A Longitudinal View of HTTP Video Streaming Performance

Louis Plissonneau
Orange Labs
louis.plissonneau@orange.com

Ernst Biersack
Eurecom
erbi@eurecom.fr

ABSTRACT

This paper investigates HTTP streaming traffic from an ISP perspective. As streaming traffic now represents nearly half of the residential Internet traffic, understanding its characteristics is important. We focus on two major video sharing sites, YouTube and DailyMotion.

We use ten packet traces from a residential ISP network, five for ADSL and five for FTTH customers, captured between 2008 and 2011. Covering a time span of four years allows us to identify changes in the service infrastructure of some providers.

From the packet traces, we infer for each streaming flow the video characteristics, such as duration and encoding rate, as well as TCP flow characteristics. Using additional information from the BGP routing tables allows us to identify the originating Autonomous System (AS). With this data, we can uncover: the server side distribution policy, the impact of the serving AS on the flow characteristics and the impact of the reception quality on user behavior.

A unique aspect of our work is how to measure the reception quality of the video and its impact on the viewing behavior. We see that not even half of the videos are fully downloaded. For short videos of 3 minutes or less, users stop downloading at any point, while for videos longer than 3 minutes, users either stop downloading early on or fully download the video. When the reception quality deteriorates, fewer videos are fully downloaded, and the decision to interrupt download is taken earlier.

We conclude that *(i)* the video sharing sites have a major control over the delivery of the video and its reception quality through DNS resolution and server side streaming policy, and *(ii)* that only half of the videos are fully downloaded and that this fraction dramatically drops when the video reception quality is bad.

Categories and Subject Descriptors

H.5.1 [**Information Systems Applications**]: Multimedia Information Systems

Keywords

Measurement, Performance, HTTP Streaming

1. INTRODUCTION

In this study, we analyse HTTP streaming traffic. Nowadays, Web driven content represents about half of the Internet traffic due to the surge of many video sharing sites, and the decrease of P2P [8, 14, 15, 16].

The main video sharing sites in Europe are YouTube, DailyMotion and MegaVideo. They provide a free online video sharing service, which is very popular for sharing user generated content and also video clips (videos are at most 10 minutes long on YouTube). Moreover, the streaming traffic is interactive in the sense that the user is actually watching the video during download and not after download completion as it is the case in P2P. The data is transmitted using TCP, which is not designed for interactive usage, but for elastic traffic.

We analyse the streaming traffic from an ISP perspective: this means that our main focus is on the ISP customer's perception of the Internet. The difficulty in HTTP streaming traffic analysis is that not only the network characteristics and the TCP congestion control mechanisms play a role in the user's viewing experience, but also the video sharing site itself.

We study different points that impact the video stream at flow level and relate them to user perceived interruptions. We use passive packet captures and vary important factors such as: the network access type (ADSL[1] *vs.* FTTH[2]), the video sharing site (mainly YouTube *vs.* DailyMotion) and the time of the day (lightly loaded afternoon *vs.* highly loaded evenings).

HTTP streaming works as follows: when a user wants to watch a video on a video sharing site, he first selects the video, *e.g.* by browsing the site portal or by receiving a direct link. Then at least 3 HTTP sessions (over TCP) are established: *(i)* download of the embedding web page; *(ii)* download of the video player (only once in a session); *(iii)* download of the video itself. The success of these video sharing sites comes from the fact that the user can start watching the video after a very small buffering period (typically several seconds). The rest of the video is downloaded while watching, therefore its name progressive download (PDL).

[1]ADSL (Asymmetric Digital Subscriber Line) is the main Internet access type for European residential customers.

[2]FTTH (Fiber To The Home) is a technology offering access rates up to 100 Mb/s, which is currently being deployed in Europe.

A user can also abandon downloading and watching if she is not interested any more.

We start by reviewing related work in Sect. 2 where we also show in what sense our work differs from previous studies. As the main user interest and most of the volume generated by HTTP streaming comes from the actual video download, we focus on this part and we explain in Sect. 3 how we identify the video flows. Then, we give some general information on video sharing sites in Sect. 4. In Sect. 5, we focus on the flow level network indicators to assess the reception quality of the videos watched by the users. Finally in Sect. 6, we highlight the impact of streaming quality on the user *download behavior*. We conclude the paper in Sect. 7 with a summary and an outlook.

2. RELATED WORK

Most related work on video sharing sites focuses on YouTube, which is the most prominent video sharing site. There is no previous work to compare YouTube with its competitors such as DailyMotion.

2.1 Characterisation of YouTube Videos

Many studies have tried to find out the characteristics of YouTube videos compared to *e.g.* Web traffic or traditional streaming video sites (real time over UDP and not PDL). In [4], the YouTube video meta-information are crawled to derive many characteristics on the video contents and its evolution with video age (*e.g.* popularity). This information is used to evaluate opportunities of P2P distribution and caching.

In [5], the authors use a long term crawl of the YouTube site to derive global characteristics of YouTube videos such as the links between *related* YouTube videos form a small-world network. Using the properties of this graph and the video size distribution, they show that P2P distribution needs to be specifically adapted to distribute YouTube videos.

In [11], University campus traces are used to gather information on YouTube video characteristics and are complemented with a crawl of most popular files on YouTube. Temporal locality of videos and transfer characteristics are analyzed, and the opportunities for network providers and for service providers are studied. Another work of same authors [10] uses the same campus traces to characterize user sessions on YouTube showing that the think time and data transfered by YouTube users are actually longer than for Web traffic.

2.2 YouTube CDN Architecture

Some recent papers study the global architecture of the YouTube CDN[3]. In [1], the authors explain with Tier-1 NetFlow statistics some of the load-balancing policies used by YouTube and use these measurements to figure out traffic dynamics outside the ISP network. This method is used to evaluate different load-balancing and routing policies. Even if the methodology still holds, the data collected for this work was taken before heavy changes in YouTube infrastructure in the second half of 2008 (two years after Google bought YouTube).

The same authors study the YouTube server selection strategy [2]. Using PlanetLab nodes to probe and measure YouTube video transfers, this analysis shows that YouTube is using many different cache servers hosted inside their network or by other ISPs. Some of the load-balancing techniques used by YouTube are also revealed in this paper.

In the same vein, the authors of [21] use recent traces from different countries and access type (university campus *vs.* ADSL and FTTH on an ISP networks) to analyse the YouTube service policy. In most cases, YouTube selects geographically close server except when these servers are heavily loaded.

The details of YouTube video streams at TCP level have been studied in [20]. This analysis of residential ISP datasets shows that the bursty nature of the YouTube video flow is responsible for most of the loss events seen.

We also would like to mention this work on load-balancing in CDNs [18] where the answers of the CDN operator to DNS queries are stored at ISP level in order to bypass recursive DNS resolution by the CDN operator and directly answer to the DNS queries of customers with an IP address chosen by the *ISP instead of the CDN operator*. The evaluation of this mechanism shows an improved performance, *e.g.* download time are reduced by up to a factor of four. This work shows that a cooperation between the CDN operator and the ISPs could not only be beneficial to these actors but also to the users. In a similar vein, the study of YouTube [21] also illustrates the importance of DNS resolution in server selection and how video sharing sites (and more generally CDN) use it to apply complex load-balancing strategies.

