**techn**alysis
R E S E A R C H

# Caching the Cloud

By Bob O'Donnell, Chief Analyst

## SUMMARY

Viewers are increasingly consuming video "on-demand" and moving away from watching content determined by a broadcaster's schedule. For content services and content delivery networks (CDNs), this new usage pattern creates new challenges—as well as opportunities—for their existing delivery infrastructure. Chief among them are quality of service (QoS) breakdowns, which are directly tied to reduced viewing times. In order to address these issues, CDNs need to look at high endurance flash-based storage solutions embedded into consumer on-premise equipment, managed edge devices, and their core networks, all working in concert to "cache the cloud." These solutions can drastically reduce or even eliminate QoS limitations, as well as open up new opportunities for customized advertising, leading to lower attrition and maximizing profitability potential.

> "By utilizing high-endurance flash-based storage solutions in conjunction with algorithms that enable intelligent prefetching of video data over the network and onto set-top boxes—in essence, 'caching the cloud'—CDNs can maximize their potential profitability."—Bob O'Donnell, Chief Analyst

# INTRODUCTION

YouTube, Netflix and Hulu have succeeded wildly in attracting viewership to on-demand video services via various business models, including freemium, Electronic Sell-Through (EST) or an "all you can consume" monthly subscription. This has spurred the content owners, broadcast networks, cable, satellite and telco Pay TV service providers ("video service providers") to innovate and get in on the act. At first, they sought to reduce the attrition in their customer base, but they quickly realized innovation can actually grow their subscriber base and increase their share of targeted viewers.

On average, consumers are watching 45 minutes per day of streamed video[1]. At the higher end data rate of 5Mb/s[2] (e.g., Netflix, full HD), this translates to 50 GB/month of downloaded video. In terms of time spent, these numbers work out to 22.3 hours per month, representing 23% of the 100 hours per month (3.3 hours per day) consumers spend looking at video-related content[3]. By 2018, this percentage is expected to more than double, and will likely reach half of all video viewership hours.

| ACTIVITIES AND TIME | |
|---|---|
| Live TV programming per week | 13.6 Hours |
| DVR/Streamed programming per week | 5.4 Hours |
| You-Tube Style Videos Per Week | 4.3 Hours |
| Total Video Content Watching Per Week | 23.3 Hours |
| % Streamed | 23.2% |
| Total Video Content Watching Per Day | 3.3 Hours |
| Total Video Content Watching Per Month | 99.9 Hours |
| Average Download Rate Per Hour | 2.2 GB/Hour |
| Total Download Per Month of Streamed Video | 50.85 GB |

These trends have magnified the video viewership competition over the Internet, resulting in more and more consumers of all demographics becoming accustomed to watching their own customized programming content. Plus, they're expecting to do so whenever, wherever and on whatever device they choose to.

As a result, content owners, broadcast networks, and video service providers have started to differentiate their services by offering more and more of the content they have rights to, not only on TVs, but on mobile phones, tablets, PCs and gaming consoles. They've also started to

---

[1] TECHnalysis Research, Worldwide Consumer Device Usage Survey, Aug. 2014
[2] Netflix
[3] TECHnalysis Research, Worldwide Consumer Device Usage Survey, Aug. 2014

differentiate on "convenience" features, such as trick play content navigation functions, and sync'ing when, where and how a viewer starts, stops and resumes watching a specific piece of content on different devices.

## THE CHALLENGE: STREAM COUNTS AND CUSTOMIZED ADS

Meeting the current customer requirements and forecasted projections for on-demand video consumption, along with the features consumers have come to expect, brings some particularly notable challenges. Essentially, it requires changing from the current broadcast-focused approach to a more IT-like data environment. Network architects at CDNs and video service providers need to make dramatic changes in network infrastructure and data delivery mechanisms in order to meet these growing new demands.

Not only will growth in on-demand video require the creation and simultaneous delivery of millions of independent streams (along with convenience features such as trick play), it will require customized insertion of personalized advertising. Expectations around advertising in VOD applications are rapidly evolving and increasing, driving yet more requirements onto the typical VOD infrastructure. These two issues—stable content delivery and dynamic ad insertion—will be key to building long-term, successful strategies for video content delivery over the Internet. They are critical for content companies, CDNs and video service providers to avoid attrition and QoS issues, and to maximize profitability.

