AWS Well-Architected 프레임워크

의 워크로드 재해 복구 AWS: 클라우드에서의 복구

의 워크로드 재해 복구 AWS: 클라우드에서의 복구: AWS Well-Architected 프레임워크

Table of Contents

의 워크로드 재해 복구 AWS: 클라우드에서의 복구

게시일: 2021년 2월 12일(문서 기록)

재해 복구는 재해를 대비하고 복구하는 프로세스입니다. 워크로드 또는 시스템이 기본 배포 위치에서 비즈니스 목표를 이행하지 못하게 하는 이벤트는 재해로 간주됩니다. 이 백서에서는 배포된 모든 워크로드에 대한 재해 복구를 계획하고 테스트하는 모범 사례를 간략하게 설명하고 AWS, 위험을 완화하고 해당 워크로드에 대한 Recovery Time Objective(RTO) 및 Recovery Point Objective(RPO)를 충족하기 위한 다양한 접근 방식을 제공합니다.

이 백서에서는 워크로드에 대한 재해 복구를 구현하는 방법을 다룹니다 AWS. 를 온프레미스 워크로드의 재해 복구 사이트로 사용하는 방법에 대한 자세한 내용은의 온프레미스 애플리케이션의 AWS 재해 복구를 참조하세요. AWS

소개

워크로드는 의도한 기능을 올바르고 일관되게 수행해야 합니다. 이를 위해서는 복원력을 설계해야 합니다. 복원력은 인프라, 서비스 또는 애플리케이션 중단으로부터 복구하고, 수요를 충족하기 위해 컴퓨팅 리소스를 동적으로 획득하고, 잘못된 구성 또는 일시적인 네트워크 문제와 같은 중단을 완화하는 워크로드의 기능입니다.

재해 복구(DR)는 복원력 전략의 중요한 부분이며 재해가 발생할 때 워크로드가 어떻게 반응하는지에 관한 것입니다(재해는 비즈니스에 심각한 부정적인 영향을 미치는 이벤트임). 이 응답은 Recovery Point Objective(RPO)라고 하는 데이터 손실을 방지하고 Recovery Time Objective(RTO)라고 하는 워크로드를 사용할 수 없는 가동 중지 시간을 줄이기 위한 워크로드의 전략을 지정하는 조직의 비즈니스 목표를 기반으로 해야 합니다. 따라서 지정된 일회성 재해 이벤트에 대한 복구 목표(RPO 및 RTO)를 충족하려면 클라우드의 워크로드 설계에 복원력을 구현해야 합니다. 이 접근 방식은 조직이 비즈니스 연속성 계획(BCP)의 일환으로 비즈니스 연속성을 유지하는 데 도움이 됩니다.

이 백서에서는 비즈니스의 재해 복구 목표를 충족하는의 아키텍처를 계획, 설계 및 구현 AWS 하는 방법을 중점적으로 다룹니다. 여기에 공유된 정보는 최고 기술 책임자(CTOs), 아키텍트, 개발자, 운영 팀 원, 위험 평가 및 완화를 담당하는 담당자와 같은 기술 역할의 사용자를 위한 것입니다.

재해 복구 및 가용성

재해 복구는 복원력 전략의 또 다른 중요한 구성 요소인 가용성과 비교할 수 있습니다. 재해 복구는 일회성 이벤트의 목표를 측정하는 반면, 가용성 목표는 일정 기간 동안의 평균 값을 측정합니다.

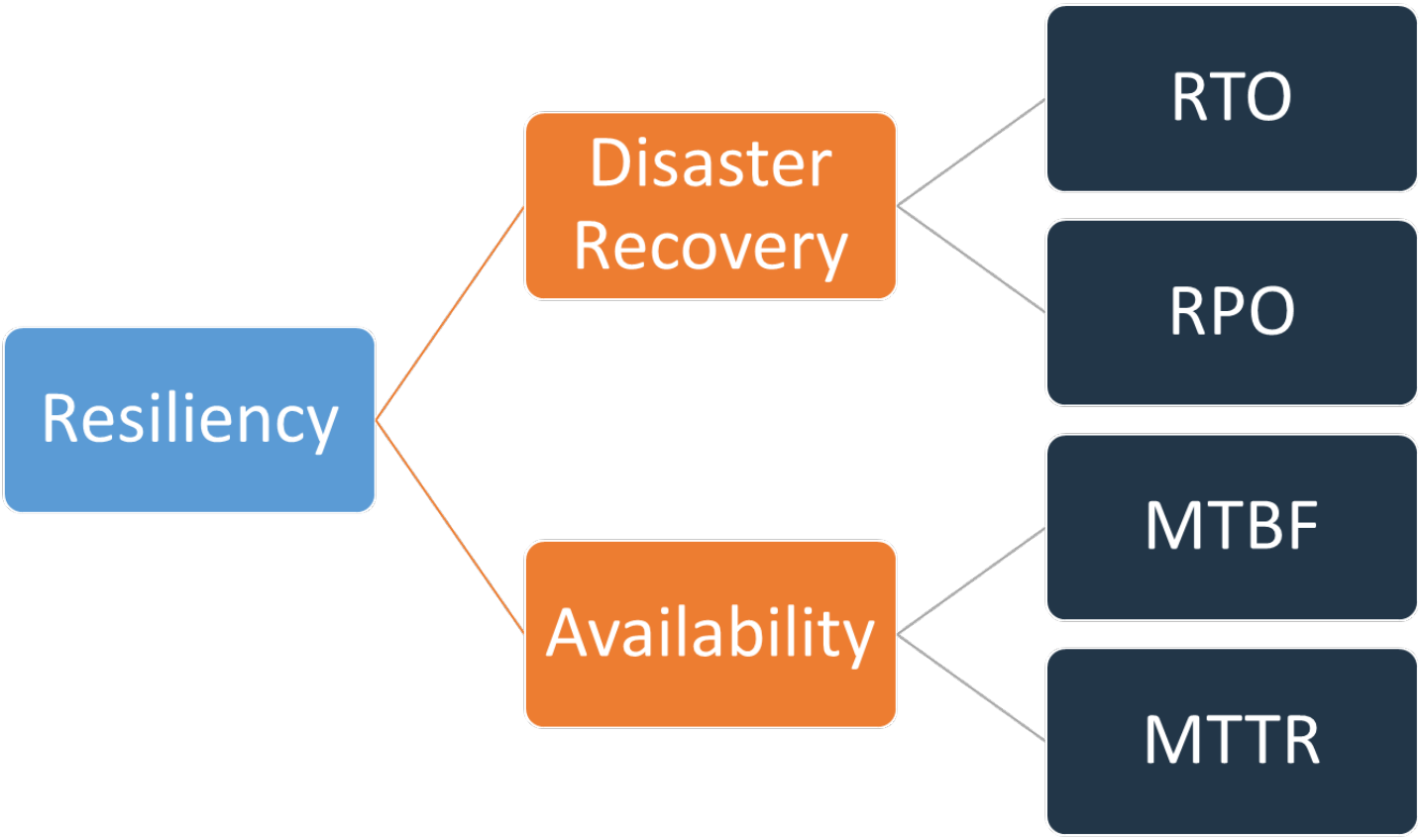

그림 1 - 복원력 목표

가용성은 평균 장애 간격(MTBF)과 평균 복구 시간(MTTR)을 사용하여 계산됩니다.

$$Availability = \frac{Available\ for\ Use\ Time}{Total\ Time} = \frac{MTBF}{MTBF + MTTR}$$

이 접근 방식을 '9'라고 하는 경우가 많으며, 99.9%의 가용성 목표를 '9'라고 합니다.

워크로드의 경우 시간 기반 접근 방식을 사용하는 대신 성공 및 실패한 요청을 계산하는 것이 더 쉬울 수 있습니다. 이 경우 다음 계산을 사용할 수 있습니다.

$$Availability = \frac{Successful\ Responses}{Valid\ Requests}$$

재해 복구는 재해 이벤트에 중점을 두는 반면, 가용성은 구성 요소 장애, 네트워크 문제, 소프트웨어 버그 및 로드 급증과 같은 소규모의 보다 일반적인 중단에 중점을 둡니다. 재해 복구의 목표는 비즈니스 연속성인 반면 가용성은 워크로드가 의도한 비즈니스 기능을 수행하는 데 사용할 수 있는 시간을 극대화하는 것과 관련이 있습니다. 둘 다 복원력 전략의 일부여야 합니다.

귀사는 Well-Architected입니까?

AWS Well-Architected Framework는 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 Console에서 무료로 사용할 수 있는 AWS Well-Architected Tool을 사용하면 각 요소에 대한 일련의 질문에 답변하여 이러한 모범 사례와 비교하여 워크로드를 검토할 수 있습니다.

이 백서에서 다루는 개념은 신뢰성 원칙 백서에 포함된 모범 사례, 특히 REL 13, “재해 복구(DR)를 어떻게 계획하나요?”에 대해 자세히 설명합니다. 이 백서의 사례를 구현한 후에는 AWS Well-Architected Tool을 사용하여 워크로드를 검토(또는 재검토)해야 합니다.

복원력을 위한 공동 책임 모델

복원력은 AWS 와 고객 간의 공동 책임입니다. 복원력의 일환으로 재해 복구 및 가용성이이 공유 모델에서 어떻게 작동하는지 이해하는 것이 중요합니다.

AWS 책임 “클라우드 복원력”

AWS는 AWS 클라우드에서 제공되는 모든 서비스를 실행하는 인프라의 복원력을 책임집니다. 이 인프라는 AWS 클라우드 서비스를 실행하는 하드웨어, 소프트웨어, 네트워킹 및 시설로 구성됩니다. AWS는 이러한 AWS 클라우드 서비스를 사용할 수 있도록 상업적으로 합리적인 노력을 기울여 서비스 가용성이 AWS 서비스 수준 계약(SLAs)을 충족하거나 초과할 수 있도록 합니다.

AWS 글로벌 클라우드 인프라는 고객이 복원력이 뛰어난 워크로드 아키텍처를 구축할 수 있도록 설계되었습니다. 각 AWS 리전은 완전히 격리되어 있으며 물리적으로 격리된 인프라 파티션인 여러 가용 영역으로 구성됩니다. 가용 영역은 워크로드 복원력에 영향을 줄 수 있는 결함을 격리하여 리전의 다른 영역에 영향을 미치지 않도록 합니다. 그러나 동시에 AWS 리전의 모든 영역은 완전히 중복된 전용 메트로 광섬유를 통해 고대역폭의 지연 시간이 짧은 네트워킹과 상호 연결되어 영역 간에 처리량이 높고 지연 시간이 짧은 네트워킹을 제공합니다. 영역 간의 모든 트래픽은 암호화됩니다. 네트워크 성능은 영역 간에 동기식 복제를 수행하기에 충분합니다. 하나의 애플리케이션이 여러 AZ에 파티셔닝되어 있는 경우 정전, 번개, 토네이도, 허리케인 등의 문제로부터 애플리케이션을 더 효과적으로 격리하고 보호할 수 있습니다.

고객 책임 “클라우드의 복원력”

사용자의 책임은 선택한 AWS 클라우드 서비스에 따라 결정됩니다. 서비스에 따라 복원력 책임의 일환으로서 고객이 수행해야 할 구성 작업의 양이 달라집니다. 예를 들어 Amazon Elastic Compute Cloud(Amazon EC2)와 같은 서비스를 사용하려면 고객이 필요한 모든 복원력 구성 및 관리 작업을 수행해야 합니다. Amazon EC2 인스턴스를 배포하는 고객은 여러 위치(예: AWS 가용 영역)에 EC2 인스턴스를 배포하고, Amazon EC2 Auto Scaling과 같은 서비스를 사용하여 자가 복구를 구현하고, 인스턴스에 설치된 애플리케이션에 대해 복원력이 뛰어난 워크로드 아키텍처 모범 사례를 사용할 책임이 있습니다. Amazon S3 및 Amazon DynamoDB와 같은 관리형 서비스의 경우 AWS는 인프라 계층, 운영 체제 및 플랫폼을 운영하며 고객은 엔드포인트에 액세스하여 데이터를 저장하고 검색합니다. 백업, 버전 관리 및 복제 전략을 포함하여 데이터의 복원력을 관리할 책임은 고객에게 있습니다.

AWS 리전의 여러 가용 영역에 워크로드를 배포하는 것은 문제를 하나의 가용 영역으로 격리하여 워크로드를 보호하도록 설계된 고가용성 전략의 일부이며, 다른 가용 영역의 중복성을 사용하여 요청을 계

속 제공합니다. 다중 AZ 아키텍처는 정전, 낙뢰, 토네이도, 지진 등과 같은 문제로부터 워크로드를 더 잘 격리하고 보호하도록 설계된 DR 전략의 일부이기도 합니다. DR 전략은 여러 AWS 리전을 사용할 수도 있습니다. 예를 들어 액티브/패시브 구성에서 활성 리전이 더 이상 요청을 처리할 수 없는 경우 워크로드에 대한 서비스가 활성 리전에서 DR 리전으로 장애 조치됩니다.

