How to Deliver Broadcast-Quality Streaming Video at a Sustainable Price

FEBRUARY 2017

www.cedexis.com

Executive Summary

Over the last three to five years, streaming video has emerged as the key disruption in the $400B global television market. The meteoric growth of online services (including especially Netflix, but also HBO Now, Hulu, BBC iPlayer, and many others) has forced the entire industry to invest heavily in keeping up with an initially unanticipated consumer demand.

With consumers routinely paying real money for subscriptions, or consuming network-level advertising loads, providers are under pressure to deliver broadcast-quality service. Doing so has required an evolution of delivery strategies: from single origin, to single CDN, through multi-CDN, to today's increasingly pervasive hybrid CDN. At every step of the way, costs have mounted, making it extremely challenging to build a self-sustaining business.

This White Paper seeks to identify the elements of consumer experience that are at the heart of building customer loyalty; demonstrates the need to build multi-provider delivery networks in order to provide world-class quality of experience; and proposes a roadmap for using data to ensure that the experience delivered to consumers is financially sustainable.

Background

Streaming video has evolved from something of a niche offering to a significant, multi-billion dollar global business. Services like Netflix, Hulu, HBO Now, and others are now generating eight and nine figure revenues, while offerings like Playstation Vue and DirecTV Now are pulling consumers away from traditional cable providers. Each of these trends brings both opportunity, and the challenge of delivering a service – across the chaotic and best-efforts-only – Internet, that meets and exceeds consumers' expectations. And those expectations are that the service will at the very least match the quality of TV delivered through satellite or cable services.

In order to achieve that level of quality, publishers must build or acquire tools that enable the tracking and dynamic optimization of video streams. Put another way, a service that is to have any hope of achieving broadcast-quality requires holistic service monitoring, coupled with active network management to counteract the challenges of Internet congestion and micro-outages. Both the monitoring and the proactive adjustments must be based on more than server load or system diagnostics – rather they must be based on the actual experience being received by the consumer. Only real user measurements (RUM) can provide the intelligence needed to understand the quality of experience; and only the actual quality of experience will define the likelihood of customer satisfaction.

In its earliest instantiation, the streaming video industry has struggled to agree upon a core set of measurements that can be broadly used to track, monitor, and establish benchmarks for quality. While a disparate array of companies has offered a range of reads on startup times, buffering (and re-buffering), and stream interruption rates, it has, to date, been tricky to compare outcomes across publisher services using different tools to track user experiences. For many publishers it has been possible to track how their service is doing – but not to put it into context; to be able, as it were, to ask the question "is that a good result?"

Defining the Metrics

An industry group, including representatives from major publishers and service providers, called the Streaming Video Alliance, is seeking to address this lack of consistency. In a recent paper, the Quality of Experience (QoE) working group proposed a series of standardized streaming metrics. By agreeing to use the same metrics, based on the same definitions, each tool and service provider can compare their outcomes, and place each service into usable context. The metrics proposed are:

- **Video Start Time (VST)**: The Video Start Time attribute refers to the time, in milliseconds, that it takes for the first frame of video to be rendered after the Play Initiation.

- **Play Failure Count**: The Play Failure Count is the absolute number of play failures observed in the period.

- **Play Failure Ratio**: Play failures ratio is play failures divided by total play attempts observed in the period.

- **Exit Before Video Start Count (EBVS Count)**: The Exit Before Video Start Count is the absolute number of abandoned streams where the video connection attempts without registered errors that have terminated before the first frame of the video has been displayed within the period.

- **Exit Before Video Start Ratio (EBVS Ratio)**: The Exit Before Video Start Ratio is the Exit Before Video Start Count divided by total play attempts observed in that period.

- **Playing Bitrate**: The Playing Bitrate attribute specifies the playback bitrate of the media being rendered, in Kbps. The value is calculated as the video stream playback bitrate + the audio stream playback bitrate.

- **Labeled Bitrate**: The Labeled Bitrate attribute is the value passed to the video player, labeling the bitrate of the media being rendered, in Kbps.

- **Client Video Player Resolution**: The Client Video Player Resolution is the size of client video player.

- **Playback Video Resolution**: The Playback Video Resolution attribute specifies the height and width resolution as well as scanning type (progressive or interlaced scan) at which the video is currently playing on the client.

- **Re-buffering Ratio**: The percentage of time that a viewer experiences re-buffering issues (i.e. when video stops playing because of buffer underflow, and not due directly to user intervention such as scrubbing or pause).

- **Re-buffering Count**: The Re-buffering Count is a count of re-buffer events that a viewer experiences during a period of video playback.

