

Streaming CDN Cheat Sheet

The five tiers, the four dashboards, and the numbers a streaming engineer should know by heart.

1. The five-tier streaming CDN hierarchy

Tier	What it does	Typical CHR	Worked load (100k viewers)
Origin	Packager + storage; canonical source	—	~10 rps in
Origin Shield	Single regional cache funnelling every edge miss	92-96%	~200 rps in
Mid-tier	Regional cache behind a group of edges	85-92%	~2,500 rps in
Edge POP	Metro-area cache, hot content	95-99%	~50,000 rps in
Player	Viewer's device, decode + render	—	50k viewers

2. The five jobs of a streaming CDN

Job	Why it matters	Metric to track
Shorten distance	Edges in the viewer's metro area cut Time to First Byte from 300-800 ms to under 50 ms.	TTFB per region
Absorb load	100k viewers × 1 seg / 2 s = 50,000 rps; CDN absorbs 95-99%.	Cache hit ratio
Resilience	Thousands of edges; stale-while-revalidate keeps content flowing during origin outages.	Per-region uptime
Edge policy	Signed URLs, token auth, geo-blocks enforced at the edge — microseconds, not round-trips.	Blocked-request rate
Billing	One bill replaces dozens of transit / peering / colo contracts. 2026 list: \$0.002-\$0.085/GB.	\$/GB, 95th-percentile

3. Four dashboards to read weekly

- Request volume per edge group — find the hot regions.
- Cache hit ratio per layer — < 80% means misconfigured cache keys.
- Origin egress bytes — should be a small fraction of edge egress.
- Per-region TTFB — pinpoints which geographies need a closer POP or a multi-CDN partner.

Worked savings — one afternoon of origin shield

50,000 viewers @ 3 Mbps — regional sports streamer
 12,500 segment fetches / sec — 1 seg every 4 s, per viewer
 Without shield: 4.5 Gbps origin egress — ~\$180K / month
 With shield: 0.36 Gbps origin egress — ~\$14.4K / month

4. Cache-hit-ratio pitfalls and the one-line fix 2.5× egress reduction — ≈ \$150K / month saved

Symptom	Fix
Per-user query string in cache key (token, session-id)	Strip from cache key; verify token at edge separately
SSAI inserts ads per viewer in the manifest	Switch to SGAI — manifest personalised, segments shared
VOD Cache-Control: max-age=60	Set days or weeks; movies don't change every minute
URL purge on every re-encode	Use versioned filenames (segment-v2.m4s) — no purge