2.3 YouTube User Experience

In a recent study [9], ISP packet data are used to derive some results on the video configurations (such as resolution, full screen playback...), or the interruption of videos at user's side. Finally, the amount of unnecessary data transfered is evaluated over fixed and mobile networks.

Our work mainly differs from this work the way we evaluate how the users watch the video (fully or not) relate to the network quality and the video duration. We are thus able to infer that the main factor for a user to interrupt the video is the quality.

2.4 Novelty of our Work

We use ten different packet traces to answer a number of important questions such as:

- Do the different video sharing sites enforce download limitations on their streams and do these limitations change over time?
- How does the YouTube CDN perform server selection for the clients of the ISP and what is the implication on the reception quality?
- How do users of video sharing sites view videos and is their viewing behavior affected by the reception quality?

Our work differs from the previous work on video sharing sites in several important aspects: *(i)* Instead of characterizing all the videos available on the YouTube servers, we analyse the characteristics of *videos actually watched by users. (ii)* We analyse video transfer characteristics to explain the performance of HTTP video streaming. *(iii)* We compare two video sharing sites, namely YouTube and DailyMotion, which is one of its popular competitors. This comparison

[3]CDN: Content Delivery Network

Table 1: Traces description

Type & Location	ADSL M	FTTH M	ADSL M	FTTH M	ADSL R	FTTH M	ADSL R	FTTH M	ADSL R	FTTH M
Date	**2008**/07	**2008**/07	2009/11	2009/11	2009/12	2009/12	2010/02	2010/02	2011/05	2011/05
Start time	20 *h*	20 *h*	20 *h*	20 *h*	20 *h*	**14 h**	20 *h*	20 *h*	20 *h*	20 *h*
Duration	1 *h* 30	1 *h*	1 *h* 20	0 *h* 38	1 *h*	0 *h* 58	1 *h*	0 *h* 28	1 *h*	0 *h* 25
Active Web/Str. users†	1121	1198	650	2502	795	2009	607	2763	944	4525
Streaming users§	109	121	96	336	113	252	74	279	153	514
Streaming videos	428	630	405	1462	334	865	258	866	470	1396
YouTube users§	41	30	49	185	47	106	46	153	45	146
YouTube videos	215	142	210	660	140	400	176	496	153	590
DailyMotion users§	25	20	16	49	12	20	13	29	6	13
DailyMotion videos	83	154	45	84	53	35	25	44	7	50

† with at least 10 flows (Web and Streaming)
§ watching at least 1 video

reveals a number of interesting differences, both w.r.t. performance aspects and the way these two video sharing sites serve requests for videos. Moreover, the distribution policies of these two sites differ a lot, leading to an interesting discussion of design choices for existing video sharing sites. *(iv)* Our traces cover the time from 2008–2011, which allows us to measure the impact at network level of the modification in the infrastructure of the YouTube CDN that was put in place in the second half of 2008. *(v)* We show that in our traces, the server chosen to stream YouTube videos is often not the closest one (in terms of RTT) or the one that assures the best video reception quality. These results are not in line with previous measurements [11, 21, 2]. *(vi)* We are the first to investigate what fraction of a video users actually download and we are also able to show that poor reception quality affects the fraction of the video downloaded.

3. TRACE CHARACTERISTICS

The main source of information for our analysis is IP packet captures taken at Broadband Access Server (BAS) level of a large European ISP. We have performed multiple packet captures at different locations. The data consists of ten approximately one hour snapshots collected on ADSL and FTTH probes from 2008–2011. The probes are equipped with dedicated capture cards (Endace DAG® card). Users[4] have been anonymised at capture time by hashing their VP/VC (ATM identifier) for ADSL and the MAC address of OLT (Optical Line Termination) for FTTH. Note that the capture reports of the cards ensure that *no packets have been lost during the capture.*

To focus on streaming flows, we first filter on the `content-type` field of HTTP header using the same regexp as in [15]. We also remove all non-video flows such as embedded player download and advertisement contents by filtering out the keyword `player` in the resource name or respectively well know advertisement URLs. We process packet traces with tools to extract flow information including RTTs and losses.

We have a specific tool to process streaming traffic that extracts relevant information about the content (mainly URL, size, and duration of the video) out of the HTTP and video headers.

[4]IP address is not used because it is not sufficient to identify users [15].

We have enhanced this data with information from BGP routing tables collected at the time of capture at the ISP level, which allows us to accurately map the IP addresses of streaming servers onto their Autonomous System (AS). Most TCP traffic indicators have been derived via an internal packet processing tool and some loss indicators have been calculated using the `tstat` software [22].

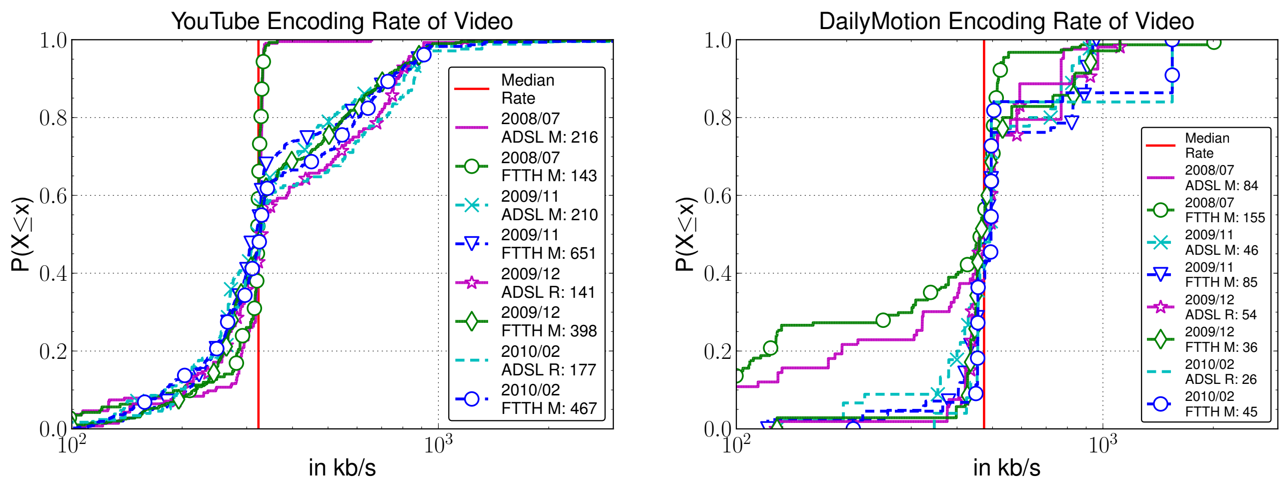
The details of the packet captures are given in Tab. 1. We have two *old* traces from July 2008, and eight traces taken in 2009, 2010 and 2011. After the acquisition of YouTube by Google, changes to the architecture of the YouTube CDN occurred in the end of 2008. We are able to see the impact of these modifications in our data (mainly the switch from the YouTube AS to the Google AS and to a new YouTube EU AS). Since then, there are constant adjustment of load repartition between these two AS. Note that FTTH M 2009/12 trace has been taken in the afternoon, whereas all the other traces have been captured in the evening, which is the period of highest network load for residential customers. Traces are captured in two geographically different locations and labelled with their access type and location indication. We label traces taken at a central site near the *Main* ISP peering point with an **M**, and with an **R** those taken at a *Regional* site. Note that the 2011 traces do not include the streaming payload, thus we were not able to recover the encoding rate for these traces. Thus we do not include these traces in streaming quality study (Sect. 6).