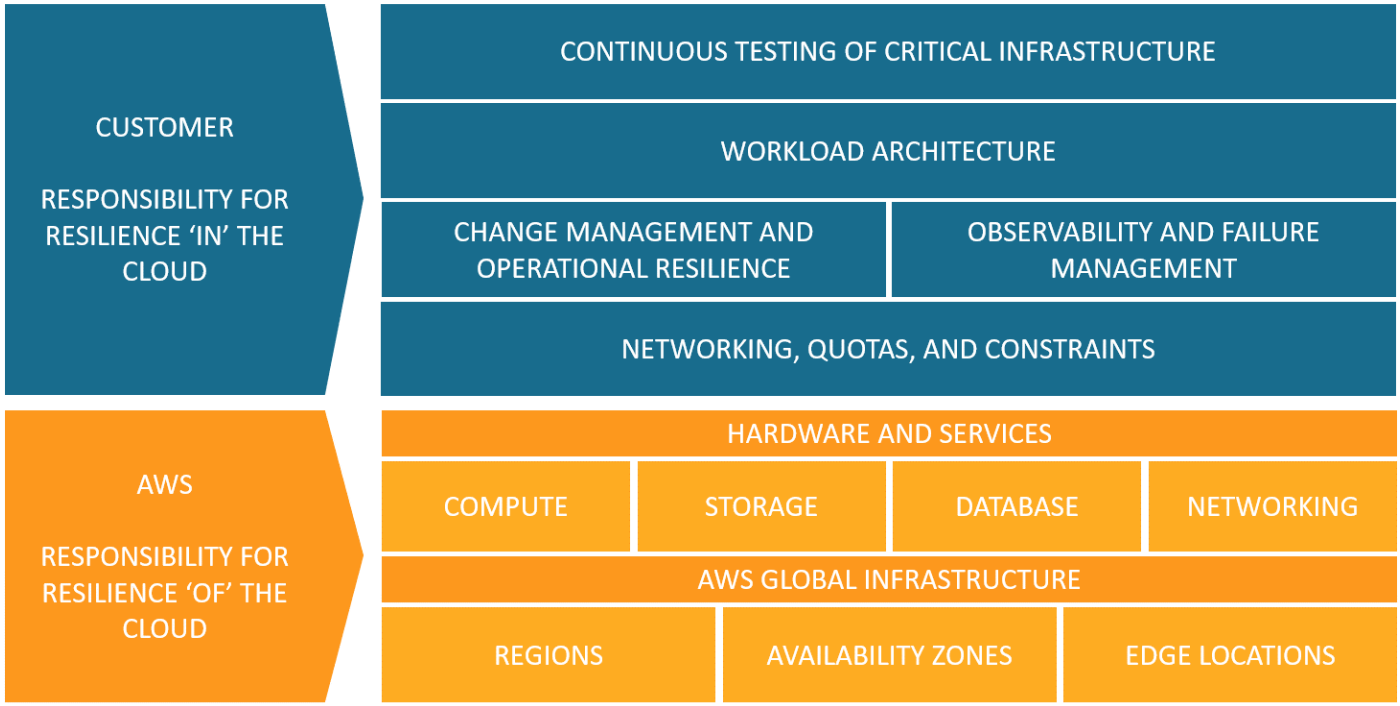

그림 2 - 복원력은 AWS와 고객 간의 공동 책임입니다.

재해란 무엇입니까?

재해 복구를 계획할 때 다음 세 가지 주요 재해 범주에 대한 계획을 평가합니다.

- 지진 또는 홍수와 같은 자연 재해
- 정전 또는 네트워크 연결과 같은 기술적 장애
- 실수로 인한 잘못된 구성, 무단/외부 당사자 액세스 또는 수정과 같은 인적 조치

이러한 각 잠재적 재해는 지역, 지역, 국가 전체, 대륙 또는 전 세계에 영향을 미칠 수도 있습니다. 재해 복구 전략을 고려할 때 재해의 특성과 지리적 영향 모두 중요합니다. 예를 들어 다중 AZ 전략을 사용하면 둘 이상의 가용 영역에 영향을 미치지 않으므로 데이터 센터 중단을 일으키는 로컬 플러딩 문제를 완화할 수 있습니다. 그러나 프로덕션 데이터를 공격하려면 다른 AWS 리전의 백업 데이터로 장애 조치하는 재해 복구 전략을 호출해야 합니다.

고가용성은 재해 복구가 아닙니다.

가용성과 재해 복구 모두 장애 모니터링, 여러 위치에 배포, 자동 장애 조치와 같은 몇 가지 동일한 모범 사례에 의존합니다. 그러나 가용성은 워크로드의 구성 요소에 중점을 두는 반면, 재해 복구는 전체 워크로드의 개별 복사본에 중점을 둡니다. 재해 복구는 가용성과 목표가 다르므로 재해로 간주되는 대규모 이벤트 후 복구 시간을 측정합니다. 가용성에 영향을 미치는 이벤트 발생 시 고가용성 아키텍처를 통해 고객의 요구 사항을 충족할 수 있으므로 먼저 워크로드가 가용성 목표를 충족하는지 확인해야 합니다. 재해 복구 전략에는 가용성과 다른 접근 방식이 필요하며, 필요한 경우 전체 워크로드를 장애 조치할 수 있도록 개별 시스템을 여러 위치에 배포하는 데 중점을 둡니다.

재해 복구 계획에서 워크로드의 가용성은 사용하는 접근 방식에 영향을 미치므로 고려해야 합니다. 한 가용 영역의 단일 Amazon EC2 인스턴스에서 실행되는 워크로드에는 고가용성이 없습니다. 로컬 플러딩 문제가 해당 가용 영역에 영향을 미치는 경우이 시나리오에서는 DR 목표를 충족하기 위해 다른 AZ로의 장애 조치가 필요합니다. 이 시나리오를 워크로드가 여러 활성 리전에 배포되고 모든 리전이 프로덕션 트래픽을 처리하는 고가용성 워크로드 배포 다중 사이트 액티브/액티브와 비교합니다. 이 경우 드물지만 대규모 재해로 인해 리전을 사용할 수 없는 경우에도 DR 전략은 모든 트래픽을 나머지 리전으로 라우팅하여 수행됩니다.

데이터에 접근하는 방법도 가용성과 재해 복구 간에 다릅니다. 고가용성(예: 다중 사이트, 활성/활성 워크로드)을 달성하기 위해 다른 사이트에 지속적으로 복제하는 스토리지 솔루션을 고려합니다. 기본 스토리지 디바이스에서 파일 또는 파일이 삭제되거나 손상된 경우 이러한 파괴적인 변경 사항을 보조 스토리지 디바이스에 복제할 수 있습니다. 이 시나리오에서는 고가용성에도 불구하고 데이터 삭제 또는 손상 시 장애 조치 기능이 손상됩니다. 대신 DR 전략의 일부로 point-in-time 백업도 필요합니다.

비즈니스 연속성 계획(BCP)

재해 복구 계획은 조직의 비즈니스 연속성 계획(BCP)의 하위 집합이어야 하며 독립 실행형 문서가 아니어야 합니다. 재해가 워크로드 이외의 비즈니스 요소에 미치는 영향으로 인해 해당 워크로드의 비즈니스 목표를 달성할 수 없는 경우 워크로드 복원을 위한 공격적인 재해 복구 목표를 유지할 필요가 없습니다. 예를 들어 지진으로 인해 eCommerce 애플리케이션에서 구매한 제품을 전송하지 못할 수 있습니다. 효과적인 DR이 워크로드를 계속 작동하더라도 BCP는 운송 요구 사항을 수용해야 합니다. DR 전략은 비즈니스 요구 사항, 우선순위 및 컨텍스트를 기반으로 해야 합니다.

비즈니스 영향 분석 및 위험 평가

비즈니스 영향 분석은 워크로드에 대한 중단의 비즈니스 영향을 정량화해야 합니다. 워크로드를 사용할 수 없는 내부 및 외부 고객에게 미치는 영향과가 비즈니스에 미치는 영향을 식별해야 합니다. 분석을 통해 워크로드를 얼마나 빨리 사용할 수 있어야 하는지, 그리고 얼마나 많은 데이터 손실을 허용할 수 있는지 확인할 수 있습니다. 그러나 복구 목표를 독립적으로 수립해서는 안 된다는 점에 유의해야 합니다. 중단 가능성과 복구 비용은 워크로드에 재해 복구를 제공하는 비즈니스 가치를 알리는 데 도움이 되는 주요 요소입니다.

비즈니스에 미치는 영향은 시간에 따라 달라질 수 있습니다. 재해 복구 계획에 이를 고려하는 것이 좋습니다. 예를 들어 급여 시스템 중단은 모든 사람이 급여를 받기 직전에 비즈니스에 매우 큰 영향을 미칠 수 있지만 모든 사람이 이미 급여를 받은 직후에는 영향이 낮을 수 있습니다.

워크로드의 기술적 구현에 대한 개요와 함께 재해 유형 및 지리적 영향에 대한 위험 평가는 각 재해 유형에 대해 발생하는 중단의 가능성을 결정합니다.

매우 중요한 워크로드의 경우 비즈니스에 미치는 영향을 최소화하기 위해 데이터 복제 및 연속 백업을 사용하여 여러 리전에 인프라를 배포하는 것이 좋습니다. 덜 중요한 워크로드의 경우 유효한 전략은 재해 복구를 전혀 수행하지 않는 것일 수 있습니다. 또한 일부 재해 시나리오의 경우 재해 발생 가능성이 낮기 때문에 정보에 입각한 결정으로 재해 복구 전략을 마련하지 않는 것도 유효합니다. AWS 리전 내의 가용 영역은 이미 서로 의미 있는 거리를 두고 설계되었으며 대부분의 일반적인 재해가 한 영역에만 영향을 미치고 다른 영역에는 영향을 미치지 않도록 위치를 신중하게 계획하고 있습니다. 따라서 AWS 리전 내의 다중 AZ 아키텍처는 이미 대부분의 위험 완화 요구 사항을 충족할 수 있습니다.

재해 복구 전략이 비즈니스에 미치는 영향과 위험을 고려하여 올바른 수준의 비즈니스 가치를 제공할 수 있도록 재해 복구 옵션 비용을 평가해야 합니다.

이 모든 정보를 사용하여 다양한 재해 시나리오 및 관련 복구 옵션의 위협, 위험, 영향 및 비용을 문서화할 수 있습니다. 이 정보는 각 워크로드의 복구 목표를 결정하는 데 사용해야 합니다.

복구 목표(RTO 및 RPO)

재해 복구(DR) 전략을 생성할 때 조직은 가장 일반적으로 Recovery Time Objective(RTO) 및 Recovery Point Objective(RPO)를 계획합니다.

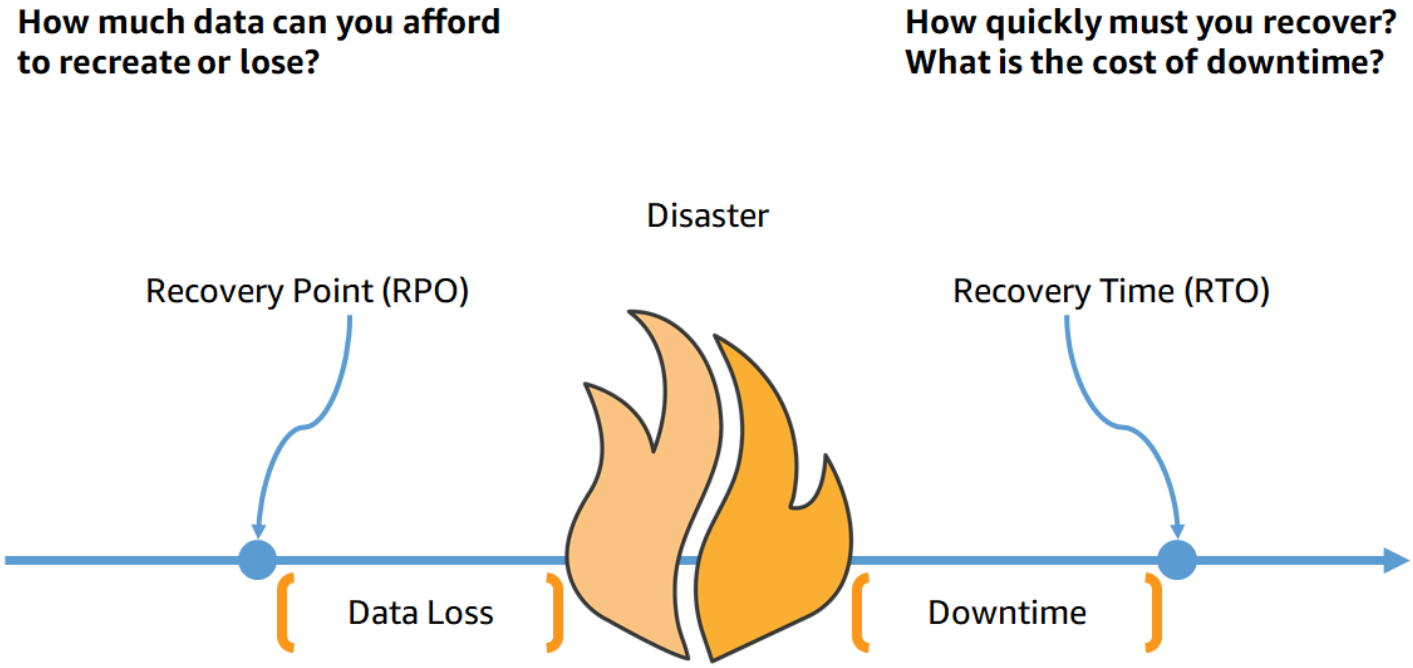

그림 3 - 복구 목표

Recovery Time Objective(RTO)는 서비스 중단과 서비스 복원 사이의 허용되는 최대 지연 시간입니다. 이 목표는 서비스를 사용할 수 없고 조직에서 정의한 허용 기간으로 간주되는 항목을 결정합니다.

