

데이터센터 재해 100% 복구를 위한 해법 찾기

신상운 / 효성인포메이션시스템 매니저

IT 재해는 누군가의 평온한 일상을 멈추게 한다. 기업의 브랜드 가치 훼손이나 사회적 혼란, 연계된 서비스 중단 등 피해 양상이나 규모도 천차만별이다. 기업이 재해 위험으로부터 IT 인프라를 보호하고 복구하기 위해 온 힘을 기울이는 것도 이 때문이다.

현재까지 가장 검증되고 안정적인 스토리지 방식의 재해복구 솔루션을 통해, 인프라 실패에 대비하는 해법을 찾아보자.

재해복구 관건은 ‘데이터 복구 시간 최소화’

데이터센터 이중화, 재해복구 인프라 구축 등 철저한 준비에도 불구하고 재해로 인한 피해가 끊이지 않는 이유는 무엇일까? 기술이 발달할수록 시스템 아키텍처가 더 복잡해지고 있으며, 재해복구 기술이 필요한 시기에 제대로 구현되지 않아 복구 시간이 오래 걸리거나, 데이터 복구 품질이 기대 수준에 미치지 못하기 때문이다. 최근에는 IT 전략 방향이 서비스 가용성과 데이터센터 내 IT 자원의 효율성 및 사용률 극대화에 맞춰지고 있다. 기업들은 보다 완벽한 데이터센터 운영을 위해 대내외 규정을 강화하고, IT 자원의 다운타임을 최소화하기 위한 최적의 솔루션을 찾고 있다.

데이터센터 재해와 관련한 중요한 기준은 ‘RTO(Recovery Time Objective: 복구 목표 시간)’다. RTO는 IT 시스템이 장애나 재해로 서비스가 중단됐을 때 서비스를 원상 복구하는 데 걸리는 최대 허용 시간을 말한다. 현재 금융권에서는 핵심 업무에 대한 RTO를 3시간 이내로, 보험회사는 24시간 이내로 엄격하게 규정하고 있다.

이러한 RTO에 영향을 미치는 중요한 요인은 RPO(Recovery Point Objective: 복구 목표 시점)다. RPO는 재해가 발생한 후 서비스를 재개하기 위해 백업 스토리지에서 복구해야 하는 파일의 최대 백업 주기를 의미한다. RPO 시간이 비는 만큼의 데이터는 수기 또는 다른 시스템에 저장된 데이터로 복구해야 하는데, 복구 절차가 복잡할수록 시간이 많이 소요되고 RTO에도 영향을 미친다. 많은 기업들이 ‘RPO 제로’를 목표로 시스템을 구축하려는 이유다.

