Working with Data Backup in Azure

Daniel Krzyczkowski
MICROSOFT MVP & SOFTWARE DEVELOPER
@DKrzyczkowski www.techmindfactory.com

Module Overview

Understanding of how Azure SQL, Cosmos DB and Storage Account store the data

Concepts of replicating data stored in the popular Azure PaaS data stores: Cosmos DB, Azure SQL and Storage Account

Enable geo-redundancy and multi-region writes in the Azure Cosmos DB

Azure PaaS Data Stores: Cosmos DB, Azure SQL, and Storage Account

Azure PaaS Data Stores

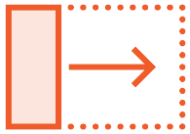
Azure SQL Database

Azure Cosmos DB

Azure Storage Account

Data Geo-replication in the Azure SQL

Azure SQL Database

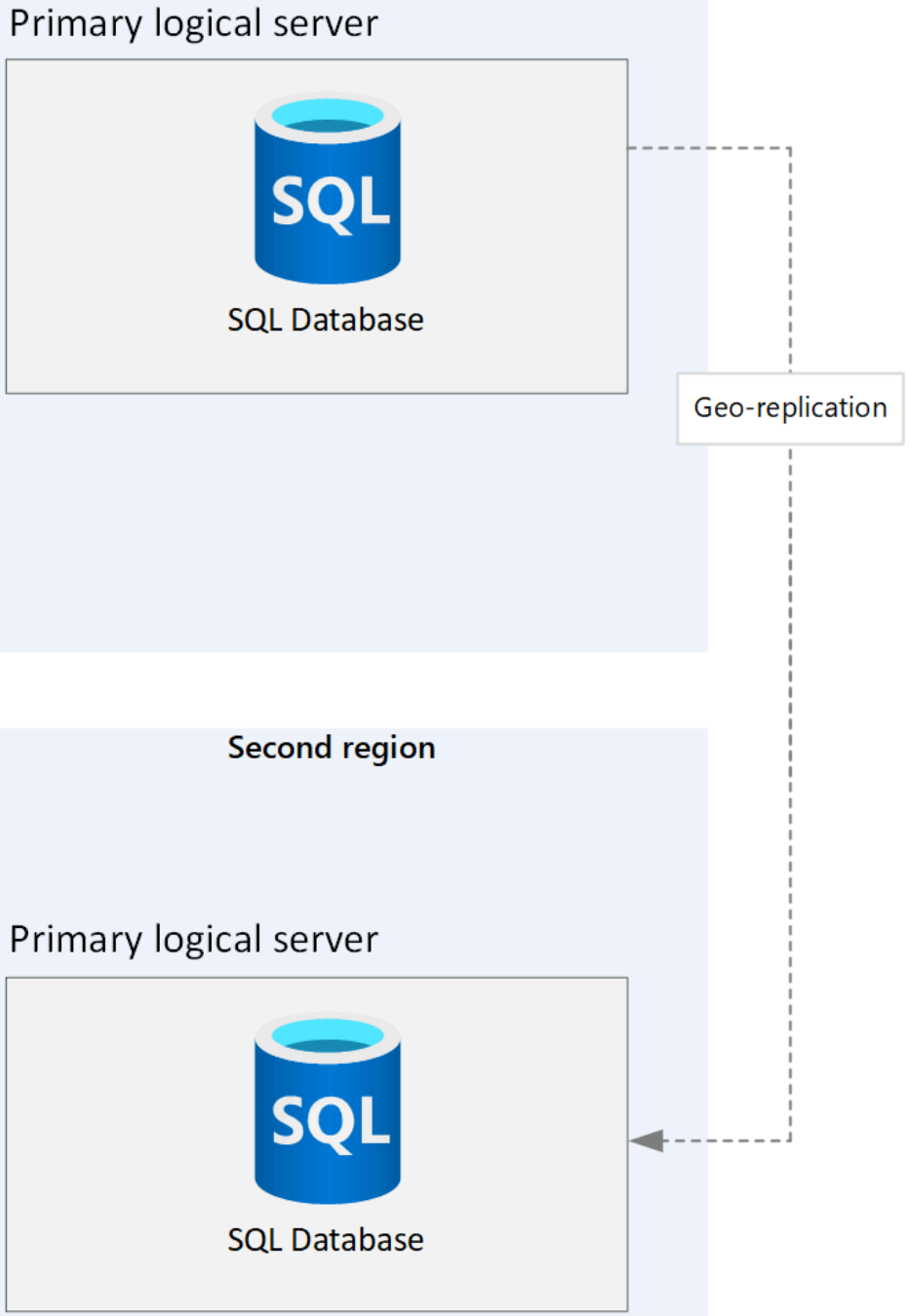


Automatically creates full database backups weekly, differential database backups every 12 hours, and transaction log backups every 5 - 10 minutes

Active geo-replication for SQL Database, which automatically replicates database changes to secondary databases in the same or different Azure region

Manual approach for backup and restore: Azure SQL Database Import/Export Service, which supports exporting databases to BACPAC files

Active Geo-replication



Allows creating readable secondary databases of individual databases on a SQL Database server in the same or different Azure region

If geo-replication is enabled, the application can initiate failover to a secondary database in a different Azure region

SQL Database Business Continuity Features

Temporal tables to restore row versions from any point in time

Built-in automated backups and Point in Time Restore to restore complete database to some point in time

Active geo-replication enables creating readable replicas and manually failover to any replica in case of a data center outage or application upgrade

Auto-failover group allows the application to automatically recover in case of a data center outage

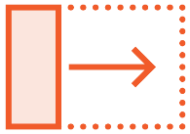
Long-term backup retention enables to keep the backups up to 10 years

Compare Geo-replication with Failover Groups

Geo-replication	Failover groups
No automatic failover	Automatic failover
No fail over multiple databases simultaneously	Fail over multiple databases simultaneously
Connection string update required after failover	No connection string update required after failover

