3. Global Infrastructure

Regions and Geographies for Azure

Region :
Grouping of multiple Datacentres (AZs).

Azure has 58 regions in 140 countries.

Geographies :
Discreet market of 2 or more regions that preserves ***data residency*** and ***compliance boundaries.***

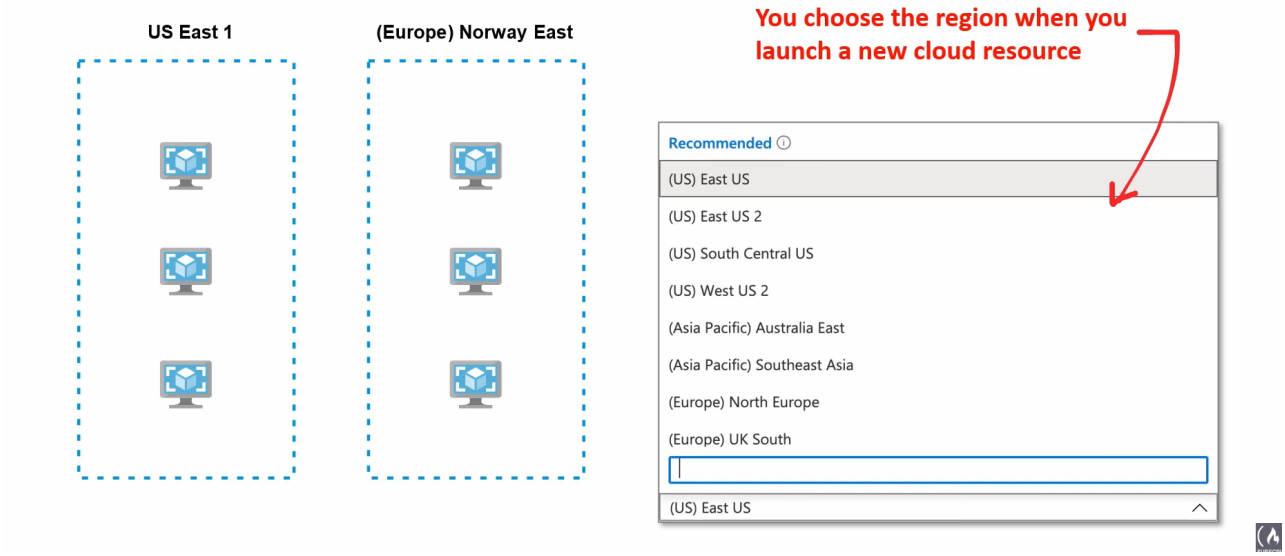
// In essence, if you want for any reason whatsoever that your data doesn't leave a certain country/group of counties, as an example, you do NOT want your data on US soil, you can choose a Geography that limits it.

Some common Azure geography examples are:

1. US
2. Azure Government (US)
3. Canada
4. Brazil
5. Mexico

Generally, how ***Regions*** work in Azure are:

> *When the client wants to launch a resource, they choose a region and a virtual machine.*



// How it looks like, in a simplified manner.

A brief map of Azure Regions is:

Paired Regions

Each region is paired with ANOTHER region around 300 miles (482 kms) away. These are synced, and ONLY ONE REGION is updated at a time, to ensure that no availability outages take place.

Some Azure Services rely on Paired Regions for DR (Disaster Recovery).

Ex: Azure Geo-Redundant Storage (GRS)

Replicates data to a secondary region ***automatically***, ensuring that data is durable even in the event of the primary region failing.

Examples of some paired regions:

Canada	Canada Central	Canada East
North America	East US	West US
Germany	Germany Central	Germany Northeast

Region Types and Service Availability

> *Not all Azure Services are available in every region. (Infrastructure limitations, Government Policy limitations, Compliance Reason)*

So, Azure has 2 Region Types:

1. **Recommended Regions** --> Provide the Broadest range of Service Capabilities. Designed to support. Designed to support multiple AZs for robust DR and HA.
2. **Alternate (Other) Regions** --> Expand Azure's Footprint within data residency boundaries where a recommended region ***ALSO*** exists. However, these have limited capabilities due to lack of multiple AZ presence, and should be used for low priority operations, and if there is virtually no other option left.