4. HTTP STREAMING CONTEXT

Due to the prominent usage of HTTP streaming [14], this traffic is important for ISPs in terms of resources required inside the ISP and at the peering points. After a brief description of the most popular video sharing sites, we evaluate the video encoding rates of the main video sharing sites as it is a key factor of the video quality and network resource consumption. Then, we briefly explain how DNS resolution works as it will be useful for the further analysis. Finally, we give an example of the distribution of the traffic across the different ASes of the YouTube CDN.

4.1 Most Popular Video Sharing Sites

The most popular video sharing sites in our traces are



(a) YouTube (b) DailyMotion

Figure 1: Video Encoding Rates for YouTube and DailyMotion

YouTube followed by DailyMotion and MegaVideo aggregating respectively 30%, 14% and 11% of the total streaming volume. The rest of the streaming volume comes from TV channels offering *replay* of their programs and porn sites. Note that the total streaming downstream volume represents about 40% of the total traffic for ADSL traces and about 30% for FTTH traces.

YouTube is the most popular video sharing site all over the world with more than 100 Million viewers per month just in US [6]. It has been bought by Google in November 2006. One of its major competitors in Europe is DailyMotion which was among the top 50 websites in the world at the time of analysis [3].

4.2 Video Encoding Rate

We are interested in video encoding rates to have an idea of the necessary mean reception rate of a flow required to watch the video without interruptions. The trend of encoding rates gives also an interesting insight into content providers choices and adoption of higher quality formats. We compute the encoding rate by dividing the content size announced in HTTP header and the content duration announced in the video header (FLV, MP4, 3GPP). Note that this is coherent with the bit-rate announced in the video header but easier to compute with multiple video formats. Video encoding rates are quite standardized inside a video sharing site (see Fig. 1).

Median encoding rates of two most popular video sharing sites are quite close: for YouTube, the median encoding rate of 330 kb/s is slightly lower than for DailyMotion, which is 470 kb/s. Encoding rates for YouTube vary quite a lot and many YouTube videos use encoding rates above or below the median rate (see Fig. 1(a)). Encoding rates of DailyMotion videos in recent years show little variance and the majority of videos have an encoding rate equal to the median rate (see Fig. 1(b)).

4.3 Domain Name System (DNS)

As explained in [1], retrieving a YouTube video begins with a connection to the YouTube web server that returns the URL of the video stored in the YouTube data center (cache server URL *e.g.* `v7.lscache1.c.youtube.com`). This URL is then resolved via a DNS lookup, which returns the IP address of a server delivering the video.

Load balancing techniques by the operator of the video sharing sites can be applied via DNS resolution: the recursive nature of DNS resolution allows the DNS server of the domain to take into account internal policies to answer with the most appropriate server IP address. If a content is highly requested and replicated (as are videos of main sharing sites), the DNS server of the video sharing site can choose to redirect the *same* URL to one of several servers. This technique can be used to balance the load but also to better take into account network path characteristics (*e.g.* return the server closest to the user). As we will see later, the same URL can even be resolved to IP addresses in different Autonomous Systems (AS), which may greatly impact the flow characteristics (see Tab. 2 and 3).

4.4 Distribution of Traffic across ASes

We present in Tab. 2 the main characteristics of the ASes providing YouTube and DailyMotion videos. The measured delay corresponds to the round trip time *from the probe towards the server and back*, also referred to as upstream RTT and defined in Sect. 5.1. We see that the *former* YouTube AS (36561) is no more used after 2008. The YouTube EU AS (43515) streams the majority of the bytes in all 2009–2010 traces, which is quite different to what was observed in previous studies [21, 2] that had identified the Google AS (15169) as the one serving most of the streams. Note that we measure for the YouTube EU AS an upstream RTT in the order of 100 ms, which corresponds to the RTT between Europe and the East Coast in the US. The Google AS, which also serves YouTube videos, has a much lower upstream RTT between 20 and 40 ms. Other ASes (Cable&Wireless and Global Crossing) are also used for stream-

Table 2: Distribution of Volumes (in percent) and delays (median value of minimal upstream RTT per flow) in milliseconds per AS for YouTube and DailyMotion

	YouTube										DailyMotion			
	GOO AS 15169		YT EU AS 43515		YT AS 36561		C&W AS 1273		GBLX AS 3549		DM AS 41690		LL AS 22822	
	Vol.	RTT	Vol.	RTT	Vol.	RTT	Vol.	RTT	Vol.	RTT	Vol.	RTT	Vol.	RTT
2008/07 ADSL M	39%	21	–	–	61%	113	–	–	–	–	67%	2	33%	14
2008/07 FTTH M	–	–	–	–	100%	114	–	–	–	–	61%	1	39%	14
2009/11 ADSL M	1%	21	90%	114	–	–	5%	21	4%	117	100%	2	–	–
2009/11 FTTH M	1%	20	91%	108	–	–	5%	215	3%	106	100%	1	–	–
2009/12 ADSL R	7%	32	93%	116	–	–	–	–	–	–	100%	14	–	–
2009/12 FTTH M	20%	20	80%	101	–	–	–	–	–	–	100%	1	–	–
2010/02 ADSL R	32%	38	56%	126	–	–	9%	29	3%	52	100%	14	–	–
2010/02 FTTH M	18%	25	60%	110	–	–	19%	24	3%	108	100%	1	–	–
2011/05 ADSL R	100%	17	–	–	–	–	–	–	–	–	100%	14	–	–
2011/05 FTTH M	100%	2	–	–	–	–	–	–	–	–	100%	2	–	–

Table 3: Distribution of number of distinct YouTube ASes per client for clients with at least 4 YouTube videos

Trace	Total # ASes	# distinct ASes per client 1	2	3	4
2008/07 ADSL M	3	33%	53%	13%	–
2008/07 FTTH M	1	100%	–	–	–
2009/11 ADSL M	3	65%	5%	30%	–
2009/11 FTTH M	4	71%	14%	12%	4%
2009/12 ADSL R	2	50%	50%	–	–
2009/12 FTTH M	2	58%	42%	–	–
2010/02 ADSL R	3	21%	53%	26%	–
2010/02 FTTH M	4	13%	53%	21%	13%
2011/05 ADSL R	1	100%	–	–	–
2011/05 FTTH M	1	100%	–	–	–

ing YouTube videos, but only marginally. In 2011, the YouTube RTT is very small and all videos ares served from the same AS: this shows the instability of the load-balancing of videos among Google ASes. Also, this table nicely illustrates the evolution of the YouTube infrastructure since Google bought YouTube.

Previous work [21] has shown that the server selected by the YouTube CDN for streaming the video is usually the closest one to the user with notable exceptions only at peak hours. For our traces, this finding does not hold since the AS that is farthest away is used to serve the majority of videos (up to 90% in terms of volume the for 2009/11 and 2009/12 traces).

In Tab. 3, we see that the same client can be directed to a different AS when requesting multiple videos, even in a timescale of one hour. As this redirection mechanism happens via DNS, the video sharing site has full control to select the AS and the server that will stream the video. We shall see in Tab. 4 that the choice of the originating AS has a significant impact on the video reception quality. Ongoing work evaluates the impact of the DNS resolver on the achieved streaming performance [13].