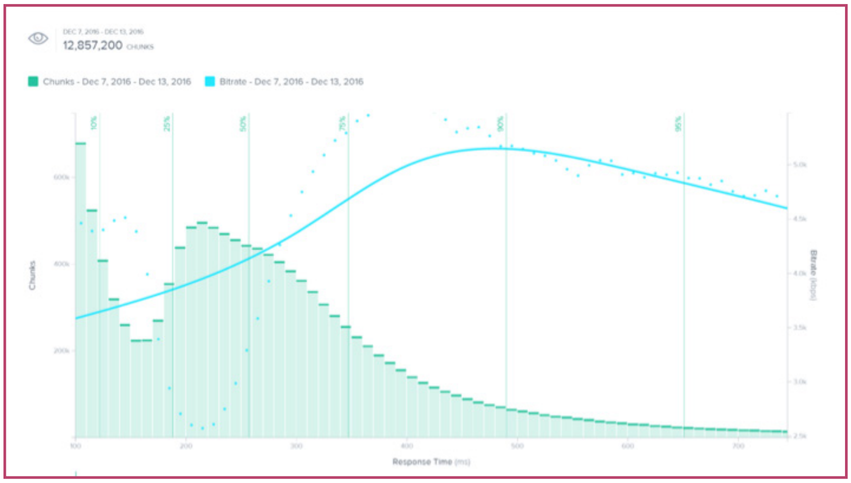
- **Dropped Frame Count**: The Dropped Frame Count is a count of dropped frames that a viewer experiences during a period of video playback.

- **Video Transition Time**: The Video Transition Time is the amount of blank screen time between one clip ending and another clip starting.

- **In-stream Error Count**: The In-Stream Error Count is the count of errors occurring during the playback experience by the player (i.e.: after the first frame is rendered).

Network metrics like these allow services to acquire an actionable view of the QoE being experienced by their consumers. Note that they are all defined as counting the experience being enjoyed by the real user, rather than by system resources: from the perspective of the consumer, the technology assets in use to make the experience possible are irrelevant.

Sometimes the insights that can be gained from tracking these metrics do an excellent job of confirming what we would expect. In the following graph, for instance, we see a high correlation between short response times and high bitrates:

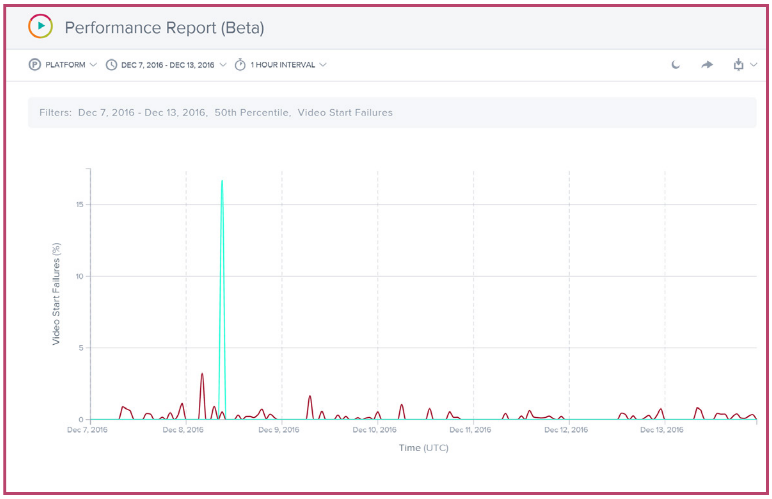


As we might expect, when there is enough free bandwidth to respond quickly, there is plenty of bandwidth to deliver high bitrates.

The Case for Combining Multiple Delivery Networks

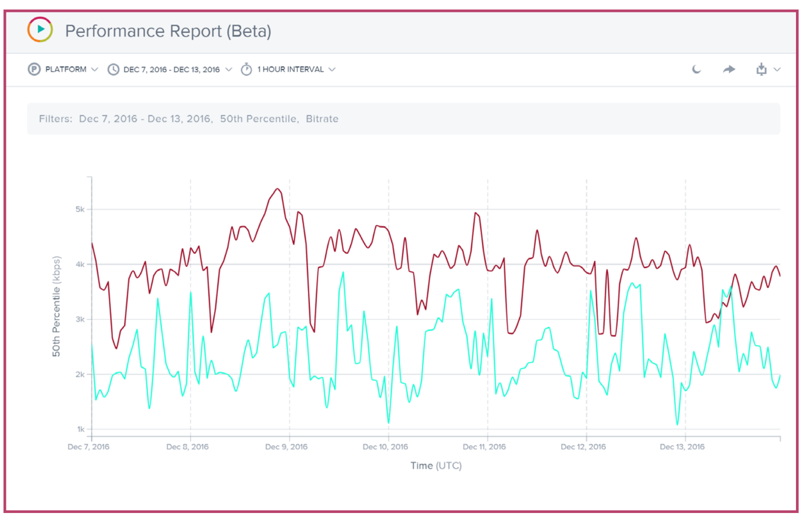
As we dig in, and particularly when combining different dimensions – particularly when narrowing our observations to particular geographies, CDNs or ISPs – we can reach confounding conclusions that challenge our business intuition.
Put another way, it is clear that no individual provider can be the top player in every geography, at every time of every day. At any point, delivery is improved by combining the capabilities of multiple providers.

