Imagine the Future

IBC Accelerator Programme 2026

Q-Stream: Quantum Secure, Network-Adaptive, Verifiable, Live Media Infrastructure

Imagine the Future

Headline Sponsor:

Google Cloud

Innovation Partner:

eit Culture & Creativity

Co-funded by the European Union

Associate Sponsors:

hp | AMD

SHURE

The Project

Champions

AP

bfbs

IET The Institution of Engineering and Technology

Participants

deviqon labs

InReality

NORSK

TESLA TECHNOLOGIES

Why does this matter to you?

Your content can be tampered with and repurposed for illicit purposes now or in the future – even if encrypted today

You want to get your content from origin to playout no matter what the connectivity conditions and too often that fails

Q Stream Alpha creates a ecosystem of trust for high value, sensitive content in hostile network and trust environments

What does Q-Stream Alpha do?

Q-Stream Alpha applies to any live contribution or even distribution feed, it provides:

- Cryptographic security, even against future quantum computing threats
- Ensures feed is not manipulated by AI between source and destination
- Adapt to hostile RF environments
- When bandwidth is most under pressure, delivers just the most important part of the video

Problems Faced Today

- Real-time Deepfakes can already be injected into live feeds with face re-enactment
- Today's live encryption won't survive tomorrow's quantum computers

Globally, governments already have ambitious targets for adopting post-quantum cryptography

Q-Stream Fights Back

Authenticating the source of live feeds by embedding C2PA Credentials

- Know that you have the correct feed, not impersonated and not manipulated

Post-quantum encryption

- Maintains Confidentiality
- Protects rushes, non-broadcast material long into the future.

5G for Contribution

- 5G is great for contribution, but can still suffer from variable conditions, low throughput or jamming.
- SRT/RIST are great, but reactive. They can fail and retransmissions add to congestion.

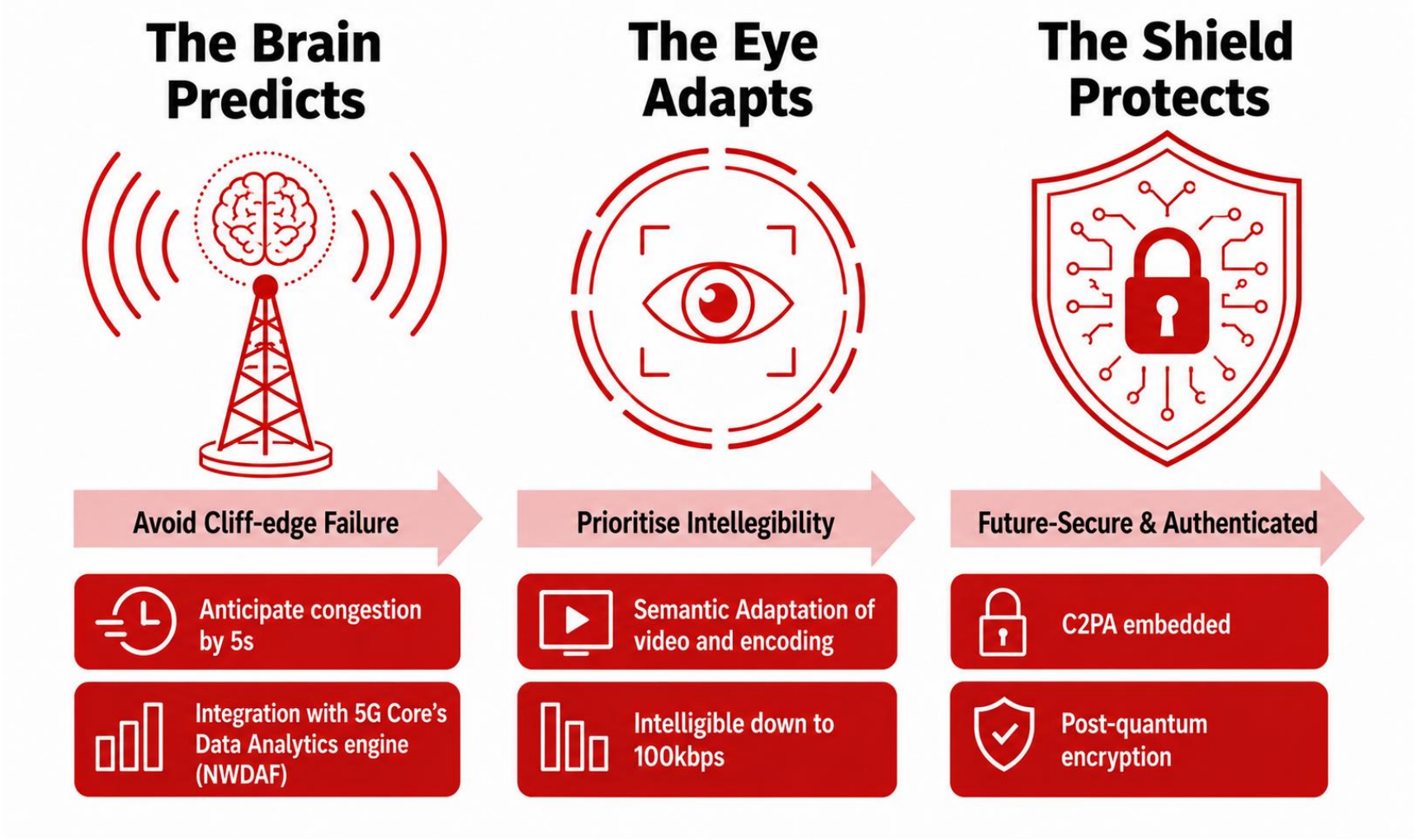
Region-of-Interest encoding guided by existing, AI-assisted semantic analysis allows reduction in bitrate.

Q-Stream Delivers:

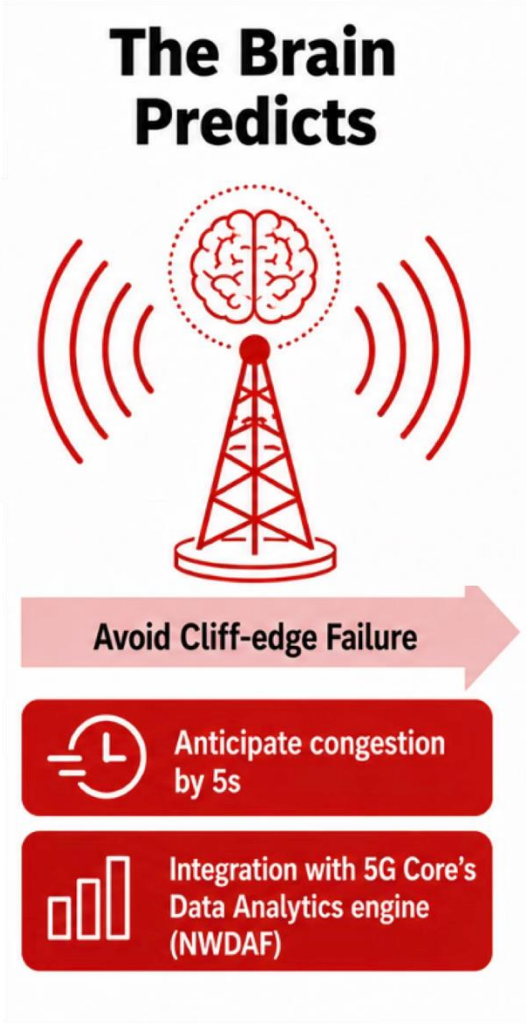
Integration with the 5G core via the Network Data Analytics Function (NWDAF) to predict congestion before it happens.

