롯을 적용할 수 없는 경우 Enterprise Support에서는 재슬롯 요청을 완료하기 위해 특정 인스턴스를 종료하거나 재시작해야 할 수 있습니다.

모든 서버는 Outpost의 EC2 용량 풀에 프로비저닝된 슬롯을 기여하며, 지정된 인스턴스 유형 및 크기의 모든 슬롯은 단일 EC2 용량 풀로 관리됩니다. 예를 들어, *m5.large*, *m5.xlarge*, *m5.2xlarge*, *m5.4xlarge* 및 슬롯이 있는 이전 이기종 *m5.8xlarge* 슬롯 서버는 이러한 슬롯을 각 인스턴스 유형 및 크기에 대해 하나의 풀인 5개의 EC2 용량 풀에 기여합니다.

N+M 서버 가용성을 위한 예비 용량을 계획할 때 서버 슬롯 및 EC2 용량 풀을 고려하는 것이 중요합니다. 는 서버가 실패하거나 성능이 저하될 때를 AWS 감지하고 사이트 방문을 예약하여 실패한 서버를 교체합니다. Outpost에서 각 인스턴스 패밀리(N+1)의 서버 중 하나 이상이 실패하는 것을 허용하도록 EC2 용량 풀을 설계해야 합니다. 이 최소 수준의 서버 가용성을 사용하면 서버에 장애가 발생하거나 서비스를 중단해야 하는 경우 같은 패밀리에 속한 나머지 서버의 예비 슬롯에서 장애가 발생하거나 성능이 저하된 인스턴스를 다시 시작할 수 있습니다.

슬롯이 균등하게 배치된 서버 또는 슬롯 레이아웃이 동일한 이기종 슬롯 서버 그룹이 있는 경우 N+M 가용성을 간단하게 계획할 수 있습니다. 모든 워크로드를 실행하는 데 필요한 서버 수(N)를 계산한 다음 장애 및 유지 관리 이벤트 발생 시 서버 가용성에 대한 요구 사항을 충족하기 위해 서버를 추가(M)하기만 하면 됩니다.

NUMA 경계로 인해 다음 슬롯 구성을 사용할 수 없습니다.

- 3 *m5.8xlarge*
- 1 *m5.16xlarge* 및 1 *m5.8xlarge*

팀에 문의하여 계획된 AWS Outposts 랙 슬롯 구성을 AWS 계정 검증합니다.

다음 그림에서는 4개의 m5.24xlarge 서버가 동일한 슬롯 레이아웃으로 이기종 슬롯됩니다. 4개의 서버는 5개의 EC2 용량 풀을 생성합니다. 각 풀은 이 4개 서버에서 실행되는 인스턴스의 N+1 가용성을 유지하기 위해 최대 사용률(75%)로 실행됩니다. 서버에 장애가 발생하는 경우 나머지 서버에서 장애가 발생한 인스턴스를 다시 시작할 수 있는 충분한 공간이 있습니다.



EC2 서버 슬롯, 실행 중인 인스턴스 및 슬롯 풀 시각화

서버가 동일하게 슬롯화되지 않는 보다 복잡한 슬롯 레이아웃의 경우 각 EC2 용량 풀에 대한 N+M 가용성을 계산해야 합니다. 다음 공식을 사용하여 실패할 수 있는 서버(특정 EC2 용량 풀에 슬롯 기여) 수를 계산하고 나머지 서버가 실행 중인 인스턴스를 전달할 수 있도록 허용할 수 있습니다.

$$M = \left\lceil \frac{poolSlots_{available}}{serverSlots_{max}} \right\rceil$$

위치:

- $poolSlots_{available}$ 는 지정된 EC2 용량 풀에서 사용 가능한 슬롯 수(풀의 총 슬롯 수에서 실행 중인 인스턴스 수를 뺀 값)입니다.
- $serverSlots_{max}$ 는 지정된 EC2 용량 풀에 서버에서 기여한 최대 슬롯 수입니다.
- M이란 장애가 발생해도 나머지 서버가 실행 중인 인스턴스를 계속 유지할 수 있는 서버의 수입니다.

예: Outpost에는 m5.2xlarge 용량 풀에 슬롯을 기여하는 서버가 세 개 있습니다. 첫 번째 서버는 슬롯 4개, 두 번째 서버는 슬롯 3개, 세 번째 서버는 슬롯 2개를 제공합니다. Outpost의 m5.2xlarge 인스턴스

스 풀은 총 용량이 슬롯 9개(4 + 3 + 2)입니다. Outpost에는 실행 중인 m5.2xlarge 인스턴스가 4개 있습니다. 몇 대의 서버에 장애가 발생해도 나머지 서버에서 실행 중인 인스턴스를 계속 사용할 수 있는 서버는 몇 대인가요?

$$poolSlots_{available} = total\ capacity - running\ instances = 9 - 4 = 5$$

$$serverSlots_{max} = \max([4, 3, 2]) = 4$$

$$M = \left\lfloor \frac{poolSlots_{available}}{serverSlots_{max}} \right\rfloor = \left\lfloor \frac{5}{4} \right\rfloor = [1.25] = 1$$

답변: 서버 중 하나를 잃어버리더라도 실행 중인 인스턴스는 나머지 서버에 계속 남아 있을 수 있습니다.