Let's take a look at some of the data that illustrates this reality, and how they might guide investment decisions. The following graphs come from an actual Cedexis client, and compare the results tracked for each of two major global CDNs. First let's take a look at Video Start Failures:



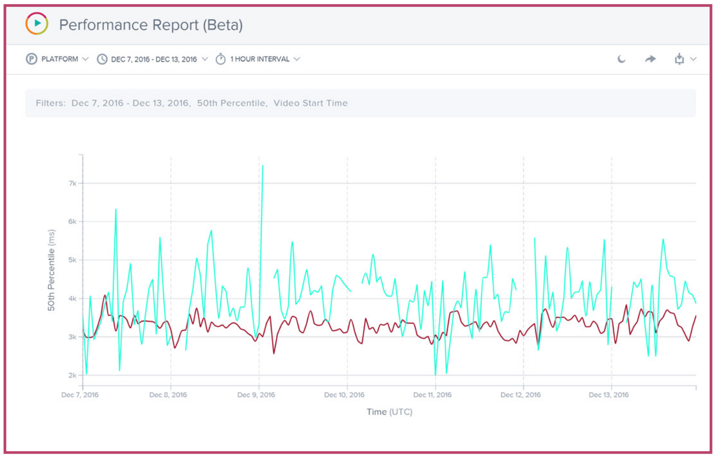
The first CDN, shown in red, is exhibiting a lot of inconsistency: while the video start failures never reach a particularly poor level, they turn up regularly. By contrast, the second CDN, shown in green, almost always has close to zero video start failures. It does have one large spike, which could have caused some consumer dissatisfaction, though, so we should dig deeper.

The next metric to look at is Playing BitRate, which will let us know which of the CDNs is delivering the highest video resolution to consumers:



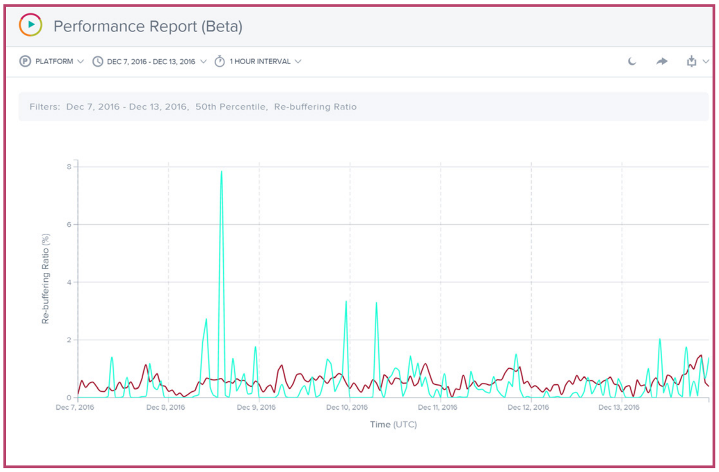
The CDN in red is consistently delivering a higher bitrate – there is only one very short period where its competitor can match it, quality for quality. Depending on the business model of the Cedexis client, the bitrate may matter more or less; but we can see that a decision on which CDN is better, based on our first two measurements, is going to be slippery.

So let's move to another measurement: Video Start Time. Generally speaking, the more quickly a video commences, the more consumers will stick with it and generate ad views or build brand loyalty to an SVOD service (Akamai coined the two second rule years ago).



Perhaps unsurprisingly, we see greater consistency with the red CDN than the green – and lower start times. Which is not to say there is no time that the CDN in green doesn't periodically have something to offer – look about two thirds of the way into the graph and you can see it providing a quicker start time.

Finally, let's take a look at the reigning champion of consumer frustrations: re-buffering, also known as those maddening moments when the video stops playing and the consumer waits for their player to refill its buffer. Research has convincingly demonstrated that nothing chases a viewer more swiftly than playback interruptions.



Strikingly, although the more consistent CDN is once again the one depicted in red, the CDN in green provides lower re-buffering rates as often as not. This may very well be connected to the earlier graph showing it delivering lower bitrates – a lower bitrate means a smaller stream, which can therefore keep the buffer filled re easily.