‘RPO 제로’를 위한 데이터 복제 방식 3가지

‘RPO 제로’를 구현하는 데 일반적인 데이터 복제 방식으로는 ‘동기식’이 꼽힌다. ‘액티브-스탠바이

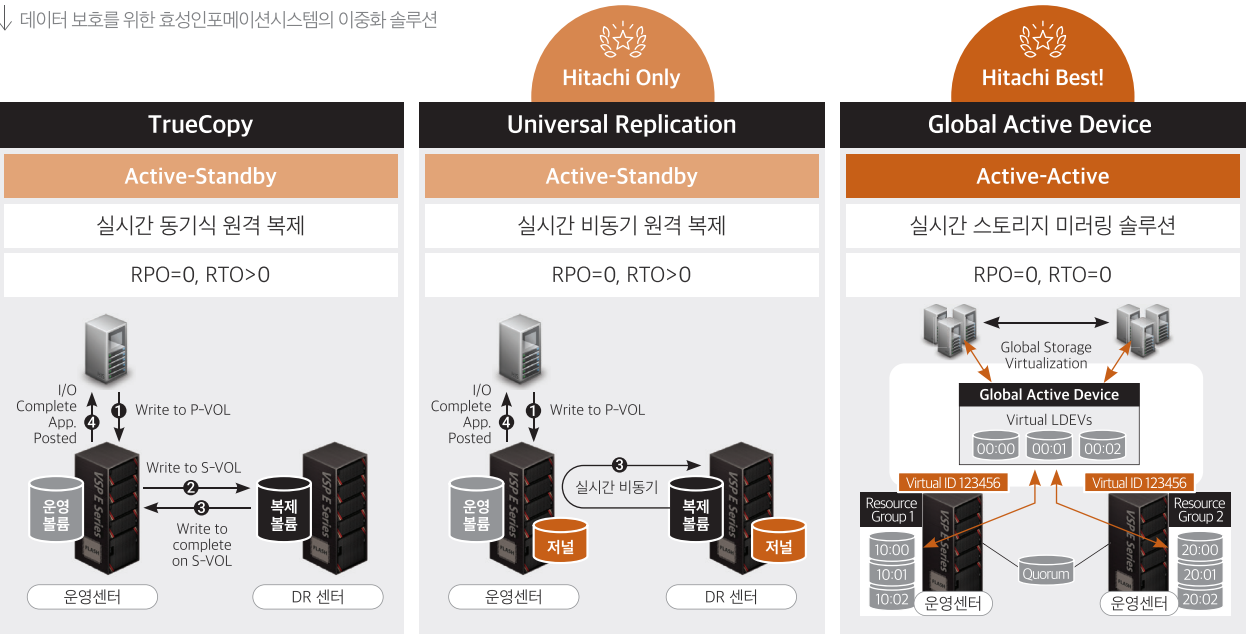
싱크(Active-Standby Sync)’라고도 하며, 하나의 서버와 별도로 분리된 2개의 스토리지를 이용해 데이터를 동기화하는 방식이다. 이 방식은 원본과 복제본이 동일한 내용으로 저장될 수 있도록, 서버에서 쓰기 요청이 들어오면 원본과 복제본에 쓰기를 완료한 후 서버에 완료 메시지를 전달한다. 이때 원본과 복제본의 쓰기가 완료될 때까지 서버 또는 애플리케이션은 대기해야 한다. 따라서 원본과 복제본의 거리에 따라 애플리케이션의 응답 시간이 길어질 수 있다. 데이터를 완벽하게 동기화하므로 데이터 복구 시간을 0으로 최소화할 수 있다는 장점이 있지만, 복제 스토리지가 대기 상태에 놓이게 돼 재해가 발생할 경우 시스템 복구 시간이 오래 걸릴 수 있다는 점이 단점으로 꼽힌다.

동기식 데이터 복제 방식의 단점을 보완한 방식이 비동기식인 ‘액티브-스탠바이 어싱크(Active-Standby Async)’다. 호스트에서 주 스토리지에 데이터를 저장하는 것과 별개로 백그라운드에서 데이터를 원격지로 복제하기 때문에, 애플리케이션의 IO 처리 응답 시간을 위해 복제 스토리지의 응답까지 기다릴 필요가 없다. 스토리지 간 거리가 서비스에 영향을 주지 않기 때문에 수백 또는 수천 km까지도 구성이 가능하다. 스토리지 벤더들은 데이터 정합성을 맞추기 위해 수초 또는 수분 단위의 간격을 두고 ‘배치 복제’하는 방식과 ‘실시간 전송’의 두 가지 방식을 사용한다. 배치 방식의 경우, 기본적으로 수초에서 10분 이상까지 데이터 차이가 발생할 수 있기 때문에 비동기식을 택할 경우에는 실시간 전송 방식으로 구성하는 게 좋다.

액티브-액티브 미러링(Active-Active Mirroring)은 완벽한 데이터 동기화 요건을 충족하면서 스토리지 단일 박스 장애에서의 시스템 복구 시간을 제로로 달성해 무중단 환경을 구성하는 방식이다. 데이터를 복제하는 방식은 동기식과 동일하지만, 원본과 복제본을 구분하지 않고, 어떤 볼륨이든 동시에 읽기와 쓰기를 모두 지원한다. 복제 역시 양방향으로 동기화해 2개의 볼륨을 마치 하나의 볼륨으로 인식해 이중화로 구성하므로, 하나의 스토리지에 장애가 발생하더라도 다른 스토리지를 통해 서비스를 지속할 수 있다. 이 방식은 가장 이상적인 재해복구 모델로 알려진 ‘액티브-액티브 데이터센터(Active-Active Datacenter)’의 기본 기술로 자리 잡고 있다.