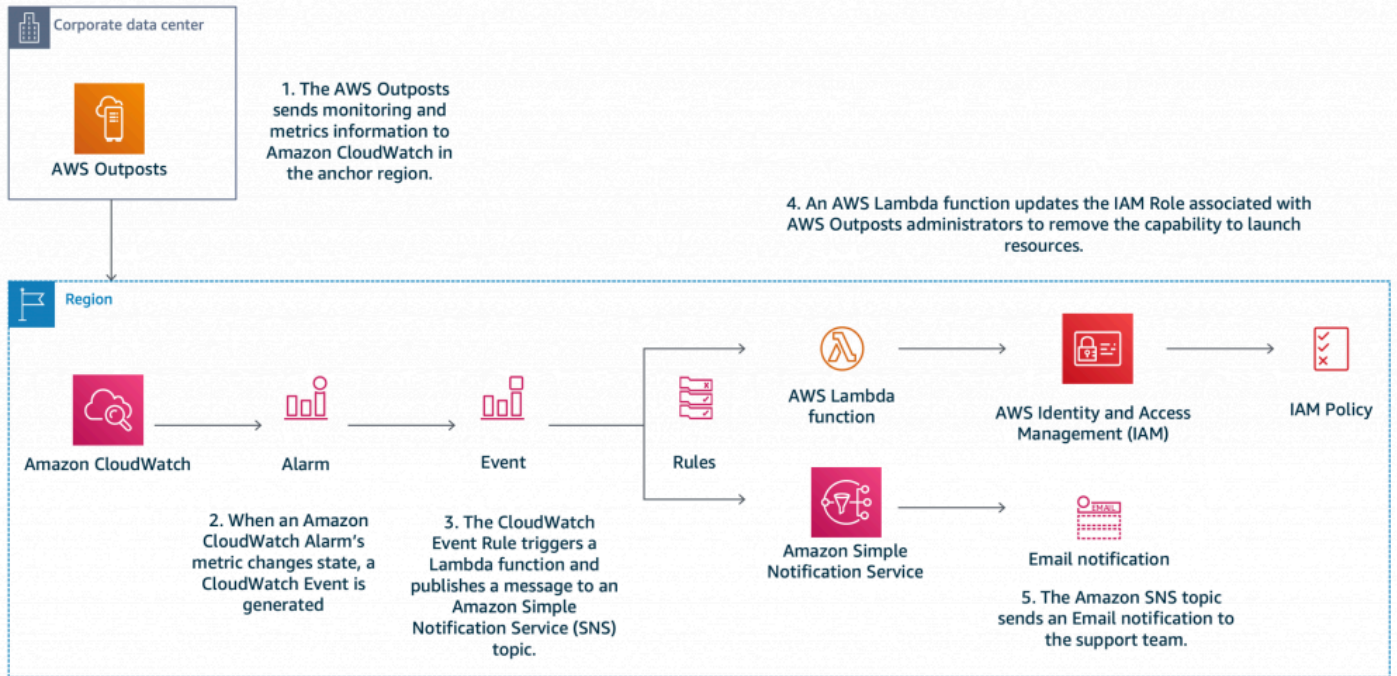
컴퓨팅 용량 계획 권장 사례:

- Outpost의 각 용량 풀에 대해 N+M 중복성을 제공하도록 컴퓨팅 EC2 용량의 크기를 조정합니다.
 - 동종 또는 동일한 이기종 슬롯 서버를 위한 N+M 서버를 배포합니다.
 - 각 EC2 용량 풀의 N+M 가용성을 계산하고 각 풀이 가용성 요구 사항을 충족하는지 확인합니다.

용량 관리

Amazon CloudWatch 지표를 통해 AWS Management Console 및 에서 Outpost EC2 인스턴스 풀 사용을 모니터링할 수 있습니다. Outpost의 슬롯 레이아웃을 검색하거나 변경하려면 Enterprise Support에 문의하세요.

동일한 [인스턴스 자동 복구](#) 및 [EC2 Auto Scaling](#) 메커니즘을 사용하여 서버 장애 및 유지 관리 이벤트의 영향을 받는 인스턴스를 복구하거나 교체합니다. Outpost 용량을 모니터링하고 관리하여 서버 장애를 수용할 수 있을 만큼 충분한 예비 용량을 항상 사용할 수 있도록 해야 합니다. [Amazon CloudWatch 및 블로그 게시물을 사용하여 AWS Outposts 용량 관리 AWS Lambda](#)에서는 Outpost 용량을 결합 AWS CloudWatch 하고 관리 AWS Lambda 하여 인스턴스 가용성을 유지하는 방법을 보여주는 실습 자습서를 제공합니다.



Amazon CloudWatch 및 를 사용한 AWS Outposts 용량 관리 AWS Lambda

컴퓨팅 용량 관리를 위한 권장 사례:

- Auto Scaling 그룹에서 EC2 인스턴스를 구성하거나 인스턴스 자동 복구를 사용하여 실패한 인스턴스를 다시 시작합니다.
- Outpost 배포의 용량 모니터링을 자동화하고 용량 경보에 대한 알림 및 자동 응답(선택 사항)을 구성합니다.

인스턴스 배치

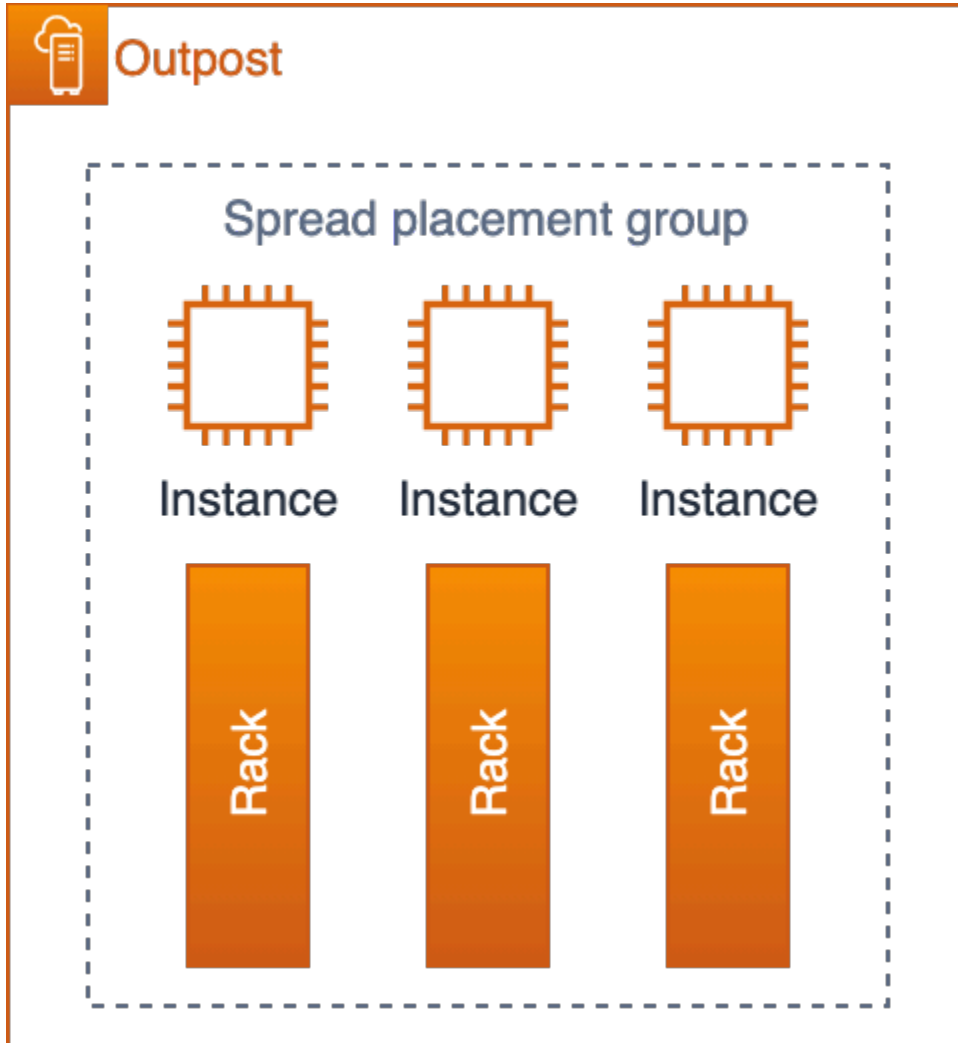
Outpost에는 유한한 수의 컴퓨팅 서버가 있습니다. 애플리케이션이 Outpost에 여러 개의 관련 인스턴스를 배포하는 경우 추가 구성 없이 인스턴스를 동일한 서버 또는 동일한 랙의 서버에 배포할 수 있습니다. 오늘날에는 동일한 인프라에서 관련 인스턴스를 실행할 때 발생하는 위험을 줄이기 위해 인스턴스를 배포하는 데 사용할 수 있는 세 가지 메커니즘이 있습니다.