### Buffering Leads to Attrition

In the case of service delivery expectations to consumers, the numbers are clear. According to Conviva's 2014 Viewer Experience Report, even an increase of just 1% in buffering will lead viewers to watch online video content for an average 11 minutes less than experiences where they don't have these glitches. For live sports, average viewing time drops from over 40 minutes to under 1 minute if they're forced to endure buffering problems. Needless to say, experiences like this could lead to dramatic drop-offs in usage, which in turn will decrease advertising opportunities and greatly increase the likelihood of cancellations and/or switching over to providers who don't suffer from these issues.

On the advertising and revenue front, CDNs, in conjunction with content providers, video service providers and ad networks, will need to develop systems that allow advertisers to deliver timely, targeted ads to the viewers of their content and insert them on the fly. One of the critical differentiators between broadcast and VOD is that advertisers want to be able to target their ads more precisely, as they typically can with web sites. This is good for ad networks, CDNs and video service providers because they can typically charge more for targeted ads. These increased prices can be justified by knowing much more about the

psychographic and social demographic profiles of viewers, as well as when and where they consume each piece of content. The challenge is that inserting more individualized ads puts even greater strains on the network infrastructure, creating more network latency challenges, resulting in unwanted video buffering events for viewers.

## THE SOLUTION: CACHING THE CLOUD

In order to avoid the glitches and video stuttering issues that are typical symptoms of buffering, CDNs and video service providers need to rethink their data center architectures and leverage technologies that can minimize potential delays most cost effectively. One of the best ways to achieve this is through the concept of "caching the cloud."

The basic idea behind "caching the cloud" is to get the video and advertising where it needs to be, before it's actually required. Simple enough in theory. In practice, it turns out there are several different ways to do this, many of which can be deployed simultaneously. Content can be cached, either for "network relief" or for "convenience", at various layers of the delivery network, from the network edge to set-top boxes sitting at the consumer's premise.

### Caching in the Flash

In the case of network storage, the key attributes to the solution are speed and endurance: two core aspects of flash-based solid-state storage. Flash offers customized-speed access to data, making it ideally suited for applications such as sorting, finding, playing, pausing, and moving through digital video content.

In the data centers dedicated to video applications, the speed and reliability of storage are crucial to enable a stable, high performance environment. The important IOPS (Input/Output Operations Per Second) metric determines how quickly storage devices can get access to and read the data they need to create the hundreds to thousands of video streams being requested simultaneously. CDNs and video service providers need to ensure they deploy storage solutions with the IOPS necessary for their specific needs. In addition, video streaming applications require enormous numbers of reads and writes and not all types of storage—not even all types of flash-based storage solutions —are up to the task. To maintain reliability, CDNs and video service providers need to utilize specialized types of high endurance flash storage with high P/E (Program/Erase) cycle capability, in order to ensure long-term, glitch-free operation.

## CACHING IN THE HOME

At the consumer's house, the key to making the experience of watching on-demand content completely glitch-free comes from using flash based storage solutions for both convenience and network relief.

### Caching for Network Relief—Prefetching

To avoid waiting for content to load during periods of high network congestion, video service providers can pre-fetch, during the afternoon, a program that the consumer expects to watch during peak evening consumption hours, and store the program in a high-speed, high endurance local flash "cache".

The intelligence built into many of today's set-top boxes in conjunction with the detailed data gathered and analyzed by many CDNs and video service providers permit the creation of algorithms that can be used to preload a variety of relevant content. For example, the service could preload episodes 4 and 5 of a show for which you've already watched episodes 1-3, or make smart guesses about preloading movies they think you will like based on others you've recently watched or that are on your favorites queue list. The same concept can be applied to targeted ads, which can be loaded in advance on the set-top box, enabling dynamic insertion without experiencing buffering issues. By leveraging these types of algorithms and preloading some of this content, consumers can avoid buffering issues because the content will play from the local device and won't require a potential congestion-filled trip out around the internet.

### Caching for Convenience – Trick Play

When a consumer wants to pause, rewind or fast-forward through content, otherwise known as trick play, it is very difficult to do so effectively over a network connection. Even when the connection is very fast, it places significant IOPS strains on video delivery servers in the network. By placing the streamed video data on a local, high endurance flash storage-based set-top box, CDNs can relieve the workload from those servers.