이 백서에서는 백업 및 복원, 파일럿 라이트, 웜 스탠바이, 다중 사이트 액티브/액티브의 네 가지 DR 전략을 설명합니다(클라우드의 재해 복구 옵션 참조). 다음 다이어그램에서 기업은 허용되는 최대 RTO와 서비스 복원 전략에 지출할 수 있는 한도도 결정했습니다. 비즈니스 목표를 고려할 때 DR 전략 파일럿 라이트 또는 웜 스탠바이는 RTO와 비용 기준을 모두 충족할 것입니다.

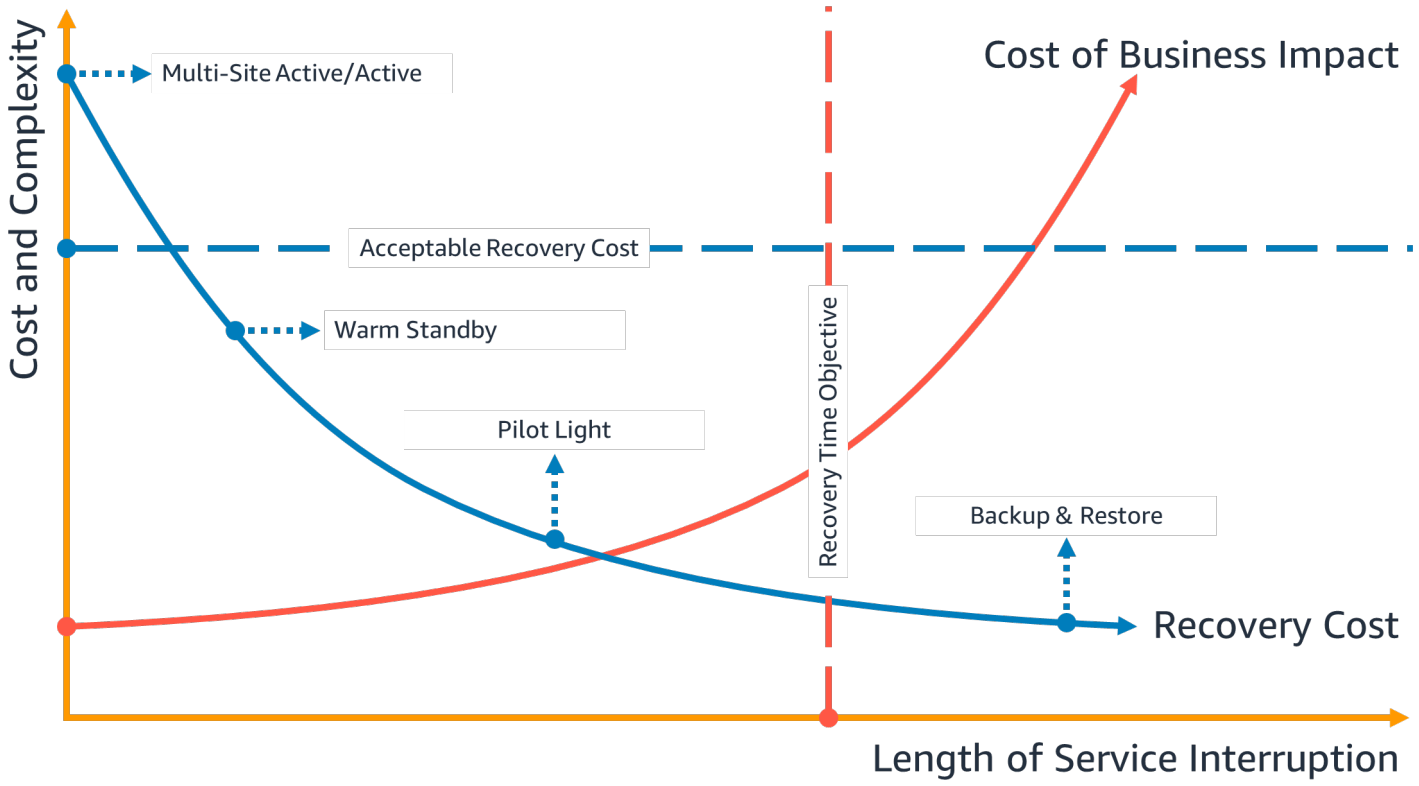

그림 4 - 복구 시간 목표

Recovery Point Objective(RPO)는 마지막 데이터 복구 시점 이후 허용되는 최대 시간입니다. 이 목표는 마지막 복구 시점과 서비스 중단 사이에 허용되는 데이터 손실로 간주되는 항목을 결정하며 조직에서 정의합니다.

다음 다이어그램에서 비즈니스는 허용되는 최대 RPO와 데이터 복구 전략에 지출할 수 있는 한도도 결정했습니다. 네 가지 DR 전략 중 파일럿 라이트 또는 웜 스탠바이 DR 전략은 RPO 및 비용 기준을 모두 충족합니다.

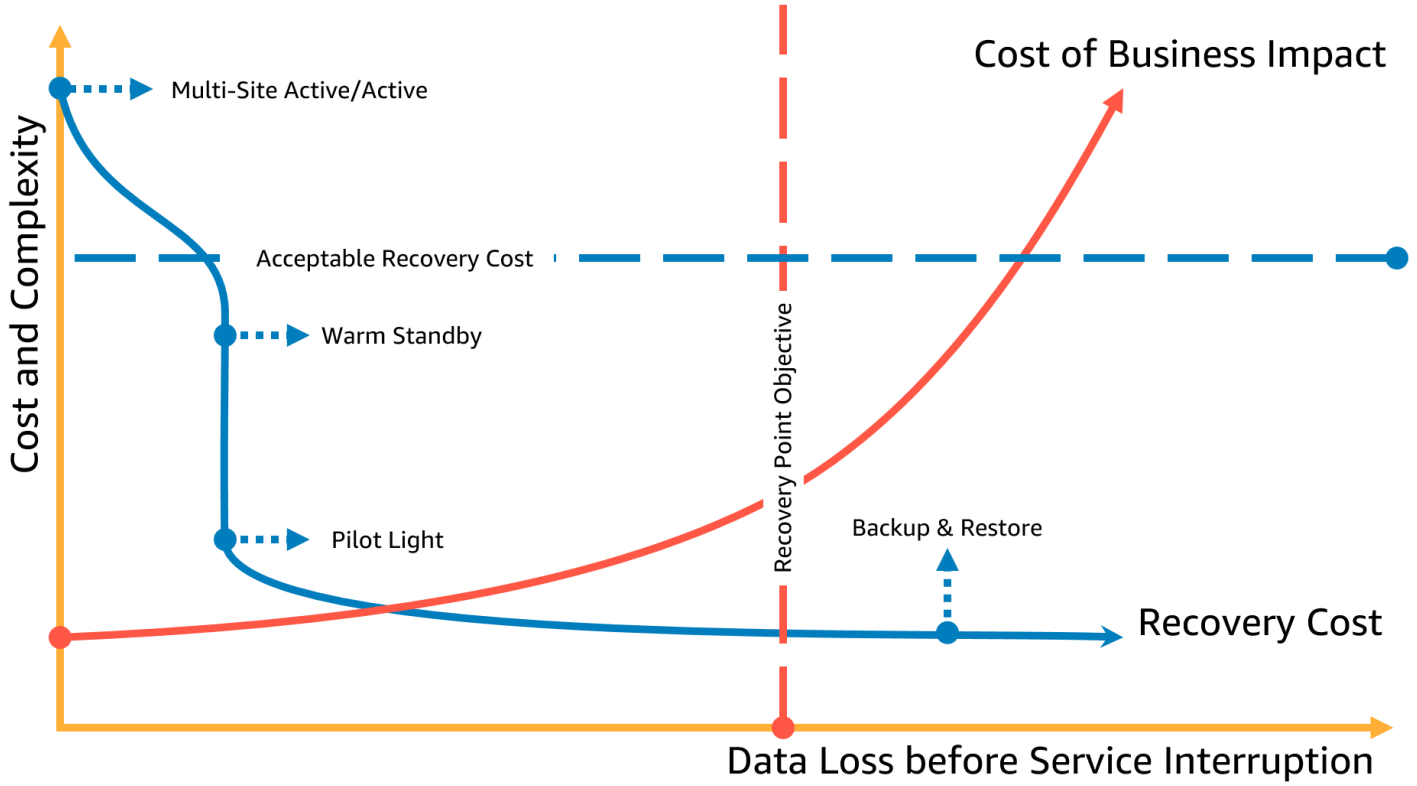

그림 5 - 복구 시점 목표

> **Note**
>
> 복구 전략 비용이 실패 또는 손실 비용보다 높은 경우 규제 요구 사항과 같은 보조 동인이 없는 한 복구 옵션을 적용해서는 안 됩니다. 이 평가를 수행할 때 다양한 비용의 복구 전략을 고려하세요.

클라우드에서는 재해 복구 방식이 다름

재해 복구 전략은 기술 혁신과 함께 진화합니다. 온프레미스 재해 복구 계획에는 테이프를 물리적으로 전송하거나 데이터를 다른 사이트로 복제하는 작업이 포함될 수 있습니다. 조직은 AWS에서 DR 목표를 달성하기 위해 이전 재해 복구 전략의 비즈니스 영향, 위험 및 비용을 재평가해야 합니다. AWS 클라우드의 재해 복구에는 기존 환경에 비해 다음과 같은 이점이 있습니다.

- 복잡성을 줄이면서 재해로부터 신속하게 복구
- 간단하고 반복 가능한 테스트를 통해 더 쉽고 자주 테스트할 수 있습니다.
- 관리 오버헤드를 줄이면 운영 부담이 줄어듭니다.
- 자동화를 통해 오류 가능성을 줄이고 복구 시간을 단축할 수 있는 기회

AWS를 사용하면 물리적 백업 데이터 센터의 고정 자본 비용을 클라우드에서 권한 있는 환경의 가변 운영 비용으로 거래할 수 있으므로 비용을 크게 줄일 수 있습니다.

많은 조직에서 온프레미스 재해 복구는 데이터 센터의 워크로드 또는 워크로드가 중단될 위험과 백업되거나 복제된 데이터를 보조 데이터 센터로 복구할 위험을 기반으로 했습니다. 조직이 AWS에 워크로드를 배포하면 잘 설계된 워크로드를 구현하고 AWS 글로벌 클라우드 인프라의 설계에 의존하여 이러한 중단의 영향을 완화할 수 있습니다. 클라우드에서 안정적이고 안전하며 효율적이고 비용 효율적인 워크로드를 설계하고 운영하기 위한 아키텍처 모범 사례에 대한 자세한 내용은 AWS Well-Architected Framework - 신뢰성 원칙 백서를 참조하세요. AWS Well-Architected Tool를 사용하여 워크로드를 주기적으로 검토하여 Well-Architected Framework의 모범 사례와 지침을 준수하는지 확인합니다. 이 도구는에서 무료로 사용할 수 있습니다AWS Management Console.

워크로드가 AWS에 있는 경우 데이터 센터 연결(액세스 기능 제외), 전원, 냉방, 소방 및 하드웨어에 대해 걱정할 필요가 없습니다. 이 모든 것이 관리되며 여러 장애 격리 가용 영역(각각 하나 이상의 개별 데이터 센터로 구성됨)에 액세스할 수 있습니다.

단일 AWS 리전

물리적 데이터 센터 하나의 중단 또는 손실로 인한 재해 이벤트의 경우 단일 AWS 리전 내의 여러 가용 영역에서 고가용성 워크로드를 구현하면 자연 재해와 기술 재해를 완화하는 데 도움이 됩니다. 이 단일 리전 내에서 데이터를 지속적으로 백업하면 데이터 손실을 초래할 수 있는 오류 또는 무단 활동과 같은 인적 위협에 대한 위험을 줄일 수 있습니다. 각 AWS 리전은 여러 가용 영역으로 구성되며, 각 가용 영역은 다른 영역의 장애로부터 격리됩니다. 각 가용 영역은 차례로 하나 이상의 개별 물리적 데이터 센터로 구성됩니다. 영향을 미치는 문제를 더 잘 격리하고 고가용성을 달성하기 위해 동일한 리전의 여

러 영역에 워크로드를 분할할 수 있습니다. 가용 영역은 물리적 중복성을 위해 설계되었으며 복원력을 제공하여 정전, 인터넷 가동 중지, 홍수 및 기타 자연 재해 발생 시에도 중단 없는 성능을 제공합니다. AWS Global Cloud Infrastructure를 참조하여 AWS가 이를 수행하는 방법을 알아보세요.

