Regions and Availability zones

Contents

Overview

Availability zones are tightly coupled (related) to *Regions*, hence this module will cover them both alongside each other.

Availability zones

An **Availability zone** (**AZ**) is a *unique* physical location within a *region*.

One or more data centres make up one *AZ* and has its own independent power supply, cooling and networking.

AZ are independent so as to maintain **zone redundancy**.

Zone redundancy ensures that a datacentre failure would be recoverable through copying everything over to another datacentre within the same *AZ*.

Zone redundancy provides protection from a single point of failure.

Each Cloud provider will be guaranteeing a certain percentage of uptime for their VMs - in most cases reaching ~99.99%.

This pledge is usually described in the **S**ervice **L**evel **A**greement (**SLA**).

An additional benefit of *AZ* is that it has good low latency replication next to high availability, which allows you to make sure that mission-critical applications are running without issues.

One of potential use cases where *AZ* comes into play would be a scenario where you would like to make sure that your database is replicated in another *AZ* for increased resilience.

This would ensure that if one *AZ* goes down due to some accident, you wouldn't lose your data as it would be available within another *AZ*.

Regions

A **region** is made up of multiple *AZ* that are interconnected with a low latency network.

In order to maximise resilience, each *region* has multiple *AZ*.

The number of *AZ* each region contains largely depends on the Cloud provider.

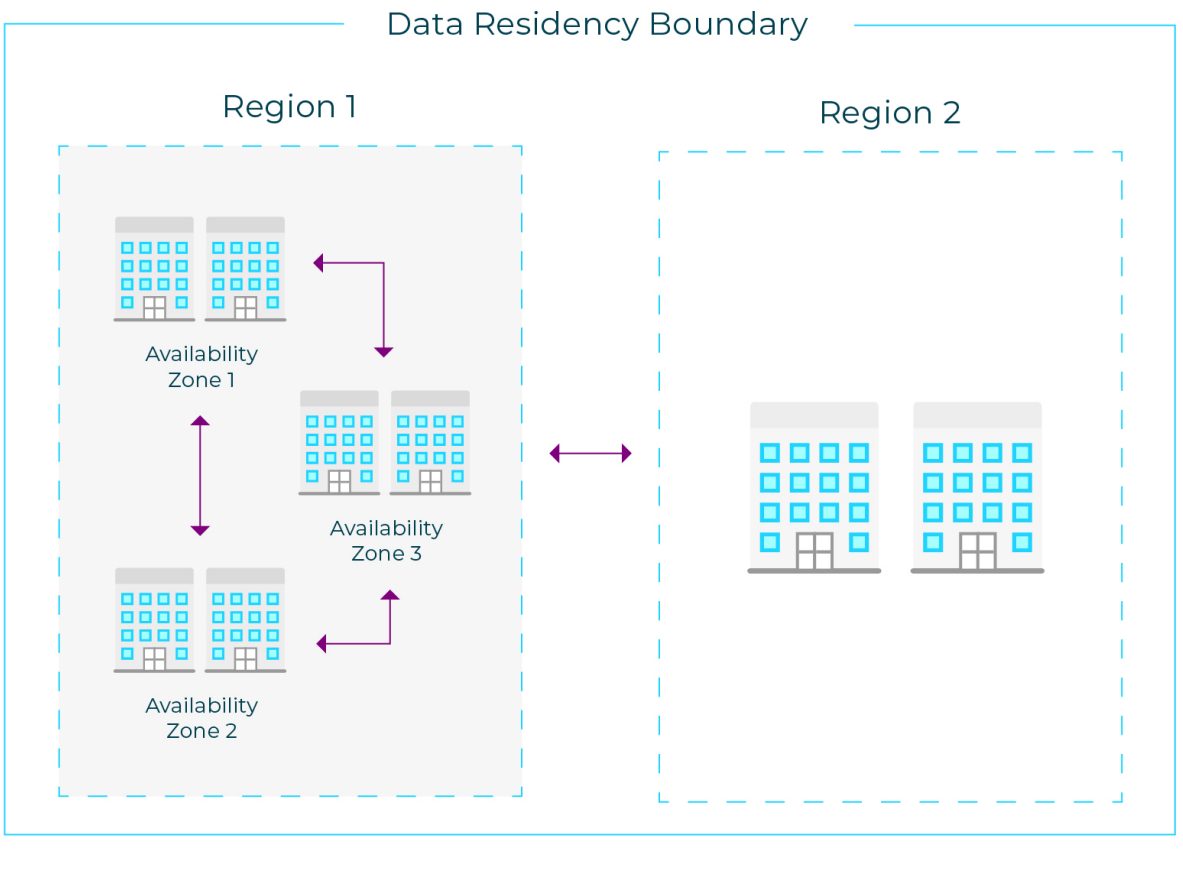
Typically, there are three *AZ* per *region*.

Regions give you the benefit of deploying your application in multiple *regions* to increase resilience and reduce latency.

In a scenario where you have half of your user base in one *region*, and another half in a different *region*, deploying to both would give several benefits:

- reduced latency
- increased resilience

Additionally, scaling could be set up for each region separately so that each region could react to their own specific demands in traffic.



Tutorial

This task is research-based.

Try answering the following questions:

Find out what the typical number of *AZ* within a region is, for the following Cloud providers:

- Google cloud
- AWS
- Azure

Which Cloud provider currently has the most regions?

Exercises

There are no exercises for this module.