Live response to predicted congestion by AI-assisted compression to reduce bandwidth gracefully rather than see a catastrophic failure.

The Three Pillars of the Project

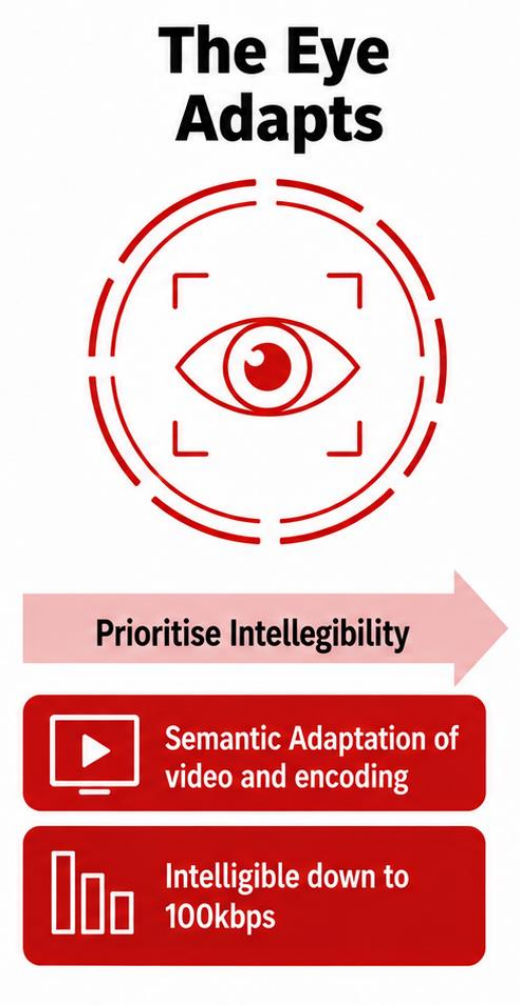


The Brain



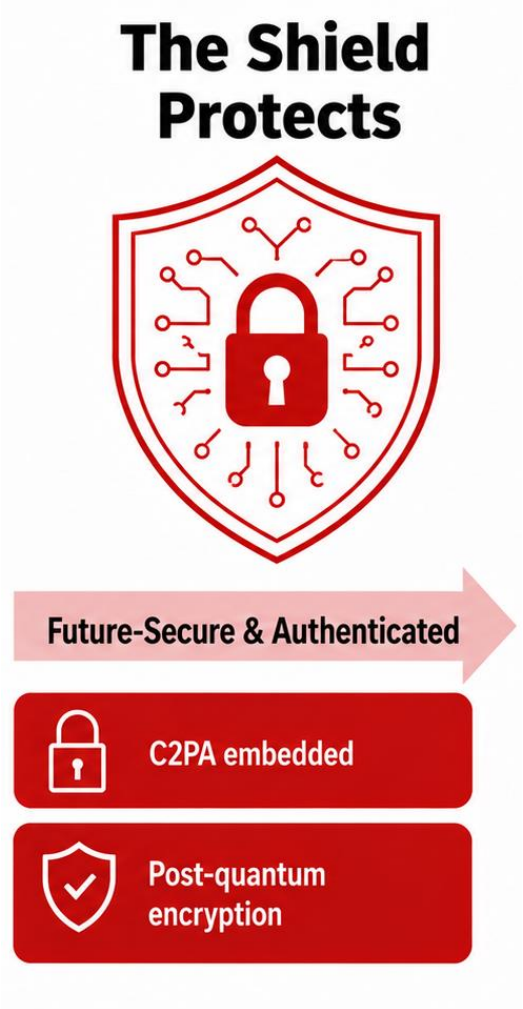
- Receives Real-Time Analytics from the 5G Core's Data Analytics engine
- Applies Machine Learning Predictive Algorithm
- Identifies changes in bandwidth before they happen
- Passes on predictions to the encoder

The Eye



- Continuously analyses video with machine learning and/or AI algorithms
- Determines what is in the video
- System chooses which component (presenter, tank, tree etc.) is most important.
- Encoder encodes down to the bandwidth predicted from The Brain’s analysis.

The Shield



- Adds C2PA Hash to Live Feed
- C2PA – Coalition for Content Provenance and Authenticity
- C2PA embeds information about how the video was generated and shows it has not been manipulated.
- Encrypts the feed in real-time using post-quantum cryptography which is hardened to future decryption by quantum computers.

Who would use Q-Stream Alpha?

News gathering

- Prioritises delivery of live feed in every way possible.
- 5G integration aligns with how news organisations' report the news.
- Region-of-interest encoding encodes just the important part of the video when bandwidth is sparse.
- Q-Stream Alpha protects confidentiality of rushes against future decryption by third parties. This protects sources and journalists.

Who would use Q-Stream Alpha?

Sports

- Robust encryption is required by many sports rights owners.
- 5G integration addresses unreliable 5G availability in and around stadia.
- Applicable to lower tier sports with less infrastructure.
- Region-of-Interest encoding helps return feeds and other non-programme video work reliably.

Who would use Q-Stream Alpha?

Defence

- Video of sensitive activity over temporary or fragile connectivity is often a key requirement in military operations
- Robust encryption is thus key – knowing it is protected in a post-Quantum world
- Network condition prediction and contingency a huge benefit
- Region-of-Interest encoding helps deliver critical aspects of battlefield video

Who would use Q-Stream Alpha?