단일 AWS 리전의 여러 가용 영역에 배포하면 단일(또는 여러) 데이터 센터의 장애로부터 워크로드를 더 잘 보호할 수 있습니다. 단일 리전 배포를 추가로 보장하기 위해 데이터 및 구성(인프라 정의 포함)을 다른 리전에 백업할 수 있습니다. 이 전략은 재해 복구 계획의 범위를 줄여 데이터 백업 및 복원만 포함합니다. 다른 AWS 리전에 백업하여 다중 리전 복원력을 활용하는 것은 다음 섹션에 설명된 다른 다중 리전 옵션에 비해 간단하고 저렴합니다. 예를 들어 Amazon Simple Storage Service(Amazon S3)에 백업하면 데이터를 즉시 검색할 수 있습니다. 그러나 데이터의 일부에 대한 DR 전략에 검색 시간(몇 분에서 몇 시간으로)에 대한 더 완화된 요구 사항이 있는 경우 Amazon Glacier 또는 Amazon Glacier Deep Archive를 사용하면 백업 및 복구 전략 비용이 크게 절감됩니다.

일부 워크로드에는 규제 데이터 레지던시 요구 사항이 있을 수 있습니다. 이는 현재 AWS 리전이 하나뿐인 로캘의 워크로드에 적용되는 경우 위에서 설명한 대로 고가용성을 위해 다중 AZ 워크로드를 설계하는 것 외에도 해당 리전 내의 AZs 별도의 위치로 사용할 수도 있습니다. 이는 해당 리전 내의 워크로드에 적용되는 데이터 레지던시 요구 사항을 해결하는 데 도움이 될 수 있습니다. 다음 섹션에 설명된 DR 전략은 여러 AWS 리전을 사용하지만 리전 대신 가용 영역을 사용하여 구현할 수도 있습니다.

여러 AWS 리전

여러 데이터 센터가 서로 상당한 거리를 잃을 위험이 포함된 재해 이벤트의 경우 AWS 내 전체 리전에 영향을 미치는 자연 재해와 기술 재해를 완화하기 위해 재해 복구 옵션을 고려해야 합니다. 다음 섹션에 설명된 모든 옵션을 다중 리전 아키텍처로 구현하여 이러한 재해로부터 보호할 수 있습니다.

클라우드의 재해 복구 옵션

AWS 내에서 사용할 수 있는 재해 복구 전략은 백업을 만드는 낮은 비용과 낮은 복잡성부터 여러 활성 리전을 사용하는 보다 복잡한 전략에 이르기까지 네 가지 접근 방식으로 광범위하게 분류할 수 있습니다. 액티브/패시브 전략은 액티브 사이트(예: AWS 리전)를 사용하여 워크로드를 호스팅하고 트래픽을 처리합니다. 패시브 사이트(예: 다른 AWS 리전)가 복구에 사용됩니다. 패시브 사이트는 장애 조치 이벤트가 트리거될 때까지 트래픽을 능동적으로 처리하지 않습니다.

재해 복구 전략이 필요한 경우 호출에 대한 확신을 가질 수 있도록 재해 복구 전략을 정기적으로 평가하고 테스트하는 것이 중요합니다. AWS Resilience Hub를 사용하여 RTO 및 RPO 목표를 충족할 가능성이 있는지 여부를 포함하여 AWS 워크로드의 복원력을 지속적으로 검증하고 추적할 수 있습니다.

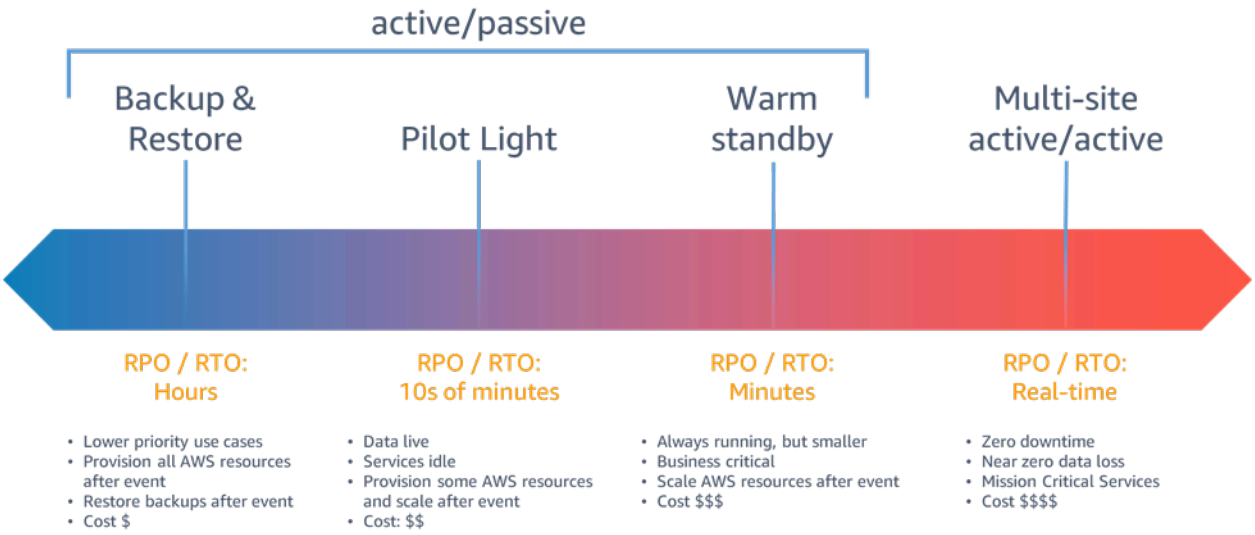

그림 6 - 재해 복구 전략

잘 설계된 고가용성 워크로드에 대한 물리적 데이터 센터 하나의 중단 또는 손실로 인한 재해 이벤트의 경우 재해 복구에 대한 백업 및 복원 접근 방식만 필요할 수 있습니다. 재해 정의가 물리적 데이터 센터의 중단 또는 손실을 넘어 리전의 중단 또는 손실을 넘어가거나 필요한 규제 요구 사항이 적용되는 경우 파일럿 라이트, 웜 스탠바이 또는 멀티사이트 액티브/액티브를 고려해야 합니다.

전략과 이를 구현할 AWS 리소스를 선택할 때는 AWS 내에서 일반적으로 서비스를 데이터 영역과 컨트롤 플레인으로 나눕니다. 데이터 영역에서는 실시간 서비스를 제공하고, 컨트롤 플레인은 환경을 구성하는 데 사용됩니다. 복원력을 극대화하려면 장애 조치 작업의 일부로 데이터 영역 작업만 사용해야 합니다. 데이터 영역은 일반적으로 컨트롤 플레인보다 가용성 설계 목표가 더 높기 때문입니다.

백업 및 복원

백업 및 복원은 데이터 손실 또는 손상을 완화하는 데 적합한 접근 방식입니다. 이 접근 방식은 다른 AWS 리전으로 데이터를 복제하여 리전 재해를 완화하거나 단일 가용 영역에 배포된 워크로드의 중복성 부족을 완화하는 데에도 사용할 수 있습니다. 데이터 외에도 복구 리전에서 인프라, 구성 및 애플리케이션 코드를 재배포해야 합니다. 인프라가 오류 없이 빠르게 재배포되도록 하려면 항상 AWS CloudFormation 또는와 같은 서비스를 사용하여 코드형 인프라(IaC)를 사용하여 배포해야 합니다AWS Cloud Development Kit (AWS CDK). IaC가 없으면 복구 리전에서 워크로드를 복원하는 것이 복잡하여 복구 시간이 늘어나고 RTO를 초과할 수 있습니다. 사용자 데이터 외에도 Amazon EC2 인스턴스를 생성하는 데 사용하는 Amazon Machine Image(AMIs 포함하여 코드와 구성도 백업해야 합니다. Amazon EC2 AWS CodePipeline를 사용하여 애플리케이션 코드 및 구성의 재배포를 자동화할 수 있습니다.

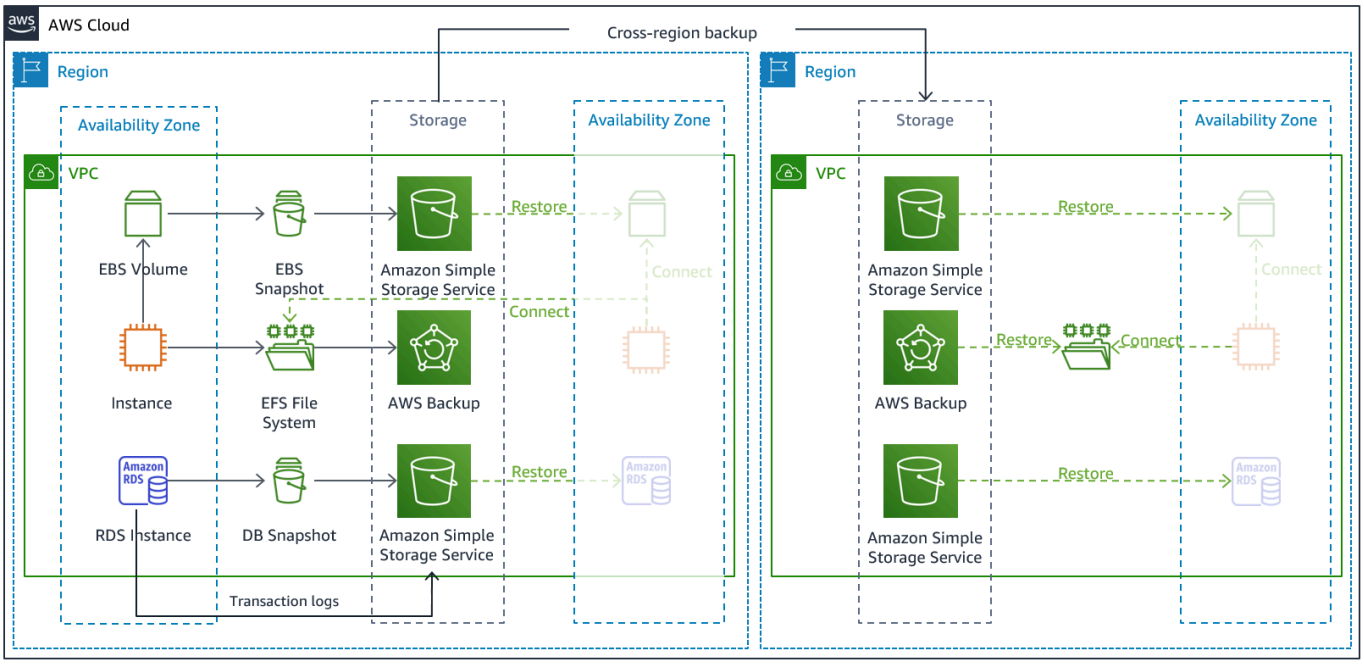

그림 7 - 백업 및 복원 아키텍처

서비스

워크로드 데이터에는 주기적으로 실행되거나 지속적인 백업 전략이 필요합니다. 백업을 실행하는 빈도에 따라 달성 가능한 복구 시점(RPO에 맞게 조정되어야 함)이 결정됩니다. 또한 백업은 백업을 가져온 시점으로 복원할 수 있는 방법을 제공해야 합니다. point-in-time으로 복구를 통한 백업은 다음 서비스 및 리소스를 통해 사용할 수 있습니다.

- Amazon Elastic Block Store(Amazon EBS) 스냅샷
- Amazon DynamoDB 백업
- Amazon RDS 스냅샷
- Amazon Aurora DB 스냅샷
- Amazon EFS 백업(사용 시 AWS Backup)
- Amazon Redshift 스냅샷
- Amazon Neptune 스냅샷
- Amazon DocumentDB
- Amazon FSx for Windows File Server, Amazon FSx for Lustre, Amazon FSx for NetApp ONTAP 및 Amazon FSx for OpenZFS

Amazon Simple Storage Service(Amazon S3)의 경우 Amazon S3 교차 리전 복제(CRR)를 사용하여 복원 지점을 선택할 수 있도록 저장된 객체에 대한 버전 관리를 제공하면서 객체를 지속적으로 DR 리전의 S3 버킷에 비동기식으로 복사할 수 있습니다. 데이터의 지속적 복제는 데이터를 백업하는 데 가장 짧은 시간(0에 가까움)이라는 이점이 있지만 데이터 손상 또는 악의적인 공격(예: 무단 데이터 삭제)과 point-in-time 백업과 같은 재해 이벤트로부터 보호하지 못할 수 있습니다. 지속적 복제는 AWS Services for Pilot Light 섹션에서 다룹니다.

AWS Backup는 다음 서비스 및 리소스에 대한 AWS 백업 기능을 구성, 예약 및 모니터링할 수 있는 중앙 위치를 제공합니다.

- Amazon Elastic Block Store(Amazon EBS) 볼륨
- Amazon EC2 인스턴스
- Amazon Relational Database Service(Amazon RDS) 데이터베이스(Amazon Aurora 데이터베이스 포함)
- Amazon DynamoDB 테이블
- Amazon Elastic File System(Amazon EFS) 파일 시스템
- AWS Storage Gateway 볼륨
- Amazon FSx for Windows File Server, Amazon FSx for Lustre, Amazon FSx for NetApp ONTAP 및 Amazon FSx for OpenZFS

AWS Backup 는 재해 복구 리전과 같은 리전 간 백업 복사를 지원합니다.