다중 Outpost 배포 - 리전의 다중 AZ 전략과 마찬가지로 Outpost를 별도의 데이터 센터에 배포하고 특정 Outpost에 애플리케이션 리소스를 배포할 수 있습니다. 이를 통해 원하는 Outpost(논리적 랙 세트)에서 인스턴스를 실행할 수 있습니다. 다중 Outpost 전략을 사용하여 랙 및 데이터 센터 장애 모드로부터 보호할 수 있으며 Outpost가 별도의 AZs 또는 리전에 고정되어 있는 경우 AZ 또는 리전 장애 모드로

부터 보호할 수도 있습니다. 다중 Outpost 아키텍처에 대한 자세한 내용은 [대규모 장애 모드](#)를 참조하세요.

Outposts의 Amazon EC2 배치 그룹(단일 Outpost 다중 랙 인스턴스 배치) - [클러스터](#), [스프레드](#) 및 [파티션](#) 전략을 사용하여 배치에 영향을 미칠 수 있습니다. 분산 및 파티션 배치 전략을 사용하면 다중 랙 Outpost의 랙 전체에 인스턴스를 배포할 수 있습니다.

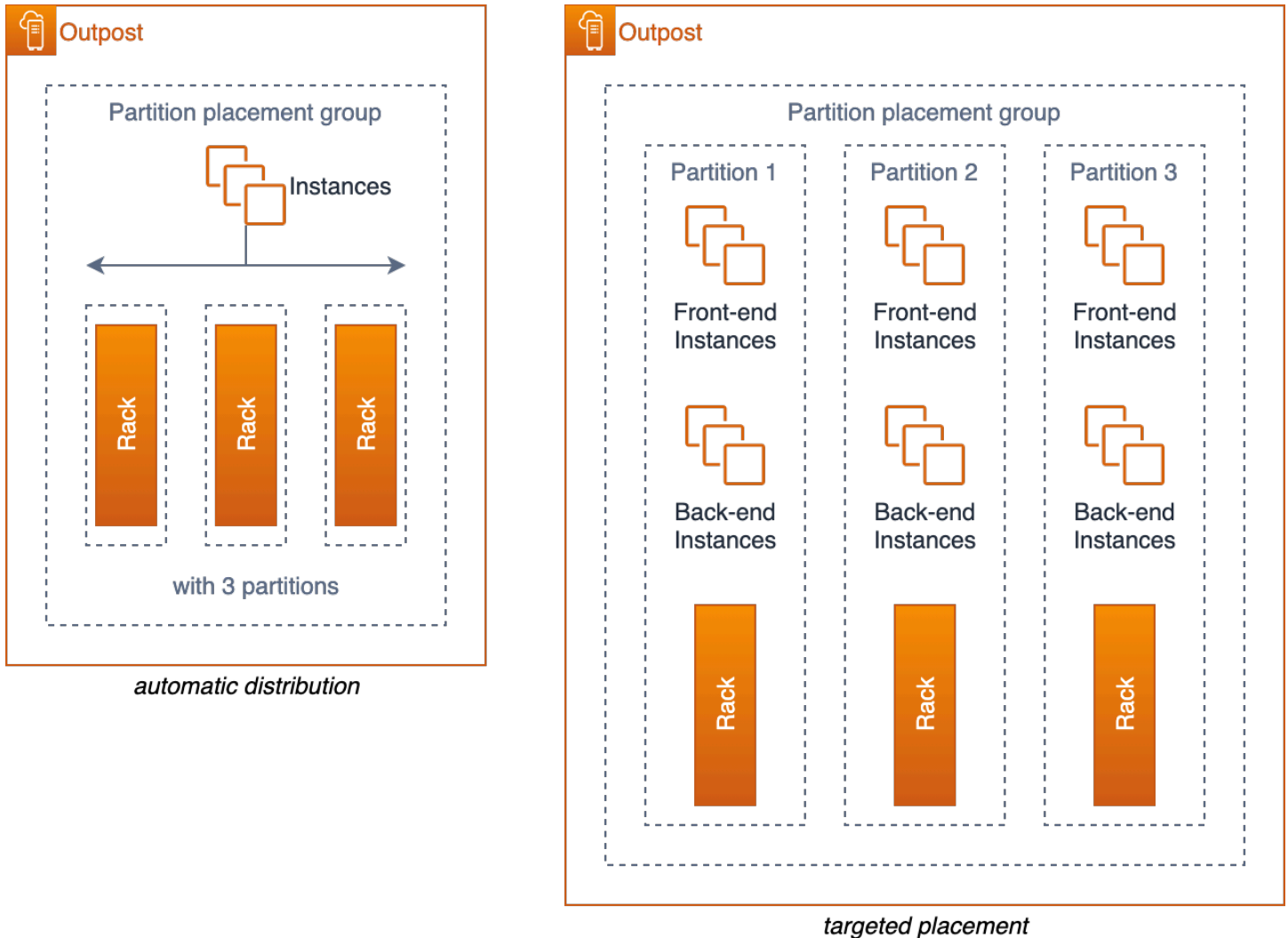
분산 배치 그룹은 단일 인스턴스를 여러 랙에 분산하여 상호 관련된 장애가 발생할 가능성을 줄이는 간단한 방법을 제공합니다. Outpost에 랙이 있는 개수만큼만 인스턴스를 그룹에 배포할 수 있습니다.



