

# Edge vs Cloud Video Analytics — Decision Guide

Where the analysis runs sets latency, bandwidth, cost shape, and privacy. Companion to the Fora Soft Learn article.

## The three tiers at a glance

Factor	On-camera	Edge server	Cloud
Latency	~10-50 ms	~30-100 ms	~200-2000 ms
Bandwidth (WAN)	metadata only	video on LAN	full video 24/7
Cost shape	CapEx, ~free	CapEx, low run	OpEx, monthly
Capability	light models	heavy models	heaviest
Privacy	strongest	strong	weakest
Best fit	fast, private	on-site heavy	elastic, central

## Decide in this order

1. Privacy / residency — biometrics or 'stays on-site'? -> edge (camera or server).
2. Latency — millisecond reaction (perimeter, safety)? -> edge.
3. Internet link & scale — can it carry every camera 24/7? If not -> edge.
4. Model heaviness — heavy or cross-site model? -> edge server or cloud.
5. Cost shape — one-time budget -> edge; elastic pay-as-you-go -> cloud.

## Worked example: 40 cameras at 2 Mbps

Cloud upload (full video): 2 Mbps x 40 = 80 Mbps sustained, 24/7 (the binding limit).

Edge upload (metadata only): ~2 Mbps for the whole site -> over 95% less.

Cloud GPU: AWS g6.xlarge ~\$0.80/hr = ~\$584/mo; ~15 streams/GPU = ~\$39/camera/mo.

40 cameras ~ 3 GPUs ~ \$1,750/mo ~ \$21,000/yr GPU alone, before bandwidth & egress.

Edge analytics: one-time hardware, then ~\$0/mo. Lines cross inside year one.

**Most serious systems blend tiers: cheap detection at the edge, heavy analysis in the cloud.**

Engineering guidance, not legal advice (GDPR Art. 9 biometrics; Chapter V transfers). Confirm with counsel.