Amazon S3 데이터에 대한 추가 재해 복구 전략으로 S3 객체 버전 관리를 활성화합니다. 객체 버전 관리는 작업 전에 원래 버전을 유지하여 삭제 또는 수정 작업의 결과로부터 S3의 데이터를 보호합니다. 객체 버전 관리는 인적 오류 유형 재해를 완화하는 데 유용할 수 있습니다. S3 복제를 사용하여 DR 리전에 데이터를 백업하는 경우 기본적으로 소스 버킷에서 객체가 삭제되면 Amazon S3는 소스 버킷에만 삭제 마커를 추가합니다. 이 접근 방식은 소스 리전의 악의적인 삭제로부터 DR 리전의 데이터를 보호합니다.

데이터 외에도 워크로드를 재배포하고 Recovery Time Objective(RTO)를 충족하는 데 필요한 구성과 인프라도 백업해야 합니다. AWS CloudFormation는 코드형 인프라(IaC)를 제공하며 워크로드의 모든 AWS 리소스를 정의할 수 있으므로 여러 AWS 계정 및 AWS 리전에 안정적으로 배포하고 재배포할 수 있습니다. 워크로드에서 사용하는 Amazon EC2 인스턴스를 Amazon Machine Image(AMIs. AMI는 인스턴스 루트 볼륨의 스냅샷과 인스턴스에 연결된 기타 EBS 볼륨에서 생성됩니다. 이 AMI를 사용하여 복원된 버전의 EC2 인스턴스를 시작할 수 있습니다. AMI는 리전 내에서 또는 리전 간에 복사할 수 있습니다. 또는 AWS Backup를 사용하여 계정 간 및 다른 AWS 리전으로 백업을 복사할 수 있습니다. 교차 계정 백업 기능은 내부자 위협 또는 계정 손상이 포함된 재해 이벤트로부터 보호하는 데 도움이 됩니다. AWS Backup 또한는 인스턴스의 개별 EBS 볼륨 외에도 EC2 백업을 위한 추가 기능을 추가하여 인스턴스 유형, 구성된 Virtual Private Cloud(VPC), 보안 그룹, IAM 역할, 모니터링 구성 및 태그와 같은 메타데이터 AWS Backup 를 저장하고 추적합니다. 그러나이 추가 메타데이터는 EC2 백업을 동일한 AWS 리전으로 복원할 때만 사용됩니다.

백업으로 재해 복구 리전에 저장된 모든 데이터는 장애 조치 시 복원해야 합니다.는 복원 기능을 AWS Backup 제공하지만 현재 예약된 복원 또는 자동 복원을 활성화하지 않습니다. AWS SDK를 사용하여 API를 호출APIs AWS Backup. 백업이 완료될 때마다 이를 정기적으로 반복되는 작업으로 설정하거나 복원을 트리거할 수 있습니다. 다음 그림은 Amazon Simple Notification Service(Amazon SNS) 및를 사용한 자동 복원의 예를 보여줍니다AWS Lambda. 백업에서 데이터 복원은 컨트롤 플레인 작업이므로 예약된 주기적 데이터 복원을 구현하는 것이 좋습니다. 재해 발생 시이 작업을 사용할 수 없는 경우에도 최근 백업에서 작업 가능한 데이터 스토어가 생성됩니다.

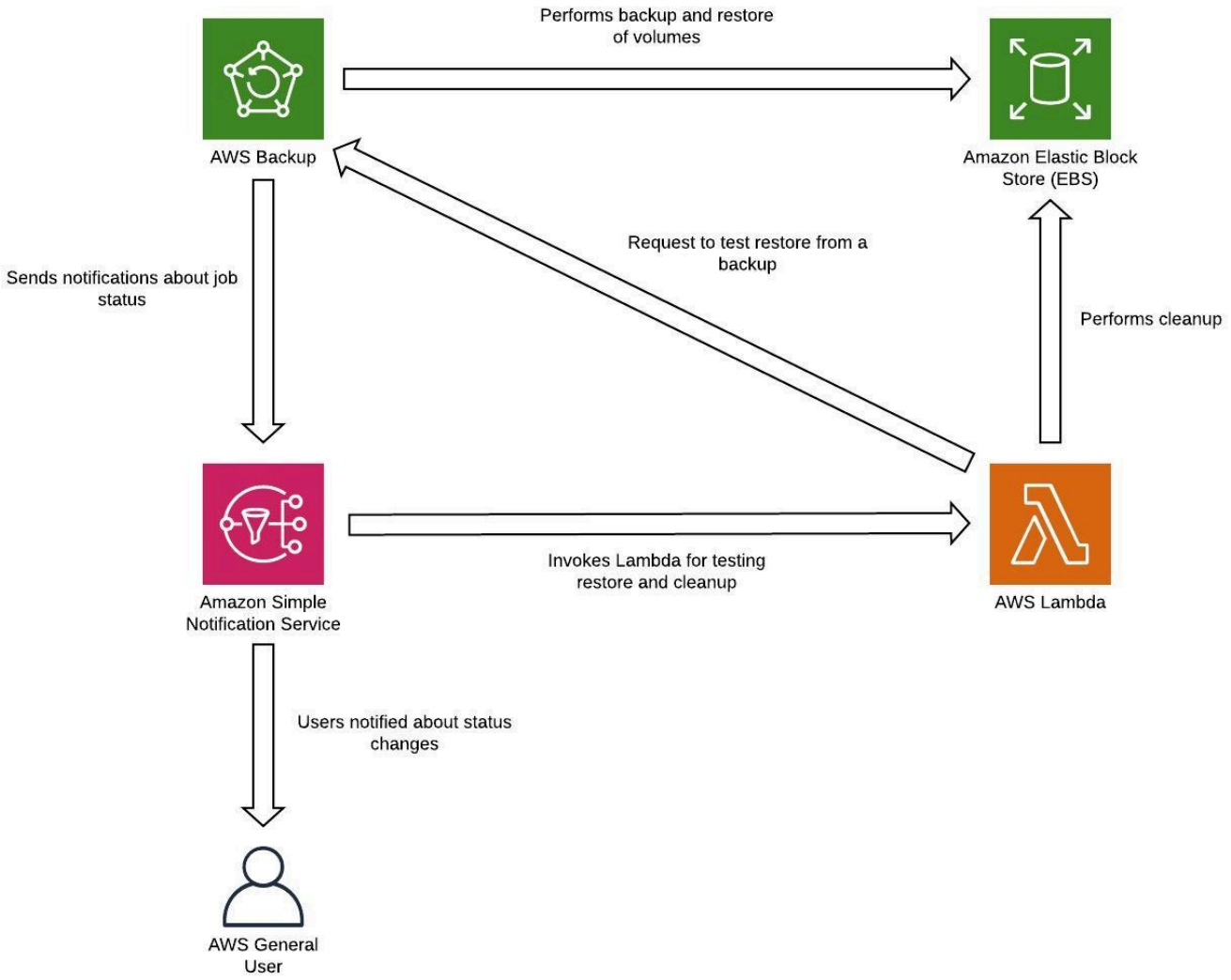

그림 8 - 백업 복원 및 테스트

> **Note**
>
> 백업 전략에는 백업 테스트가 포함되어야 합니다. 자세한 내용은 재해 복구 테스트 섹션을 참조하세요. 구현에 대한 실습 데모는 AWS Well-Architected Lab: 데이터 백업 및 복원 테스트를 참조하세요.

파일럿 라이트

파일럿 라이트 접근 방식을 사용하면 한 리전에서 다른 리전으로 데이터를 복제하고 코어 워크로드 인프라의 사본을 프로비저닝할 수 있습니다. 데이터베이스 및 객체 스토리지 등 데이터 복제 및 백업을

지원하는 데 필요한 리소스가 항상 실행됩니다. 애플리케이션 서버와 같은 다른 요소는 애플리케이션 코드 및 구성과 함께 로드되지만 "끄기"되며 테스트 중에 또는 재해 복구 장애 조치가 호출될 때만 사용됩니다. 클라우드에서는 리소스가 필요하지 않을 때 리소스를 프로비저닝 해제하고 프로비저닝할 수 있는 유연성이 있습니다. “끄기”의 모범 사례는 리소스를 배포하지 않고 필요한 경우 리소스를 배포(“켜기”)하는 구성 및 기능을 생성하는 것입니다. 백업 및 복원 접근 방식과 달리 코어 인프라는 항상 사용할 수 있으며 애플리케이션 서버를 켜고 확장하여 전체 규모의 프로덕션 환경을 신속하게 프로비저닝할 수 있습니다.

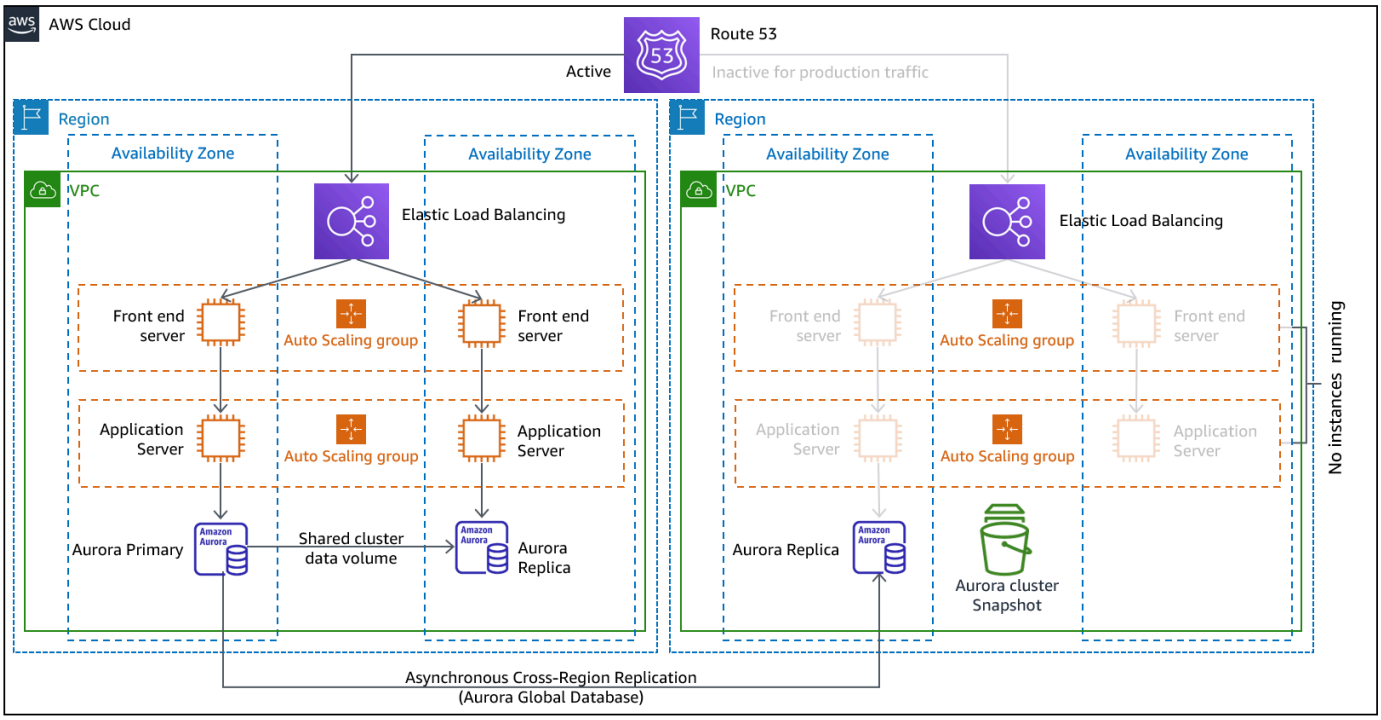

그림 9 - 파일럿 조명 아키텍처

파일럿 라이트 접근 방식은 활성 리소스를 최소화하여 재해 복구의 지속적인 비용을 최소화하고 핵심 인프라 요구 사항이 모두 적용되므로 재해 발생 시 복구를 간소화합니다. 이 복구 옵션을 사용하려면 배포 접근 방식을 변경해야 합니다. 각 리전에서 핵심 인프라를 변경하고 워크로드(구성, 코드) 변경 사항을 각 리전에 동시에 배포해야 합니다. 이 단계는 배포를 자동화하고 코드형 인프라(IaC)를 사용하여 여러 계정 및 리전에 인프라를 배포하여 간소화할 수 있습니다(기본 리전으로의 전체 인프라 배포 및 DR 리전으로의 축소/전환된 인프라 배포). 최고 수준의 리소스 및 보안 격리를 제공하려면 리전별로 다른 계정을 사용하는 것이 좋습니다(피해된 보안 인증 정보도 재해 복구 계획의 일부인 경우).

이 접근 방식을 사용하면 데이터 재해도 완화해야 합니다. 지속적인 데이터 복제는 일부 유형의 재해로부터 보호해 주지만, 전략에 저장된 데이터의 버전 관리 또는 특정 시점 복구 옵션이 포함되지 않은 이상 데이터 손상 또는 중단으로부터 보호해 주지는 않습니다. 재해 리전에서 복제된 데이터를 백업하여 동일한 리전에서 point-in-time 백업을 생성할 수 있습니다.

서비스

백업 및 복원 섹션에서 다루는 AWS 서비스를 사용하여 point-in-time 백업을 생성하는 것 외에도 파일럿 라이트 전략에 대해 다음 서비스도 고려하세요.

파일럿 라이트의 경우 DR 리전의 라이브 데이터베이스 및 데이터 스토어에 대한 지속적인 데이터 복제는 낮은 RPO에 가장 적합합니다(앞에서 설명한 point-in-time 백업 외에 사용하는 경우). AWS는 다음 서비스 및 리소스를 사용하여 데이터에 대해 지속적인 교차 리전 비동기 데이터 복제를 제공합니다.

