

Chapter 7. Azure 스토리지

발표자 | 신도인

2023.12.27



Azure 스토리지는 다양한 유형의 데이터를 다룰 수 있는 저장소입니다

- 데이터의 대표적인 3가지 유형인 **정형/비정형/반정형**의 데이터를 다룰 수 있습니다.
- 주요 프로그래밍 언어와 연동이 쉽고, Azure 셸 환경에서 **스크립트로 관리**할 수 있습니다.

정형 / 비정형 / 반정형 데이터 예시

유형	설명	예시
정형	행과 열에 의해 데이터의 속성이 구분되는 유형	스프레드시트, RDBMS 테이블
비정형	구조화되어 있지 않은 개체로 존재하는 유형	문서, 동영상, 이미지, 음원
반정형	메타데이터와 데이터가 함께 제공되는 유형	JSON, XML, HTML, YAML

Azure 스토리지에서 사용 가능한 프로그래밍 언어

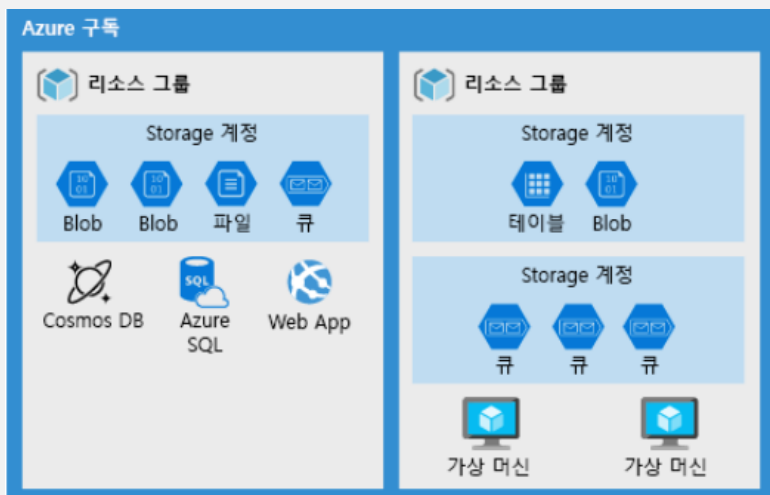


Storage Account

Azure 스토리지는 스토리지 계정을 통해 관리합니다

- 하나의 스토리지 계정으로 여러 개의 Azure 스토리지 서비스를 사용할 수 있도록 인증 및 권한을 부여합니다.
- 예기치 못한 이벤트에 대해 데이터를 보호하는 중복 옵션도 스토리지 계정별로 관리됩니다.

스토리지 계정 다이어그램



사용 가능한 중복(복제) 옵션 종류

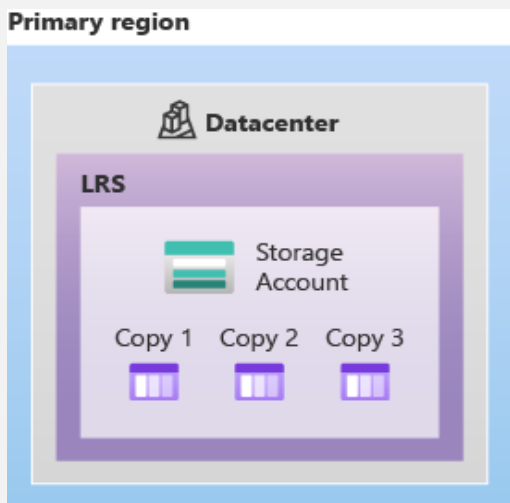
주 지역	LRS (로컬 중복 저장소)
	ZRS (영역 중복 저장소)
보조 지역	GRS (지역 중복 저장소)
	GZRS (지역 영역 중복 저장소)
	RA-GRS (보조 지역에서 읽기 권한이 주어진 지역 중복 저장소)
	RA-GZRS (보조 지역에서 읽기 권한이 주어진 지역 영역 중복 저장소)

Redundancy Option

LRS는 주 지역의 데이터 센터에 3개의 복제본을 생성하는 옵션입니다

- LRS(Locally Redundant Storage)는 소프트웨어 오류로부터 데이터를 보호하지만, 하드웨어에 발생한 재해에 대해서는 보호할 수 없습니다.
- 따라서 가장 저렴한 옵션이지만, 고가용성이나 높은 내구성이 필요한 경우엔 권장하지 않습니다.