Fundamentally, the message is that neither CDN is consistently superior to the other. And without being able to track the actual user experience in clearly-defined ways, we would never have known that fact. At any given moment in time, we now know, either might be preferable. This matches Cedexis' ongoing findings (found at cedexis.com) that the fastest CDN provider for any publisher is…some combination of at least two. In any given geographical location there is normally a dominant player, who is best situated to deliver great service for the majority of the traffic; but it will periodically be challenged by traffic volume and congestion, and need back up from a secondary partner. Streaming video providers whose audience is global will likely find that, to deliver acceptable service to all their consumers, they will need more than two. To get a real-time, or historical, view of how traffic is moving around the Internet, take a look at Radar Live, which is the quickest way to see where congestion and micro-outages are impacting QoE.

Using Data to Deliver Quality Cost-Effectively

With the growing availability of reliable, easily-implementable private cloud technology (including caching servers from vendors like Varnish Software), it is now often worthwhile for publishers with far-flung audiences to consider building their own self-managed CDNs, or private clouds. Such a strategy is known as Hybrid CDN, and is finding increasing traction, particularly among publishers seeking to not only deliver amazing quality, but also do so at a cost that is tightly managed. Using holistic monitoring, publishers can identify the geographical locations (or other audience subsets) that are underserved by all the CDNs, and establish their own Points of Presence (POPs) nearby – then clearly demonstrate the improvement in QoE using the same monitoring.

And if monitoring can help publishers to make business decisions around CDN provider selection, its use can also revolutionize global traffic management. A regular Global Server Load Balancer (GSLB) will re-distribute Internet traffic between server resources based on metrics like server load and packet delivery; essentially using the health of the network elements as a proxy for Quality of Service. But a global traffic manager that can switch between delivery pathways based on the actual QoE at the consumer level will not just protect the technical back end – it will actively ensure that the end-user is satisfied. And satisfied customers are repeat customers.

The Math of Optimizing for Quality and Cost

In the early going, most online video providers focus on delivering a high-quality product to their viewers – high quality is necessary, if insufficient, to building and sustaining a viable audience.

As time passes, however, it becomes clear that to deliver profitability, companies must not only increase revenues, but also reduce costs. The imperative to deliver high quality video experiences cannot be abandoned – but delivery is a large line item in any spreadsheet.

The good news is that the delivery networks improve every day, delivering ever more reliable availability, up-time, and throughput. What this means is that, with the right global server load balancer (GSLB), providers can start to update their traffic shaping strategy: they can now set thresholds above which any provider is acceptable, then select from qualifiers, in real-time, based on cost and contractual terms.

For instance, imagine a company with three CDNs under contract. For the sake of simplicity, let's assume they pay a predictable per-Gigabyte fee of

- **CDN A**: $0.08
- **CDN B**: $0.06
- **CDN C**: $0.03

In any circumstance where any two, or all, can pass a QoS bar (such as a minimum latency level, or throughput), the GSLB should choose C over B or A, and B over A.

Similarly, having solved quality, companies can also start to unwind overage and bursting charges, by instructing their GSLB to direct traffic across providers in such a way that all committed bandwidth is consumed before additional cost is added for new capacity (sudden spikes in traffic, when not combined with an intelligent GSLB, tend to consume unbudgeted capacity on the primary CDN, adding unnecessary costs).

In short, while delivering broadcast quality is a fundamental goal of any streaming video enterprise, the data clearly show that an intelligent GSLB can leverage the same data platform to manage costs downward.

Conclusion

Today's streaming video consumer expects a payoff on their investment of subscription fees, and/or ad consumption: they anticipate a broadcast-quality experience. The data clearly demonstrate that, to guarantee such an experience, it is necessary to use multiple delivery providers, balanced by an intelligent GSLB. Using the same data empowers any business not only to meet and exceed consumer expectations, but to being the key process of managing delivery costs downward, helping to build sustainability and profitability.

To understand the extent of the variability in CDN delivery, please visit the Radar Live site, where you can see outages (from sustained, newsworthy crashes, to short micro-outages that otherwise go unnoticed while they enrage your audience) as they happen, around the world. To see metrics for your own site or service, join the Radar Community: add one simple JavaScript and gain access to powerful reporting, and learn how QoE is impacting your business.

Portland, Oregon
421 SW 6th Ave, Suite 700
Portland, OR 97204
+1 855 CEDEXIS (233-3947)

Paris, France
15-17 rue Vulpian
75013 Paris, France
+33 (0)1 79 755 253

visit cedexis.com or email sales@cedexis.com.