- Amazon Simple Storage Service(Amazon S3) 복제
- Amazon RDS 읽기 전용 복제본
- Amazon Aurora 글로벌 데이터베이스
- Amazon DynamoDB 글로벌 테이블
- Amazon DocumentDB 글로벌 클러스터
- Amazon ElastiCache용 글로벌 데이터 스토어(Redis OSS)

지속적 복제를 사용하면 DR 리전에서 거의 즉시 데이터 버전을 사용할 수 있습니다. 실제 복제 시간은 S3 객체용 S3 Replication Time Control(S3 RTC)과 같은 서비스 기능과 Amazon Aurora 글로벌 데이터베이스의 관리 기능을 사용하여 모니터링할 수 있습니다. S3

재해 복구 리전에서 읽기/쓰기 워크로드를 실행하기 위해 장애 조치하는 경우 RDS 읽기 전용 복제본을 기본 인스턴스로 승격해야 합니다. Aurora 이외의 DB 인스턴스의 경우 프로세스를 완료하는 데 몇 분 정도 걸리며 재부팅은 프로세스의 일부입니다. RDS를 사용한 교차 리전 복제(CRR) 및 장애 조치의 경우 Amazon Aurora 글로벌 데이터베이스를 사용하면 몇 가지 이점이 있습니다. 글로벌 데이터베이스는 전용 인프라를 사용하여 데이터베이스를 완전히 사용할 수 있게 해주며, 일반적으로 1초 미만의 지연 시간으로 보조 리전에 복제할 수 있습니다(AWS 리전 내에서는 100밀리초 미만). Amazon Aurora 글로벌 데이터베이스를 사용하면 기본 리전에서 성능 저하 또는 중단이 발생하는 경우, 리전이 완전히 중단되더라도 1분 이내에 읽기/쓰기 책임을 맡도록 보조 리전 중 하나를 승격할 수 있습니다. 또한 모든 보조 클러스터의 RPO 지연 시간을 모니터링하여 하나 이상의 보조 클러스터가 대상 RPO 기간 내에 유지되도록 Aurora를 구성할 수 있습니다.

리소스가 적거나 작은 코어 워크로드 인프라의 축소된 버전을 DR 리전에 배포해야 합니다. AWS CloudFormation를 사용하면 인프라를 정의하고 AWS 계정 및 AWS 리전 전체에 일관되게 배포할 수 있습니다.는 사전 정의된 가상 파라미터를 AWS CloudFormation 사용하여 배포된 AWS 계정과 AWS 리전을 식별합니다. 따라서 CloudFormation 템플릿에서 조건 로직을 구현하여 DR 리전에 인프라의 축소된 버전만 배포할 수 있습니다. EC2 인스턴스 배포의 경우 Amazon Machine Image(AMI)는 하드웨어 구성 및 설치된 소프트웨어와 같은 정보를 제공합니다. 필요한 AMIs를 생성하는 Image Builder

파이프라인을 구현하고 이를 기본 리전과 백업 리전 모두에 복사할 수 있습니다. 이렇게 하면 재해 발생 시 이러한 골든 AMIs가 새 리전에서 워크로드를 재배포하거나 스케일 아웃하는 데 필요한 모든 것을 확보할 수 있습니다. Amazon EC2 인스턴스는 축소된 구성(기본 리전보다 적은 인스턴스)으로 배포됩니다. 프로덕션 트래픽을 지원하기 위해 인프라를 확장하려면 웜 스탠바이 섹션의 Amazon EC2 Auto Scaling을 참조하세요.

파일럿 라이트와 같은 액티브/패시브 구성의 경우, 모든 트래픽은 처음에 기본 리전으로 이동하고 기본 리전을 더 이상 사용할 수 없는 경우 재해 복구 리전으로 전환됩니다. 이 장애 조치 작업은 자동 또는 수동으로 시작할 수 있습니다. 상태 확인 또는 경보에 따라 자동으로 시작된 장애 조치는 주의해서 사용해야 합니다. 여기에 설명된 모범 사례를 사용하더라도 복구 시간과 복구 시점은 0보다 커져 가용성과 데이터가 약간 손실됩니다. 필요하지 않을 때 장애 조치하면(거짓 경보) 이러한 손실이 발생합니다. 따라서 수동으로 시작된 장애 조치를 자주 사용합니다. 이 경우에도 여전히 장애 조치 단계는 자동화하여 수동 시작은 버튼을 누르는 것 정도가 되도록 해야 합니다.

AWS 서비스를 사용할 때 고려해야 할 몇 가지 트래픽 관리 옵션이 있습니다.

한 가지 옵션은 Amazon Route 53을 사용하는 것입니다. Amazon Route 53을 사용하면 하나 이상의 AWS 리전에 있는 여러 IP 엔드포인트를 Route 53 도메인 이름과 연결할 수 있습니다. 그런 다음 해당 도메인 이름으로 트래픽을 적절한 엔드포인트로 라우팅할 수 있습니다. 장애 조치 시 트래픽을 복구 엔드포인트로 전환하고 기본 엔드포인트에서 다른 곳으로 전환해야 합니다. Amazon Route 53 상태 확인은 이러한 엔드포인트를 모니터링합니다. 이러한 상태 확인을 사용하면 자동으로 시작된 DNS 장애 조치를 구성하여 트래픽이 데이터 영역에서 수행되는 매우 안정적인 작업인 정상 엔드포인트로만 전송되도록 할 수 있습니다. 수동으로 시작된 장애 조치를 사용하여 이를 구현하려면 Amazon Application Recovery Controller(ARC)를 사용할 수 있습니다. ARC를 사용하면 실제로 상태를 확인하지 않고 사용자가 완전히 제어할 수 있는 켜기/끄기 스위치 역할을 하는 Route 53 상태 확인을 생성할 수 있습니다. AWS CLI 또는 AWS SDK를 사용하면 가용성이 높은이 데이터 영역 API를 사용하여 장애 조치를 스크립트로 작성할 수 있습니다. 스크립트는 이러한 전환(Route 53 상태 확인)을 전환하여 Route 53에 기본 리전 대신 복구 리전으로 트래픽을 전송하도록 지시합니다. 일부에서 사용한 수동 시작 장애 조치를 위한 또 다른 옵션은 가중치 기반 라우팅 정책을 사용하고 모든 트래픽이 복구 리전으로 이동하도록 기본 및 복구 리전의 가중치를 변경하는 것입니다. 그러나이 작업은 컨트롤 플레인 작업이므로 Amazon Application Recovery Controller(ARC)를 사용하여 데이터 영역이 접근하는 만큼 복원력이 뛰어나지 않습니다.

또 다른 옵션은를 사용하는 것입니다AWS Global Accelerator. AnyCast IP를 사용하면 하나 이상의 AWS 리전에 있는 여러 엔드포인트를 동일한 정적 퍼블릭 IP 주소 또는 주소와 연결할 수 있습니다. AWS Global Accelerator 그런 다음는 트래픽을 해당 주소와 연결된 적절한 엔드포인트로 라우팅합니다. Global Accelerator 상태 확인은 엔드포인트를 모니터링합니다. 이러한 상태 확인을 사용하여는 애플리케이션의 상태를 AWS Global Accelerator 확인하고 사용자 트래픽을 정상 애플리케이션 엔드포

인트로 자동으로 라우팅합니다. 수동으로 시작된 장애 조치의 경우 트래픽 다이얼을 사용하여 트래픽을 수신하는 엔드포인트를 조정할 수 있지만 이는 컨트롤 플레인 작업입니다. Global Accelerator는 광범위한 AWS 엣지 네트워크를 사용하여 가능한 한 빨리 AWS 네트워크 백본에 트래픽을 배치하므로 애플리케이션 엔드포인트에 대한 지연 시간을 줄입니다. 또한 Global Accelerator는 DNS 시스템(예: Route 53)에서 발생할 수 있는 캐싱 문제를 방지합니다.

Amazon CloudFront는 오리진 장애 조치를 제공합니다. 여기서 기본 엔드포인트에 대한 지정된 요청이 실패하면 CloudFront는 요청을 보조 엔드포인트로 라우팅합니다. 앞서 설명한 장애 조치 작업과 달리 모든 후속 요청은 여전히 기본 엔드포인트로 이동하며 각 요청마다 장애 조치가 수행됩니다.

AWS Elastic Disaster Recovery

AWS Elastic Disaster Recovery(DRS)는 기본 서버의 블록 수준 복제를 AWS 사용하여 모든 소스에서로 서버 호스팅 애플리케이션 및 서버 호스팅 데이터베이스를 지속적으로 복제합니다. Elastic Disaster Recovery를 사용하면 온프레미스 또는 다른 클라우드 공급자 및 해당 환경에서 호스팅되는 워크로드의 AWS 클라우드 재해 복구 대상으로의 리전을 사용할 수 있습니다. EC2에서 AWS 호스팅되는 애플리케이션 및 데이터베이스(즉, RDS 아님)로만 구성된 호스팅 워크로드의 재해 복구에도 사용할 수 있습니다. Elastic Disaster Recovery는 파일럿 라이트 전략을 사용하여 스테이징 영역으로 사용되는 Amazon Virtual Private Cloud(Amazon VPC)에서 데이터 사본과 “스위치 오프” 리소스를 유지합니다. 장애 조치 이벤트가 트리거되면 스테이징된 리소스가 복구 위치로 사용되는 대상 Amazon VPC에서 전체 용량 배포를 자동으로 생성하는 데 사용됩니다.

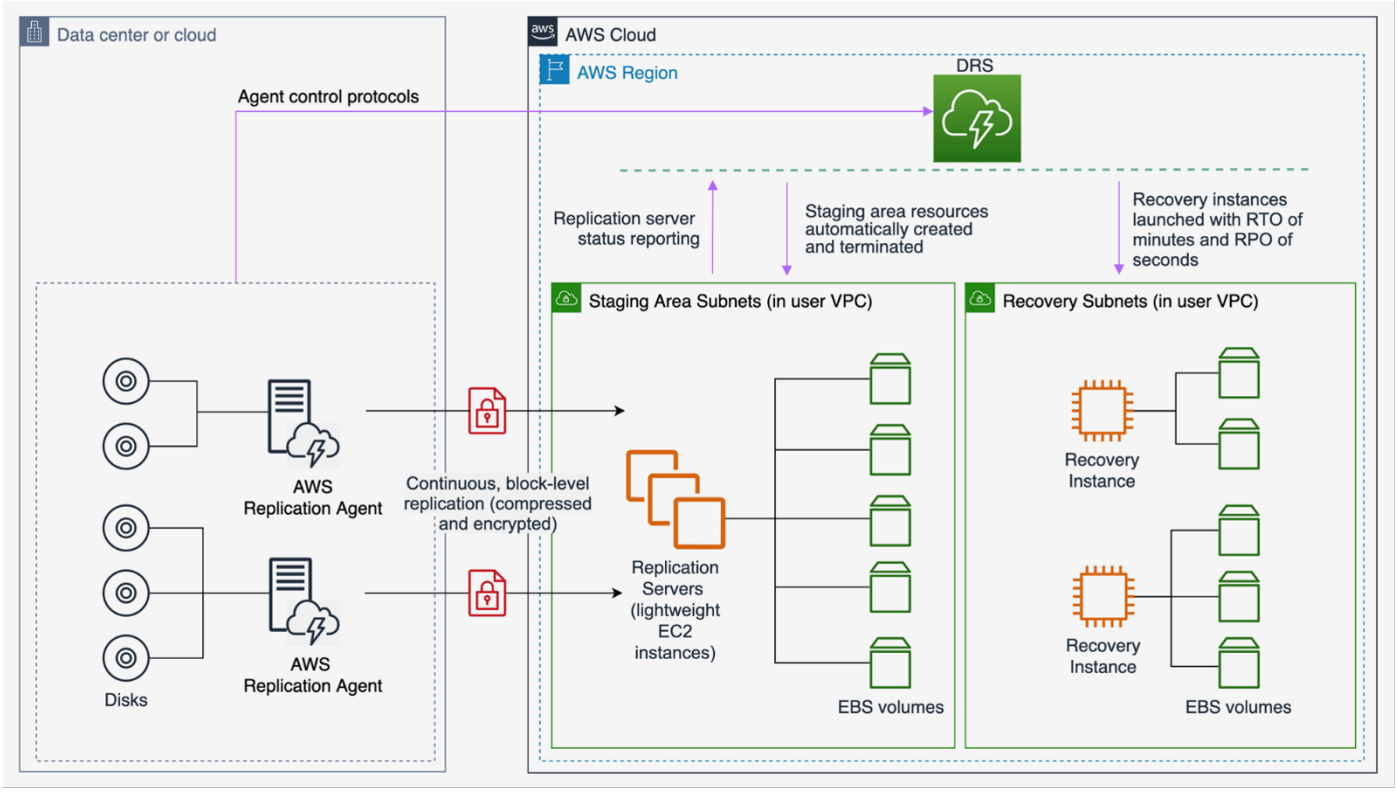

그림 10 - AWS Elastic Disaster Recovery 아키텍처

웜 대기

웜 대기 접근 방식에는 스케일 다운되었지만 완전히 기능하는 프로덕션 환경의 복사본이 다른 리전에 복사됩니다. 이 접근법은 파일럿 라이트의 개념을 확대하고 복구 시간을 단축합니다. 워크로드가 다른 리전에서 상시 실행되기 때문입니다. 또한이 접근 방식을 사용하면 테스트를 더 쉽게 수행하거나 지속적인 테스트를 구현하여 재해 복구 능력에 대한 신뢰도를 높일 수 있습니다.

