AWS - Global - Infra

- The core components of the AWS global infrastructure are Regions, Availability Zones, Edge Locations, and Regional Edge Caches. These components work together to provide a reliable, scalable, and globally distributed cloud infrastructure.

🌐 AWS Global Infrastructure – Quick Notes

1. Regions

- Separate **geographic areas** with multiple data centers.
- Provide **redundancy** and **disaster recovery**.
- Examples: us-east-1 (N. Virginia), eu-west-1 (Ireland).

2. Availability Zones (AZs)

- Each Region has **multiple AZs** (typically 3+).
- AZs are **isolated** but connected with **low-latency links**.
- Enable **fault-tolerant**, **highly available** architectures.

3. Edge Locations

- **Global Points of Presence (PoPs)** used by CloudFront.
- Cache content **closer to users** for **reduced latency**.
- Ideal for static content delivery and dynamic acceleration.

4. Regional Edge Caches

- **Larger caches** located between edge locations and origin.
- Help serve **less frequently accessed** content faster.
- Reduce pressure on origin servers and improve latency.

🛰️ **Specialized AWS Infrastructure**

5. Local Zones

- **Extend AWS services** to major metro areas.
- Support **ultra-low latency** use cases (e.g. gaming, media).
- Still managed by AWS, but closer to end-users.

6. Wavelength Zones

- AWS infrastructure **inside 5G networks** (with telecom providers).
- Enables **ultra-low latency (single-digit ms)** for 5G apps.
- Great for real-time use cases like AR/VR, IoT, autonomous driving.

7. AWS Outposts

- **Bring AWS to your data center** (fully managed hardware).
- Run AWS services **on-premises** with local low-latency needs.
- Ideal for **hybrid cloud** or **data residency** requirements.