Enterprise – Financial Fair Disclosure

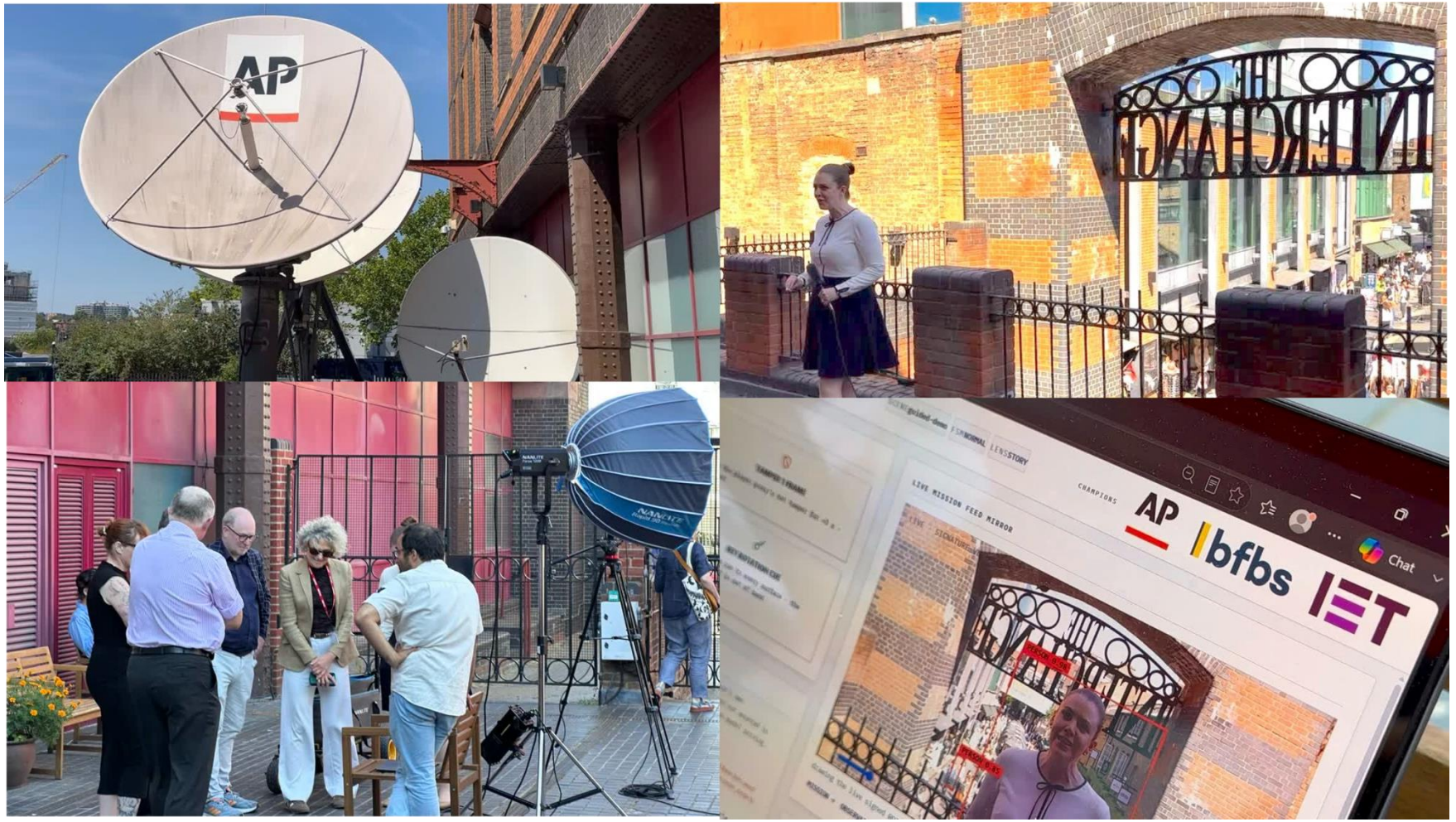
- Live video and coordinated data release
- Commercially sensitive and legally mandated
- Robust encryption is key – knowing the financial information is from the appropriate source
- Open standards so end users can verify authenticity
- Excellent connectivity from venues around the world is *not* a given
- Prioritise most important content (audio over video, video optimised with region of interest)

Q-Stream Alpha in the Real World

POC at The Associated Press in London

Proof-of-Concept Day

#IBCAccelerators2026



POC: Live News Reporter from BFBS



Reporter Sian Grzeszczyk-Melbourne described on a live camera feed the demonstration.

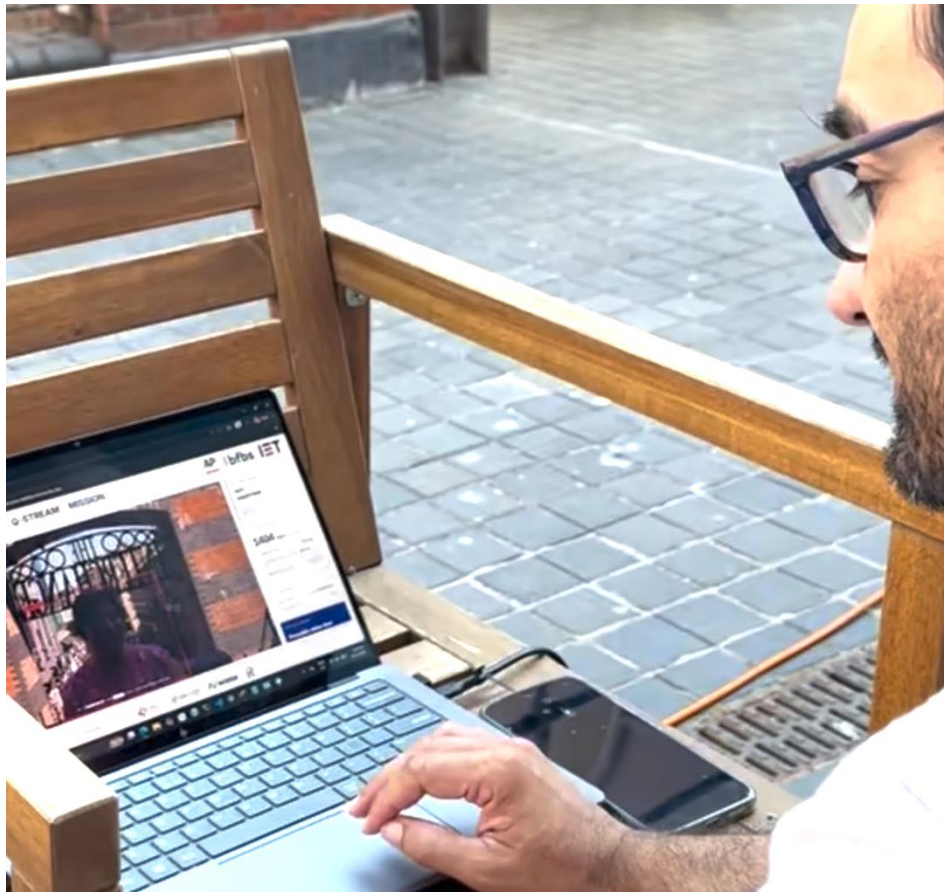
She also discussed her experience reporting live in hostile environments and the frustration when it doesn’t work.

POC: On-Site Live Workflow

An HP Laptop processed and encoded the live video from the encoder locally.

The laptop ran the full, interactive workflow including

- Semantic Analysis
- 5G Simulation & Prediction
- Encoding + C2PA and encryption
- C2PA verification & Playback