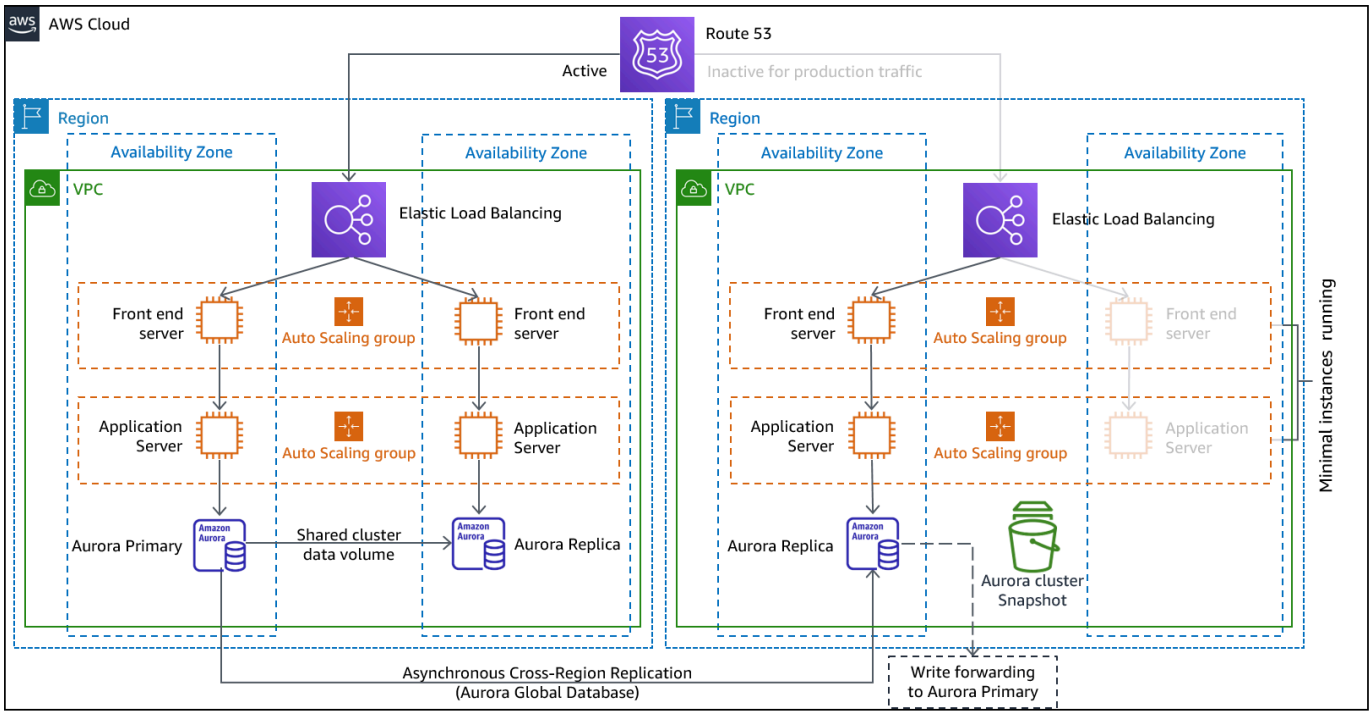

그림 11 - 웜 스탠바이 아키텍처

참고: 파일럿 라이트와 웜 스탠바이의 차이점을 이해하기 어려운 경우가 있습니다. 둘 다 기본 리전 자산의 복사본이 있는 DR 리전의 환경을 포함합니다. 파일럿 라이트는 먼저 추가 조치를 취하지 않고 요청을 처리할 수 없는 반면, 웜 스탠바이는 트래픽(감소된 용량 수준에서)을 즉시 처리할 수 있습니다. 파일럿 라이트 접근 방식을 사용하려면 서버를 “켜고”, 추가(코어가 아닌) 인프라를 배포하고 확장해야 하는 반면, 웜 스탠바이를 사용하려면 확장만 하면 됩니다(모든 것이 이미 배포되어 실행 중임). RTO 및 RPO 요구 사항을 사용하여 이러한 접근 방식 중에서 선택할 수 있습니다.

서비스

백업 및 복원과 파일럿 라이트에 적용되는 모든 AWS 서비스는 데이터 백업, 데이터 복제, 액티브/패시브 트래픽 라우팅, EC2 인스턴스를 포함한 인프라 배포를 위한 웜 스탠바이에도 사용됩니다.

Amazon EC2 Auto Scaling은 AWS 리전 내의 Amazon EC2 인스턴스, Amazon ECS 작업, Amazon DynamoDB 처리량 및 Amazon Aurora 복제본을 포함한 리소스를 확장하는 데 사용됩니다. Amazon EC2 Auto Scaling은 AWS 리전 내의 가용 영역 간에 EC2 인스턴스 배포를 확장하여 해당 리전 내에서 복원력을 제공합니다. Auto Scaling을 사용하면 파일럿 라이트 또는 웜 스탠바이 전략의 일환으로 DR 리전을 전체 프로덕션 기능으로 확장할 수 있습니다. 예를 들어 EC2의 경우 Auto Scaling 그룹에서 원하는 용량 설정을 늘립니다. 를 통해 AWS Management Console, AWS SDK를 통해 자동으로 또는 원하는 새 용량 값을 사용하여 AWS CloudFormation 템플릿을 재배포하여이 설정을 수동으로 조정할 수 있습니다. AWS CloudFormation 파라미터를 사용하여 CloudFormation 템플릿을 더 쉽게 재배포할 수 있습니다. 프로덕션 용량으로 확장하는 것을 제한하지 않도록 DR 리전의 서비스 할당량이 충분히 높게 설정되어 있는지 확인합니다.

Auto Scaling은 컨트롤 플레인 활동이므로 이에 의존하면 전체 복구 전략의 복원력이 저하됩니다. 이는 절충입니다. 복구 리전이 배포된 전체 프로덕션 로드를 처리할 수 있도록 충분한 용량을 프로비저닝하도록 선택할 수 있습니다. 이 정적 안정성 구성을 상시 대기라고 합니다(다음 섹션 참조). 또는 더 적은 리소스를 프로비저닝하여 비용이 절감되지만 Auto Scaling에 종속되도록 선택할 수 있습니다. 일부 DR 구현은 초기 트래픽을 처리하기에 충분한 리소스를 배포하여 낮은 RTO를 보장한 다음 Auto Scaling을 사용하여 후속 트래픽을 증가시킵니다.

다중 사이트 액티브/액티브

다중 사이트 액티브/액티브 또는 핫 스탠바이 액티브/패시브 전략의 일환으로 여러 리전에서 워크로드를 동시에 실행할 수 있습니다. 다중 사이트 액티브/액티브는 배포된 모든 리전의 트래픽을 처리하는 반면, 핫 스탠바이는 단일 리전의 트래픽만 제공하고 다른 리전(들)은 재해 복구에만 사용됩니다. 다중 사이트 액티브/액티브 접근 방식을 사용하면 사용자는 워크로드가 배포된 모든 리전의 워크로드에 액세스할 수 있습니다. 이 접근 방식은 재해 복구에 대한 가장 복잡하고 비용이 많이 드는 접근 방식이지만 올바른 기술 선택 및 구현을 통해 대부분의 재해에 대해 복구 시간을 거의 0으로 줄일 수 있습니다(그러나 데이터 손상은 백업에 의존해야 할 수 있으며, 이는 일반적으로 0이 아닌 복구 시점이 될 수 있음). 핫 스탠바이는 사용자가 단일 리전으로만 이동하고 DR 리전이 트래픽을 받지 않는 액티브/패시브 구성을 사용합니다. 대부분의 고객은 두 번째 리전에서 전체 환경을 구축하려는 경우 액티브/액티브 환경을 사용하는 것이 좋습니다. 또는 두 리전을 모두 사용하여 사용자 트래픽을 처리하지 않으려는 경우 웜 스탠바이는 보다 경제적이고 운영상 덜 복잡한 접근 방식을 제공합니다.

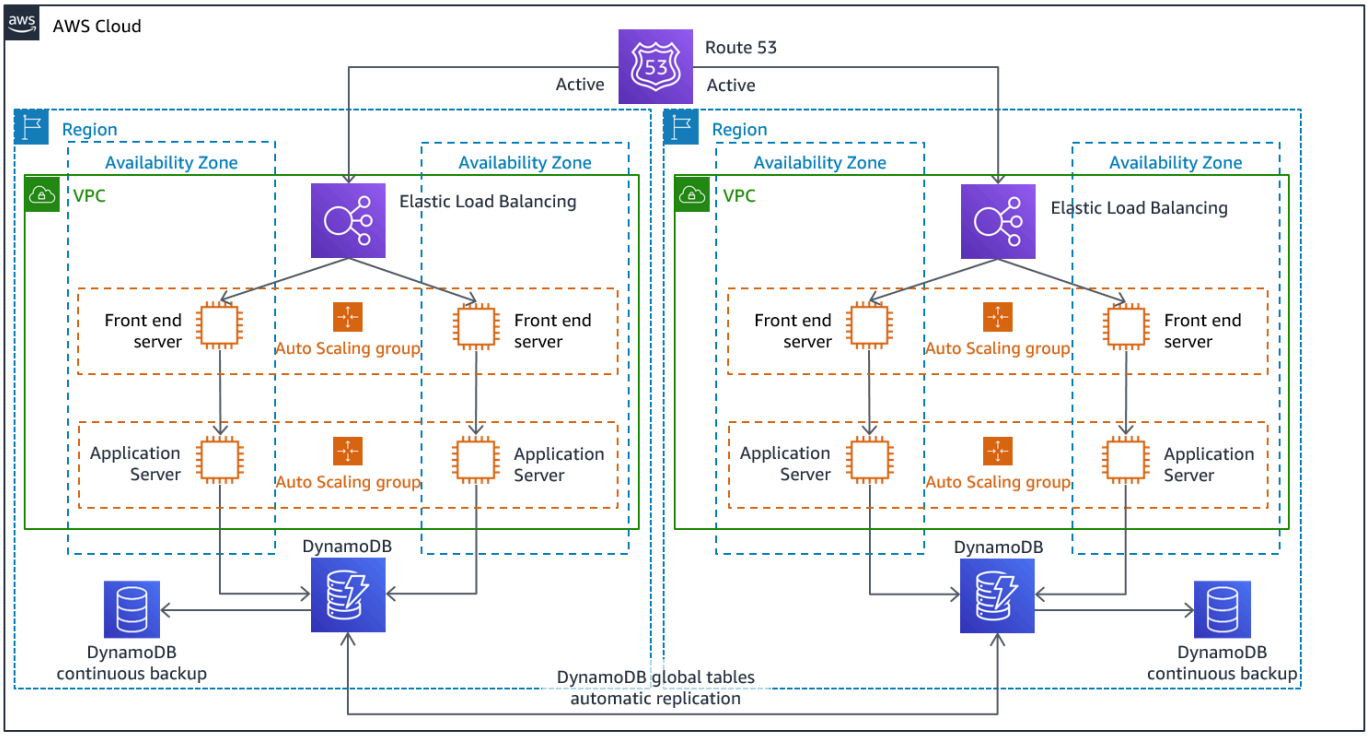

그림 12 - 다중 사이트 액티브/액티브 아키텍처(핫 스탠바이의 경우 하나의 액티브 경로를 비활성으로 변경)

다중 사이트 액티브/액티브에서는 워크로드가 둘 이상의 리전에서 실행되므로이 시나리오에서 장애 조치와 같은 것은 없습니다. 이 경우 재해 복구 테스트는 워크로드가 리전 손실에 반응하는 방식에 중점을 둡니다. 트래픽이 장애가 발생한 리전에서 멀리 라우팅되나요? 다른 리전(들)이 모든 트래픽을 처리할 수 있습니까? 데이터 재해에 대한 테스트도 필요합니다. 백업 및 복구는 여전히 필요하며 정기적으로 테스트해야 합니다. 또한 데이터 손상, 삭제 또는 난독화와 관련된 데이터 재해의 복구 시간은 항상 0보다 크고 복구 시점은 항상 재해가 발견되기 전의 특정 시점입니다. 복구 시간을 거의 0으로 유지하기 위해 다중 사이트 액티브/액티브(또는 핫 스탠바이) 접근 방식의 추가 복잡성과 비용이 필요한 경우 보안을 유지하고 인적 재해를 완화하기 위해 인적 오류를 방지하기 위해 추가 노력을 기울여야 합니다.

서비스

백업 및 복원, 파일럿 라이트, 웜 스탠바이가 적용되는 모든 AWS 서비스는 point-in-time 데이터 백업, 데이터 복제, 액티브/액티브 트래픽 라우팅, EC2 인스턴스를 포함한 인프라 배포 및 규모 조정에도 사용됩니다.