In addition, by installing high-speed, high-endurance flash storage-based set-top boxes into their top customers' households, video service providers can simplify the network upgrade planning process. Instead of having to upgrade entire swaths of their network when subscription densities, either current or projected, will not meet the ROI expectations of such an upgrade, they can target these specific households as they roll out new services by leveraging the improved capabilities of these devices.

# THE HDD CHALLENGE

One obvious question that needs to be answered is why choose flash over traditional hard disk drives (HDDs)? After all, HDDs are more cost effective on a $/GB metric compared to flash-based storage solutions. While this is certainly true on storage devices where content is infrequently accessed, the advantages of HDDs quickly disappear in applications with heavy usage of random reads and writes, such as on-demand video delivery.

For frequently accessed content, HDDs cannot deliver on the IOPS necessary to address the high throughput and random nature of the streams that are being transferred from network to consumer on-premise equipment. Plus, the size, power footprint and reliability metrics all point to flash-based storage as more attractive when considering total cost of ownership.

In home-based set-top boxes, flash-based solutions offer a better match to the expected evolution of on-demand video services. Many video service providers are moving to a Cloud DVR infrastructure, where they are going to stream all content to the consumer's premise. For those types of cloud-based services, a smaller footprint, likely lower capacity but higher speed, higher reliability and higher endurance storage solution is what will be demanded, and those attributes are a perfect match for flash storage.

## Why High Endurance Flash?

Another important issue to address is the need for high endurance vs. standard flash. On average, set-top boxes are on for 6 hours per day at a data rate of 10Mbit/second, including HD content being recorded and pre-fetching of other content. Taking into account a set-top box's average circulating lifetime of 7 years in a service provider's customer base, that works out to nearly 66TB of video being written to it over its lifetime.

| ENDURANCE AND LIFETIME | |
| --- | --- |
| Set-top Box On Time Per Day | 6 Hours |
| Recording Rate | 10 Mb/Sec |
| Set-top Box Lifetime | 7 Years |
| Total Lifetime Storage | 65.8 TB |
| Average Flash Set-Top Box Capacity | 8 GB |
| Total Flash Rewrites Over Lifetime | 8,421 |

If that average set-top box has 8GB of flash storage, the memory in it will need to endure more than 8,400 writes over its entire lifetime. In comparison, retail-grade flash devices are based on cost-optimized, three bits per cell flash bit density with an endurance of less than 200 cycles. While this is acceptable for all but the most demanding mobile phone and camera applications, write intensive video applications require much higher endurance. As a result,

these video solutions need a unique type of flash that is more commonly found in enterprise grade solutions. For video service providers, the set-top boxes they loan out to consumers must last much longer or they risk expensive repeated inquiries from unhappy or frustrated consumers who have experienced a feature that has failed, as well as the added expense of replacing the flash storage device.

## CONCLUSION

The move to online video delivery is afoot and with that exciting development comes both a variety of opportunities and several important new challenges. For content providers and CDNs, the opportunity is to deliver a more highly targeted service that can match the unique interests and needs of all their individual customer subscribers. At the same time, there's also the chance for greater advertising revenues, leveraging the kind of targeted ads now in regular use on the web.

Despite this increase in flexibility and choice, however, basic customer demands have not changed. Having been trained on traditional broadcasting delivery, they expect their video experience to be immediate and glitch-free. If they don't get that level of quality, research shows that they will reduce their viewing and/or switch to different providers. Eyeballs follow quality and if the quality of a video service suffers, people will move.

In order to avoid attrition and maximize their profitability potential, CDNs need to create infrastructures that can deliver a "buffer-free" experience all the way from the network down to the set-top box. One of the key ways to achieve this is by leveraging high-speed, long-lasting, reliable storage both in their data centers and on their client devices. By utilizing flash-based storage solutions in conjunction with algorithms that enable intelligent pre-fetching of video data over the network and onto these devices—in essence, "caching the cloud"—CDNs can deliver on their potential. They can avoid expensive costs associated with the attrition of customers to other services and they can maximize their customized advertising opportunity. In addition, by investing intelligently in storage, they can avoid more brute force capital acquisition costs that simply attempt to expand capacity with traditional tools.

With the dramatic increases in on-demand video delivery that are expected to occur over the next few years, the time is right for video service providers to shift from their broadcast mentality to a more IT-like perspective and start integrating high endurance flash based storage solutions into their data centers and set-top boxes to meet those evolving needs today.