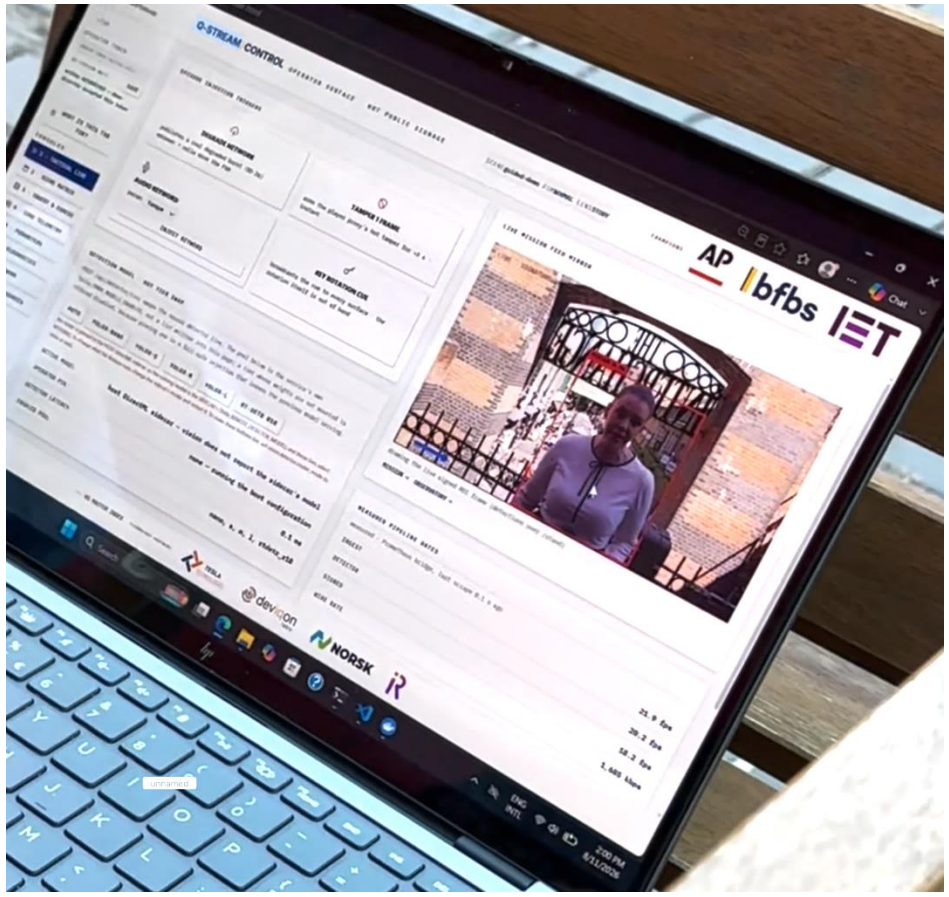
A History of Making the News

The Associated Press's offices in Camden, London have been home to a news organisation for 40 years making it a perfect setting for testing this new technology.





Proof-of-Concept Successes

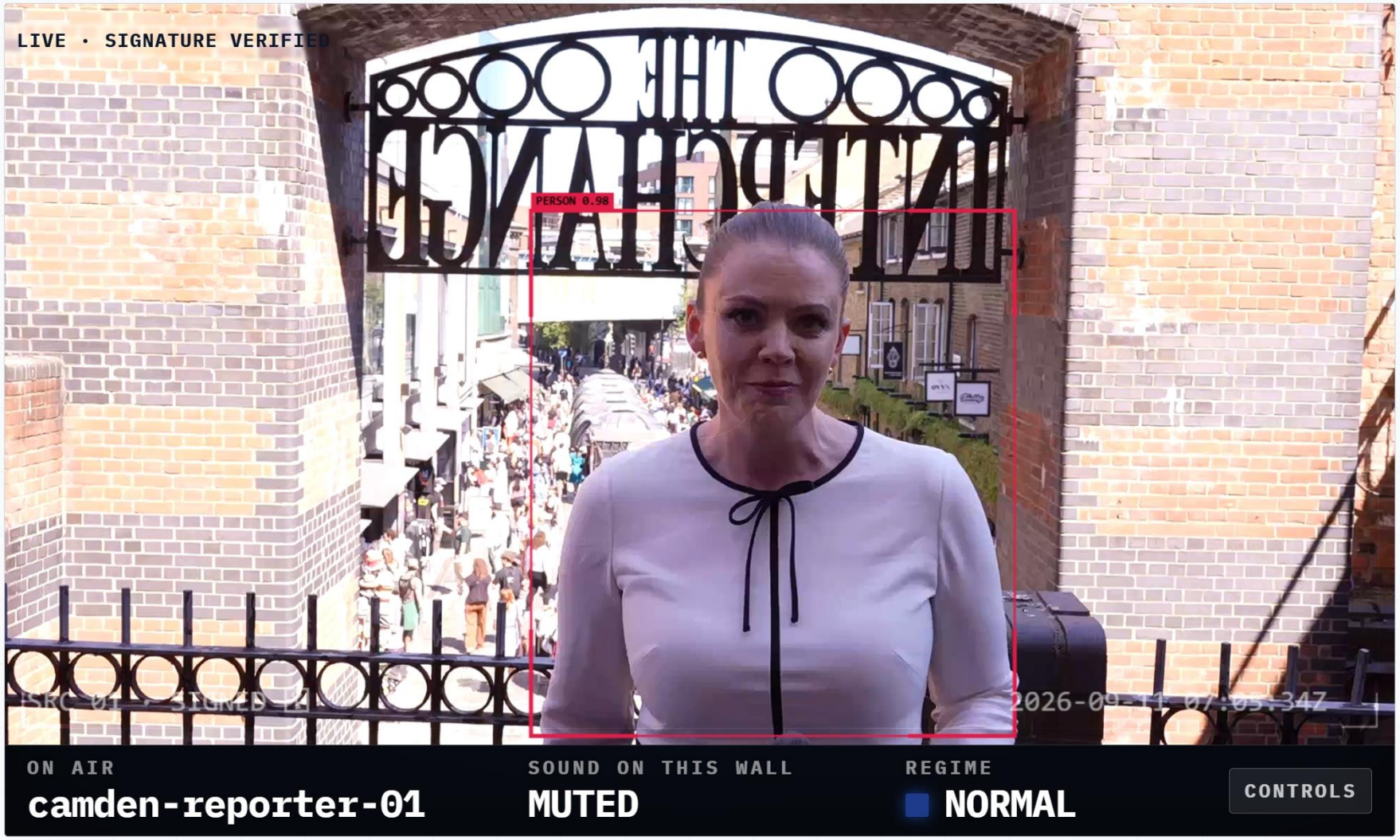


We believe this was the first ever demonstration of:

- Live broadcast contribution with post-quantum encryption
- Live broadcast with live, dynamic, C2PA
- Live detection of C2PA presence/corruption

Important achievement:

→ Everything ran locally on a single laptop



INTEGRITY

VERIFIED

ML-DSA-44 · BLAKE3
PINNED CERTIFICATE

0 · WHERE IT COMES FROM

RECORDING

A file on this stand, played out and signed as it is encoded — not live.