EC2 3개의 랙이 있는 Outpost의 분산 배치 그룹

파티션 배치 그룹을 사용하여 여러 랙에 인스턴스를 분산할 수도 있습니다. 자동 배포를 사용하여 그룹 내 파티션에 인스턴스를 분산하거나 선택한 대상 파티션에 인스턴스를 배포할 수 있습니다. 대상 파티션에 인스턴스를 배포하면 선택한 리소스를 동일한 랙에 배포하고 다른 리소스는 랙 전체에 분산할 수

있습니다. 예를 들어 랙이 3개인 논리적 Outpost가 있는 경우 세 개의 파티션으로 구성된 파티션 배치 그룹을 만들면 랙 전체에 리소스를 분배할 수 있습니다.



EC2 랙 3개가 있는 Outpost의 파티션 배치 그룹

크리에이티브 서버 슬롯팅 - 단일 랙 Outpost가 있거나 Outpost에서 사용 중인 서비스가 배치 그룹을 지원하지 않는 경우, 크리에이티브 슬롯팅을 사용하여 인스턴스가 동일한 물리적 서버에 배포되지 않도록 할 수 있습니다. 관련 인스턴스의 EC2 인스턴스 크기가 동일한 경우 서버를 슬롯화하여 각 서버에 구성된 해당 크기의 슬롯 수를 제한하여 서버 전체에 슬롯을 분산할 수 있습니다. 서버 슬롯팅은 단일 서버에서 실행할 수 있는 해당 크기의 인스턴스 수를 제한합니다.

그림 13에 표시된 슬롯팅 레이아웃을 예로 들어 보겠습니다. 이 슬롯 레이아웃으로 구성된 Outpost에 세 개의 m5.4xlarge 인스턴스를 배포해야 하는 경우 EC2는 각 인스턴스를 별도의 서버에 배치하며 슬롯 구성이 변경되어 서버에 추가 m5.4xlarge 슬롯이 열리지 않는 한 이러한 인스턴스가 동일한 서버에서 실행될 가능성은 없습니다.

컴퓨팅 인스턴스 배치 권장 사례:

- Outposts의 Amazon EC2 배치 그룹을 사용하여 단일 Outpost 내의 랙에서 인스턴스 배치를 제어할 수 있습니다.
- 단일 중간 또는 대형 Outpost 랙이 있는 Outpost를 주문하는 대신 용량을 두 개의 소형 또는 중간 랙으로 분할하여 EC2 배치 그룹 기능을 활용하여 랙 간에 인스턴스를 배포할 수 있도록 하는 것이 좋습니다.

스토리지

AWS Outposts 랙 서비스는 세 가지 스토리지 유형을 제공합니다.

- 지원되는 인스턴스 유형의 EC2 인스턴스 [스토리지](#)
- 영구 [블록 스토리지](#)를 위한 [Amazon Elastic Block Store\(EBS\) gp2 볼륨](#)
- 로컬 객체 스토리지를 위한 [Outpost의 Amazon 심플 스토리지 서비스\(S3 on Outposts\)](#)

인스턴스 스토리지는 지원되는 서버(C5d, M5d, R5d, G4dn, I3en)에서 제공됩니다. 리전과 마찬가지로 인스턴스 스토어의 데이터는 실행 중인 [인스턴스의 수명](#) 동안만 지속됩니다.

Outposts EBS 볼륨 및 S3 on Outposts 객체 스토리지는 AWS Outposts 랙 관리형 서비스의 일부로 제공됩니다. Outpost 스토리지 풀의 용량 관리는 고객의 책임입니다. 고객은 Outpost. AWS configures Outpost를 주문할 때 EBS 및 S3 스토리지에 대한 스토리지 요구 사항을 지정합니다. 요청된 스토리지 용량을 제공하는 데 필요한 스토리지 서버 수로 Outpost를 구성합니다. AWS는 EBS 및 S3 on Outposts 스토리지 서비스의 가용성을 책임집니다. Outpost에 고가용성 스토리지 서비스를 제공할 수 있을 만큼 충분한 스토리지 서버가 프로비저닝되어 있습니다. 스토리지 서버 한 대에 장애가 발생해도 서비스가 중단되거나 데이터가 손실되어서는 안 됩니다.

AWS Management Console 및 [CloudWatch 지표](#)를 사용하여 Outpost EBS 및 [S3 on Outposts 용량 사용률](#)을 모니터링할 수 있습니다.

