

## Standard AWS Regions & Availability Zones (baseline)

- **Region** = A physical geographic area (e.g., `us-east-1`, `ap-south-1`).
- **Availability Zone (AZ)** = An isolated data center within a Region, usually 2–6 per Region.

Everything else (Local Zone, Outpost, Wavelength, Edge Location) is an *extension* of these.

---

## Local Zones

- **What:** Mini–AWS data centers placed **closer to large metro areas** that are *not* near an AWS Region.
- **Why:** To reduce latency for workloads (gaming, media, healthcare) that need **single-digit millisecond latency** to end-users.
- **Example:** If the nearest AWS Region is 400 miles away, a Local Zone in your city gives you compute + storage closer.
- **Services:** Usually subset (EC2, EBS, sometimes RDS, etc.).

Think: a “satellite” of a Region, closer to users in a city.

---

## Edge Locations

- **What:** Data centers that host **Amazon CloudFront CDN and caching services**.
- **Why:** To deliver **static content faster** (images, videos, API responses).
- **Example:** You open a website; images load from the nearest Edge Location instead of the origin server.
- **Services:** Mainly **CloudFront, Lambda@Edge, Route 53 DNS**.

Think: *good for content delivery, not general-purpose compute*.

---

## Wavelength Zones

- **What:** AWS infrastructure embedded inside **telecom 5G networks** (at the telco's edge).
- **Why:** For **ultra-low latency (sub-10ms)** apps like AR/VR, self-driving cars, IoT, or gaming.
- **Example:** A 5G user streams an AR app that needs real-time rendering. The app runs in a Wavelength Zone inside the mobile network.
- **Services:** Subset of EC2, EBS, VPC — optimized for latency.

Think: *AWS inside a 5G carrier's network.*

---

## Quick Comparison Table

Feature	Local Zone	Edge Location	Wavelength
<b>Purpose</b>	Low-latency compute near cities	CDN & DNS caching	Ultra-low latency 5G apps
<b>Services</b>	EC2, EBS, RDS (limited set)	CloudFront, Lambda@Edge, Route 53	EC2, EBS, VPC (limited)
<b>Where it lives</b>	In metro areas, outside AWS Regions	Globally distributed near users	Inside telecom 5G networks
<b>Latency target</b>	Single-digit ms to city users	Milliseconds (for content delivery)	<10ms for mobile users
<b>Best for</b>	Apps needing compute near users	Faster content delivery	Real-time mobile/edge apps

---

Memory tip:

- **Edge** = *caching content (CDN).*
- **Local** = *extra compute capacity near cities.*
- **Wavelength** = *compute inside 5G networks for real-time apps.*