In the case of DailyMotion, almost all videos are served by the DailyMotion AS (41690) which has a median delay of $2\,ms$ over all traces (resulting in a total RTT of $42\,ms$ on ADSL and $7\,ms$ on FTTH). The only exception is found in our 2008 traces where about 1/3 of the videos were coming from the LimeLight AS (22822) with a median delay of $14\,ms$.

5. FLOW PERFORMANCE INDICATORS

In this section, we measure various metrics such as RTT, mean rate, or loss rate in order to understand the performance experienced by the flows. One of the novelties of our analysis is that we compute all these metrics for each different AS that host servers of the video sharing site, which allows us to reveal the existence of considerable performance differences between different ASes of the *same* video sharing site. In all the graphs, the number after the label in the legend indicates the number of data samples (videos).

5.1 Round Trip Time

Round Trip Time (RTT) is defined as the time between the emission of a data packet and the reception of its acknowledgement. In order to get an idea of the distance between the client and the server, we use the minimum of all the RTT measures of a flow. As the probe that captures the packets is located between the customers and the server, we separate the RTT in two parts:

upstream RTT delay from the probe towards the server (in the Internet) and back;

downstream RTT delay from the probe towards the local user and back.

As the infrastructure between the probe and the remote site is the same on different access types, this allows to compare

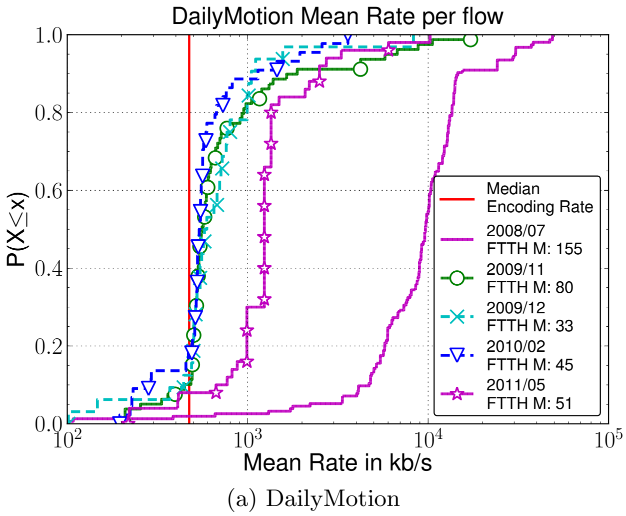


Figure 2: Mean Flow Rate of Videos for FTTH M traces

the distance to streaming servers across the two different access types, ADSL and FTTH. Note that *real* RTT between the client and the server is the sum of the upstream and downstream RTT.

As the probes are quite close to the clients, downstream RTTs are very short and very stable (also for non-streaming flows). The CDF of downstream RTTs (between the BAS and the customer), not shown here, reveals that:

- almost all ADSL video flows have approximately $40\,ms$ of downstream RTT;
- whereas FTTH video flows have downstream RTTs between $1\,ms$ to $5\,ms$.

For upstream RTTs, we often have a CDF that is multimodal, which can be explained by looking at the AS of the flows. The RTT of the two main ASes used by YouTube differ in their upstream RTT by almost one order of magnitude (around $110\,ms$ for YouTube EU AS *vs.* $20\,ms$ for Google AS). Also for each AS, the RTTs are very stable showing little variance. Note that the results are similar for the all the traces.

5.2 Mean Flow Rates

In this section, we focus on the mean flow rate of video transfers, which is defined as:

$$\text{mean flow rate} = \frac{\text{total flow volume}}{\text{total flow duration}}$$

Mean flow rate is an important metric as it is related to the user perceived quality as we will see later (*cf.* Tab. 4).

In Fig. 2, we plot the CDF of the mean flow rate of YouTube and DailyMotion. We also plot the median video encoding rate for each site to be able to compare the reception rates with the standard encoding rate. As we are interested in server side limitations, we only plot the mean rates for FTTH M traces, as they are much less likely to be limited by their access speed. Flow mean rates are generally not very high: few videos achieve rates above $1\,Mb/s$ even for FTTH traces.

5.2.1 DailyMotion Mean Flow Rates

As for DailyMotion in Fig. 2(a), we have very homogeneous mean rates in all traces that show a large accumulation point just above the median video encoding rate at $500\,kb/s$, except for 2008 and 2011 traces. Thus, there is a mean rate limit for DailyMotion videos set slightly above the median video encoding rate. Note that in 2011, this mean rate limit is clearly higher: this is probably due to the default high quality policy for DailyMotion videos. While such a choice of the rate limit should allow for a correct reception (and viewing) quality for most of the videos, the reception can be very sensitive to any network problem that may cause the reception rate to fall below the encoding rate for some limited time. In the FTTH M trace of 2008, we see that the mean rate limit originally was higher at about $12\,Mb/s$.

Such modifications in the rate limitation policies made by the video sharing sites are usually not known in advance to the ISP.

5.2.2 YouTube Mean Flow Rates

In Fig. 2(b), we can see that the policy concerning the mean rate limitation of YouTube has evolved over time. For the 2008/07 trace, there is a sharp mean rate limit at $1.2\,Mb/s$ that has been previously observed [17].

Such a limitation of mean rate (and peak rate not shown here), as in the case of YouTube, was *most likely* implemented using a well-known open-source rate limiter, the Token Bucket Filter over Hierarchical Token Bucket (HTB [12]) with two buckets (one limiting peak rate and one limiting mean rate). Note that YouTube uses a **new distribution policy** since 09/2010: there is a high rate delivery of the beginning of the video (or of the requested part in case of a jump inside the video) followed by a delivery rate equal to the encoding rate of video. In the 2011 traces, the mean rate resulting of this new distribution policy (as shown by 2011/05 FTTH trace in Fig. 2(b)) has a similar shape and is higher in this case (due to lower losses see Sect. 5.3.1).

The FTTH M 2009/12 afternoon trace achieves average

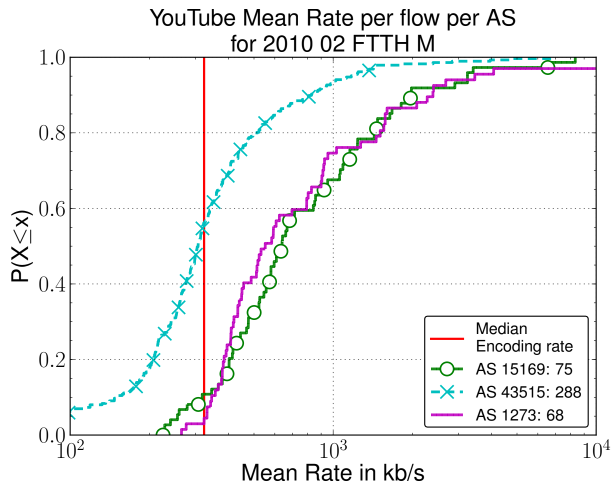


Figure 3: Reception Rate of YouTube videos per serving AS for FTTH M 2010/11 trace

flow bit-rates superior to the median video encoding rate for 95% of the videos. As for mean rate, the shape of the graphs does not allow to infer any mean rate limitation.

For the traces taken in the evening, around 40% of the videos achieve a mean reception rate that is inferior to the median encoding rate. The curves are concave with no clear limitation. As we will see later (in Tab. 4), such low reception rates result in bad reception quality. In 2011, the mean rates are very good also in the evenings: probably due to the infrastructure change (see Tab. 2).

