in28minutes

Getting Started with Regions and Zones

Scenario - Application in One Data Center

Scenario: Imagine that your application is deployed in a data center in London

Challenges:

- **Challenge 1** : Slow access for users from other parts of the world (**High Latency** ❗)
- **Challenge 2** : What if the data center crashes?
  - Your application goes down (**Low Availability** ❗)

Scenario - Application in Multiple data centers

Scenario: Let's **add in one more data center** in London

Challenges:

- **Challenge 1** : Slow access for users from other parts of the world
- **Challenge 2 (SOLVED)** : What if one data center crashes?
  - Your application is **still available** from the other data center
- **Challenge 3** : What if **entire region** of London is unavailable?
  - Your application goes down

Scenario - Add a New Region

Scenario: Let's add a new region - Mumbai

Challenges:

- **Challenge 1 (PARTLY SOLVED)** : Slow access for users from other parts of the world
  - You can solve this by adding deployments for your applications in other regions
- **Challenge 2 (SOLVED)** : What if one data center crashes?
  - Your application is still live from the other data centers
- **Challenge 3 (SOLVED)** : What if entire region of London is unavailable?
  - Your application is served from Mumbai

Regions

The Challenge: Imagine a startup in India setting up data centers in different regions around the world

- **Would that be easy?**

Cloud Makes It Easy: Cloud Providers provide Regions all around the world

Region : A specific geographical location to host your resources

- **Examples**: New York, London, Mumbai, Sydney

Choosing The Right Region - Factors

1 – **Compliance** ! : Be aware about regulations

- **Why**: Some countries require citizens data to stay within borders

2 – **Latency**: Be close to your users

- **Goal**: Choose a region closest to the majority of your users

3 - **Services Available**: Does the region support the services you need?

- **Note** – New services often launch in specific regions first

4 – **Pricing**: Check pricing in different regions

- **Why**: Costs can vary between regions for same resources

Why Use Multiple Regions?

Low Latency – Serve users faster by placing resources closer to them

Global Footprint – Expand services and reach customers worldwide

Data Residency – Store data in specific geographic locations to meet regulations

High Availability – Reduce downtime by running apps in multiple regions

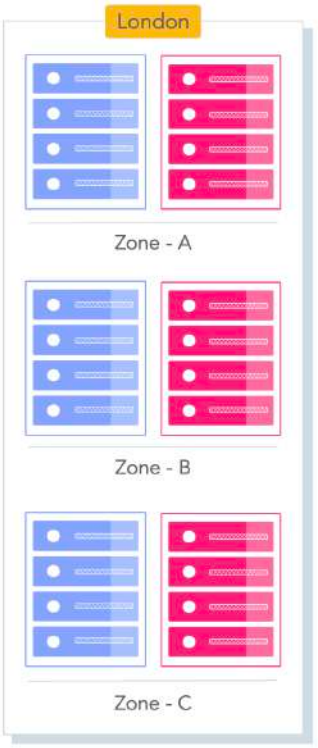
Disaster Recovery – Protect against regional failures (e.g., natural disasters)

Zone(s) or AZ(s) - Availability Zones

Scenario: How do you achieve high availability within a single Region?

Solution: Use Zones (Availability Zones)

- **Zones**: Isolated locations within a Region
- **Independent**: Each Zone has independent power, n/w & connectivity
- **High Speed Link**: Zones in a region are connected through low-latency links



Advantage: Increased availability and fault tolerance within same region

- Survive the failure of a complete data center

Regions vs. Zones - Quick Summary

Feature	Region	Zone/AZ
Definition	A separate geographic area (e.g., London, Mumbai)	An isolated location *within* a Region
Distance	Regions are separated by large geographic distances	Zones are connected by ultra-low latency links within the same region
Failure Scope	Deploying in multiple regions protects against massive disasters (earthquakes, floods) affecting a whole region	Deploying in multiple zones in a region protects against local failures (power outage, fire) in a specific building
Primary Goal	Disaster Recovery, Data Residency, Low Latency for global users	High Availability within the same region

Common Misconceptions – Regions and Zones

Misconception	Reality
Regions are only for large enterprises	Cloud makes regions accessible even to startups
All regions cost the same	Costs vary from region to region - based on local taxes and costs of running a data center in that specific region
Using one zone is enough for high availability	Zone failure can still bring apps down
All regions have same number of zones	Number of zones can vary by region
All zones have just one data center	Nope - Some zones can have more than one data centers

Regions and Zones – Scenarios

Scenario	Best Practice
Design highly available applications within a Region	Multi-Zone Deployment
Global SaaS application with users worldwide	Deploy across multiple regions
Survive data center failure	Distribute across Zones
Survive entire city or regional outage	Distribute across Regions
Meet data residency laws	Select specific compliant Region