

NAT, STUN, TURN Deployment Checklist

Companion to: NAT, Firewalls, STUN, TURN, ICE — How WebRTC Reaches a Phone

1 · Ports & transports

Open every transport the spec defines, then verify in production.

- ☐ STUN: UDP/3478 open inbound and outbound at every TURN host.
- ☐ TURN: UDP/3478, TCP/3478, TLS over TCP/5349 all reachable.
- ☐ turns://host:443 over TLS — the universal corporate-firewall fallback.
- ☐ Verify a real client picks UDP first, falls back to TLS-on-443 cleanly.
- ☐ Block direct egress from the TURN host except to the public internet.

2 · Credentials (RFC 8656 §9.2)

Short-lived HMAC credentials minted server-side per call.

- ☐ Issue ephemeral usernames (e.g. unixtime:userid) on the application server.
- ☐ Compute password = `base64(HMAC-SHA1(username, shared_secret))`.
- ☐ Set credential lifetime ≤ 24 h, ideally minutes for sensitive verticals.
- ☐ Rotate the TURN shared secret on a regular cadence; keep an N-1 grace window.
- ☐ Never expose long-lived TURN credentials in client-side code or HTML attributes.

3 · Geographic placement

Latency added by TURN equals client-to-TURN RTT, doubled.

- ☐ Deploy at least one TURN cluster per continent your users live on.
- ☐ Use anycast IPs or geo-DNS to steer clients to the nearest cluster.
- ☐ Target a P75 client-to-TURN RTT of ≤ 40 ms in each region.
- ☐ Watch the long tail: rural mobile, Starlink, and CGNAT users may add 80–120 ms.
- ☐ Cap allocation lifetime; reclaim idle slots after ≤ 10 minutes of silence.

4 · Observability KPIs

Instrument every call and every TURN allocation; ship the dashboard.

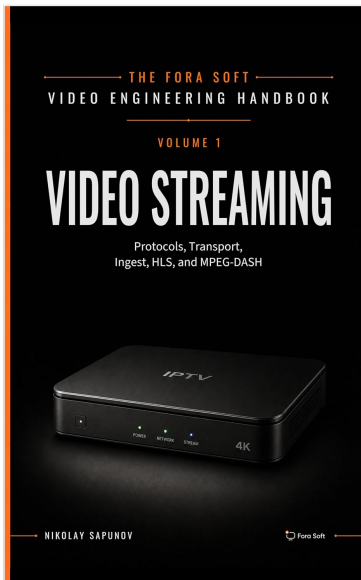
- ☐ ICE-connected rate per network type (Wi-Fi / 4G / 5G / corporate).
- ☐ TURN-relay fraction — expect 15–20 % for consumer, ~ 100 % for AI / IoT.
- ☐ Median time to selected candidate — target ≤ 500 ms on Trickle ICE.
- ☐ TURN bandwidth per call — relay traffic = $2 \times$ call bitrate.
- ☐ Consent freshness failures (RFC 7675) — alert at 0.5 % session drop rate.

Worked TURN bandwidth model — 1,000 concurrent calls, 20 % relayed

$200 \text{ relayed calls} \times 3.2 \text{ Mbps per call (1.6 Mbps} \times 2 \text{ directions)} = 1.28 \text{ Gbps sustained at TURN. Monthly egress} \approx 410 \text{ TB at a 12.5 \% busy-hour ratio.}$

Cost on AWS EC2 egress at blended $\sim \$0.054/\text{GB}$ above 150 TB: $\approx \$22,000 / \text{month}$. Same throughput on a colocated bare-metal cluster with $\$0.50\text{--}\1.00 per Mbps/month transit: $\approx \$700\text{--}\$1,400 / \text{month}$. The 20 $\times$ gap is why every WebRTC product at meaningful scale uses bare metal or a TURN provider whose pricing reflects bare-metal economics.

Controlling specs: RFC 8489 (STUN, Feb 2020) · RFC 8656 (TURN, Feb 2020) · RFC 8445 (ICE, Jul 2018) · RFC 8838 (Trickle ICE, Jan 2021) · RFC 4787 (NAT behavioural requirements) · RFC 7675 (consent freshness).



GO DEEPER · VOLUME 2 OF 7

Video Streaming: A Complete Guide to Video Delivery

From Network Protocols and Transport to Ingest, HLS, and MPEG-DASH

By Nikolay Sapunov, CEO at Fora Soft · 1,124 pages · Kindle & paperback on Amazon

[Read it on Amazon →](#)
amazon.com/dp/B0HH8F4R82

This checklist covers ports, credentials, TURN placement and the KPIs to watch when WebRTC calls must cross NATs. Volume 2 explains the network underneath in depth: NAT behaviour, TCP versus UDP, congestion control, QUIC and HTTP/3, and the glass-to-glass latency budget.

WHAT YOU WILL FIND INSIDE

- NAT types and why calls fail behind them
- TCP versus UDP for real-time media
- Congestion control and what the network does to a stream
- Latency budgets from capture to screen

THE FULL HANDBOOK · 7 VOLUMES · 9,000+ PAGES · ALL ON AMAZON

Vol. 1 · Video Encoding — pixels and color to H.264, HEVC, AV1, quality metrics and AI

Vol. 2 · Video Streaming I — network transport, contribution and ingest, HLS, MPEG-DASH, CMAF

Vol. 3 · Video Streaming II — ABR, CDNs, players, WebRTC, DRM and the economics of delivery

Vol. 4 · Audio for Video I — PCM, codecs, loudness, accessibility, real-time voice processing

Vol. 5 · Audio for Video II — jitter buffers, lip sync, Atmos, quality metrics, neural codecs

Vol. 6 · AI for Video I — models and inference, computer vision, speech, search, generative video

Vol. 7 · AI for Video II — real-time pipelines, agents, serving, cost, compliance, verticals

Calls failing behind corporate firewalls or carrier NAT?

We will review your STUN/TURN setup, credential flow and relay costs, and size a TURN cluster for your traffic and regions. 250+ video products shipped since 2005.

[Book a 30-min call →](#)
[WhatsApp →](#)
[Email us →](#)