LRS 다이어그램



상황별 LRS 적용 예시

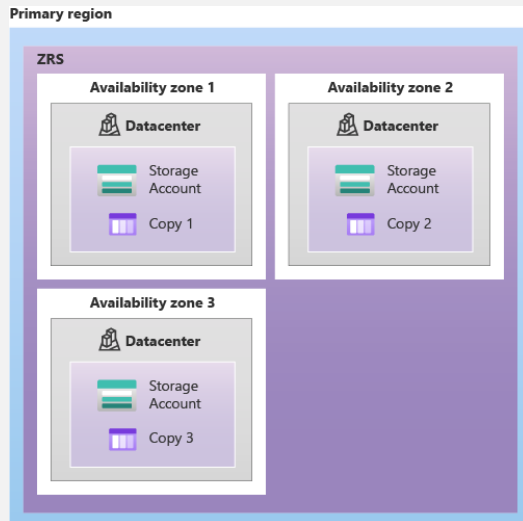
- 데이터 손실이 발생하더라도 쉽게 복구할 수 있는 데이터를 저장하는 경우
- 데이터의 특성 상 특정한 국가 또는 지역 내에서만 사용이 제한되는 경우
- 사용자가 직접 디스크의 복제 및 가용성을 설정하는 Azure 비관리 디스크를 사용하는 경우

Redundancy Option

ZRS는 주 지역의 가용성 영역에 3개의 복제본을 생성하는 옵션입니다

- ZSR(Zone Redundant Storage)는 한 가용성 영역에 물리적인 피해가 있더라도 다른 가용성 영역에서 작업이 가능합니다.
- 일시적으로 접근이 불가능 할 때 접근 요청이 들어오면, 지수 백오프와 같은 기법을 사용하여 동기적으로 요청을 처리 및 반환합니다.

ZRS 다이어그램



상황별 ZRS 적용 예시

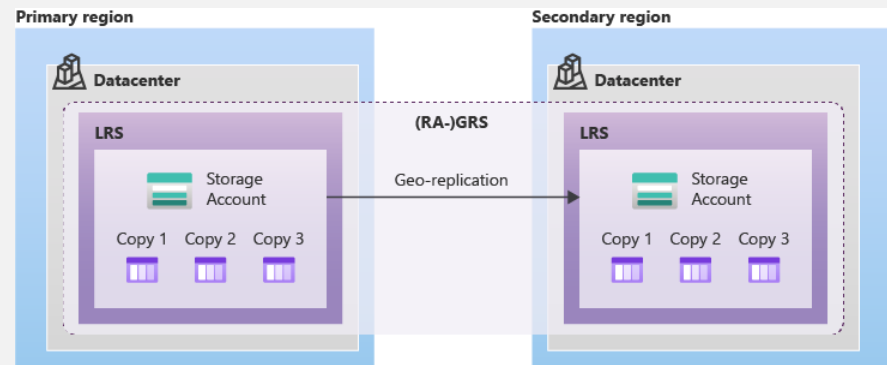
- 고가용성 및 높은 내구성이 필요한 데이터를 저장하는 경우
- 데이터의 특성 상 특정한 국가 또는 지역 내에서만 사용이 제한되는 경우

Redundancy Option

GRS는 주 지역과 보조 지역에 3개의 복제본을 생성하는 옵션입니다

- GRS(Geo Redundant Storage)는 주 지역에는 LRS 방식으로 동기 복제하며 동시에 비동기로 보조 지역에 복제합니다.
- 이때, 보조 지역에 비동기로 복제가 완료되면 LRS 방식으로 보조 지역 내에서 동기 복제가 일어납니다.

GRS 다이어그램



상황별 GRS 적용 예시

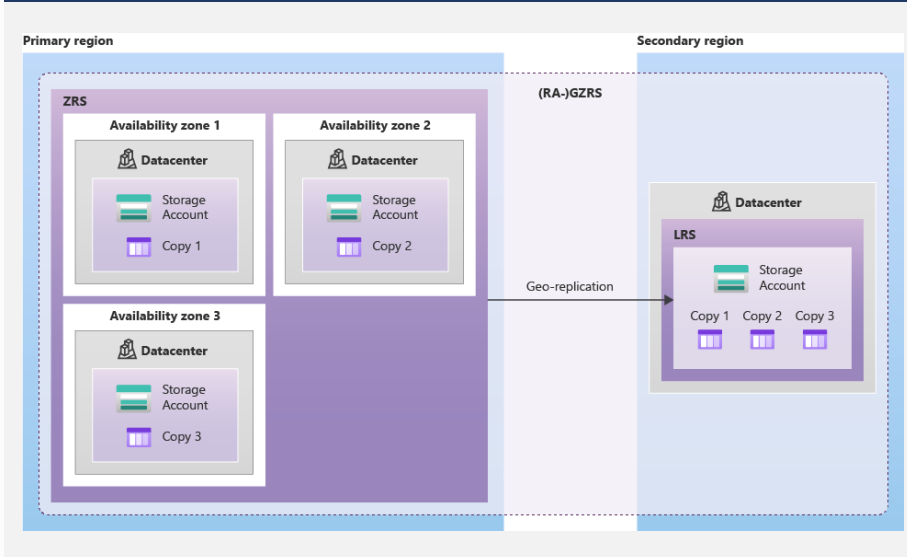
- 고가용성 및 높은 내구성이 필요한 데이터를 저장하는 경우
- 재해 중 전체 지역 중단 또는 복구 불가능한 주 지역 재해 발생에 대비하고 싶은 경우