1 · AT THE LENS

2 · MotionTracked

Detection source, signed inside the frame

2 · THE SOUND

BOUND · AUDIO

audio_tag = 1 — a window is signed into every frame

3 · IN THIS BROWSER

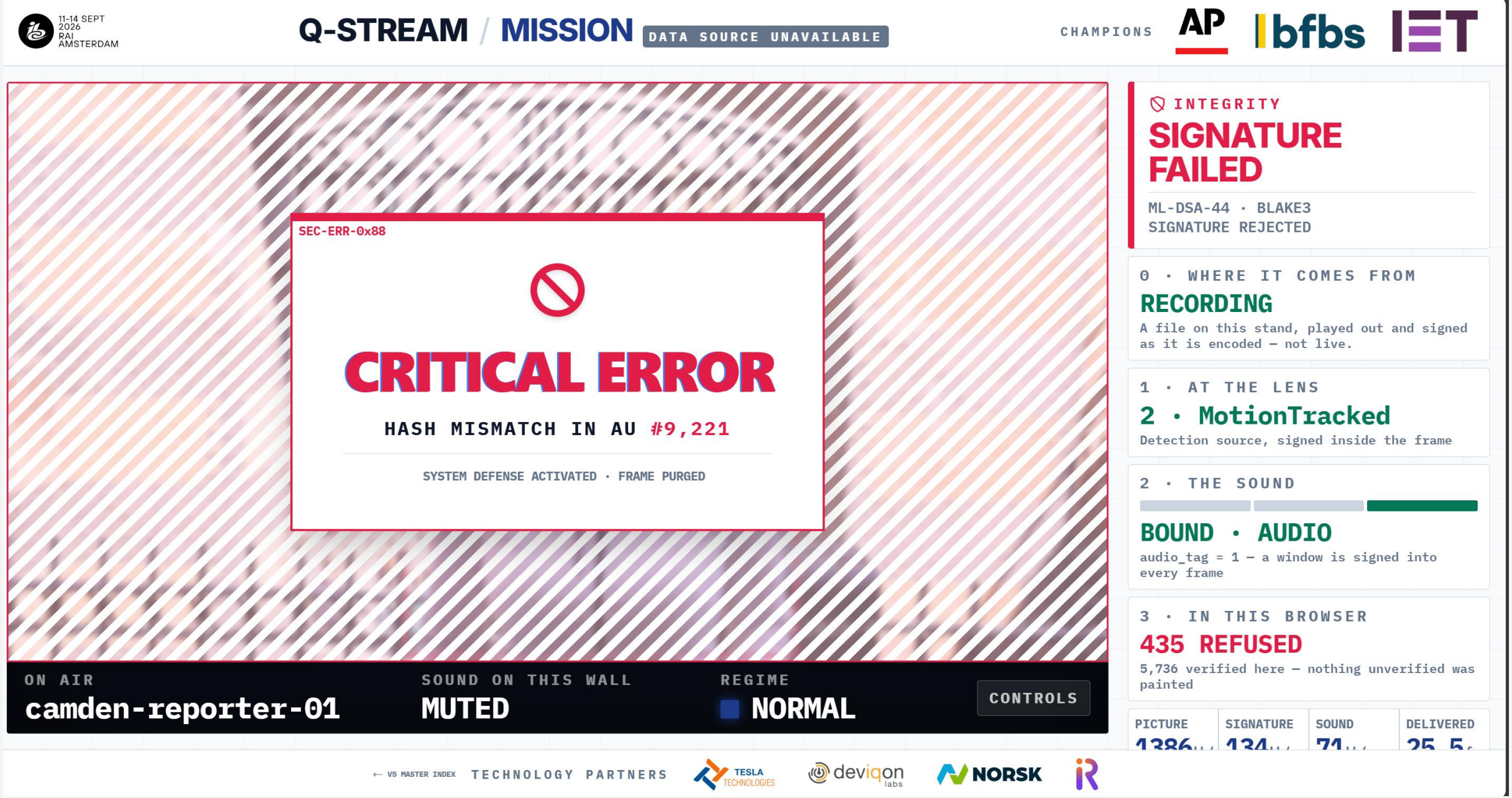
2,195 frames verified

1 refused and never painted — verified on this machine — no server is trusted

PICTURE	SIGNATURE	SOUND	DELIVERED
1616 kb/s	136 kb/s	67 kb/s	24.0 fps

1819 kb/s on the wire · signature 8%

Normal mode


11-14 SEPT 2026 RAI AMSTERDAM
Q-STREAM / MISSION
DATA SOURCE UNAVAILABLE
CHAMPIONS
AP
bfbs
IET
SEC-ERR-0x88
CRITICAL ERROR
HASH MISMATCH IN AU #9,221
SYSTEM DEFENSE ACTIVATED · FRAME PURGED
ON AIR
camden-reporter-01
SOUND ON THIS WALL
MUTED
REGIME
NORMAL
CONTROLS
INTEGRITY
SIGNATURE FAILED
ML-DSA-44 · BLAKE3
SIGNATURE REJECTED
0 · WHERE IT COMES FROM
RECORDING
A file on this stand, played out and signed as it is encoded — not live.
1 · AT THE LENS
2 · MotionTracked
Detection source, signed inside the frame
2 · THE SOUND
BOUND · AUDIO
audio_tag = 1 — a window is signed into every frame
3 · IN THIS BROWSER
435 REFUSED
5,736 verified here — nothing unverified was painted
PICTURE
SIGNATURE
SOUND
DELIVERED
← V5 MASTER INDEX
TECHNOLOGY PARTNERS
TESLA TECHNOLOGIES
deviqon labs
NORSK

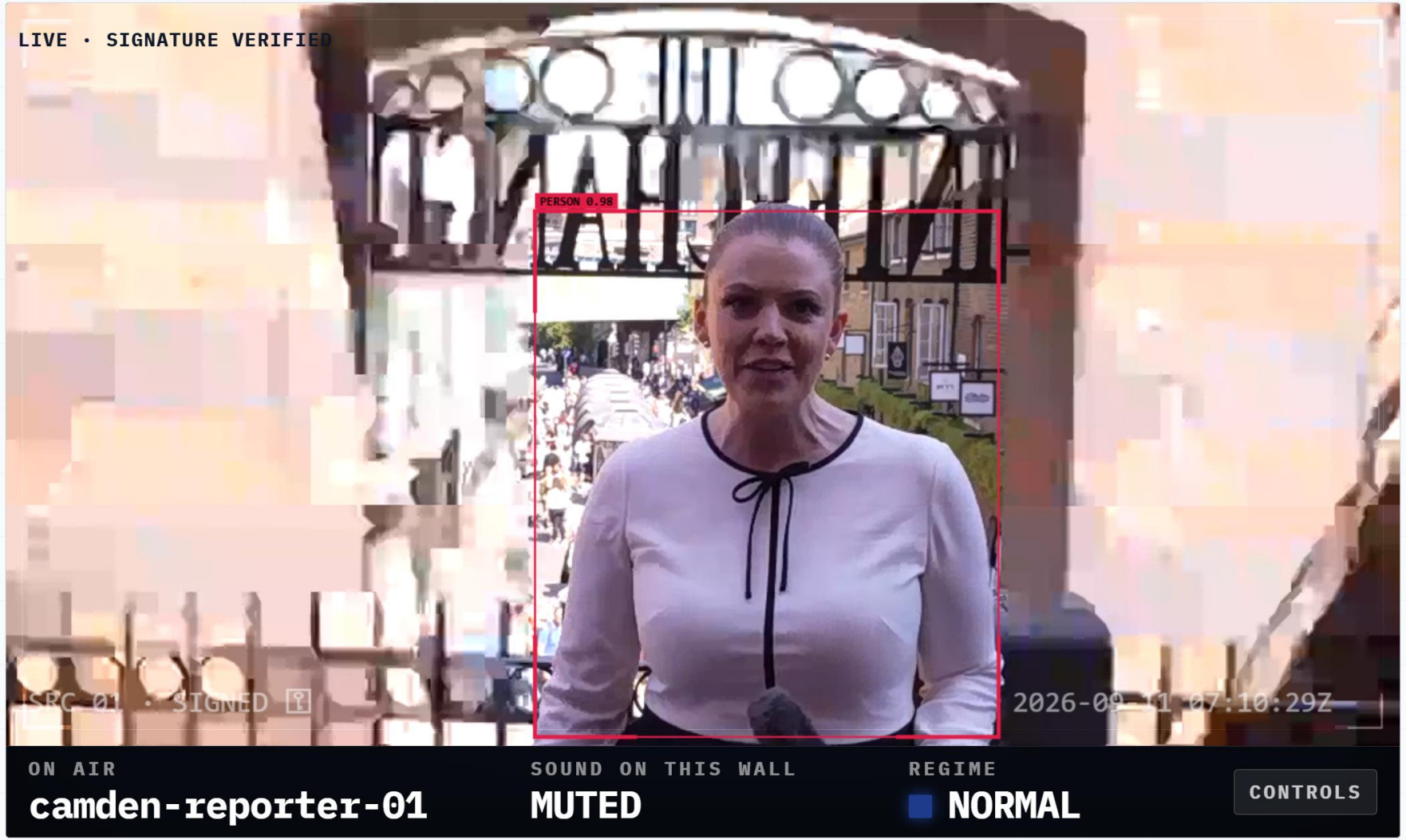
Tamper mode

11-14 SEPT 2026 RAI AMSTERDAM

Q-STREAM / MISSION

DATA SOURCE UNAVAILABLE

CHAMPIONS AP bfbs IET



INTEGRITY

VERIFIED

ML-DSA-44 · BLAKE3
PINNED CERTIFICATE

0 · WHERE IT COMES FROM

RECORDING

A file on this stand, played out and signed as it is encoded — not live.