In Fig. 3, we closer look at the different ASes used by YouTube in 2010. There is no indication for server side limitation of the mean rate, as it was the case in 2008. The CDF of the mean rates is concave for all ASes. Even if the shape of the CDF for mean rates is similar among ASes, the YouTube EU AS (43515) clearly has lower mean rates: in most of the cases, almost 50% of the videos achieve a mean reception rate below the median encoding rate. We will come back to this point when we discuss in detail the video quality (see Sect. 6.2).

Studying the achieved mean rates has allowed us to understand the distribution policies used by the two main video sharing sites. We have seen that videos from the same video sharing site achieve very different mean rates (independently of the ISP policy) depending on the AS delivering the video.

5.3 Loss Rate

There are different ways to estimate losses at packet level, depending on *where* the loss happens.

retransmitted packets, *i.e.* packets carrying a sequence number already seen; for the flows in downstream direction this allows to measure **access loss**;

out of order packets, *i.e.* packets with an unexpected sequence number ($<$ min{*seq. nb.*} or $>$ max{*seq. nb.*} + *pkt_size*) but not retransmitted; for the flows in downstream direction this allows to measure **Backbone loss**.

As the probe is located at the BAS level, all packets from/to the customers of the ISP *must* pass through, which ensures that our measures are not biased by multiple paths taken by the packets.

As for video streaming most of the data are transmitted from server towards the client, we focus on losses between the server and the BAS, which is referred to as Backbone loss. The access loss rates of most of the flows are below 1% (details are not shown here for space reasons).

In Fig. 4, we look at Backbone loss for YouTube and DailyMotion.

5.3.1 *YouTube Loss Rate*

In the case of YouTube, it is interesting to understand how the AS connectivity to the ISP can greatly influence the loss rate. The CDF of the **Backbone loss rate**, which is defined as the ratio of the number of packets lost in Backbone to the total number of packets, is shown in Fig. 4(a). If we focus on the 1% loss region, in all traces (except the 2008 traces and the 2009/12 afternoon FTTH M trace) between $60 - 80\%$ of the flows experience more than 1% packet loss along the path from the server to the capture point. For a TCP connection, the throughput achieved is inversely proportional the square root of its loss rate. Accordingly, the mean flow rate of all the YouTube flows with more than 1% Backbone loss is only $285\,kb/s$ (including FTTH flows), whereas the median encoding rate is $330\,kb/s$ (Fig. 1(a)).

A threshold of 2% on loss rates allows us to discriminate traces in Fig. 4(a). For example, 2009/12 FTTH M and 2008/07 ADSL M are the only traces where the large majority of flows have less than 2% loss rate. We will see in Tab. 4 that these are also the only traces with consistently good reception quality.

Note that 2011 traces have similar loss rates than other traces.

5.3.2 *DailyMotion Loss Rate*

We show in Fig. 4(b) the CDF of Backbone loss rates for DailyMotion. Most flows see less than 1% upstream loss rate. We also plot the 2% loss rate, which is adequate to discriminate DailyMotion videos according to their reception quality. Indeed, both of the ADSL R traces (2009/12 and 2010/02) encounter much more losses (above the threshold of 2%), and we shall see in Tab. 4 that these are exactly the traces where a lot of videos experience a bad reception quality.

6. USER BEHAVIOR STUDY

In this section, we want to study how users view videos and also how users adapt their viewing behavior in response to bad reception quality.

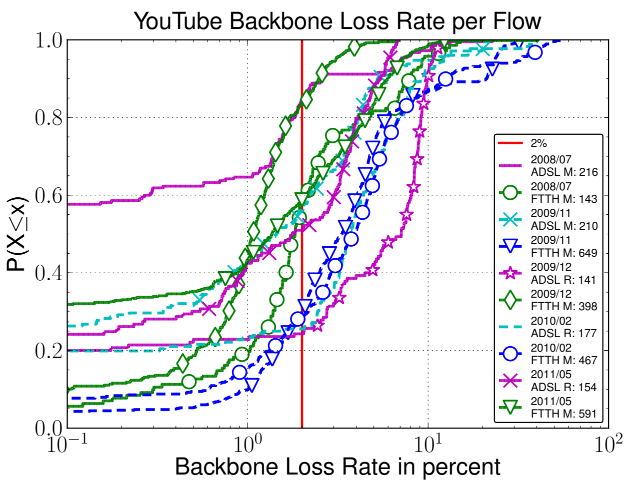
First, we globally measure how much of a video the users download. Then, we define a simple metric for user experience to differentiate videos with good reception quality from others. We then evaluate the fraction of *wasted* bytes during streaming transfer. Finally, we relate this indicator to the fraction of video downloaded to the fact that a user has completely downloaded the video, and to the video length.

6.1 Downloaded Duration

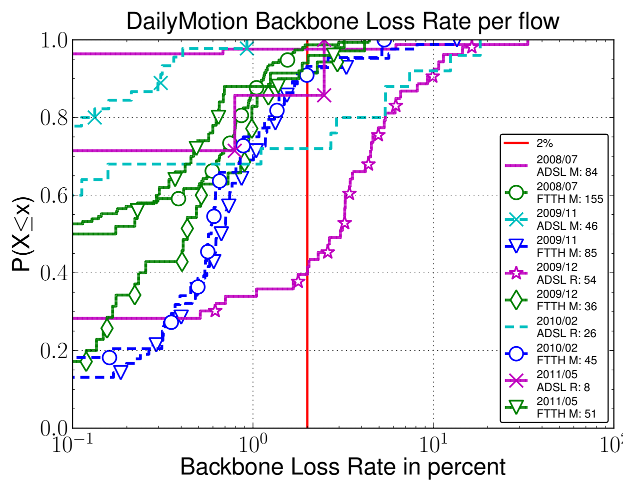
Ideally, we would like to know how much of a video the user is actually watching. However, as the video interactions (like pausing/resuming the video) are not transmitted to the server, we cannot retrieve them at network level. Instead, we approximate how much of a video a user watches by how

Table 4: Fraction of Videos with Bad Reception Quality (normalized rate ≤ 1)

	YouTube					DailyMotion	
Trace	YT EU AS 43515	GOO AS 15169	YT AS 36561	C&W AS 1273	GBLX AS 3549	DM AS 41690	LL AS 22822
2008/07 ADSL M	–	1%	7%	–	–	4%	49%
2008/07 FTTH M	–	–	18%	–	–	1%	3%
2009/11 ADSL M	49%	0%	–	47%	50%	11%	–
2009/11 FTTH M	34%	0%	–	88%	5%	12%	–
2009/12 ADSL R	74%	50%	–	–	–	30%	–
2009/12 FTTH M	6%	10%	–	–	–	15%	–
2010/02 ADSL R	68%	45%	–	56%	80%	20%	–
2010/02 FTTH M	52%	17%	–	1%	69%	8%	–