> *General Availability (GA) --> When a service is considered ready to be made public/ available to everyone. Transitioning out of Beta testing but before production level in a sense.*

Azure's Services based on ***Availability*** are grouped into 3 categories:

1. **Foundational** : When declared GA, a ***foundational*** cloud service is made available IMMEDIATELY OR IN 12 MONTHS time in Recommended and Alternate Regions. ***Ex: AWS Elastic Compute Cloud (EC2) provides basic cloud computation services.***
2. **Mainstream** : When declared GA, a ***mainstream*** Cloud service becomes available IMMEDIATELY OR IN 12 MONTHS time in ***Recommended Regions***. Availability in Alternate Regions depends on Customer Demand. ***Ex: Google's GSuite***
3. **Specialised** : Made available in both Recommended and Alternate Regions based on Customer Demand. ***Ex: IBM's WATSON AI model***

Special Regions

Azure has specialised regions to meet compliance or legal reasons. Often in collaboration with Governments.

Ex1: 3 Disclosed Azure Government Regions in US are:

1. US DoD Central
2. US Gov Virginia
3. US Gov Iowa

Ex2: Microsoft + 21vianet (in-partnership) disclosed Government Regions in China:

1. China East
2. China North

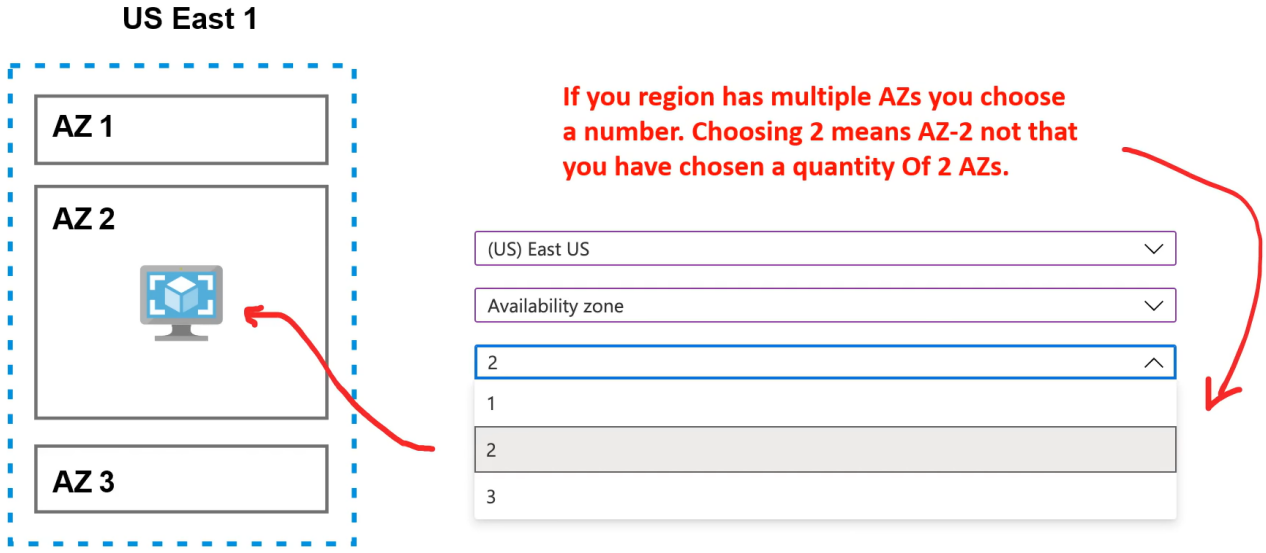
Availability Zones (AZs)

> *An Availability Zone is a PHYSICAL LOCATION made up of one or more Datacentres.*

Also,
AZ == Fault Domain + Update Domain
// more on this in a bit.

