



Multi-Video CDN System

Liang Ma, Haowei Yuan
Pinterest

Feb 2024, MVH '24

Agenda

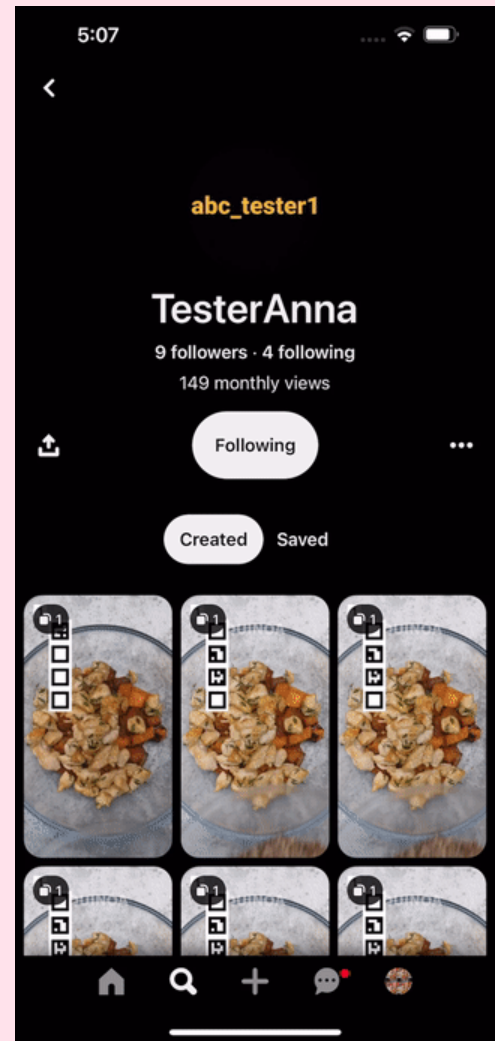
- ❑ Pinterest video overview
- ❑ Multi-CDN GSLB and challenges
- ❑ Future work

Video in Pinterest Product

- Global Monthly Active Users (MAUs): **498 million** (Pinterest, Global analysis, Q4 2023)
- We support both organic video and ads video in large-scale, on iOS, Android and Web.
- Video content on Pinterest has grown **245%** year over year (Pinterest internal data; Global; Q1 2022 vs Q1 2023).
- New videos ingested to the platform every day (Pinterest internal data; Global; Q3 2023).

Video Experience

- Playback Performance
- Perceived Quality



Video Playback Performance

Good Video View (GVV)

- Measure video playback performance
- components: start up latency, rebuffering and player fatal errors

Definition:

Key Component	Condition	App
Player Fatal Error	False	All
Rebuffering Rate	< X %	All
Startup Latency Threshold	< Y ms	Native (iPhone & Android Mobile)
	< Z ms	Web
Video Watched Duration	> 0 ms	All

Performance (GVV) by CDN

Table 1. GVV by CDN (US)

CDN	good_video_view	percentage
vendor_a	74.6%	medium
vendor_b	77.2%	high
vendor_c	72.4%	low

Table 2. GVV by CDN (India)

CDN	good_video_view	percentage
vendor_a	70.3%	medium
vendor_b	70.5%	medium
vendor_c	69.2%	low

One day data from Android devices, aggregated by CDN vendors:

- Different vendors have different GVV rates
- Same vendor may have different GVV rates in different region/country
- Same vendor may have different percentage of traffic in different region/country

Good Video View by CDN - cont'd

Table 3. Video start up time by CDN (US)

CDN	startup_latency_ms(p50)	p90	p95
vendor_a	206	731	1101
vendor_b	195	665	1010
vendor_c	211	840	1301

Table 4. Video start up time by CDN (India)

CDN	startup_latency_ms(p50)	p90	p95
vendor_a	237	881	1380
vendor_b	236	914	1368
vendor_c	244	1024	1769

Let's look into start up time which is a major component of GVV:

- vendor b has better start up latency in the US
- vendor b and vendor a have similar performance in India

Multi-CDN GSLB

GSLB (Global Server Load Balancing) for CDN selection

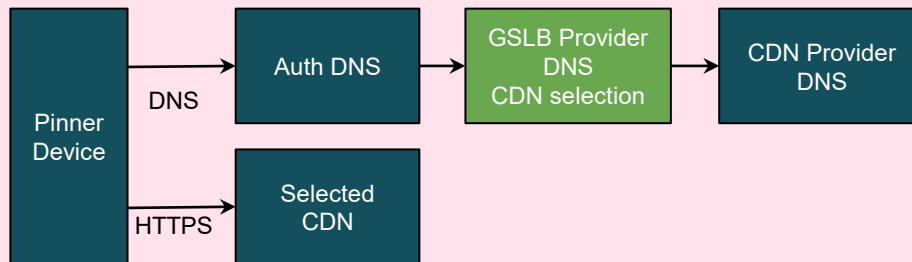
- Availability
- Performance
- Costs
- Real User Monitoring (RUM)