Redundancy Option

GZRS는 가용성 영역과 보조 지역에 3개의 복제본을 생성하는 옵션입니다

- GZRS(Geo Zone Redundant Storage)는 주 지역의 가용성 영역에는 ZRS 방식으로 동기 복제하며 동시에 비동기로 보조 지역에 복제합니다.
- 이때, 보조 지역에 비동기로 복제가 완료되면 LRS 방식으로 보조 지역 내에서 동기 복제가 일어납니다.

GZRS 다이어그램



상황별 GZRS 적용 예시

- 최고 수준의 고가용성 및 내구성이 필요한 데이터 저장하는 경우
- 재해 중 전체 지역 중단 또는 복구 불가능한 주 지역 재해 발생에 대비하고 싶은 경우
- 가용성 영역 사용 불능 또는 복구 불능에 대비하고 싶은 경우

Practice

실전 연습 1 – 스토리지 계정 만들기

- Azure Portal 검색창에서 “스토리지 계정” 을 검색하여 선택
- 블레이드에서 “만들기” 버튼을 클릭하여 스토리지 계정 생성
- 스토리지 계정 이름은 고유하게 작명
- 중복 옵션은 “LRS(로컬 중복 스토리지)”로 설정
- 고급 탭에서 “개별 컨테이너에 대한 익명 액세스 허용” 체크
- 배포(=생성)가 완료되면, 스토리지 계정 리소스로 이동하여 서비스 메뉴의 “보안 + 네트워킹 섹션 아래의 액세스키 메뉴” 선택하여 확인

Azure Storage Service

Azure 스토리지 서비스는 HTTP/HTTPS로 어디서나 접근 가능합니다

- Azure 스토리지 서비스는 컨테이너(Blob), 파일 공유, 테이블, 큐 서비스가 있으며 각 서비스마다 고유한 엔드포인트를 가집니다.
- 따라서 각 서비스에 HTTP/HTTPS를 통해 어디서나 접근이 가능합니다.

Azure 스토리지 서비스 종류



서비스별 엔드포인트 예시

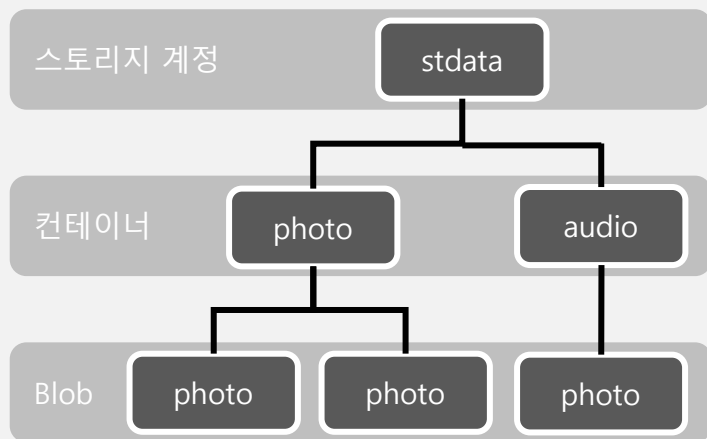
- Blob 서비스(컨테이너)
`http://<스토리지_계정_이름>.blob.core.windows.net/`
- 파일 서비스
`http://<스토리지_계정_이름>.file.core.windows.net/`
- 큐 서비스
`http://<스토리지_계정_이름>.queue.core.windows.net/`
- 테이블 서비스
`http://<스토리지_계정_이름>.table.core.windows.net/`

Azure Storage Service

Blob은 구조화되지 않은 데이터를 대규모로 저장하는 서비스입니다

- 컨테이너 서비스라고 부르기도 하며, Blob 데이터를 일정 기준으로 구분 및 관리하는 단위는 컨테이너입니다.
- 주로 분석하기 위한 데이터들을 1차적으로 적재하는 용도로 사용되며, Blob 버전 관리도 수행합니다.