A ***Region*** will generally have 3 AZs. For HA (High Availability)

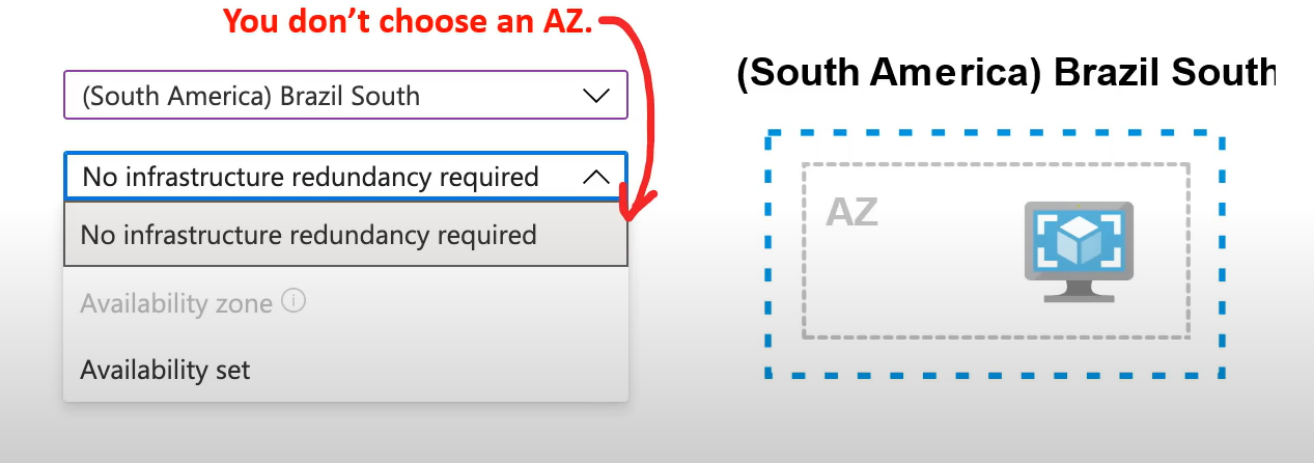
AZs in a region will be isolated from each other, but still close enough to provide minimal latency.



// A simplified version of how to select AZs.

List of Azure Regions that HAVE a minimum of 3 AZs:

1. Central US
2. East US 2
3. West US 2
4. West Europe
5. France Central
6. North Europe
7. Southeast Asia



// How it looks like when no other AZs are available (in Alternate(other) Regions)

Fault and Update Domains

As mentioned before, an Availability Zone in Azure is a combination of a Fault Domain and an Update Domain.

Fault Domain

A logical grouping of hardware to avoid a single point of failure within an Availability Zone.

Basically, a group of VMs that share a Common Power Source and Network Switch.