DNS-based GSLB decision delivery

- Simple to manage
- Short TTL to allow fast traffic shift

Request flow

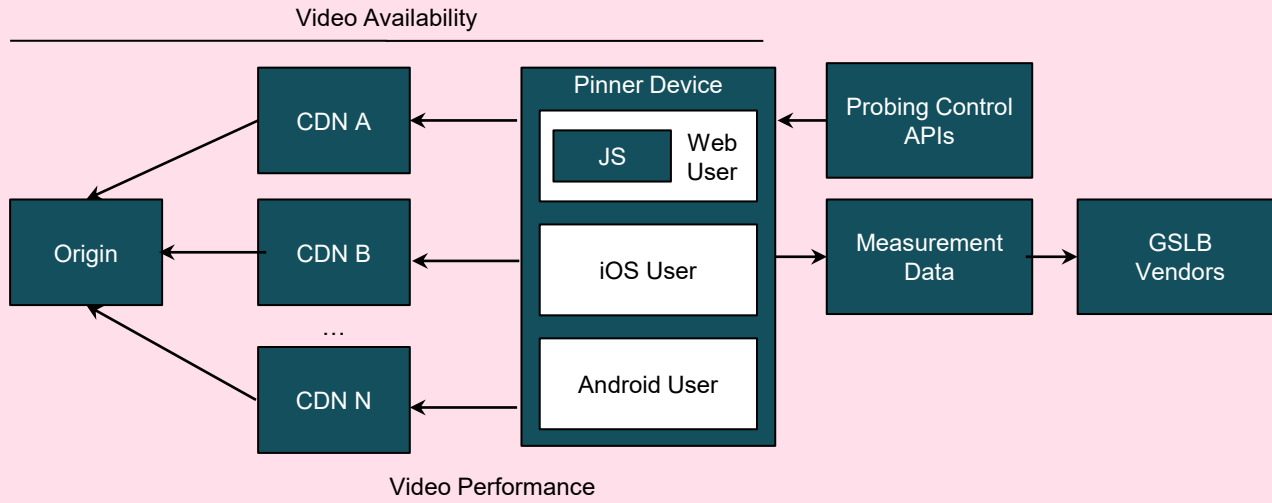
- DNS resolved via CNAME chain, obtaining CDN IP
- HTTPS connected to selected CDN IP



Real User Monitoring

RUM metric data collection

- We measure **availability**, **latency**, and **bandwidth** from Pinner devices
 - Availability: Pinner to CDN and CDN to origin (in case of cache misses)
 - Latency and Bandwidth: Pinner to CDN
- Limiting probing overhead
 - Probing is performed on sampled clients (controlled on server side)



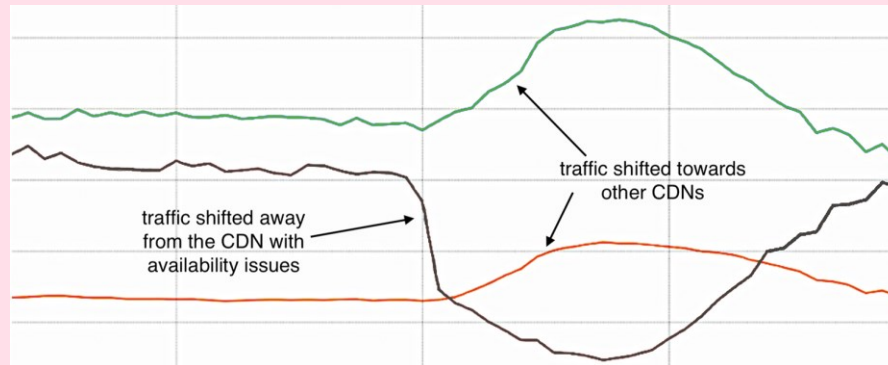
GSLB for CDN Selection Details

Approach

- RUM data aggregated at **country/region/ASN** levels
- CDN selection logics
 - First, **filter out unavailable CDNs** (availability, zero rating)
 - Then, **prefer better performing CDN(s)** based on specified thresholds
 - Lastly, **weighted allocation** when performances are similar
- Tuning performance thresholds and CDN allocation weights
 - **cost** is an important factor as video has highest bandwidth

GSLB traffic shift example

- Traffic shifted away from a CDN that had temporary availability issues



traffic distribution when one CDN had an outage

GSLB Challenges

Quantifying CDN selection quality

- What are the ideal CDN selections?
 - A/B testing alternative CDN selections

RUM signals

- Currently uses only availability and latency
- Bandwidth & video signals not considered yet

Multiple GSLB providers

- Pros: improves reliability
- Cons: inconsistent decisions
 - Quantifying impact of the inconsistency issue

Multi-CDN Challenges

Configuration management

- Same feature set, multiple configurations
 - Continuous monitoring
- Unified access log ingestion

Staged feature rollout in CDNs

- Inconsistent behaviors when traffic shifts among CDNs
- Example: HTTP/3 for video traffic
 - Overhead added when shifts from h3-enabled to h3-disabled CDNs
 - Solution: lowered **max-age** in **alt-svc** header
 - More in [our blog post](#)
 - <https://bit.ly/pins-h3>

Future Work

Future work

- Continue improving video performance and perceived video quality
- Evaluate CDN selection decision quality
- Provide video realtime performance/quality signals to GSLB so it could adjust accordingly
- Generate CDN selection maps in house
 - Integrate with GSLB DNS
 - Alternative GSLB decision delivery methods (e.g., API)
 - Allowing client-side video CDN selection

Q&A