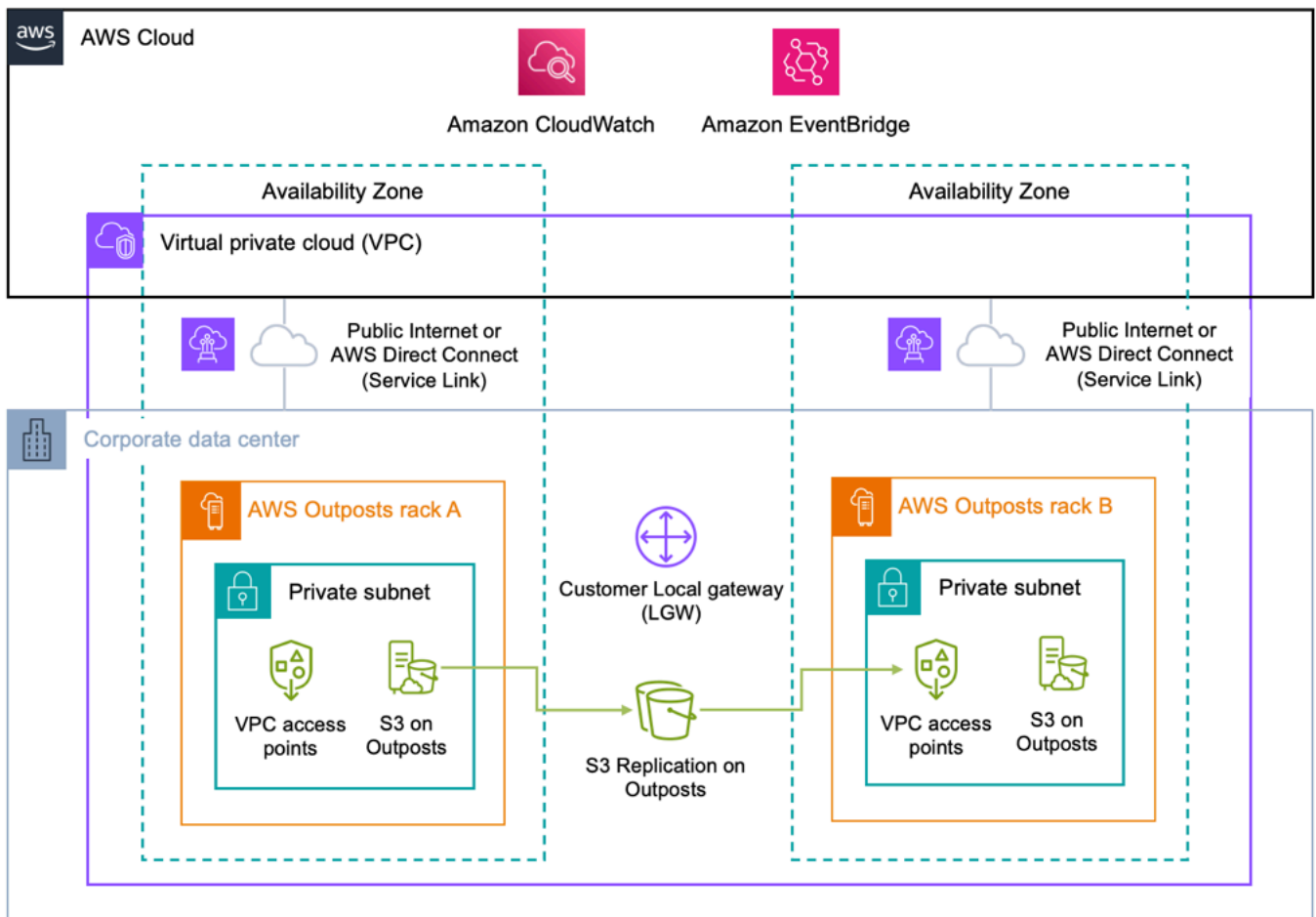
데이터 보호

For EBS Volumes: AWS Outposts rack은 EBS 볼륨 스냅샷을 지원하여 블록 스토리지 데이터를 보호하는 간단하고 안전한 데이터 보호 메커니즘을 제공합니다. 스냅샷은 point-in-time EBS 볼륨의 증분 백업입니다. 기본적으로 Outpost의 [Amazon EBS 볼륨 스냅샷](#)은 리전의 Amazon S3에 저장됩니다. Outposts가 S3 on Outposts 용량으로 구성된 경우 [EBS Local Snapshots on Outposts](#)를 사용하여 S3 on Outposts 스토리지를 사용하여 Outpost에 로컬로 스냅샷을 저장할 수 있습니다.

S3 on Outposts 버킷의 경우(데이터 레지던시 사용 사례):

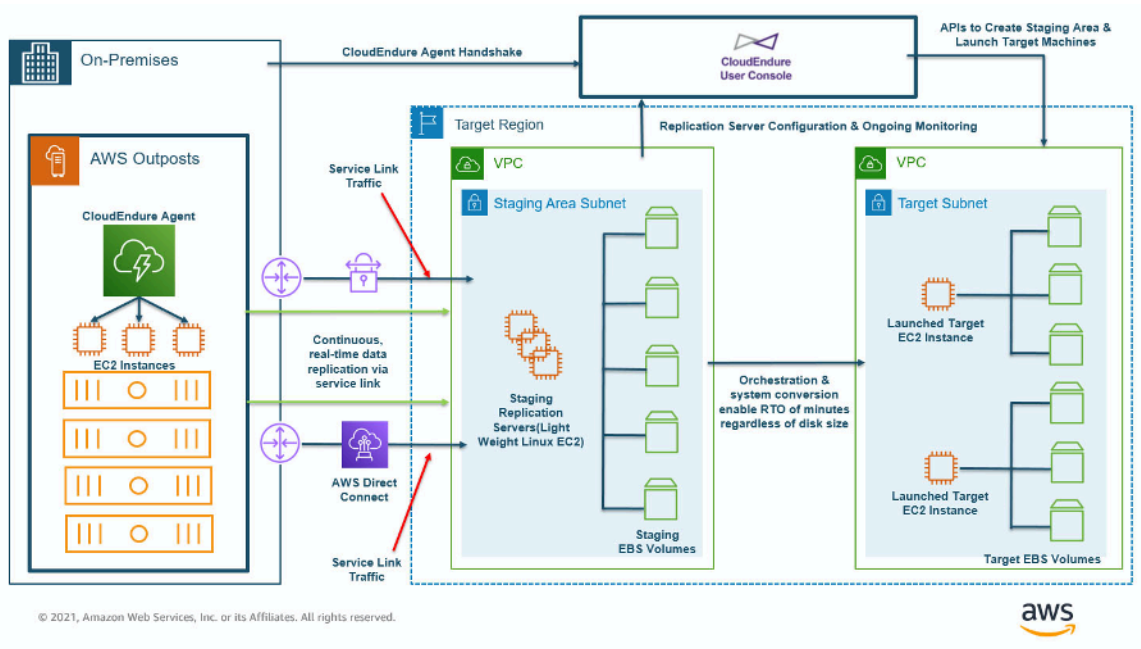
- [Outposts의 S3 버전 관리](#)를 사용하여 모든 변경 사항과 객체 기록을 저장할 수 있습니다. S3 버전 관리를 활성화하면 동일 버킷 내에 여러 개의 개별 객체 복제본을 저장합니다. S3 버전 관리를 사용하여 Outposts 버킷에 저장된 모든 버전의 객체를 모두 보존, 검색 및 복원할 수 있습니다. S3 버전 관리는 의도치 않은 사용자 작업 및 애플리케이션 장애로부터 복구하는 데 도움이 됩니다.
- [Outpost의 S3 복제](#)를 사용하여 S3 객체를 다른 Outpost나 동일한 Outpost의 다른 버킷에 자동으로 복제하는 복제 규칙을 생성 및 구성할 수 있습니다. 복제 중에 S3 on Outposts 객체는 고객의 로컬 게이트웨이(LGW)를 통해 전송되며 객체는 로 다시 이동하지 않습니다 AWS 리전. Outpost의 S3 복제 기능은 데이터 중복성 및 규정 준수 요구 사항을 해결하기 위해 특정 데이터 경계 내에서 데이터를 자동으로 복제하는 쉽고 유연한 방법을 제공합니다.