1 · AT THE LENS

2 · MotionTracked

Detection source, signed inside the frame

2 · THE SOUND

BOUND · AUDIO

audio_tag = 1 — a window is signed into every frame

3 · IN THIS BROWSER

77 frames verified

4 refused and never painted — verified on this machine — no server is trusted

PICTURE	SIGNATURE	SOUND	DELIVERED
687 kb/s	153 kb/s	75 kb/s	27.0 fps

915 kb/s on the wire · signature 18%

← V5 MASTER INDEX

TECHNOLOGY PARTNERS TESLA TECHNOLOGIES deviqon labs NORSK iR

Survival mode

PROVENANCE: DOWN

11-14 SEPT 2026 RAI AMSTERDAM

Q-STREAM / CONTROL

OPERATOR SURFACE · NOT PUBLIC SIGNAGE

SCENE guided-demo | FSM NORMAL | LENS STORY

CHAMPIONS

- LIVE
- ACTIVE ALERTS 3

OPERATOR TOKEN

Bearer token (writes only)

SET (SESSION ONLY) SAVE

TAMPER accepted by demo-director (HTTP 200)

WHAT IS THIS TAB FOR?

CONSOLES

- 1 · TACTICAL LIVE
- 2 · SCENE MATRIX
- 3 · INGEST & EGRESS
- 4 · CORE TELEMETRY
- 5 · PARAMETERS
- 6 · DIAGNOSTICS
- 7 · NETWORK
- 8 · VIDEO SOURCES

ABORT SEQUENCE

SYSTEM RESET

← return to index

EPISODE INJECTION TRIGGERS

DEGRADE NETWORK

publishes a real degraded burst (DD-26) · aurasec → celia move the FSM

TAMPER 1 FRAME

arms the player proxy's hot tamper for ~3 s · instant

AUDIO KEYWORD

param: tanque

INJECT KEYWORD

KEY ROTATION CUE

broadcasts the cue to every surface · the rotation itself is out of band

DETECTION MODEL · HOT TIER SWAP

POST /api/detector/tier swaps the neural detector live. The pool below is the service's own Tesla_ONNX_MODELS_ENABLED, not a list written into this page; a tier whose weights are not mounted is offered disabled, because pinning one is a fail-safe rejection that leaves the previous model serving.

AUTO | YOLOX-NANO | YOLOX-S | YOLOX-M | YOLOX-L | RT-DETR R18

the router is pinned to the HOST DirectML sidecar, so the running model is the SIDECAR's (Tesla_REMOTE_DETECTOR_MODEL) and these tiers select NOTHING. To change what the Radeon runs, change the sidecar's model and restart it. To make these buttons live, set vision.detector.router_mode to onnx or auto.

ACTIVE MODEL	host DirectML sidecar — vision does not report the sidecar's model
OPERATOR PIN	none — running the boot configuration
DETECTOR LATENCY	0.2 ms
ENABLED POOL	nano, s, m, l, rtdetr_r18

LIVE MISSION FEED MIRROR



drawing the live signed ROI frame (detections over /stand)

MISSION → OBSERVATORY →

MEASURED PIPELINE RATES

measured — Prometheus bridge, last scrape 1.5 s ago

INGEST	24 fps
DETECTOR	23.9 fps
SIGNED	23.9 fps

EN ES NL

← V5 MASTER INDEX

TECHNOLOGY PARTNERS

TESLA TECHNOLOGIES | deviqon labs | NORSK

Control Plane

Technical overview

Subtitle

The Q-Stream Alpha Stack

#IBCAccelerators2026

Nineteen edge microservices in four planes on one MQTT bus — and every box below runs on the laptop in front of you.

HOST PLANE — WHAT NO CONTAINER CAN REACH (WINDOWS: DIRECTSHOW, DIRECTML)

- **capture-agent :8091** — DirectShow enumeration and mode negotiation
- → **capture-host.ps1** — dshow bridge: H.264 / MPEG-TS on :9500+
- **detector-sidecar :19091** — DirectML on the host iGPU — worth +12.5 fps

DATA PLANE — THE SIGNED PIXELS

- **frame-producer** — ten source kinds into one shm slot
- → **tesla-vision-ai** — Semantic Switch, foveation, rate governor
- → **c2pa-ingest-camera** — the lens: signs every frame at source
- → **c2pa-egress** — verify-then-relay, C2PA SEI inject, RTSP
- → **player + headless_verify** — two verifiers that share no code

ONE shared-memory slot · a frozen 3,401,000-byte contract · a four-state handshake that IS the backpressure — no queue to grow, no buffer to tune, no unbounded latency

CONTROL PLANE — JSON OVER MQTT + THE 3GPP SBI · NETSHAPER TC1–TC7 ECHOES ONTO THE SAME BUS

- **aurasec-predictor** — forecasts the QoS diff-edge, horizon 2 s
- → **celia-orchestrator** — eBPF GTP-U tap, lock-free FSM, bus hub
- → **5g-af-gateway** — the only 3GPP speaker — TS 29.514
- → **5G Core PCF / NEF** — policy authorisation over HTTP/2 + mTLS

PRESENTATION PLANE — THE BOOTH

- **stand-web · nginx :8090** — baked v5 assets behind a locked CSP
- → **demo-director** — bus + direct scrape into one WebSocket
- → **stand v5 surfaces** — operator console, Sources panel, mission wall

The control plane sets the bounds; the data plane spends inside them. No control-plane message can change a byte that has already been signed — which is why a change in resilience cannot silently weaken integrity.

The Signed Pixel Path

#IBCAccelerators2026

Four states, one slot, and one ordering decision that makes the whole thing possible.

status 0 — PendingVision
frame-producer writes raw YUV420 into the slot

→

status 1 — PendingCrypto
tesla-vision-ai detects, foveates and encodes H.264 in place

→

status 2 — PendingEgress
c2pa-ingest-camera hashes the access unit and asks the chip to sign it

→

status 3 — Complete
pushed onto the Noise_NNpsk0 wire; the slot is released

Foveation happens BEFORE the lens signs

- So the C2PA signature covers the video that is actually delivered — the foveated bytes, not a pristine master nobody ever sees.
- Compression happens exactly once. Saliency is the spatial half of the quantisation decision, never a post-encode pass.
- Re-encoding after signing would cost rate-distortion, double the latency and break the byte-binding. So we never do it.

One slot. No queue.

- The handshake IS the backpressure. There is nowhere for frames to accumulate, which is why this pipeline has no unbounded latency.
- clip and pattern sources WAIT on a busy slot — order and completeness: every frame of the master is signed.
- live sources DROP and count — freshness, because the picture must be of now and not of the backlog.

A designed trade, not an inconsistency — and the drop is a counter, not a silence.

C2PA: what the signature actually covers

Not a hash of the picture. A hash of the picture together with its identity, its freshness and its provenance.