Data Geo-replication in the Azure Cosmos DB

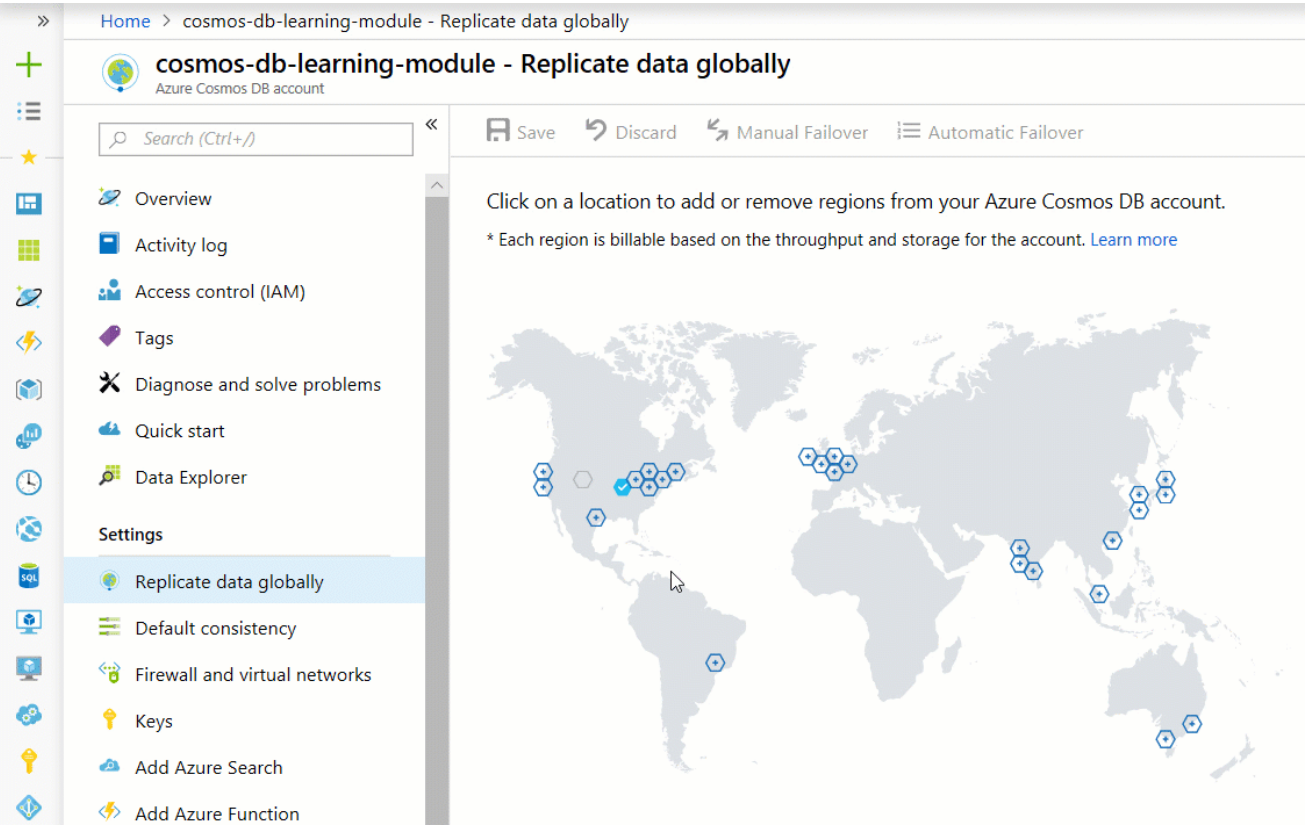
Azure Cosmos DB



Backups are stored separately in another storage service and are replicated globally to