또한 Outpost의 S3 복제 기능은 버킷 간 객체 복제 상태를 모니터링하기 위한 상세한 지표 및 알림을 제공합니다. Amazon CloudWatch를 사용하여 소스 버킷과 대상 Outposts 버킷 간의 보류 중인 바이트, 보류 중인 작업 및 복제 지연 시간을 추적하여 복제 진행 상황을 모니터링할 수 있습니다 CloudWatch. 복제 실패 이벤트를 수신하도록 Amazon EventBridge 규칙을 설정하여 구성 문제를 빠르게 진단하고 수정할 수도 있습니다.



S3 on Outposts 버킷(비 데이터 레지던시 사용 사례)에서 AWS 리전: [AWS DataSync](#)를 사용하여 Outpost와 리전 간의 S3 on Outposts 데이터 전송을 자동화할 수 있습니다. 를 DataSync 사용하면 전송할 대상, 전송 시기 및 사용할 대역폭을 선택할 수 있습니다. 온프레미스 S3 on Outposts 버킷을 AWS 리전 의 S3 버킷에 백업하면 99.999999999%(11 9's)의 데이터 내구성과 추가 스토리지 계층(표준, 간헐적 액세스 및 Glacier)을 활용하여 리전의 S3 서비스에서 제공되는 비용 최적화를 위해 사용할 수 있습니다.

인스턴스 복제: [CloudEndure](#)를 사용하여 온프레미스 시스템의 개별 인스턴스를 Outpost로, Outpost에서 리전으로, 리전에서 Outpost로 또는 한 Outpost에서 다른 Outpost로 복제할 수 있습니다. 블로그 게시물을 [AWS Outposts 사용한 의 Architecting for DR CloudEndure](#)은 이러한 각 시나리오와 를 사용하여 솔루션을 설계하는 방법을 설명합니다 CloudEndure.



Outpost에서 해당 리전으로의 재해 복구(DR)

AWS Outposts 랙을 CloudEndure 대상으로 사용(복제 대상)하려면 S3 on Outposts 스토리지가 필요합니다.

데이터 보호를 위한 권장 사례:

- EBS 스냅샷을 사용하여 리전의 Amazon S3 또는 Outposts의 S3에 대한 블록 스토리지 볼륨의 백업을 생성합니다 point-in-time.
- S3 on Outposts 객체 버전 관리를 사용하여 객체의 여러 버전과 기록을 유지할 수 있습니다.
- Outposts의 S3 복제를 사용하여 객체 데이터를 다른 Outpost에 자동으로 복제할 수 있습니다.
- 비 데이터 상주 사용 사례의 경우 AWS DataSync 를 사용하여 S3 on Outpost에 저장된 객체를 리전의 Amazon S3에 백업합니다.
- CloudEndure 를 사용하여 온프레미스 시스템, 논리적 Outposts 및 리전 간에 인스턴스를 복제합니다.