앞서 설명한 액티브/패시브 시나리오(파일럿 라이트 및 웜 스탠바이)의 경우, Amazon Route 53와를 모두 사용하여 네트워크 트래픽을 액티브 리전으로 라우팅할 AWS Global Accelerator 수 있습니다. 여기

에서 활성/활성 전략의 경우 이러한 두 서비스 모두 어떤 사용자가 어떤 활성 리전 엔드포인트로 이동할지 결정하는 정책의 정의도 활성화합니다. 를 AWS Global Accelerator 사용하여 각 애플리케이션 엔드포인트로 전송되는 트래픽의 비율을 제어하도록 트래픽 다이얼을 설정합니다. Amazon Route 53는 이 백분율 접근 방식과 지리 근접성 및 지연 시간 기반 정책을 포함하여 사용 가능한 기타 여러 정책을 지원합니다. Global Accelerator는 AWS 엣지 서버의 광범위한 네트워크를 자동으로 활용하여 가능한 한 빨리 AWS 네트워크 백본에 트래픽을 온보딩하므로 요청 지연 시간이 줄어듭니다.

이 전략을 사용한 비동기식 데이터 복제는 거의 0에 가까운 RPO를 활성화합니다. Amazon Aurora 글로벌 데이터베이스와 같은 AWS 서비스는 전용 인프라를 사용하여 데이터베이스를 완전히 사용할 수 있게 해주고, 일반적으로 1초 미만의 지연 시간으로 최대 5개의 보조 리전에 복제할 수 있습니다. 액티브/패시브 전략의 경우 쓰기는 기본 리전에만 발생합니다. 액티브/액티브의 차이점은 각 액티브 리전에 대한 쓰기와의 데이터 일관성이 처리되는 방식을 설계하는 것입니다. 로컬 읽기라고 하는 가장 가까운 리전에서 제공할 사용자 읽기를 설계하는 것이 일반적입니다. 쓰기에는 다음과 같은 몇 가지 옵션이 있습니다.

- 쓰기 글로벌 전략은 모든 쓰기를 단일 리전으로 라우팅합니다. 해당 리전에 장애가 발생하면 다른 리전이 쓰기를 수락하도록 승격됩니다. Aurora 글로벌 데이터베이스는 리전 간 읽기 전용 복제본과의 동기화를 지원하므로 글로벌 쓰기에 적합하며, 보조 리전 중 하나를 승격하여 1분 이내에 읽기/쓰기 책임을 맡을 수 있습니다. 또한 Aurora는 Aurora 글로벌 데이터베이스의 보조 클러스터가 기본 클러스터에 대한 쓰기 작업을 수행하는 SQL 문을 전달할 수 있도록 쓰기 전달을 지원합니다.
- 쓰기 로컬 전략은 쓰기를 가장 가까운 리전으로 라우팅합니다(읽기처럼). Amazon DynamoDB 글로벌 테이블은 이러한 전략을 활성화하여 글로벌 테이블이 배포된 모든 리전에서 읽기 및 쓰기를 허용합니다. Amazon DynamoDB 글로벌 테이블은 동시 업데이트 간의 마지막 라이터 성공 조정을 사용합니다.
- 쓰기 분할 전략은 쓰기 충돌을 방지하기 위해 파티션 키(예: 사용자 ID)를 기반으로 특정 리전에 쓰기를 할당합니다. 이 경우 양방향으로 구성된 Amazon S3 복제를 사용할 수 있으며, 현재 두 리전 간의 복제를 지원합니다. 이 접근 방식을 구현할 때는 버킷 A와 B 모두에서 복제본 수정 동기화를 활성화하여 복제된 객체에 객체 액세스 제어 목록(ACLs), 객체 태그 또는 객체 잠금과 같은 복제본 메타데이터 변경 사항을 복제해야 합니다. 활성 리전의 버킷 간에 삭제 마커를 복제할지 여부를 구성할 수도 있습니다. 복제 외에도 데이터 손상 또는 파괴 이벤트로부터 보호하기 위해 전략에는 point-in-time 백업도 포함되어야 합니다.

AWS CloudFormation 는 여러 AWS 리전의 AWS 계정 간에 일관되게 배포된 인프라를 적용하는 강력한 도구입니다. AWS CloudFormation StackSets는 단일 작업으로 여러 계정 및 리전에서 CloudFormation 스택을 생성, 업데이트 또는 삭제할 수 있도록 하여이 기능을 확장합니다. AWS CloudFormation 는 YAML 또는 JSON을 사용하여 코드형 인프라를 정의하지만를 AWS Cloud

Development Kit (AWS CDK) 사용하면 익숙한 프로그래밍 언어를 사용하여 코드형 인프라를 정의할 수 있습니다. 코드는 AWS에서 리소스를 배포하는 데 사용되는 CloudFormation으로 변환됩니다.

탐지

워크로드가 제공해야 하는 비즈니스 성과를 제공하지 않는지 최대한 빨리 파악하는 것이 중요합니다. 이렇게 하면 재해를 신속하게 선언하고 인시던트에서 복구할 수 있습니다. 공격적인 복구 목표의 경우 이 응답 시간과 적절한 정보가 결합되어 복구 목표를 달성하는 데 매우 중요합니다. 복구 시간 목표가 1시간인 경우 인시던트를 감지하고, 적절한 담당자에게 알리고, 에스컬레이션 프로세스에 참여하고, 예상 복구 시간(DR 계획을 실행하지 않고)에 대한 정보(있는 경우)를 평가하고, 재해를 선언하고, 1시간 이내에 복구해야 합니다.

> **Note**
>
> RTO가 위험에 처하더라도 이해관계자가 DR을 호출하지 않기로 결정한 경우 DR 계획 및 목표를 재평가합니다. DR 계획을 호출하지 않기로 한 결정은 계획이 부적절하거나 실행에 대한 신뢰도가 부족하기 때문일 수 있습니다.

비즈니스 가치를 제공하는 현실적이고 달성 가능한 목표를 제공하려면 계획 및 목표에 인시던트 감지, 알림, 에스컬레이션, 검색 및 선언을 고려하는 것이 중요합니다.

AWS는 서비스 상태 대시보드에 서비스 가용성에 대한 up-to-the-minute 정보를 게시합니다. 언제든지를 확인하여 현재 상태 정보를 가져오거나 RSS 피드를 구독하여 각 개별 서비스의 중단에 대한 알림을 받을 수 있습니다. 서비스 상태 대시보드에 표시되지 않은 서비스 중 하나에 실시간 운영 문제가 발생하는 경우 지원 요청을 생성할 수 있습니다.

는 계정에 영향을 미칠 수 있는 AWS Health 이벤트에 대한 정보를 AWS Health Dashboard 제공합니다. 이 정보는 최근 이벤트와 예정된 이벤트를 범주별로 보여 주는 대시보드 및 지난 90일간의 모든 이벤트를 보여 주는 전체 이벤트 로그의 두 가지 방법으로 표시됩니다.

가장 엄격한 RTO 요구 사항의 경우 상태 확인을 기반으로 자동 장애 조치를 구현할 수 있습니다. 사용자 경험을 대표하고 주요 성능 지표를 기반으로 상태 확인을 설계합니다. 심층 상태 확인은 워크로드의 주요 기능을 연습하고 얕은 하트비트 확인을 넘어섭니다. 여러 신호를 기반으로 심층 상태 확인을 사용합니다. 할 필요가 없을 때 장애 조치할 경우 가용성 위험이 발생하므로 거짓 경보를 트리거하지 않도록이 접근 방식에 주의하세요.

재해 복구 테스트

재해 복구 구현을 테스트하여 구현을 검증하고 워크로드의 DR 리전에 대한 장애 조치를 정기적으로 테스트하여 RTO 및 RPO가 충족되는지 확인합니다.

피해야 할 패턴은 거의 실행되지 않는 복구 경로를 개발하는 것입니다. 읽기 전용 쿼리에 사용되는 보조 데이터 스토어를 예로 들 수 있습니다. 데이터 스토어에 데이터를 쓸 때 기본 스토어에서 장애가 발생하면 보조 데이터 스토어로 장애 조치를 진행할 수 있습니다. 이 장애 조치를 자주 테스트하지 않으면 보조 데이터 스토어의 기능에 대한 가정이 잘못될 수 있습니다. 마지막으로 테스트했을 때 충분했을 수 있는 보조의 용량은이 시나리오에서 더 이상 부하를 견딜 수 없거나 보조 리전의 서비스 할당량으로 충분하지 않을 수 있습니다.

경험에 따르면 자주 테스트하는 경로만이 유일하게 작동하는 오류 복구 방법입니다. 이것이 복구 경로 수가 적은 것이 가장 좋은 이유입니다.

복구 패턴을 설정하고 정기적으로 테스트할 수 있습니다. 복잡하거나 중요한 복구 경로가 있는 경우에도 복구 경로가 작동하는지 검증하기 위해 프로덕션 환경에서 해당 실패를 정기적으로 실행해야 합니다.

DR 리전에서 구성 드리프트를 관리합니다. 인프라, 데이터 및 구성이 DR 리전에서 필요에 따라 이루어지는지 확인합니다. 예를 들어 AMIs 및 서비스 할당량이 up-to-date 상태인지 확인합니다.

AWS Config를 사용하여 AWS 리소스 구성을 지속적으로 모니터링하고 기록할 수 있습니다. AWS Config 는 드리프트를 감지하고 AWS Systems Manager Automation을 트리거하여 드리프트를 수정하고 경보를 발생시킬 수 있습니다.는 배포한 스택의 드리프트를 추가로 감지할 AWS CloudFormation 수 있습니다.

결론

고객은 클라우드에서 애플리케이션의 가용성에 대한 책임이 있습니다. 재해가 무엇인지 정의하고 이 정의와 비즈니스 성과에 미칠 수 있는 영향을 반영하는 재해 복구 계획을 수립하는 것이 중요합니다. 영향 분석 및 위험 평가를 기반으로 Recovery Time Objective(RTO) 및 Recovery Point Objective(RPO)를 생성한 다음 재해를 완화할 적절한 아키텍처를 선택합니다. 재해 탐지가 가능하고 시기 적절한지 확인합니다. 목표가 언제 위험한지 파악하는 것이 중요합니다. 계획이 있는지 확인하고 테스트를 통해 계획을 검증합니다. 신뢰도 부족 또는 재해 복구 목표 미충족으로 인해 검증되지 않은 재해 복구 계획이 구현되지 않습니다.

기여자

다음은 이 문서의 기여자입니다.

- Alex Livingstone, AWS Enterprise Support의 프랙티스 리드 클라우드 운영
- Seth Eliot, Amazon Web Services의 보안 주체 신뢰성 솔루션 아키텍트

참조 자료

자세한 내용은 다음을 참조하세요.

- AWS 아키텍처 센터
- 신뢰성 원칙, AWS Well-Architected Framework
- 재해 복구 계획 체크리스트
- 상태 확인 구현
- AWS의 재해 복구(DR) 아키텍처, 1부: 클라우드의 복구 전략
- AWS의 재해 복구(DR) 아키텍처, 파트 II: 빠른 복구를 통한 백업 및 복원
- AWS의 재해 복구(DR) 아키텍처, 파트 III: 파일럿 라이트 및 웜 스탠바이
- AWS의 재해 복구(DR) 아키텍처, 4부: 다중 사이트 활성/활성
- Creating Disaster Recovery Mechanisms Using Amazon Route 53
- Minimizing Dependencies in a Disaster Recovery Plan
- AWS Well-Architected Disaster Recovery Labs 실습
- AWS 솔루션 구현: 다중 리전 애플리케이션 아키텍처
- AWS re:Invent 2018: 다중 리전 활성-활성 애플리케이션을 위한 아키텍처 패턴(ARC209-R2)

문서 이력

이 백서에 대한 업데이트 알림을 받으려면 RSS 피드를 구독하면 됩니다.

변경 사항	설명	날짜
마이너 업데이트	전체적으로 버그 수정 및 여러 사소한 변경 사항이 있습니다.	2022년 4월 1일
백서 업데이트	사소한 편집 업데이트.	2022년 3월 21일
백서 업데이트	데이터 영역 및 컨트롤 플레인에 대한 정보가 추가되었습니다. 액티브/패시브 장애 조치를 구현하는 방법에 대한 자세한 내용을 추가했습니다. CloudEndure 재해 복구를 AWS Elastic Disaster Recovery로 대체했습니다.	2022년 2월 17일
마이너 업데이트	AWS Well-Architected Tool 정보가 추가되었습니다.	2022년 2월 11일
최초 게시	백서가 처음 게시되었습니다.	2021년 2월 12일

고지 사항

고객은 본 문서의 정보를 독립적으로 평가할 책임이 있습니다. 이 문서: (a) 정보 제공만을 목적으로 하고, (b) 현행 AWS 제품 제공 및 관행을 나타내며, (c) AWS와 그 계열사, 공급업체 또는 라이선스 제공자로부터 어떠한 약정이나 보증도 하지 않습니다. AWS 제품 또는 서비스는 명시적이든 묵시적이든 어떠한 종류의 보증, 진술 또는 조건 없이 “있는 그대로” 제공됩니다. 고객에 대한 AWS의 책임 및 채무는 AWS 계약에 준거합니다. 본 문서는 AWS와 고객 간의 어떠한 계약도 구성하지 않으며 이를 변경하지도 않습니다.

AWS 용어집

최신 AWS 용어는 AWS 용어집 참조의 AWS 용어집을 참조하세요.

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