protect against regional disasters

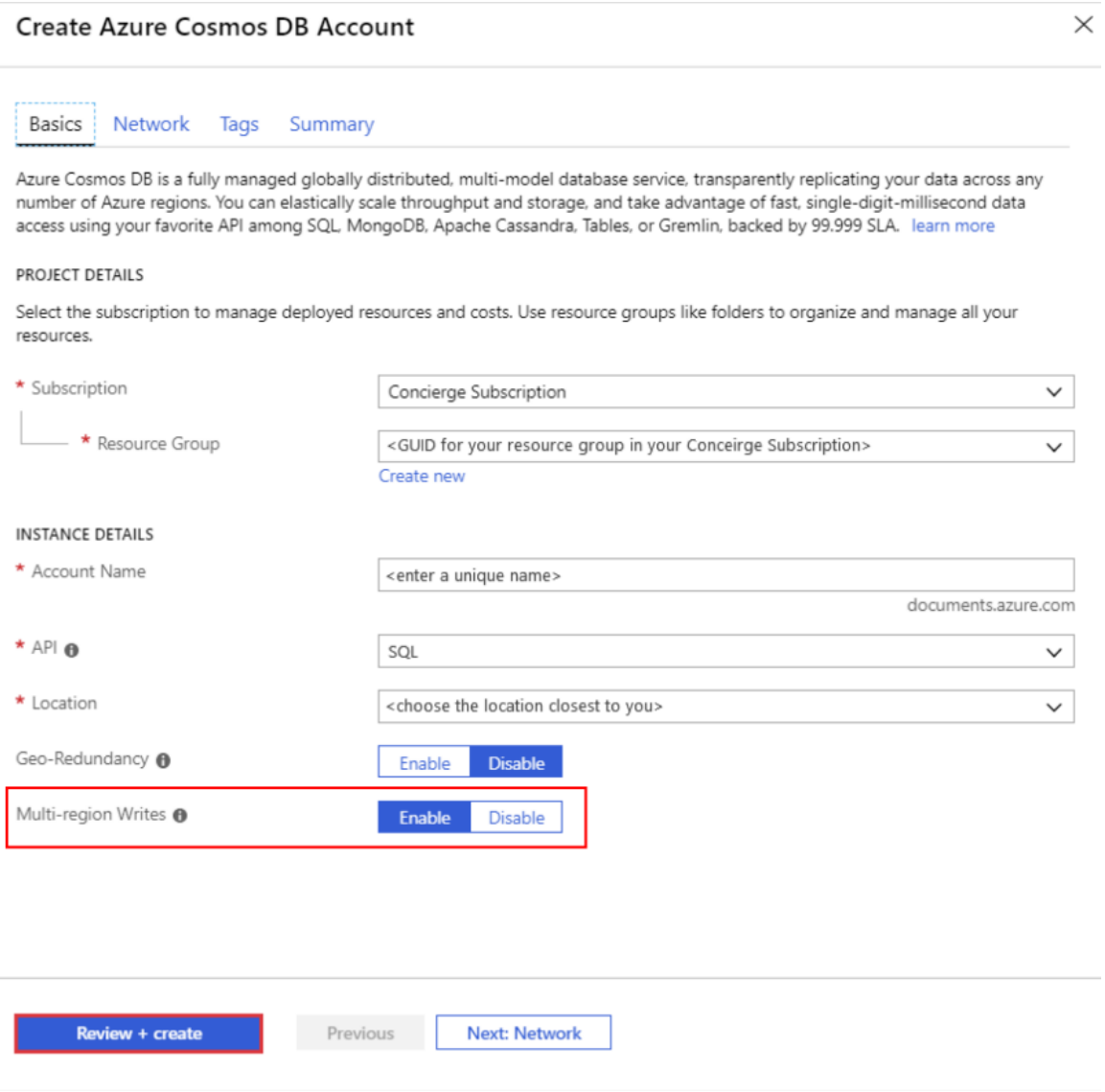
Automatically takes a backup of a database every 4 hours and at any point of time, only the latest 2 backups are stored