One digest, signed at the lens

```
digest = BLAKE3(
    "tactical c2pa frame-sig v4-fwd"
    | frame_id          u64 LE
    | timestamp_us      u64 LE
    | threat_level      u8
    | detection_source  u8
    | BLAKE3(video)
)
```

Signed with ML-DSA-44 — FIPS 204, a 2,420-byte signature over a 1,312-byte public key, produced by a sealed and mlocked signing chip.

Five bindings in one hash

- Identity — the domain string IS the scheme. A verifier applies the scheme it was configured with and never one read off the wire: that is the downgrade shape which killed SSLv3.
- Freshness — frame_id and timestamp are signed, and the edge challenges the lens with a fresh nonce per connection, so a replayed session is refused.
- The pixels — BLAKE3 over the exact H.264 access unit that leaves the encoder.
- The AI — detection_source rides inside the signature, so a frame that shed neural inference cannot borrow neural credibility.
- The sound — av-sig v5 binds the audio window by BLAKE3 in the same signature; absence is bound too, so stripping the track is detectable.

Post-quantum provenance that fits on the wire

#IBCAccelerators2026

ML-DSA-44 is a fixed 2,420 bytes, and it used to ride every single frame. Here is what that cost us.

PROVENANCE SHARE OF THE WIRE
67.1% → 26.1%
measured on ten seconds of live wire, not modelled

C2PA SEI PER FRAME
2,496 B → 609 B
4.1× less provenance on the link

PICTURE BITRATE
79 → 291 kbps
3.7× more picture, on less total wire

QUANTISER IN THE ROI
41.27 → 26.56
14.7 QP better where the subject is

The problem, measured

- Walking the Annex-B NAL types over ten seconds of live wire: the C2PA SEI was 67.1 % of every byte we sent. Provenance 445 kbps against 218 kbps of picture.
- Structural, not waste — the signature is a fixed size and it was fully verified on every frame.
- Every encoder gain we had ever fought for was being fought inside the one third of the wire the video actually got.

The forward chain, shipped

- One ML-DSA-44 anchor every six frames. The frames between carry a 32-byte BLAKE3 commitment to their successor.
- A link verifies on arrival with ONE BLAKE3 against a commitment its predecessor already carried — no signature, no waiting.
- Forging a link means breaking BLAKE3; forging the root means breaking ML-DSA-44. The claim did not weaken — only the bill.

Total wire 500 → 394 kbps — and the picture still got 3.7× more bits.

The arithmetic predicted 601 bytes per frame for a six-frame interval; the wire carried 609 — a 1.3 % prediction of an interval the formula was not derived from, and it is now asserted in the tests.

C2PA: fail-closed, and annunciated

Three independent verifiers, and a refusal an operator can actually read.

Edge Defense, zero-trust browser

- c2pa-egress verifies every frame at ingest and refuses the invalid ones, so a tampered frame never reaches a consumer as valid.
- The WebSocket proxy is a dumb gateway. All crypto runs in the page, against a key pinned out of band.
- headless_verify shares no code with the browser. Agreement between independent implementations is evidence; agreement with your own copy is a tautology.
- Four hand-written implementations of the chain walk are held together by frozen known-answer vectors behind SHA256 pins.

Loss is not tampering

- A gap in the signed frame ids is amber: the frames are refused, the chain drops, and it re-roots at the next anchor in about 250 ms.
- A contiguous digest mismatch is red, sticky and compromising.
- The split was paid for in the field: a 1250 kbps cap sheds about six frames in 360 by construction, and every one of them used to pin the wall to INVALID.

  Nothing is painted from an unverified frame either way. An attacker who skips ids to earn the softer label gains the label and loses the frames.

Keys rotate hot on an established session with a cross-certificate, and the edge enforces anti-rollback — a rotation can roll forward, never back.

What we can compress, and by how much

#IBCAccelerators2026

The envelope is retargeted live over 5 Hz telemetry — x264_encoder_reconfig, no reopen, no stream restart.

ADAPTIVE ENVELOPE
2500 → 80 kbps
a 31× range, traversed live and continuously

SURVIVAL vs NORMAL
< 55 %
of the p75 NORMAL baseline — gated by test T7

SUBJECT vs BACKGROUND
48 QP
core-to-background spread at absolute ROI priority

DECODED SHARPNESS
2–3×
subject over background, by Laplacian variance (T8)

The law, and the ladder it walks

- target = clamp(0.7 × bandwidth, mode bounds). NORMAL spans 300–2500 kbps at uniform quality; SURVIVAL spans 80–600 kbps with saliency foveation on top.
- A live episode reads: 2107 → 1757 → 1484 tracking the ramp, then 600 → 441 → 371 → 301 → 231 tracking the collapse, then 600 → 2500 on recovery.
- Lab measurement across both flips: 150–250 KB/s in NORMAL down to 30–70 KB/s in SURVIVAL, and 392 of 392 captured frames decode clean.

Where the extra ratio comes from

- Geometry — SURVIVAL encodes 1024×576 against a curve calibrated on 1920×1088, so the same rate buys 3.54× the bits per macroblock.
- Absolute ROI priority — below a 900 kbps grant the salient half of the offset map deepens, reaching −24 at the 600 kbps floor. A person stays recognisable at 80 kbps.
- Pre-degradation efficiency — content-aware capped VBR enters SURVIVAL with 1.4–2× more quality per bit.
- And the wire budget subtracts measured provenance before the encoder is ever handed a target.

With no region of interest on screen the map holds at maximum background sacrifice for up to 72 frames — about three seconds at 24 fps — banking bits against a threat, and then hands the frame back to plain adaptive quantisation.

Foveation: attested QP vs delivered QP

#IBCAccelerators2026

Two numbers that must always be quoted together, because either one alone gets misread.

WHAT THE SIGNATURE ATTESTS

~15 QP

of separation in the signed, quantised QP map — auditable and exact

WHAT THE WIRE DELIVERS

~1.7 QP

0.95–1.33 by profile — what an eye in front of the screen actually sees

Why the gap exists

- x264 adaptive quantisation and mbtree re-derive their own per-macroblock decisions on top of our offset map. The encoder honours the map as a hint, not as an instruction.
- The attested figure answers "what does the signature cover". The wire figure answers "what does a visitor see". Both are true, and they are different questions.

What bounds it, measured

- The salient fraction reads 0.30–0.57 on the booth clip. With more than half the frame classified as subject there is very little background left to sell — and no offset depth changes that.
- Run-to-run dispersion of the ROI quantiser is about 1.1 QP. Anything smaller is not a finding, and that is the floor we report against.

Under the forward chain the delivered gap rose from 0.30 to 1.45 QP on less total wire. The saving reached the picture, which is the only test that counts.

The Semantic Switch

Audio context selects the model; the model decides where the bits go — hot-swapped mid-stream, and unable to lie.

ONNX EXPERTS DECLARED
11
ten baked and sha256-pinned; the eleventh reports absent rather than faked

BEST TIER, MEASURED
0.505 mAP
rtdetr_r18 at 23.12 fps — beats YOLOX-l on both axes, NMS-free

DETECTOR BUDGET
80 ms
shed to a lighter backend over it, climb back on sustained headroom

MEAN ROI CONFIDENCE
0.90
agreement fusion plus the per-ROI verifier

Hot-swap, with a fail-safe

- A spoken keyword routes to an expert and the active ONNX session is swapped without interrupting the stream.
- If the new model cannot load, the previous engine keeps serving and the error is logged. The stream never goes blind on a bad swap.
- The manifest is data, never code: a bad tensor layout or a missing default expert is rejected at load rather than panicking mid-mission.
- A keyword can never talk the system into a degraded posture. Sound improves what the encoder attends to; only measured telemetry can move the mode.