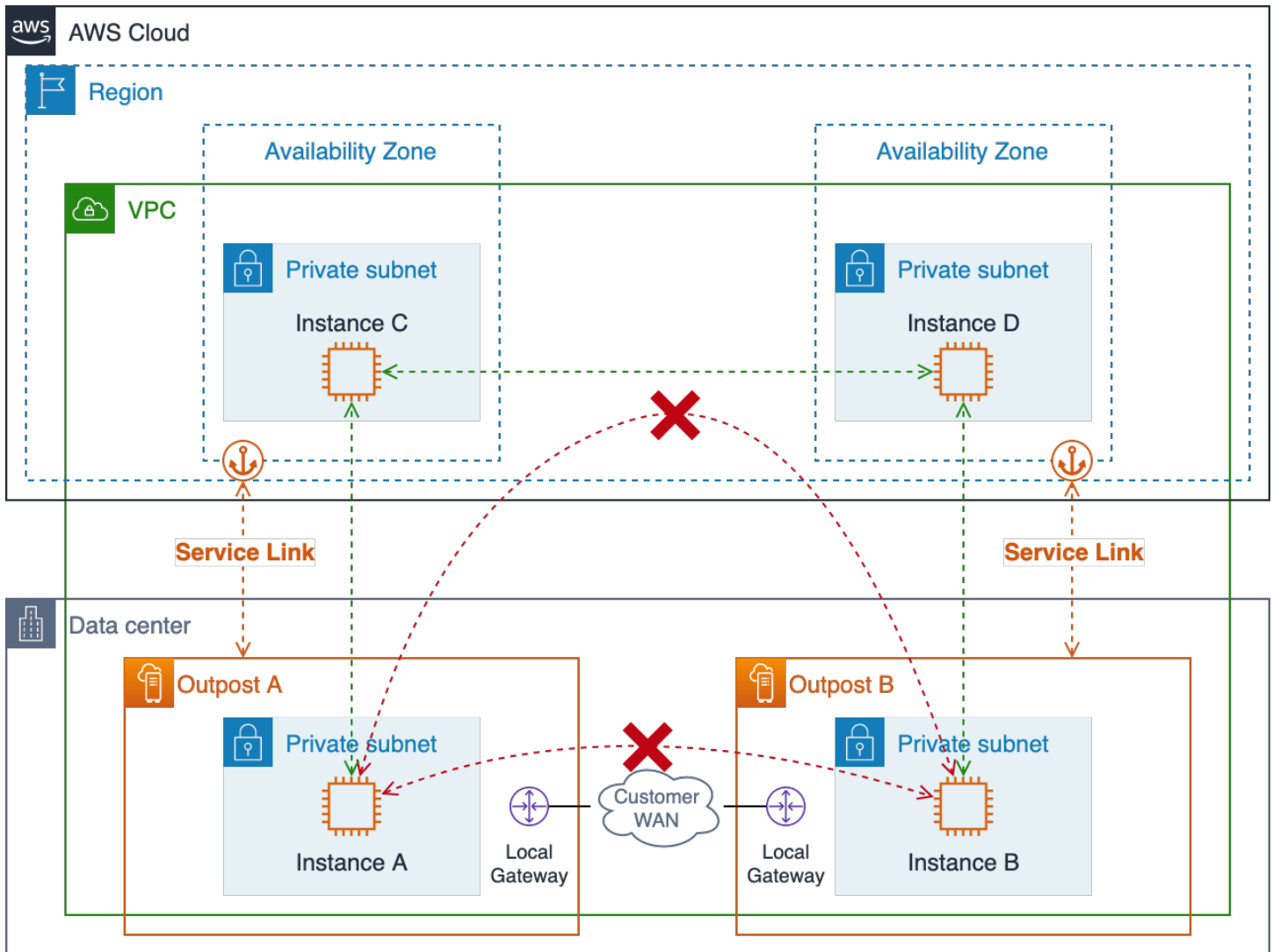
더 큰 장애 모드

랙, 데이터 센터, 가용 영역(AZ) 또는 리전 장애와 같은 더 큰 장애 모드를 완화하기 위해 HA 아키텍처를 설계하려면 독립적인 전원 및 WAN 연결이 가능한 별도의 데이터 센터에 충분한 인프라 용량을 갖춘 여러 Outpost를 배포해야 합니다. Outposts를 여러 리전 내 AWS 리전 또는 여러 리전에 걸쳐 다양

한 가용 영역(AZs)에 고정합니다. 또한 동기 또는 비동기 데이터 복제 및 워크로드 트래픽 리디렉션을 지원하기 위해 위치 간에 복원 가능하고 충분한 site-to-site 연결을 프로비저닝해야 합니다. 애플리케이션 아키텍처에 따라 전역적으로 사용 가능한 [Amazon Route 53](#) DNS 및 리전별로 사용 가능한 [Elastic Load Balancing](#) 서비스를 사용하여 트래픽을 원하는 위치로 보내고 대규모 장애가 발생할 경우 생존 위치로의 트래픽 리디렉션을 자동화할 수 있습니다.

여러 Outpost에서 애플리케이션 워크로드를 설계하고 배포할 때 주의해야 하는 네트워킹 제한 사항이 있습니다. 두 개의 개별 Outpost에 있는 리소스는 리전을 통해 트래픽을 전송함으로써 서로 통신할 수 없습니다. 동일한 내에 배포된 두 개의 별도 Outposts에 있는 리소스는 고객 네트워크를 통해 서로 통신할 수 없습니다. 서로 다른 에 배포된 두 개의 개별 Outpost에 있는 리소스는 고객 네트워크를 통해 서로 통신할 VPCs 수 있습니다.

다음 두 그림은 차단된 네트워크 경로와 성공한 네트워크 경로를 보여줍니다.