When the container or database is deleted, Azure Cosmos DB retains the existing snapshots of a given container or database for 30 days



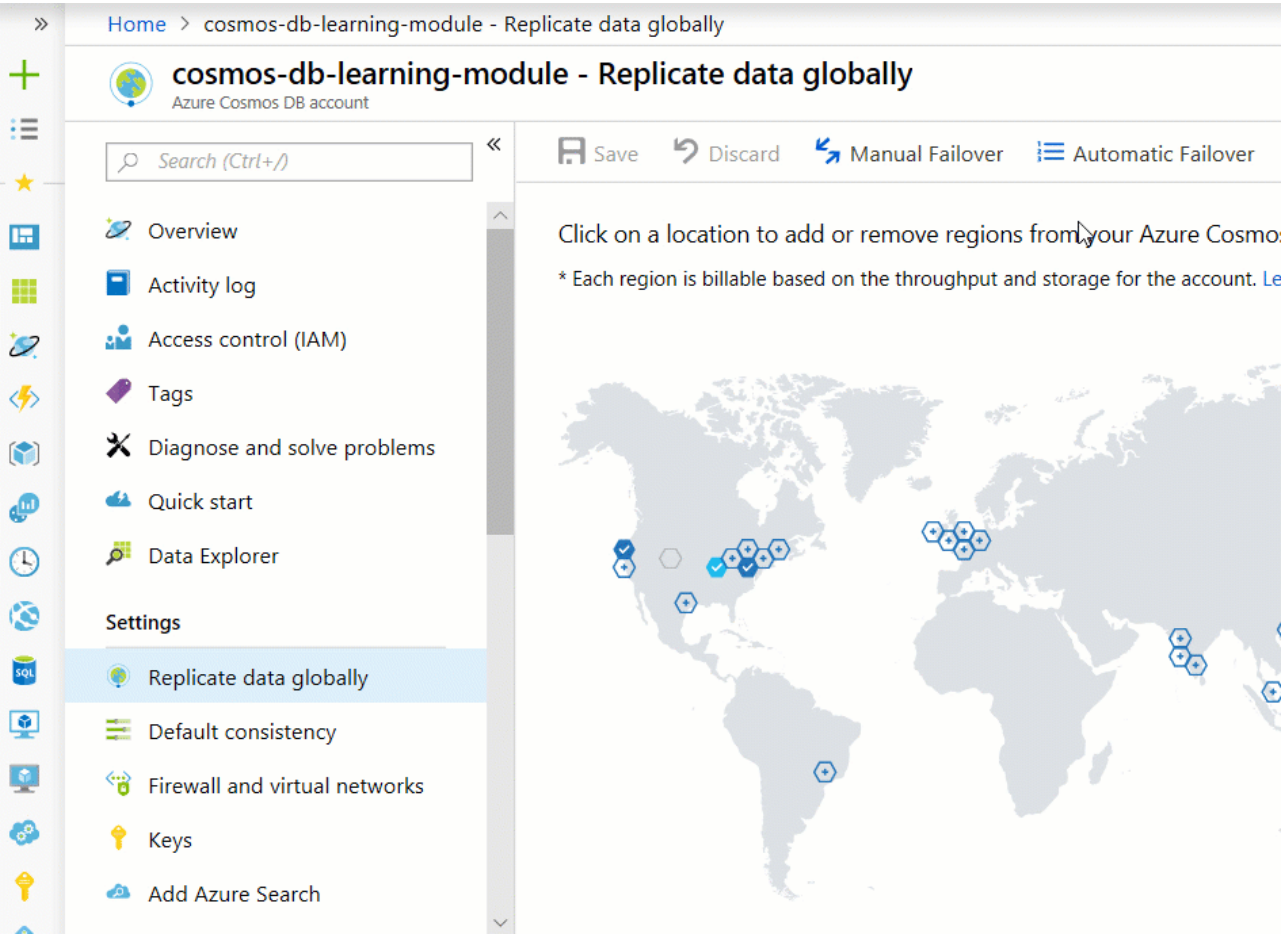
Global Distribution Basics

Azure Cosmos DB ensures that when additional region is added, the data is **available for operations within 30 minutes**, assuming data size is 100 TBs or less



Multi-master Support

When the account is replicated in multiple regions, each region is a master region that equally participates in a write-anywhere model

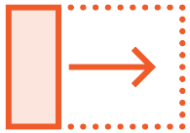


Read Region Priorities

Read regions can be prioritized by drag and drop with **Automatic Failover** option

Data Geo-replication in the Azure Storage Account

Azure Storage Account



Data in the Azure Storage Account is always replicated to ensure durability and high availability

Azure automatically stores Azure Storage data three times within different fault domains in the same Azure region

When geo-replication is enabled, the data is stored three additional times in a different region

Locally Redundant Storage (LRS)

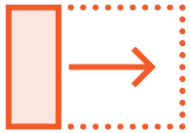
Replicates the data three times within a single data center

When datacenter-level disaster occurs, all replicas in a storage account using LRS may be lost or unrecoverable

Recommended approach:

- Use Zone-Redundant Storage (ZRS)
- Use Geo-Redundant Storage (GRS)

Zone-Redundant Storage (ZRS)

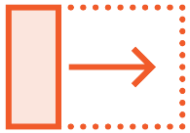


Replicates the data synchronously across three storage clusters in a single region

Each storage cluster is physically separated from the others and is located in its own availability zone

When data is stored in a storage account using ZRS replication, data access and management is still possible if one of availability zones becomes unavailable

Geo-Redundant Storage (GRS)



Replicates the data to a secondary region that is hundreds of miles away from the primary region

When storage account has GRS enabled, data is durable even in the case of a complete regional outage

Data is replicated to the secondary region asynchronously, meaning there is a delay between when data written to the primary region is written to the secondary region

Important!

An account failover usually involves some data loss. It is important to understand the implications of initiating an account failover

Demo

Enable geo-redundancy and multi-region writes in the Azure Cosmos DB

- Create Azure Cosmos DB instance
- Enable geo-redundancy
- Enable multiple writes

Summary

Azure PaaS Data Stores description: Cosmos DB, Azure SQL and Storage Account

Data geo-replication concepts for the Azure PaaS data stores: Cosmos DB, Azure SQL and Storage Account

Geo-redundancy and multi-region writes in the Azure Cosmos DB