효성인포메이션시스템의 스토리지 이중화 솔루션은 세 가지 방식을 모두 지원한다. 동기화 방식의 트루카피(TrueCopy), 비동기 방식의 유니버설 리플리케이터(Universal Replicator), 액티브-액티브 미러링 방식을 지원하는 GAD(Global Active Device) 등 데이터센터 인프라 환경에 최적화된 이중화 솔루션으로, 미션 크리티컬한 애플리케이션을 위해 뛰어난 확장성과 데이터 무결성을 보장한다.

↓ 데이터 보호를 위한 효성인포메이션시스템의 이중화 솔루션



차세대 DR 기술 ‘GAD’ 적용된 액티브-액티브 데이터센터 모델

다양한 재해복구 솔루션 중에서 가장 이상적인 모델은 실시간 액티브-액티브 데이터센터 모델이라고 할 수 있다. 효성인포메이션시스템은 GAD 솔루션을 통해 완벽한 스토리지 액티브-액티브 구성을 제공한다. GAD는 다른 2개의 스토리지를 하나의 볼륨처럼 관리하는 미러링 기법을 활용해, 장애가 발생하더라도 서비스 무중단을 보장하는 차세대 DR 기술이다.

그런데 스토리지에서 데이터 미러링을 구현했다고 해서 액티브-액티브 데이터센터 환경을 구축했다고 볼 수는 없다. 모든 인프라 레이어에 대한 기술 검토와 적용이 필요하다. 서버 자원의 경우, 센터 구간에서의 스트레치드 클러스터링 기술과 WEB, WAS 부하 분산, 그리고 가상 서버의 모빌리티를 위한 방안을 포함해야 한다. 또 가장 중요한 부분인 네트워크는 센터 간 트래픽에 대한 분산과 라우팅을 위한 네트워크 확장, 방화벽 정책 등 데이터를 어떻게 효과적으로 빠르게 전송할 것인지 고민해야 한다.

최근 CPU, 메모리 등의 파워가 향상되고 전송 알고리즘이 발달하면서 네트워크 하드웨어와 소프트웨어 역시 데이터 전송 성능이 비약적으로 높아졌지만, 빛을 이용해 데이터를 전송하는 광통신 방식이기 때문에 빛의 속도에 제약받을 수밖에 없다. 이에 복제 거리에 따라 전송 시간이 증가하므로, 현재 수십 km의 장거리 복제를 위해서는 비동기 방식을 채택하는 경우가 많다.

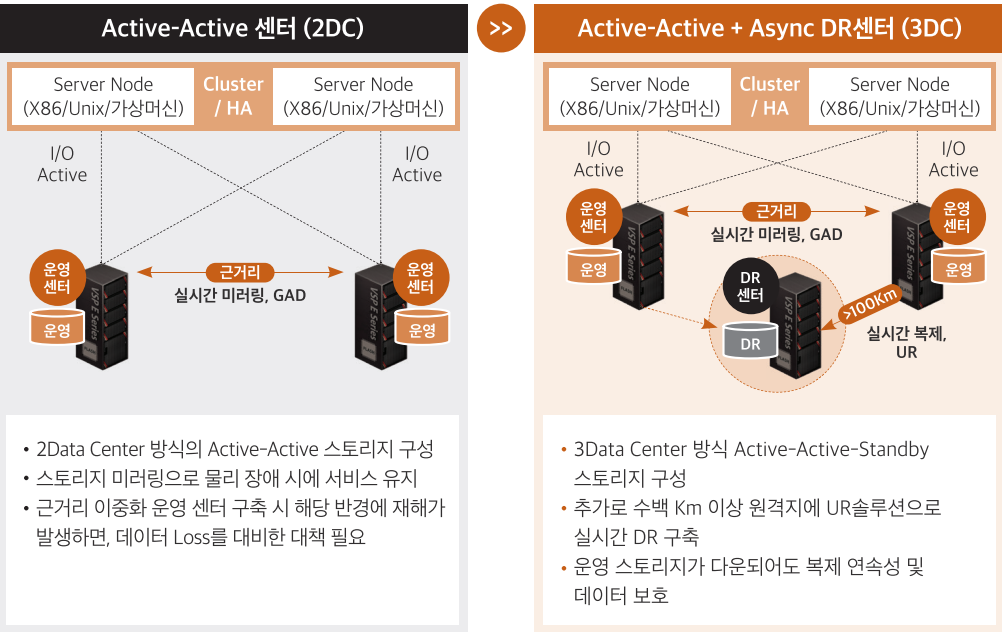
원격지 센터 간에 GAD 솔루션으로 이중화한 스토리지와 스트레치드 클러스터로 서버를 구성하면, 최대 수백 km 떨어진 데이터센터 간에도 액티브-액티브 구성이 가능하다. 하지만 네트워크 광통신의 한계를 보완하기 위해 수 km 이내의 근거리에 액티브-액티브 데이터센터를 구축하는 사례가 많다. 이 경우에는 단일 지점의 데이터센터 재해는 충분히 대응할 수 있지만, 특정 지역에 걸쳐 홍수, 지진 등의 재해가 발생하면 두 곳의 센터가 모두 안전하지 않을 수도 있다.