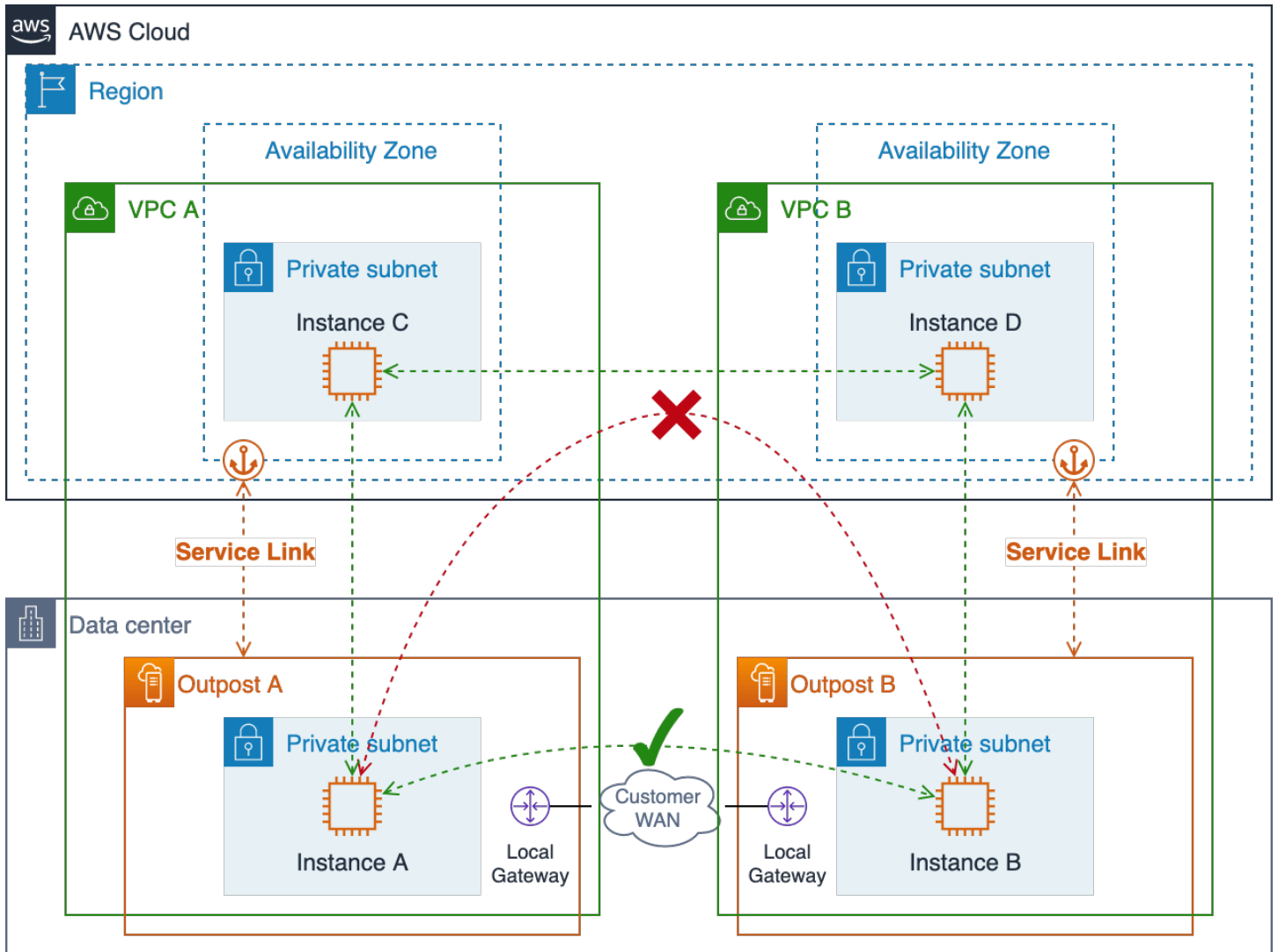


단일 VPC 다중 아웃포스트 네트워크 경로

Outpost-to-Outpost 리전을 통과하는 트래픽은 안티 패턴이므로 차단됩니다. 이러한 트래픽에는 양방향으로 송신 요금이 발생하며 고객 간에 트래픽을 단순히 라우팅하는 것보다 지연 시간이 훨씬 더 길 수 있습니다WAN.

동일한 의 여러 Outpost에 있는 리소스는 서로 통신VPC할 수 없습니다. 동일한 의 Outpost 간 트래픽 VPC은 항상 차단될 리전을 통한 로컬 VPC CIDR 경로를 따릅니다.

로컬 온프레미스 및 WAN 네트워크 간에 트래픽을 라우팅VPCs Outpost-to-Outpost할 수 있도록 별도의 의 를 사용하여 여러 Outpost에 리소스를 배포해야 합니다.



다중VPC 아웃포스트 네트워크 경로

더 큰 장애 모드로부터 보호하기 위한 권장 방법:

- 여러 AZs 및 리전에 고정된 여러 Outposts를 배포합니다.

- 다중 Outpost 배포의 각 OutpostVPCs에 대해 별도의 를 사용합니다.

결론

AWS Outposts 랙을 사용하면 Amazon , Amazon , EBS Amazon S3 on Outposts, Amazon EC2, ECS Amazon , Amazon 와 같은 친숙한 AWS 도구와 서비스를 사용하여 가용성이 높은 온프레미스 애플리케이션을 구축, 관리 EKS 및 확장할 수 있습니다 RDS. 워크로드는 로컬에서 실행되고, 클라이언트에 서비스를 제공하고, 온프레미스 네트워크의 애플리케이션 및 시스템에 액세스하고, AWS 리전의 전체 서비스 세트에 액세스할 수 있습니다. Outposts는 온프레미스 시스템에 짧은 지연 시간으로 액세스해야 하는 워크로드, 로컬 데이터 처리, 데이터 레지던시 및 로컬 시스템 상호 의존성이 있는 애플리케이션의 마이그레이션에 적합합니다.

Outpost 배포에 적절한 전력, 공간, 냉각 및 에 대한 복원력 있는 연결을 제공 AWS 리전하면 가용성이 높은 단일 데이터 센터 서비스를 구축할 수 있습니다. 또한 가용성과 복원력을 높이기 위해 여러 Outpost를 배포하고 논리적 및 지리적 경계를 넘어 애플리케이션을 배포할 수 있습니다.

Outposts 랙은 온프레미스 컴퓨팅, 스토리지 및 애플리케이션 네트워킹 풀 구축의 차별화되지 않은 과도한 부담을 제거하고 AWS 글로벌 인프라의 범위를 데이터 센터 및 코로케이션 시설로 확장할 수 있습니다. 이제 애플리케이션을 현대화하고, 애플리케이션 배포를 간소화하고, IT 서비스가 비즈니스에 미치는 영향을 높이는 데 시간과 에너지를 집중할 수 있습니다.

기여자

다음은 이 문서의 기여자입니다.

- Mallory Gershenfeld, S3 on Outposts, Amazon Web Services
- Chris Lunsford, AWS Outposts Amazon Web Services 선임 전문가 솔루션 아키텍트
- Rohan Mathews, 수석 아키텍트 AWS Outposts, Amazon Web Services

문서 이력

이 백서의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독하세요.

변경 사항	설명	날짜
마이너 업데이트	용량 계획에 슬롯 지침을 추가했습니다.	2024년 2월 9일
마이너 업데이트	최초 게시 이후 출시된 기능을 반영하도록 업데이트되었습니다.	2023년 7월 19일
마이너 업데이트	고가용성 네트워크 연결에 대한 권장 방법을 업데이트했습니다.	2023년 6월 29일
최초 게시	백서가 처음 게시되었습니다.	2021년 8월 12일

Note

RSS 업데이트를 구독하려면 사용 중인 브라우저에 대해 RSS 플러그인이 활성화되어 있어야 합니다.

고지 사항

고객은 본 문서의 정보를 독립적으로 평가할 책임이 있습니다. 이 문서: (a) 정보 제공용이며, (b) 예고 없이 변경될 수 있는 현재 AWS 제품 제안 및 관행을 나타내며, (c) AWS 및 그 계열사, 공급업체 또는 라이선스 제공자로부터 어떠한 약속 또는 보장도 생성하지 않습니다. AWS 제품 또는 서비스는 명시적이든 묵시적이든 어떤 종류의 보증, 표현 또는 조건 없이 “있는 그대로” 제공됩니다. 고객에 AWS 대한 책임과 책임은 AWS 계약에 의해 관리되며, 이 문서는 AWS 와 고객 간의 계약의 일부이거나 수정하지 않습니다.

© 2023 Amazon Web Services, Inc. 또는 계열사. All rights reserved.

AWS 용어집

최신 AWS 용어는 AWS 용어집 참조 의 [AWS 용어집](#)을 참조하세요.

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.