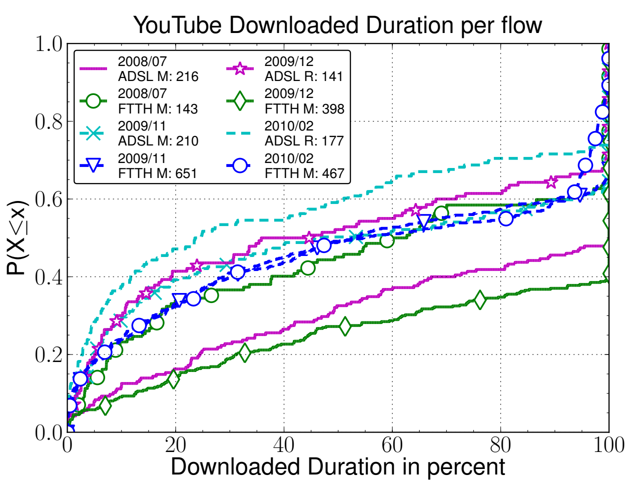


(a) YouTube

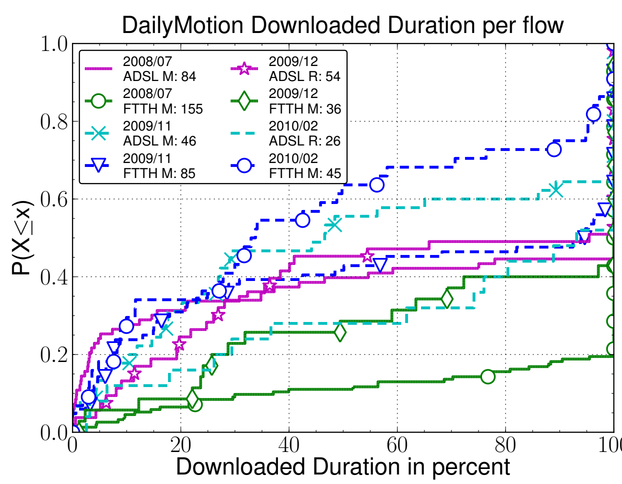


(b) DailyMotion

Figure 4: Backbone Loss Rates



(a) YouTube



(b) DailyMotion

Figure 5: Fraction of Video Downloaded per Trace

much of the video she has downloaded, which provides us with an *upper bound* on how much she has watched. For instance, the fact that a video has been fully downloaded does *not* mean that the user watched the video completely (if the video was paused and never resumed).

We define the downloaded duration as:

$$\text{downloaded duration} = \frac{\text{size of downloaded flow}}{\text{video encoding rate}}$$

We have also checked that **the distribution of video length does not change for our different traces**. Also, the distribution of video durations for the videos watched by the ISP clients matches the size of videos as seen in [11]: the most frequent durations are videos of 3–4 minutes videos (most likely video clips).

In Fig. 5, we plot the CDF of the fraction of the video downloaded for each trace for YouTube and DailyMotion. We see that the distribution can vary a lot among traces. Globally, *not more than* $40 - 60\%$ *of the videos are completely downloaded*. Such behavior seems to indicate that progressive download induces users to "browse" videos without necessarily watching each video to the end.

Focusing on YouTube in Fig. 5(a), we first notice two traces with a much higher fraction of videos that are completely downloaded: the 2008/07 ADSL M and 2009/12 FTTH M, which are the traces with the lowest loss rates (see Fig. 4(a)). In all the other traces, the fraction of videos that is fully downloaded is 40% or less.

As for DailyMotion (see Fig. 4(b)), the difference in download behavior among the traces is more pronounced than for YouTube. We have no good explanation why this is the case, except that there are fewer samples in each trace for Fig. 4(b). The FTTH M 2010/02 trace has the largest number of completely downloaded videos (80%), which is significantly more than what we have seen for YouTube.

Since many of videos are not downloaded, and thus not watched, until the end, we want to understand the reasons: Is it lack of interest, bad reception quality, or video duration?

6.2 Simple User Experience Metric

To "measure" the user experience we want to know if the video was interrupted during playback. First, we define the normalized rate for each video as:

$$\text{normalized rate} = \frac{\text{mean flow rate}}{\text{video encoding rate}}$$

In lack of a better metric, we say a video has *good reception quality* if its normalized rate is above 1, and *bad reception quality* otherwise. We admit that this is quite a crude measure. However, we have done several controlled lab experiments under different network conditions. We have recorded both, the packet traces and the occurrences of video playout interruptions, and have found that the normalized rate is a reasonable indicator for the reception quality. We report in Tab. 4 the video quality for YouTube and DailyMotion per AS streaming the video.

What is striking is that for YouTube, the reception quality depends a lot on the AS that serves the video. Many videos coming from the YouTube EU AS (43515) have a bad reception quality. If we relate this to the traffic distribution given in Tab. 2, we see that the AS that serves most of the YouTube videos for this particular ISP is the one providing the worst performance. YouTube videos coming from other ASes usually have a good reception quality.

The afternoon trace FTTH M 2009/12 is the only one with a good reception quality for streams served from the YouTube EU AS. This makes us conclude that in the evening hours there are not sufficient bandwidth resources along the path from the YouTube EU AS to the ISP. Note that the situation has much improved in 2011 (see Fig. 4(a)) where all the requests are served from a single AS.

In the case of DailyMotion, the reception quality among traces is much more uniform. In the 2008 ADSL trace, the LimeLight AS (22822) had much lower reception quality than the DailyMotion one. For DailyMotion, the time of day has no impact on reception quality as the afternoon FTTH M 2009/12 trace does not have better performance than the other FTTH traces.

The case of the two ADSL R traces is worth considering separately: for both traces, the reception quality of a large number of videos coming from either the YouTube EU AS or from DailyMotion is bad. Since videos being served by the other ASes are not particularly affected, these seems to indicate that some of links internal to the ISP are congested.

6.3 Wasted Bytes

We compute how much bytes have been downloaded but not watched with this formula:

$$\text{fraction of wasted bytes} = 1 - \frac{\text{Video Encoding}}{\text{Mean Download Rate}}$$

We compute the fraction of wasted download only for "good" flows (*i.e.* the ones with higher download throughput than encoding rate).

In Fig. 6(a), we see that the 2008 traces have the highest waste: indeed at that time, there were much higher throughput limits for YouTube (resulting in a median of about 70% of waste). We note a median fraction of wasted bytes at about 35% except for the 2010 afternoon FTTH trace with almost 50% of wasted bytes. Then traces with higher loss rates result in fewer wasted bytes (both ADSL R traces).

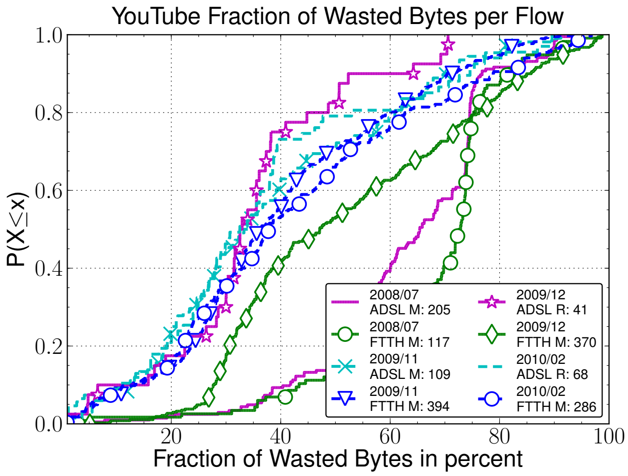
In the case of DailyMotion with a mean flow rate just above encoding rate (see Sect. 5.3.2), the fraction of wasted bytes is much lower ($0-20\%$) except for the 2008 traces (with more than 80% of wasted bytes for all flows). Note that as the distribution policy limits the rate, there is no difference in wasted bytes between ADSL and FTTH traces. This was not the case in 2008 with unlimited transfer rates.