최악의 상황에 대비한 시나리오, ‘3데이터센터 복제 방식’

이러한 최악의 상황에 대비할 수 있는 방안으로 ‘3데이터센터 복제 방식’이 사용되고 있다. 수백 km 이상의 거리에 비동기 데이터 동기화 솔루션을 적용한 액티브-액티브 어싱크 복제 방식을 구축해 데이터센터를 삼중화하는 것이다. 이렇게 하면 거리에 제약받지 않고 광범위한 지역의 재해에도 완벽하게 대비할 수 있다.

데이터를 삼중화하는 ‘3데이터센터 복제 방식’을 구현하면, 정전, 침수, 화재 등의 재해뿐 아니라 전쟁과 같은 상황에서도 데이터를 온전하게 동기화해 보호할 수 있다. 근거리에 있는 2개의 데이터센터 중 하나가 다운되면, 근거리 백업센터에서 장거리에 위치한 제3의 데이터센터로 데이터 복제 서비스를 유지하기 때문이다.

↓ GAD 환경에서의 재해복구 구성



3데이터센터 구성 방식은 과거부터 사용되어 왔다. 그러나 동기식과 비동기식을 결합해 사용하던 과거와 달리, 최근에는 액티브-액티브 미러링 방식과 비동기 방식을 혼합해 구성하는 사례가 늘고 있다.

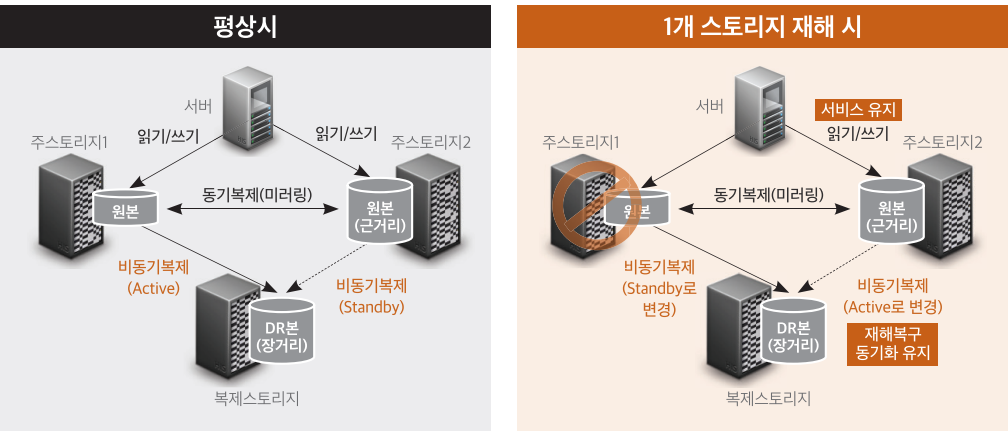
3데이터센터 기술의 핵심은 삼중화된 동기화 복제본 중, 하나의 스토리지 볼륨에서 장애가 발생하더라도 나머지 스토리지에서 서비스를 얼마나 빠르게 정상화할 수 있느냐에 달려 있다. 평상시에는 주 볼륨 3개가 액티브-액티브 미러링으로 이중화된 상태로 있으면서 주 볼륨 중에서 주 스토리지 1번을 통해 원거리의 복제 스토리지와 동기화한다. 이때 주 스토리지 2번과 복제 스토리지는 복제 회선을 스탠바이로 구성해 놓기 때문에, 주 스토리지 1번에 장애가 발생할 경우 주 스토리지 2번이 서비스를 계속 이어갈 수 있다. 동기화 복제의 경우는 주 스토리지 2번이 서비스를 이어받아 동기화를 지속한다. 즉, 단일 스토리지에 장애가 발생할 경우 나머지 스토리지 2대를 통해 데이터 동기화가 중단 없이 이루어지므로 매우 강력한 재해복구 대비 방안이라고 할 수 있다.