컨테이너 서비스의 구성요소 관계



공용 액세스 수준

- 익명 액세스 없음: 기본값, 요청 시 권한이 필요함
- 익명 읽기 전용 액세스: 인증 절차 없이 데이터 읽기 가능, 컨테이너 내의 Blob 목록은 확인 불가
- 익명 읽기 액세스: 인증 절차 없이 데이터 및 목록 읽기 가능, 권한 설정 및 메타데이터 확인 불가

Azure Storage Service

Blob은 데이터에 대한 여러 액세스 계층을 지원합니다

- Blob은 Hot, Cool, Cold, Archive 액세스 계층을 지원합니다.
 - Hot: 개체를 자주 읽고 쓰는 경우, 최소 보관 기간 없음
 - Cool: 자주 액세스 하지 않는 대량의 데이터, 최소 보관 기간 30일
 - Cold: 자주 액세스 하지 않는 대량의 데이터, 최소 보관 기간 90일
 - Archive: 데이터 액세스에 대기 시간이 소요되어도 상관없는 경우, 최소 보관 기간 180일
- 언제든지 Blob 액세스 계층을 변경 및 적용할 수 있습니다.

Blob은 컨테이너를 비동기식으로 복제합니다

- 개체 복제 시 원본 계정의 Blob에 대한 스냅샷은 대상 계정에 복제되지 않습니다.
- 또한, Archive 계층에 있을 때는 복제가 불가능하며, 다른 계층 간 복제는 가능합니다.

Blob 수명 주기는 계층 전환으로 관리합니다

- 주로 Blob을 Cool 스토리지 계층으로 전환하여 성능 및 비용을 최적화합니다.
- Hot 계층에서 Cool 계층으로 전환 / Hot 계층에서 Archive 계층으로의 전환은 즉시 적용됩니다.

Practice

실전 연습 2 – 컨테이너 스토리지 만들기

- 앞서 생성한 스토리지 계정의 개요 블레이드에서 속성 탭 아래의 “[Blob service](#)” 선택
- 컨테이너 블레이드에서 “[만들기](#)” 버튼을 클릭하여 새 컨테이너 생성
- 공용 액세스 수준은 기본값인 “[프라이빗\(익명 액세스 없음\)](#)”으로 설정
- 생성이 완료되면, 개요 탭의 “[URL을 복사](#)”해서 브라우저 주소 창에 입력하여 접속
- 접속이 안되는 경우는 컨테이너 서비스 메뉴의 개요를 선택하고 액세스 수준 변경 팝업에서 “[Blob\(Blob에 대한 익명 읽기 전용 액세스\)](#)”로 변경

파일 공유는 SMB 및 NFS 프로토콜로 공유 스토리지를 제공하는 서비스입니다

- 이전에 살펴봤던 Blob과 달리 디렉토리 구조를 가지기 때문에 리소스 간 파일 공유를 편리하게 할 수 있습니다.
- SAS(Shared Access Signature) 토큰을 포함하는 URL을 제공하기 때문에 어디서나 안전하게 파일을 공유할 수 있습니다.

파일 공유 서비스 사용 시나리오

- 온프레미스 파일 서버 또는 NAS 디바이스를 대체하거나 보완하려는 경우
- 많은 가상 머신에서 액세스해야 하는 개발 및 디버깅 도구를 저장하려는 경우
- Azure 파일 동기화를 통해 파일 공유 서비스에 저장된 데이터를 다른 서버에 복제하는 경우

파일 공유 서비스에서의 파일 동기화

- 저대역폭에서도 파일 동기화를 효율적으로 처리 가능
- 효율적인 동기화를 위해 클라우드 계층화도 활용함
- 파일 동기화 에이전트를 통해 로컬에 등록된 공유 파일을 Azure 파일 공유 서비스에 동기화 가능

Practice

실전 연습 3 - 파일 공유 만들고 연결하기

테이블은 키-값 쌍으로 구조화된 반정형 데이터용 저장소입니다

- 테이블에 적재된 데이터들은 NoSQL 처럼 쿼리문을 통해 검색 및 추출할 수 있습니다.
- 데이터의 속성의 집합이 늘어나더라도 키-값 형태로 유연하게 적용이 가능합니다.

실전 연습 4 – 테이블 스토리지 만들고 샘플 데이터 입력하기

Azure Storage Service

큐는 메시지를 통해 비동기 작업의 안정성을 높여주는 서비스입니다

- 큐는 다양한 리소스 간 데이터를 안전하게 전달하고 비동기적으로 처리할 수 있는 중간 매체로 활용됩니다.
- 큐가 중간 매체로써 동작하기 때문에 연결된 리소스들이 서로 독립적으로 동작하며 서비스의 확장성을 높여줍니다.

실전 연습 5 – 큐 스토리지 만들고 메시지 추가하기

실전 연습 6 – Azure Storage Explorer 사용하기