Honest provenance tiers

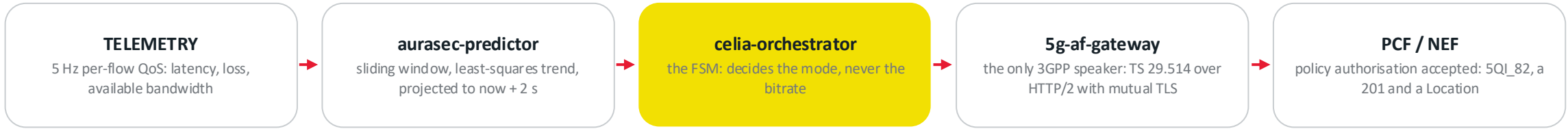
- detection_source names the tier that produced the regions of interest — neural, classical or motion — inside the signature itself.
- So the signed provenance never claims neural AI when only background subtraction ran, and the browser renders the weaker tiers muted instead of blocking a legitimate frame.

The defect it caught: a 40 ms budget retune shed neural inference permanently. The picture looked no worse — the provenance byte was the only thing that said so.

5G: predict, do not react

#IBCAccelerators2026

The loop forecasts the cliff-edge and asks the 5G Core for a priority slot before the video stalls — not after.



The forecast, calibrated

- It fires predictively when a worsening trend would breach a hard limit inside the horizon, and reactively when a limit is already breached.
- Calibration is committed with the model: threshold 0.4411, recall 0.797, PR-AUC 0.823, ROC-AUC 0.895.
- 131 tests on the predictor alone — including that it predicts the cliff before the breach, and the split between the ONNX and the heuristic thresholds.
- On recovery the gateway deletes the Location it was given and celia detaches the eBPF filter. The loop closes at both ends.

Measured on real tc netem

- 2105.2 kbps clear, 329.4 kbps under a 300 Kbit cap, 1634.8 kbps recovered — read from the qdisc own byte counter inside the namespace.
- netshaper carries TC1 to TC7 impairments as loadable scenarios and echoes onto the bus what it actually applied, not what it was asked for.

An earlier triple of 14.7, 0.0 and 22.0 frames per second was withdrawn: its middle term was a starved metrics scrape, not a dead link. The correction is published next to the number.

Is it really 5G? The honest answer

#IBCAccelerators2026

The same code runs in three shapes. What differs is whether anything downstream of the PCF enforces the grant.

	THE BOOTH STAND	5g-lab · synthetic	5g-lab · full
telemetry source	**netshaper echo + director**	synthetic feed	chaos on a real UPF uplink
the forecast (aurasec)	**real**	real	real
the FSM (celia)	**real**	real	real
foveation response	**real — measured on the wire**	real	real
the 3GPP SBI request	**real HTTP/2 + mTLS**	real	real, to Open5GS PCF
a 5QI actually enforced	**no — nothing schedules it**	no	**YES, by the core**
GTP-U datapath, eBPF	**none**	none	UERANSIM UE/gNB to UPF
real tc netem impairment	**yes, on the distribution hop**	optional	yes, on the uplink

Every arrow in the loop is real code in all three shapes. At this stand the policy request is genuinely made, genuinely mutually authenticated and genuinely answered — and no radio scheduler acts on it, because there is no radio. The booth demonstrates the loop; the lab demonstrates the enforcement. We would rather say that than let you assume it.

Performance: one laptop, nothing removed

#IBCAccelerators2026

AMD Ryzen AI MAX PRO 390 + Radeon 8050S. Ten seconds of delivered RTSP with -c copy — the wire, not a re-encode.

PROFILE	safe	fast-cpu	fast-gpu — shipped
detector	ONNX auto tier	rtdetr_r18, 12 ORT threads	**DirectML sidecar**
delivered fps	14.6	17.6	**30.1**
whole vision frame	45.14 ms	38.49 ms	**11.45 ms**
deep pass (the detector)	32.94 ms	26.87 ms	**0.31 ms**
Lucas-Kanade	10.86 ms	10.28 ms	**9.64 ms**
H.264 encode	10.76 ms	10.75 ms	**12.66 ms**
jitter above the floor	123 ms	169 ms	**20 ms**
p95 above the floor	+271 ms	+585 ms	**+51 ms**

tesla-vision-ai is roughly 90 % of every frame budget, so its cost table IS the pipeline frame-rate table. Hoisting the Lucas-Kanade pyramid took the CPU-only reference host from 4.10 to 12.65 fps — 3.08×, with no accuracy trade-off. Resident memory was flat across 18 samples on top of 20.5 hours of continuous uptime, with no leak in any service sampled.

How we know it works

Around 2,700 automated tests, four independent implementations of the chain, and four rules that were each learned the hard way.

AUTOMATED TESTS
2,681
across eleven counted components, plus the JS and PowerShell suites

CI CRATE GATES
10
plus the shell and Node suites, guarded by a test of the pipeline itself

CHAIN IMPLEMENTATIONS
4
hand-written, held together only by frozen known-answer vectors

CONTROL PARAMETERS
375
one registry, confirm gates on the dangerous ones

The four rules we test by

- A green test is not evidence. Prove the test can fail: break the mechanism and watch the red appear.
- Not-exercised is not passed. Exit 2 is a distinct verdict, because on a booth morning "passed" and "never ran" differ by a blank screen.
- Converged is not working. Every control needs a witness — a counter that moves because of that parameter and nothing else.
- Two verifiers, no shared code. Agreement between independent implementations is evidence; agreement with your own copy is a tautology.

What that discipline actually found

- Of sixteen controls measured, three converged with a flat witness and were reported cosmetic — fields written to and read back by nothing.
- The independent verifier had no chain support at all: 302 frames, zero verified, on a stream the edge was logging link by link.
- A CI gate handed its image build the wrong context, so the step that proves the crypto library is pinned to a tag and a commit had never once run.

All three are written down, with dates and ids, in the documents that back every number in this deck.

Standards, not a private protocol

Every layer here is somebody else's published contract — which is what makes it interoperable, and auditable.

C2PA
Coalition for Content Provenance and Authenticity — hard binding by hash, carried as H.264 SEI

FIPS 204 · ML-DSA-44
a 2,420-byte signature over a 1,312-byte key; liboqs pinned to a tag AND a commit

3GPP TS 29.514
policy authorisation over HTTP/2 with mutual TLS, answered by Open5GS in the lab

ONNX Runtime
Vitis AI, CUDA, VAAPI, DirectML or CPU — chosen by probe, with a safe fallback

RTSP · MPEG-TS · SRT
ordinary contribution and distribution transports; the SRT tee is off by default

MQTT + Prometheus
one JSON control bus, and a /metrics endpoint on every single service

What we believe is genuinely new

- Live contribution in which every access unit is covered by a post-quantum signature, verified twice by code that shares nothing — and in which the provenance itself declares which class of detector produced the regions of interest.
- Foveated, bandwidth-adaptive compression underneath that chain, with no re-encode and no broken byte-binding, driven by a 3GPP policy request that answers a forecast rather than an alarm.

All of it on one laptop, at 30.1 fps delivered.

Every figure in this section is measured, dated and traceable to a verification document — including the ones that are not flattering.

Thank You