Catching the Invisible

A Deep Learning Case Study in High-Velocity CDN Leeching

Anand Vijayan | Product Lead, Media CDN

Google Cloud

From Stream Capture to Infrastructure Hijacking

LEGACY PIRACY (EXTERNAL)

Pirates capture video output and pay to host/distribute it on their own offshore servers.

Stack Impact: Zero

MODERN PIRACY (INTERNAL LEECHING)

Pirates harvest valid manifest URLs. Your edge servers deliver the bytes to unauthorized users.

Stack Impact: Maximum

The Economics of Infrastructure Abuse

$29.2B*

Global Annual Impact of Egress Abuse

- **CDN Leeching:** Leverages existing delivery capacity without revenue attribution.
- **Infrastructure Transfer:** Unauthorized manifest exploitation transfers underlying expenses to the provider.
- **Egress Inflation:** Artificial traffic spikes disrupt legitimate traffic prioritization.

[Lost Revenue] +
[Egress Fees] +
[Compute Waste]

} **Piracy Tax**

* U.S. Chamber of Commerce https://www.uschamber.com/intellectual-property/quick-take-your-primer-digital-piracy-and-its-impact-the-us-economy

The Validation Gap

PAYLOAD SECURITY

Payload Encryption (DRM)

DRM recognizes the content bits, but it is agnostic to the delivery path. Unauthorized entities successfully hijack egress even when content is encrypted.

Does not prevent delivery resource theft.

DELIVERY SECURITY

Session Validation

Moving to Token/Dual-Token Authentication determines session validity in real-time.

> **The Static Conflict:** CG-NAT and Private Relay render traditional IP-filtering obsolete for session hardening.

Strategic Pivot: Profiling Normalcy

Transitioning from tracking who the user is to how the session behaves in flight.

Current: Identity Tracking

Relies on reputation scores and identifying bad patterns from known actors. Susceptible to IP rotation and rotating proxies.

New: Behavioral Analysis

Analyzes variables like network hopping and total usage to identify outlying sessions in real time

Intent Verification via CMCD

The Player's "Heartbeat"

By leveraging the Common Media Client Data (CMCD) open standard, we use real-time player telemetry to advocate for the user's legitimacy.

> "Don't just block; Verify intent. This protects the consumer experience by proving the session is bound to a physical player, not a headless script."

- **Transparency:** No proprietary black-box signals.
- **Resilience:** Bots cannot easily simulate player buffer dynamics.

Signal	Legitimate Human	Headless Bot
Buffer (bl)	Steady/Cyclic	0 or Static
Throughput	Consistent w/ Bitrate	Max/Bursty
Intent	Playback Flow	Data Harvest

CMCD allows us to see the 'Why' behind the egress request.

"Whack-a-Mole" as a Service (WhaaS)

Rule Engine Fatigue

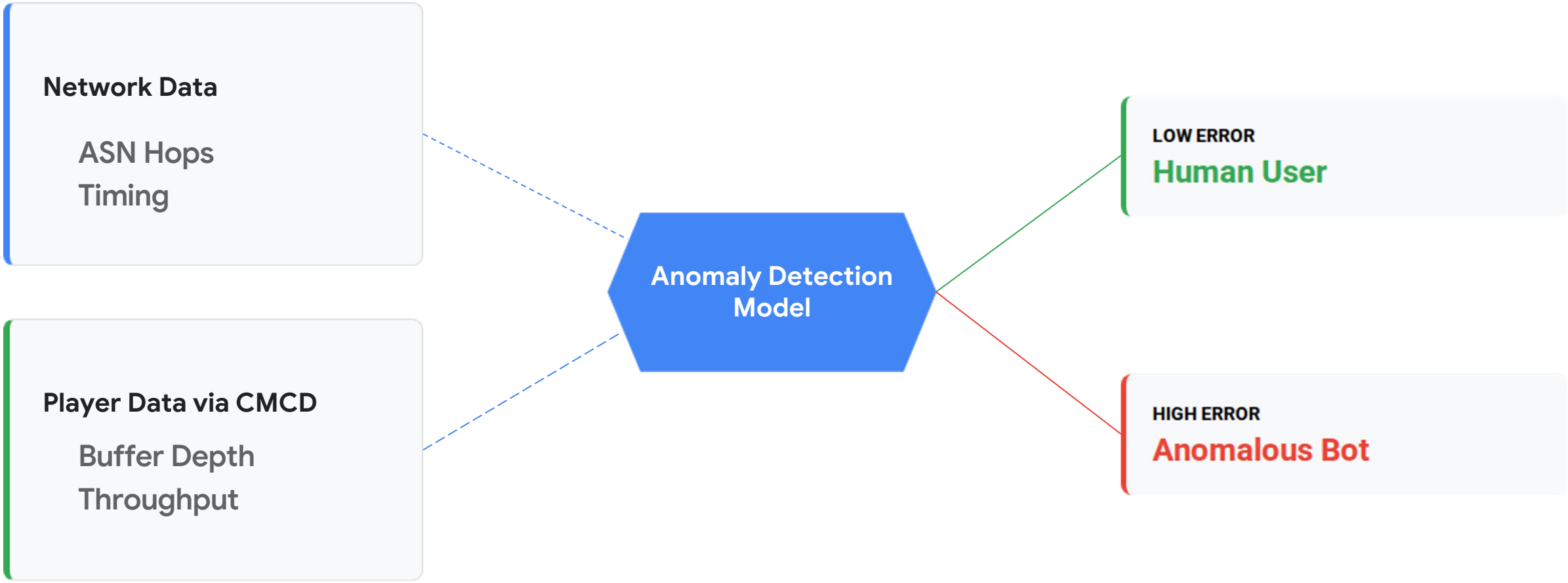
Traditional hard-coded logic is unsustainable for large-scale events:

- ⚠ Hard to morph threats in real-time.
- ⚠ Creates reactive "War Room" scenarios.
- ⚠ Unsustainable operational overhead.

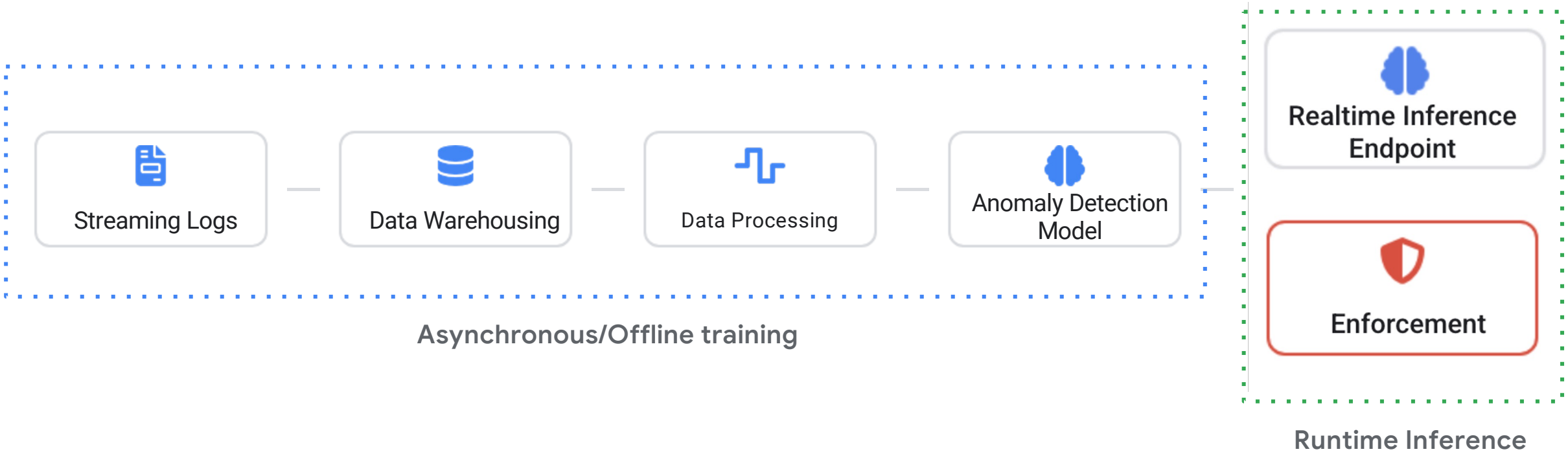
The Driver for ML

Evaluating unsupervised models to construct legitimate sessions.

Anomaly detection model: The Trust Engine



Asynchronous Training, Runtime Inference



Real-time sequential processing for low latency behavioral enforcement.

Operational Calibration

>98%

Precision

<0.001%

False Positives

50%

Recall Rate

Aim for high precision. Prioritize legitimate subscriber continuity over aggressive detection coverage.

Deployment Stages

PHASE 01

Synthesized Baseline

Global deployment establishing broad behavioral boundaries for valid media fetch rhythms across the network.

PHASE 02

Workload specific Tuning

Targeted calibration for specific workload profiles and unique network conditions to optimize efficacy.

It is challenging to stop sophisticated syndicates from extracting content keys. **But you can absolutely stop paying the cloud bill to deliver them.**

Thank you

Gracias Merci Danke 谢谢 धन्यवाद Grazie ありがとう