Trying to limit the amount of wasted bytes may be detrimental to the quality. However, as we see in the case of DailyMotion, many bytes are saved and the reception quality is better than YouTube one. This seems to indicate that with a carefully calibrated distribution policy, waste can be kept very low while the reception quality is good.

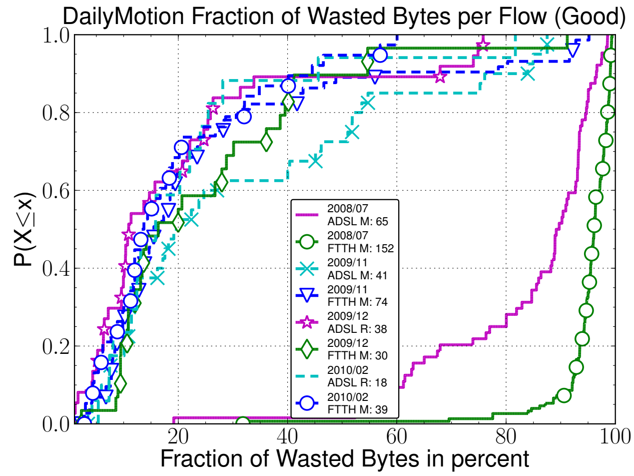
6.4 How do Users watch Videos

In this section, we want to understand why users decide to interrupt video downloads. We first analyse the downloaded duration in function of content duration. Then, by discriminating on the reception quality, we are able to see that videos with bad reception quality have much shorter downloaded durations than others. Moreover, the decision of interrupting the download is taken very quickly for videos with bad quality.

6.4.1 Relation of Video Length to Reception Quality



(a) YouTube



(b) DailyMotion

Figure 6: Fraction of Wasted Bytes for Flows with Good Reception Quality

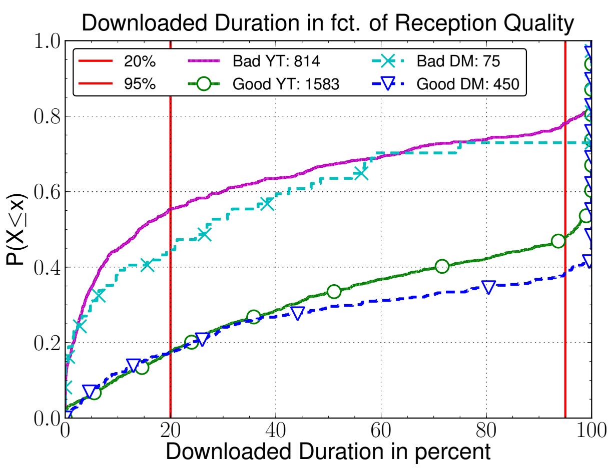


Figure 8: Fraction of Video Downloaded as function of Video Reception Quality

In Fig. 7, we plot for each video the fraction of the video downloaded in function of the video length.

We see that videos with good reception quality have more complete downloads than videos with bad reception quality. To analyse the graph, we first look at the videos that have download durations of less than 3 minutes, which make up the majority of the videos, independently of the reception quality. The usage for YouTube seems to be either to download less than 3 minutes or to download the video completely. In case of good reception quality, about three times as many videos longer than 3 minutes are fully downloaded (28%) as when the reception quality is bad (11%).

6.4.2 *Relation of Download Duration to Reception Quality*

As far as the downloaded duration is concerned, we can see in Fig. 7 that in case of good reception quality, 34% of the videos have a downloaded duration of 3 minutes or more, while in the case of bad reception quality their share drops to only 15%.

In Fig. 8, we relate the *fraction* of the video downloaded to the video quality. Again, we clearly see the impact of the reception quality on the downloaded portion of the video. The results for both sites, YouTube and DailyMotion are very similar. We have clearly two zones:

- completely downloaded videos ($\geq 95\%$);
- videos for which only a small portion has been downloaded ($\leq 20\%$).

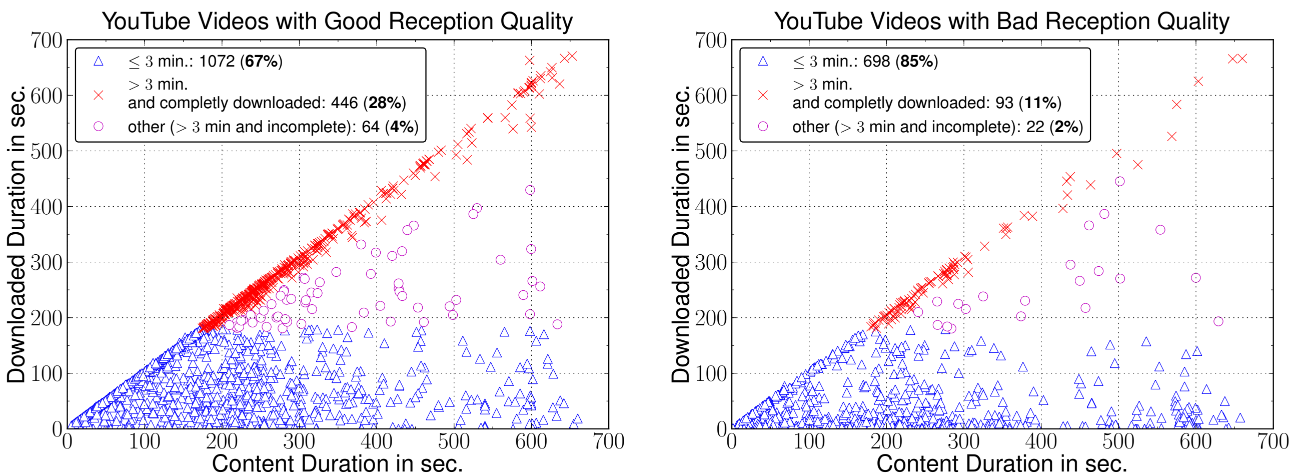
In case of bad reception quality, very few videos are completely downloaded. Moreover, the decision to stop downloading a video is taken quickly (in the first 20% of the video duration).

In case of good reception quality, about half of all videos are completely downloaded and the decision to interrupt download is not taken right from the beginning but at any point during the viewing.

We have seen in Tab. 3 that in the case of YouTube, the same user, when requesting multiple videos, will be served with high probability from machines that are located in different ASes. This observation leads us to carry out one more analysis in order to validate that there exists a positive correlation between video reception quality and the fraction of the video downloaded.

For the FTTH M 2010/02 trace, we take all clients that meet either one of the following two conditions: clients having received at least one video with good reception quality *(i)* from both AS 43515 and AS 1273, or *(ii)* from both AS 43515 and AS 15169.

In Fig. 9, we plot for all the clients that meet condition *(i)* or *(ii)* the fraction of video downloaded as function of the reception quality for the three different YouTube ASes. We see that independently of the AS that serves the video, the fraction of the video downloaded is much higher for videos with good reception quality than for bad videos.



