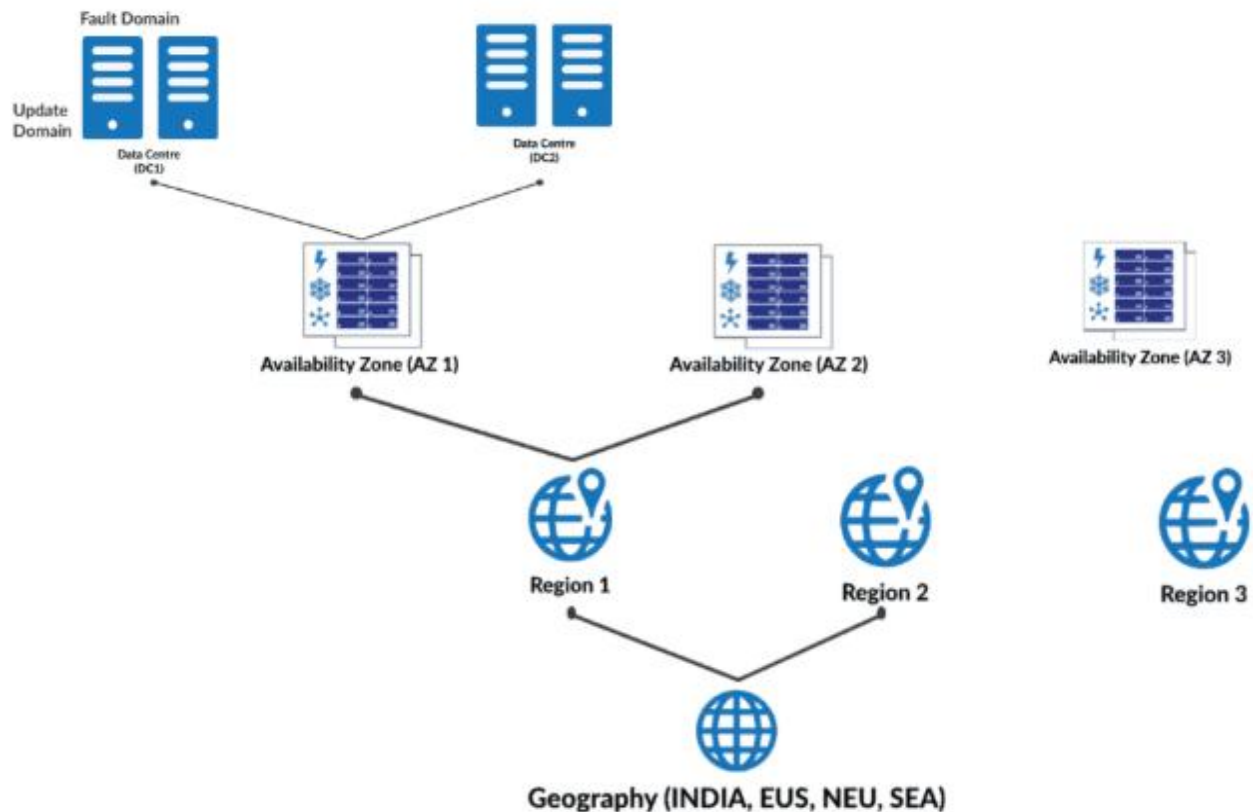


Azure Global Infrastructure

1. Introduction

Microsoft Azure's global infrastructure is designed to provide highly available, resilient, scalable, and secure cloud services across the world. It is composed of a strategic hierarchy: Geographies, Regions, Availability Zones, and Data Centers. Understanding this architecture is essential for building enterprise-grade cloud solutions. Azure Global Infrastructure is a massive network of data centres strategically situated across the world that are intended to offer high availability, reliability, and performance for cloud services. It features regions, availability zones, and fault domains, which allow for more robust and fault-tolerant installations.





2. Azure Geographies

Definition:

Azure Geographies are discrete markets that preserve data residency and compliance boundaries. They align with geopolitical boundaries and ensure that customer data stays within specific regional confines.

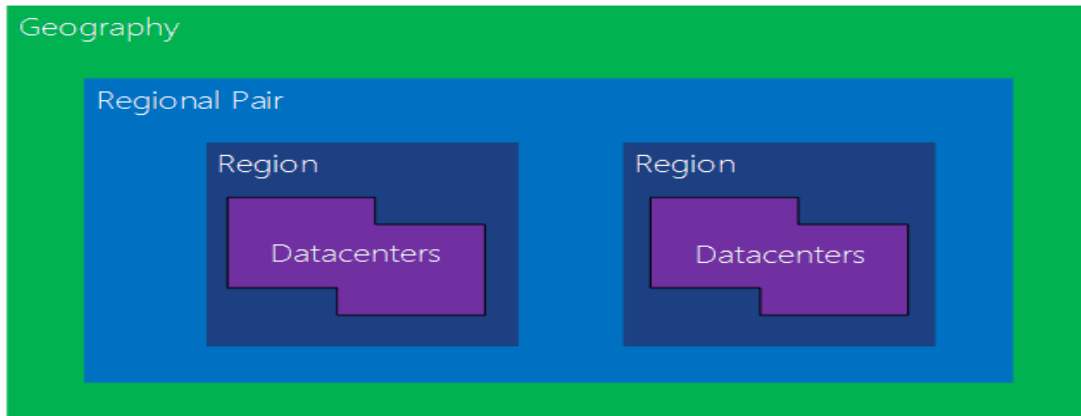
Purpose:

- Compliance with local regulations and data residency requirements.
- Disaster recovery and data sovereignty planning.
- Serving multinational corporations with region-specific deployment needs.

Examples of Azure Geographies:

- United States
- European Union
- United Kingdom
- Asia Pacific
- Middle East & Africa

Each geography typically contains two or more regions for high availability.



3. Azure Regions

Definition:

A region is a set of datacenters deployed within a specific geographic area, connected through a dedicated low-latency network.

Purpose:

- Allow customers to deploy applications close to their users.
- Enable redundancy and failover strategies.
- Support regulatory compliance and data residency.

Key Points:

- Over 60+ Azure regions worldwide.
- Each region is paired with another region in the same geography for regional pair failover (e.g., East US & West US).
- Regions offer Azure services based on availability and compliance certifications.

Types of Regions:

- Public regions – available to all customers.

- Sovereign regions – e.g., Azure Government, Azure China (operated under partnership with local government).
- Specialized regions – built for special use cases like military or government operations.

Geography	Regional Pair A	Regional Pair B
Canada	Canada Central	Canada East
China	China North	China East
India	Central India	South India
Japan	Japan East	Japan West
North America	East US	West US

4. Availability Zones (AZs)

Definition:

Availability Zones are physically separate locations within an Azure region, each with its own power, cooling, and networking.

Purpose:

- High availability and fault tolerance.
- Mitigate the impact of datacenter-level failures.
- Enable zone-redundant services like Azure Kubernetes Service, Azure SQL Database, etc.

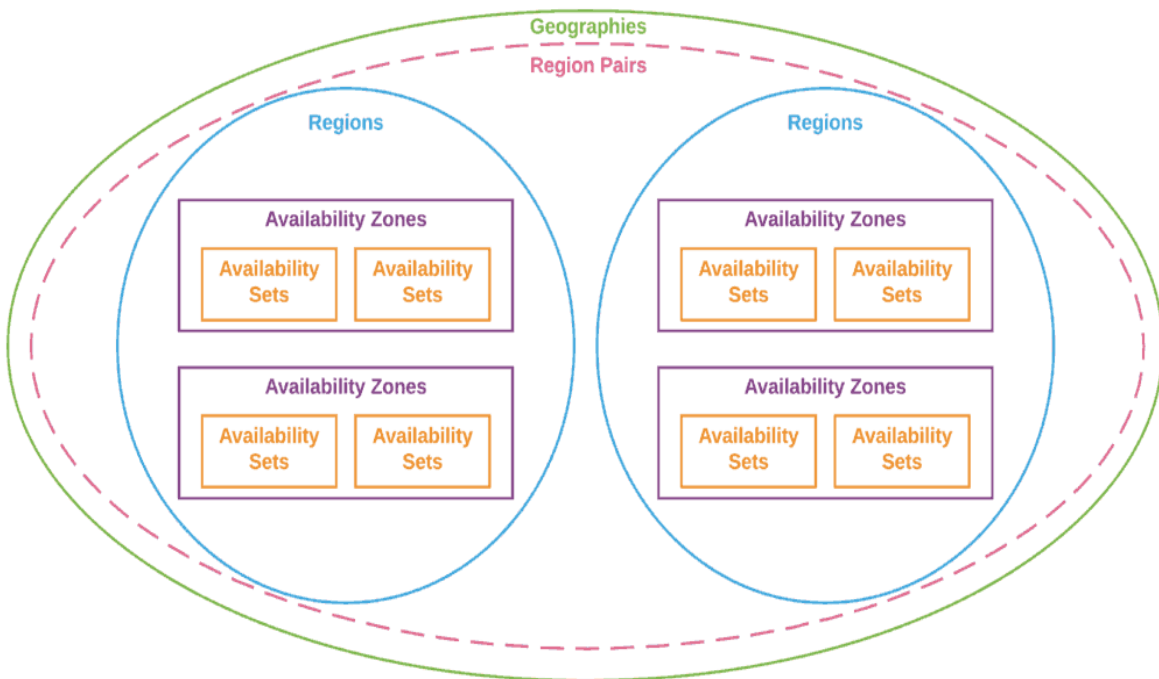
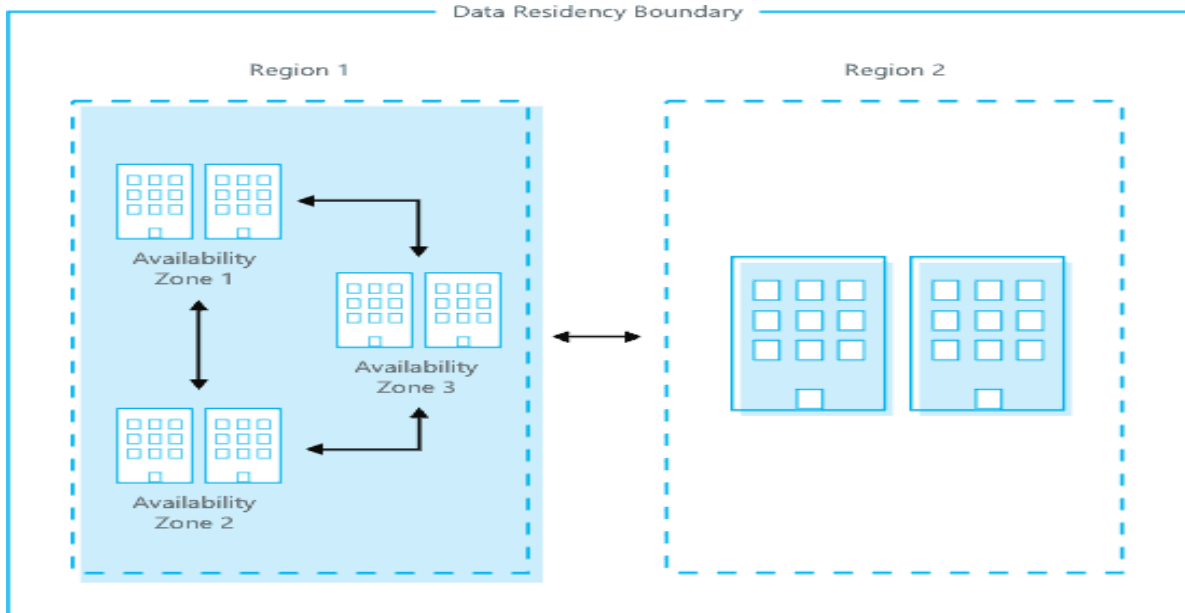
Architecture:

- Each region with AZs has a minimum of 3 separate zones.
- AZs are interconnected with high-bandwidth, low-latency fiber.

Usage Scenarios:

- Deploy virtual machines (VMs) across zones for resilience.
- Set up zone-redundant storage.

- Create multi-zone services to ensure 99.99% uptime SLA.



Azure regions with availability zones

The largest global presence of any cloud service, Azure is introducing new regions and availability zones quickly. Availability zones are presently supported in the following areas.

Americas	Europe	Middle East	Africa	Asia Pacific
Brazil	France Central	Qatar Central*	South Africa North	Australia East
Canada Central	Germany West Central	UAE North		Central India
Central US	North Europe			Japan East
East US	Norway East			Korea Central
East US 2	UK South			Southeast Asia
South Central US	West Europe			East Asia
US Gov Virginia	Sweden Central			China North 3
West US 2	Switzerland North			
West Us 3				

5. Azure Data Centers

Definition:

Data centers are physical facilities that house the computing, storage, and networking resources Azure provides. A distinct physical structure called an Azure data center houses hundreds of real servers and has its own power, cooling, and networking systems. These data centers are dispersed throughout the world. There are more than 160 Azure datacenters as of the time of this course recording.

China East, China East 2, China East 3, East Asia, Central India, South India, West India, Japan East, Japan West, Australia Central, Australia Central 2, Australia East, Australia Southeast, China North, China North 2, China North 3, East Asia, Japan East, Japan West ,etc.

Key Features:

- Energy-efficient designs with sustainability goals.
- Equipped with state-of-the-art security (multi-layered physical security, access control, CCTV, etc.).

- High-performance infrastructure: latest CPUs, GPUs, storage arrays, etc.

Security & Compliance:

- Compliant with ISO 27001, SOC 1/2/3, GDPR, HIPAA, etc.
 - Redundant power and network connections to ensure business continuity.
-

6. Interconnectivity and Networking

- Azure Backbone Network: One of the largest global networks, with more than 180,000 miles of fiber and global Points of Presence (PoPs).
 - ExpressRoute: Enables private connections between on-premises and Azure.
 - Virtual WANs & Hubs: Connect regional resources seamlessly.
-

7. Resilience & High Availability Design

Azure's infrastructure is built with redundancy at every level:

- Region Pairs for disaster recovery.
 - Availability Zones for fault isolation.
 - Data replication strategies for business continuity and geo-redundancy.
-

8. Sustainability and Innovation

Microsoft is committed to:

- 100% renewable energy in all data centers by 2025.
- Being carbon negative by 2030.
- Water-positive and zero waste by 2030.