물론 액티브-액티브 미러링 방식과 비동기 방식을 혼합해 구성할 경우에도 주 데이터센터와 근거리 데이터센터에 동시에 장애가 발생하면 문제가 생길 수 있다. 이 경우, 장거리 데이터센터에서 서비스를 지속해야 하는데, 일반적으로는 데이터베이스의 재기동 또는 최신 데이터 동기화 시간 등으로 인해 서비스 전환 시간이 과다하게 소요된다. 실시간 비동기 방식으로 데이터를 동기화하면 데이터 복구 시간을 최소화할 수 있지만, 시스템 복구 시간은 여전히 문제가 된다.

이와 같은 상황을 극복하려면 재해복구 시스템을 설계할 때 업무의 특성을 먼저 고려해야 한다. 데이터 변경이 적은 편인 애플리케이션 엔진, 웹서비스, WAS 서비스 등을 실시간 비동기 방식으로 구성하면 데이터를 충분히 동기화해 서비스를 신속하게 기동시킬 수 있다. 데이터 변경이 많은 편인 데이터베이스는 시스템 복구 시 가장 많은 시간이 필요하다. 데이터베이스를 오픈 상태에 두고 로그 시핑(Log shipping)¹⁾해 데이터를 동기화하면 시스템 복구 시간을 획기적으로 단축할 수 있다. 이러한 구성을 통해 주 데이터센터와 근거리 데이터센터 두 곳 모두에서 장애가 발생하더라도 재해복구 시간을 최소화할 수 있게 된다.

1) 로그 시핑(Log shipping): 주 스토리지의 트랜잭션 로그 조각을 규칙적으로 복제 스토리지로 옮겨, 주 스토리지에 장애가 발생할 경우 복제 스토리지를 운영해 가용성을 향상하는 기능

↓ 3데이터센터 복제 중 재해 시나리오



고객 특성 반영한 재해복구 시스템 활용 방안

제조기업의 경우 공장 가동률이 무엇보다 중요하기 때문에 무중단 서비스가 꼭 필요하다. 대규모 제조기업인 A사는 수년 전 태풍으로 인한 침수로 전산실 건물에 전원 공급이 중단되면서 서비스가 중단된 경험이 있다. 이후 약 2km 남짓 떨어져 있는 다른 건물에 이중화된 시스템을 구축하는 ‘액티브-액티브 2데이터센터’ 방식을 채택해 문제를 해결했다. 현재는 효성인포메이션시스템의 GAD를 이용해 스토리지 데이터 미러링 방식을 구현, 특정 건물에 장애가 발생하더라도 서비스에는 아무런 지장이 없도록 했다. 두 건물이 동일한 네트워크로 묶여 있기 때문에 DB 클러스터링은 일반적인 구성 방식을 도입했다. 그 외 대부분의 애플리케이션은 이중화 액티브 또는 스탠바이 서버를 분리된 건물에 구성해 재해 발생 시 테이크 오버(take-over)를 받아 서비스하고 있다.

제1금융권의 경우 중요한 시스템을 삼중화하는 구성 방식을 선호한다. 금융기업인 B사는 VM웨어 기반의 프라이빗 클라우드로 3데이터센터를 구축했다. 스토리지는 효성인포메이션시스템의 액티브-액티브 미러링 솔루션인 GAD를 이용해 이중화하고, 추가로 약 30km 정도 떨어진 위치의 데이터센터에는 스탠바이 동기화 방식을 운영 중이다. 2개의 사이트는 액티브-액티브 상태로 유지되고 있으며, VM웨어의 사이트 리커버리 매니저(SRM)를 통해 유사시에 원격지의 데이터센터에서 자동으로 복구 시스템을 운영할 수 있도록 구성했다.

완벽한 IT 시스템이라도 장애 및 재해로 인한 서비스 중단 위험이 늘 존재하기 때문에 재해복구를 위한 체계적 프로세스 정립이 반드시 필요하다. 효성인포메이션시스템은 다양한 재해복구 솔루션 포트폴리오를 기반으로 검증된 제품과 20년 이상의 구축 경험을 보유하고 있다.