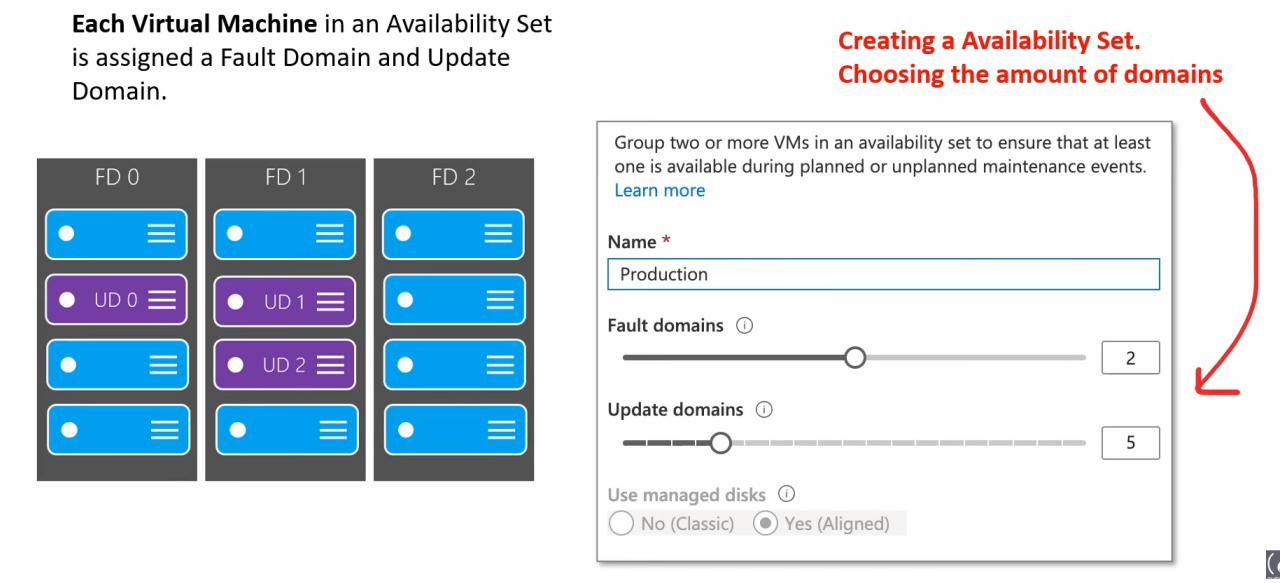
This helps with preventing other servers from going down even if part of the datacentre fails.

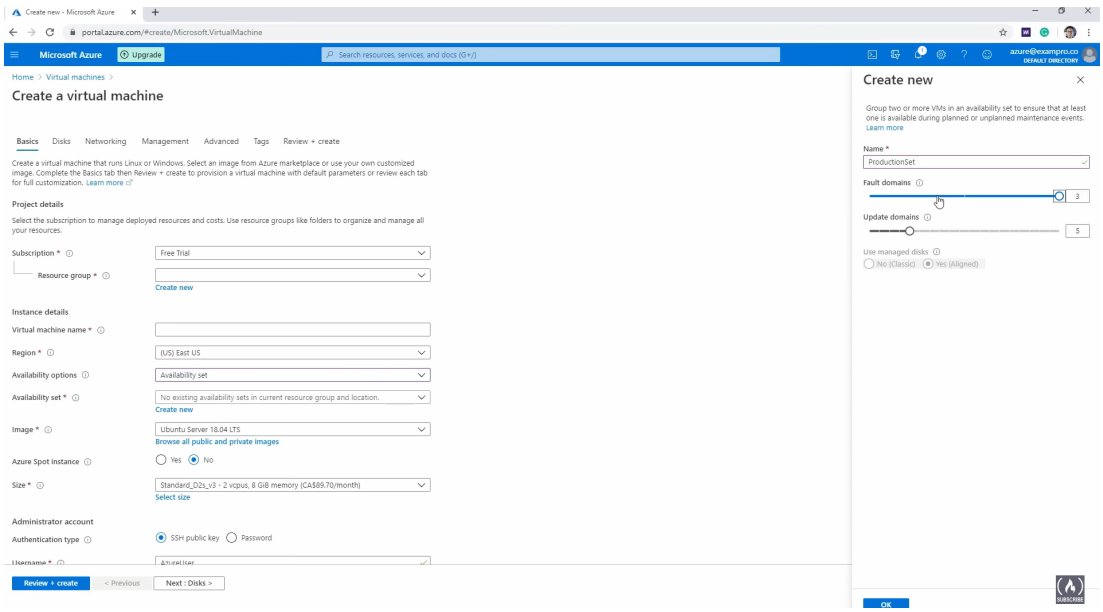
Update Domain

No two domains will be updated at the exact same time, and this allows workload shifting to the Update Domain, while the primary one is updated, and vice-versa, preventing the services from going offline and ensuring HA.

Availability Sets

A logical grouping in Azure that can be used to ensure that the VMs placed in ON AVAILABILITY SET are in different fault/update domains to avoid downtime.





// How it looks like in the actual GUI (The one above it is a more distilled explanation, and the one below is how it actually appears.)