**Azure physical infrastructure**

The core architectural components of Azure may be broken down into two main groupings: the physical infrastructure, and the management infrastructure.

## Physical infrastructure

The physical infrastructure for Azure starts with datacenters. Conceptually, the datacenters are the same as large corporate datacenters. They’re facilities with resources arranged in racks, with dedicated power, cooling, and networking infrastructure.

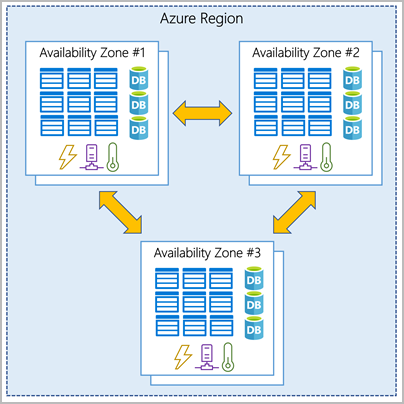
Datacenters are grouped into Azure Regions or Azure Availability Zones that are designed to help you achieve resiliency and reliability for your business-critical workloads.

### **Regions**

A region is a geographical area on the planet that contains at least one, but potentially multiple datacenters that are nearby and networked together with a low-latency network.

### **Availability Zones**

Availability zones are physically separate datacenters within an Azure region. Each availability zone is made up of one or more datacenters equipped with independent power, cooling, and networking. An availability zone is set up to be an isolation boundary. If one zone goes down, the other continues working. Availability zones are connected through high-speed, private fiber-optic networks.



Note: To ensure resiliency, a minimum of three separate availability zones are present in all availability zone-enabled regions. However, not all Azure Regions currently support availability zones.

**Use availability zones in your apps**

use availability zones to run mission-critical applications and build high-availability into your application architecture by co-locating **your compute, storage, networking, and data resources** within an availability zone and replicating in other availability zones. Keep in mind that there could be a **cost** to duplicating your services and transferring data between availability zones.

Availability zones are primarily for VMs, managed disks, load balancers, and SQL databases. Azure services that support availability zones fall into three categories:

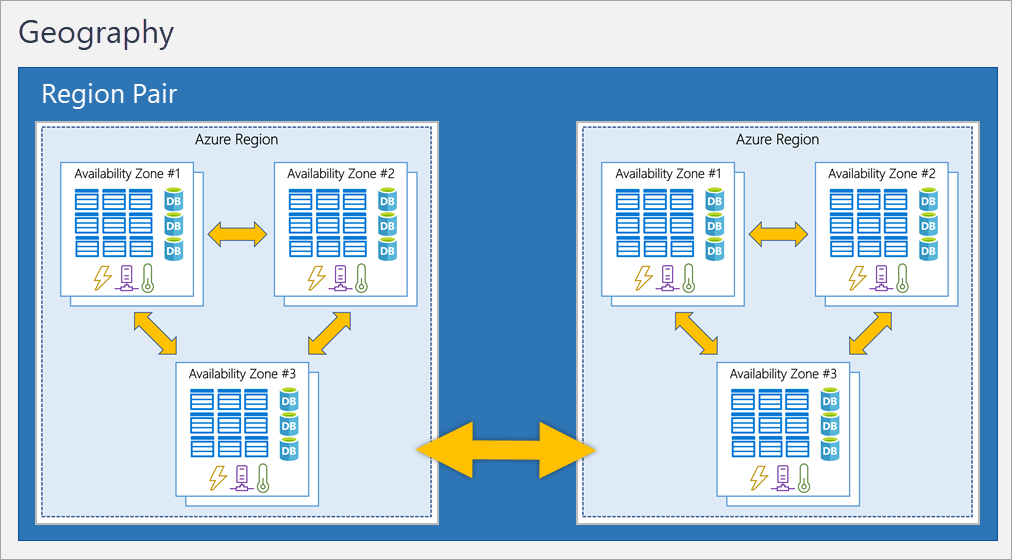
* Zonal services: You pin the resource to a specific zone (for example, VMs, managed disks, IP addresses).
* Zone-redundant services: The platform replicates automatically across zones (for example, zone-redundant storage, SQL Database).
* Non-regional services: Services are always available from Azure geographies and are resilient to zone-wide outages as well as region-wide outages.

**Region pairs**

Most Azure regions are paired with another region within the same geography (such as US, Europe, or Asia) at least 300 miles away.

* approach allows for **the replication of resources** across a geography that helps **reduce** the likelihood of interruptions because of events such as **natural disasters, civil unrest, power outages, or physical network outages** that affect an entire region. For example, if a region in a pair was affected by a natural disaster, services would automatically fail over to the other region in its region pair.

Note: Not all Azure services automatically replicate data or automatically fall back from a failed region to cross-replicate to another enabled region.



Because the pair of regions are directly connected and far enough apart to be isolated from regional disasters, **you can use them to provide reliable services and data redundancy**.

**Additional advantages of region pairs:**

* If an extensive Azure outage occurs, one region out of every pair is prioritized to make sure at least one is restored as quickly as possible for applications hosted in that region pair.
* Planned Azure updates are rolled out to paired regions one region at a time to minimize downtime and risk of application outage.
* Data continues to reside within the same geography as its pair (except for Brazil South) for tax- and law-enforcement jurisdiction purposes.

### **Sovereign Regions**

In addition to regular regions, Azure also has sovereign regions. Sovereign regions are instances of Azure that are isolated from the main instance of Azure. You may need to use a sovereign region for **compliance or legal purposes**.

# Azure management infrastructure

## **Azure resources and resource groups**

A resource is the **basic building block** of Azure. Anything you create, provision, deploy, etc. is a resource. Virtual Machines (VMs), virtual networks, databases, cognitive services, etc. are all considered resources within Azure.

Rules of Resource groups:

* While a resource group can contain many resources, a single resource can only be in one resource group at a time.
* Some resources may be moved between resource groups, but when you move a resource to a new group, it will no longer be associated with the **former group**.
* Additionally, resource groups **can't be nested**, meaning you can’t put resource group B inside of resource group A.
* If you grant or deny access to a resource group, you’ve granted or denied access to all the resources within the resource group.

## **Azure management groups**

The final piece is the management group. Resources are gathered into resource groups, and resource groups are gathered into subscriptions.

Management groups can be nested.

Important facts about management groups:

* 10,000 management groups can be supported in a single directory.
* A management group tree can support up to six levels of depth. This limit doesn't include the root level or the subscription level.
* Each management group and subscription can support only one **parent.**

# Azure virtual machines

VMs provide infrastructure as a service (IaaS).

VMs are an ideal choice when you need:

* Total control over the operating system (OS).
* The ability to run custom software.
* To use custom hosting configurations.

## **Scale VMs in Azure**

Azure can also manage the grouping of VMs for you with features such as scale sets and availability sets.

### Virtual machine scale sets

Virtual machine scale sets let you create and manage a group of **identical**, **load**-**balanced** VMs.

Scale sets allow you to centrally manage, configure, and update a large number of VMs in minutes. The number of VM instances can automatically increase or decrease in response to demand.

Virtual machine scale sets also automatically deploy a **load balancer** to make sure that your resources are being used efficiently. With virtual machine scale sets, you can build large-scale services for areas such as **compute, big data, and container workloads**.

### Virtual machine availability sets