CDN (Content Delivery Network) – Interview Summary

What is a CDN?

A CDN (Content Delivery Network) is a distributed network of servers that deliver content (images, videos, static assets) to users based on their location. It improves latency, reduces origin load, and provides security features like DDoS protection.

When to Use a CDN

• Serving static assets (images, CSS, JS).
• Large media files (videos, streaming).
• Global users needing low latency.
• DDoS protection and TLS offload.

Whiteboard Sketch

[User] → [DNS] → [CDN Edge POP] → (cache miss) → [Origin Shield] → [Origin Servers].

Impact (Back-of-the-envelope Numbers)

Formulas:
- Total egress (Gbps) ≈ (RPS × avg_size_KB / 1024) × 8 / 1000
- Origin egress = total_egress × (1 – hit_rate)
- Latency reduction (%) = (origin_latency – edge_latency) / origin_latency × 100

Scenario A – Images:
- 100k RPS, 100 KB each = 78 Gbps total.
- At 85% hit rate → 11.7 Gbps origin load.
- At 95% hit rate → 3.9 Gbps origin load.

Scenario B – Video:
- 2M users × 5 Mbps = 10 Tbps total.
- At 90% hit rate → 1 Tbps origin load.

Latency Example

- Origin RTT = 120 ms vs Edge RTT = 20 ms.
- Reduction = 83% improvement.

Operational Targets

• Hit rate goal: 80–95%.
• Use TTLs, cache key normalization, origin shield.
• Pre-warm for launches.

Security and Trade-offs

• DDoS protection, TLS termination, WAF.
• Trade-offs: TTL vs freshness, personalization complexity, CDN cost vs savings.

Quick Unit Conversions

• 1 KB = 1024 B, 1 MB = 1024 KB, 1 GB = 1024 MB.
• Throughput: MB/s × 8 = Mb/s; divide by 1000 ≈ Gbps.
• Example: 100 KB × 100k RPS = 78 Gbps.

CDN Flow Diagram