(a) Good reception quality videos (b) Bad reception quality videos

Figure 7: Fraction of Video Downloaded in function of Video Length for YouTube

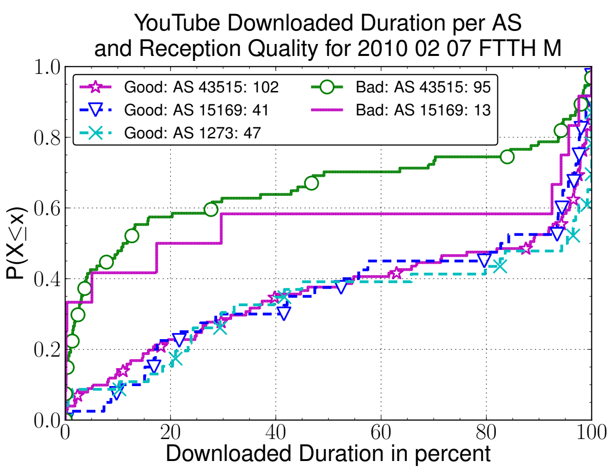


Figure 9: Fraction of Video Downloaded as function of Video Reception Quality for Trace FTTH M 2010/02 for YouTube ASes

7. CONCLUSION

We have carried out a detailed analysis of HTTP video streaming based on the actual videos downloaded by the clients of an ISP. We have used ten different traces captured over four years, at two geographically different capture sites and for two different access technologies and considered two of the main video sharing sites.

We have seen that the video sharing sites have a major control over the delivery of the video and its reception quality through DNS redirection and server side streaming policy. Also, the AS chosen to serve the video is often not the one that assures the best video reception quality.

YouTube.

Our traces, which cover a period of four years (2008–2011) allow to reconstruct the evolution of the Youtube server infrastructure from a situation when videos where first served from a single AS (AS 36561 in 2008), then by several different ASs in 2009–2010, and now in 2011 to all videos being served from a single AS. The selection of the AS serving the video is done via DNS and is under the full control of YouTube. When multiple ASs were used in 2009–2010, the choice of the AS had a big impact on the reception quality. Also, the YouTube server selection did not seem to apply the usual metrics such as proximity. In the case of the YouTube EU AS, which served most of the videos in 2009–2010, the RTTs and loss were high and the reception quality of many videos bad. In 2011 there is a consolidation of all servers into a single AS in 2011 that has a well provisioned network path from the servers to the clients.

DailyMotion.

DailyMotion imposes a mean rate limitation that is slightly above the median encoding rate. This distribution policy allows us to evaluate how a carefully chosen distribution policy can reduce the amount of wasted bytes (downloaded but not watched) without degradation of quality.

Since 2009, there is only one AS that serves all requests assuring for most of the videos a good reception quality.

Viewing Behavior.

This paper is the first to look at the influence of the reception quality on the user viewing behavior in the context of HTTP streaming. We use the normalized reception rate as a simple indicator of video reception quality. We see that videos with bad reception quality are rarely fully downloaded and that bad reception quality results in reduced viewing durations. What is equally interesting is that even when the reception quality is good, only half of the videos are fully downloaded, which indicates that both, the reception quality and the interest in the content, impact the fraction of the video downloaded.

Future Work.

In the future, we want to evaluate more precisely the video reception quality. Two approaches are possible: *(i)* instrumenting the end-user; *(ii)* modeling the video player. As the first option is intrusive, it may be difficult to obtain a representative number of samples. However [7] reports a large scale study of a commercial video player that monitors video reception quality at end-user. This data comprises short and long video on demand viewing as well as some live streaming events. They show that time spent in buffering has a large impact on user behavior in all types of videos. From an ISP perspective, user-generated video streaming sites (*à la* YouTube, DailyMotion and MegaVideo) are the most interesting because they generate most bytes. In this case, the monitoring of the video player is not possible. However, we have already prototyped a tool [19] to crawl the videos on these sites and monitor the end-user perceived quality. We plan to utilize this tool to monitor and compare calibrated Internet accesses on the long term, and also ask volunteers to run it in order to draw general results on video streaming quality according to ISPs and DNSs.

8. REFERENCES

[1] Vijay Kumar Adhikari, Sourabh Jain, and Zhi-Li Zhang. YouTube Traffic Dynamics and Its Interplay with a Tier-1 ISP: An ISP Perspective. In *IMC*, 2010.

[2] Vijay Kumar Adhikari, Sourabh Jain, and Zhi Li Zhang. Where Do You "Tube"? Uncovering YouTube Server Selection Strategy. In *IEEE ICCCN*, 2011.

[3] Alexa. http://www.alexa.com/siteinfo/dailymotion.com.

[4] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System. In *IMC*, 2007.

[5] Xu Cheng, Cameron Dale, and Jiangchuan Liu. Understanding the Characteristics of Internet Short Video Sharing: YouTube as a Case Study. In *IMC*, 2007.

[6] comScore. http://www.comscore.com/press_events/press_releases/2009/3/youtube_surpasses_100_million_us_viewers.

[7] Florin Dobrian, Asad Awan, Ion Stoica, Vyas Sekar, Aditya Ganjam, Dilip Joseph, Jibin Zhan, and Hui Zhang. Understanding the Impact of Video Quality on User Engagement. In *SIGCOMM*, 2011.

[8] Jeffrey Erman, Alexandre Gerber, Mohammad Taghi Hajiaghayi, Dan Pei, and Oliver Spatscheck. Network-aware forward caching. In *WWW*, 2009.

[9] Alessandro Finamore, Marco Mellia, Maurizio Munafo, Ruben Torres, and Sanjay Rao. YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience. In *IMC*, 2011.

[10] Phillipa Gill, Martin Arlitt, Zongpeng Li, and Anirban Mahanti. Characterizing User Sessions on YouTube. In *ACM/SPIE MMCN*, 2008.

[11] Phillipa Gill, Martin F. Arlitt, Zongpeng Li, and Anirban Mahanti. Youtube Traffic Characterization: A View From the Edge. In *IMC*, 2007.

[12] HTB. http://luxik.cdi.cz/~devik/qos/htb/.

[13] Parikshit Juluri, Louis Plissonneau, and Deep Medhi. Pytomo: A Tool for Analyzing Playback Quality of YouTube Videos. In *ITC23*, 2011.

[14] Gregor Maier, Anja Feldmann, Vern Paxson, and Mark Allman. On Dominant Characteristics of Residential Broadband Internet Traffic. In *IMC*, 2009.

[15] Marcin Pietrzyk, Jean-Laurent Costeux, Guillaume Urvoy-Keller, and Taoufik En-Najjary. Challenging Statistical Classification for Operational Usage: the ADSL Case. In *IMC*, 2009.

[16] Marcin Pietrzyk, Louis Plissonneau, Guillaume Urvoy-Keller, and Taoufik En-Najjary. On Profiling Residential Customers. In *TMA*, 2011.

[17] Louis Plissonneau, Taoufik En-Najjary, and Guillaume Urvoy-Keller. Revisiting Web Traffic from a DSL Provider Perspective: the Case of YouTube. In *ITC SS on Network Usage and Traffic*, 2008.

[18] Ingmar Poese, Benjamin Frank, Bernhard Ager, Georgios Smaragdakis, and Anja Feldmann. Improving Content Delivery using Provider-aided Distance Information. In *IMC*, 2010.

[19] Pytomo. http://code.google.com/p/pytomo/.

[20] R. Nelson S. Alcock. Application Flow Control in YouTube Video Streams. *ACM SIGCOMM CCR*, 41(2), 2011.

[21] Ruben Torres, Alessandro Finamore, Jesse Kim, Marco Mellia, Maurizio Munafo, and Sanjay Rao. Dissecting Video Server Selection Strategies in the YouTube CDN. In *IEEE ICDCS*, 2011.

[22] Tstat. http://tstat.tlc.polito.it/.