User Impatience and Network Performance

Denis Collange, Mahdi Hajji

Orange Labs

Sophia Antipolis, France

{denis.collange, mahdi.hajji}@orange.com

Junaid Shaikh, Markus Fiedler, Patrik Arlos

School of Computing

Blekinge Institute of Technology

Karlskrona, Sweden

{junaid.junaid, markus.fiedler, patrik.arlos}@bth.se

Abstract— **In this work, we analyze from passive measurements the correlations between the user-induced interruptions of TCP connections and different end-to-end performance metrics. The aim of this study is to assess the possibility for a network operator to take into account the customers' experience for network monitoring. We first observe that the usual connection-level performance metrics of the interrupted connections are not very different, and sometimes better than those of normal connections. However, the request-level performance metrics show stronger correlations between the interruption rates and the network quality-of-service. Furthermore, we show that the user impatience could also be used to characterize the relative sensitivity of data applications to various network performance metrics.**

Keywords- Network monitoring, Quality of Experience, Quality of Service

I. INTRODUCTION

During the recent years, several subjective user studies have been conducted, in order to correlate the user satisfaction with the network Quality of Service (QoS). Unfortunately, those classical methods are expensive and consume a lot of time. Moreover, they are unable to cope with the fast evolving nature of the Internet usage. Our objective therefore is to define a relatively quicker and cheaper way to correlate the network performance and user impatience, considered as an insight into the actual users' satisfaction. Specifically, the objective of this paper is to show that the user-driven interruptions of TCP connections are actually correlated with some bad end-to-end performance metrics.

The notion of Quality of Experience (QoE) emerged during the last few years, with the intention to involve the subjective view of the end-users into the evaluation of the quality of telecommunication services. ITU-T SG12 defines QoE as the overall acceptability of a service by the end-user. Moreover, it can also be described as the measure of the satisfaction of the customers [13].

Usually, methods of assessing user satisfaction are based on the periodic polls sent to the customers, which incur delays and cost to acquire the user feedback. Moreover, this feedback is not instantaneous; it varies due to customer mood, and changes over time. So the correlations with highly varying network performance are not very precise. Another faster method is to consider the user complaints about the network services, made in the form of calls to the hotline. However, relying on this approach of acquiring the user feedback is not very effective as the disappointed users may already churn, i.e. choose another competitor in the market, instead of reporting the problems. Consequently, there is a need for a proactive method to assess QoE.

Additionally, there are other methods, which use statistical "known correlations" between objectively measurable metrics. Such methods have been applied and standardized for voice and then for video: with MOS (Mean Opinion Score) [8], PESQ (Perceptual Evaluation of Speech Quality) [9], PEVQ (Perceptual Evaluation of Video Quality) [10], and many other similar approaches. Such correlations are typically evaluated on representative panels of customers, in well-defined environments that offer full control over experimental conditions. However, such costly lab experiments do neglect the multitude of influence factors on QoE [4]. They are mainly designed for the improvement of a particular application or to predict the quality during the network planning phase [15]. The execution of these experiments takes time and meanwhile, new applications, and versions of applications, appear all the time, with new traffic characteristics and new performance requirements. Moreover, the behavior of the customers and their sensitivity to the QoS evolves with the time. Hence, there is a need for a cheap and fast method to assess in real-time the satisfaction level of the users, being good enough to detect problems, in a proactive manner.

In this paper, we propose to consider the client interruptions of TCP connections, as users bothered by end-to-end performance are more likely to "interrupt" their transfers. Some recent papers [3][5][6] require an explicit signal from the user to launch punctual traffic measurements and compare the performance with usual conditions. Such methods need the installation of tools on the users' host, and only few (expert) users actually implement them.

Considering the last flags of TCP connections to detect interruptions gives a similar signal for all the connected users without installing any tool on customer's side. We must nevertheless consider TCP interrupts cautiously keeping in mind some facts:

- the algorithm [2] we use to detect interruptions is only a heuristic, based on TCP flags exchanged in the last packets of the connection, as detailed in subsection II.B;

- users may abort a transfer for other reasons besides of the network performance. For example, some studies in the past reported that most of the streaming sessions are interrupted due to the lack of interest in the (end of the) content [16][17][18]. Nevertheless, the detection of such interruptions are still important for the network operators, as the amount of content downloaded but not watched is an overhead for the network [19];

- bad TCP implementation of the application on the client or on the server side may also appear as client TCP interruptions [11][12].

The objective of this document is then to validate the possibility of using TCP interrupts to detect bad user experience due to the degrading network performance. In the next sections, we look at the correlation between aborted TCP connections and their end-to-end performance metrics. We expect to observe an increase of the interruption rate in case of bad performance.

We first describe in Section II the capture and the method we used to detect the interrupted TCP connections. We then analyze in Subsection III.A the differences between the TCP connections according to their type of end. Subsequently, in Subsection III.B we apply a similar analysis separately on the main application types. We then analyze in Section IV the differences between normal and interrupted connections on a smaller timescale closer to user perception. We conclude on the interest of TCP interruptions for network monitoring in the last section.

II. METHODOLOGY

A. Traffic Measurement

We analyze in this document a traffic trace taken with DAG^{TM} cards during a busy hour, at 5pm on a weekday, Tuesday 27th April 2011. The probe captures the whole data traffic of a BAS (Broadband Access Server), between DSLAM (Digital Subscriber Line Access Multiplexers) and the first routers towards the Internet as shown on Figure 1. Data, TV and Voice over IP traffic are transmitted on different VLANs (Virtual Local Area Networks). Our traffic trace is limited to the data traffic. The main traffic characteristics of this capture are shown in the last line in Table I.
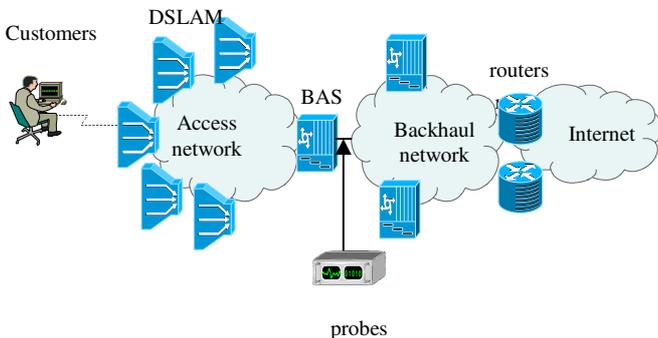

Figure 1. The architecture of the ADSL access network

B. Detection of interrupted TCP connections

Rossi et al. have proposed in [2] the following method based on the last TCP flags to infer whether the client has interrupted a connection.

Eligible connections: The "eligible" connections are by definition those respecting the three following conditions:

- the client has sent at least one packet with the RST (ReSeT) flag;

- non-empty packets have been exchanged;

- and the server has not sent any packet with the flag FIN (FINish) or the flag RST.

Interrupted connections: The connections classified as "interrupted" are by definition the eligible connections such that the delay between the last non-empty packets sent respectively by the client and by the server is smaller than the average round-trip time.

Normal connections: The normal connections are those ending with a FIN flag in each direction.

Unfinished connections: The connections without any FIN or RST flag are considered unfinished at the end of the traffic trace.

Abnormal connections: The connections that do not fall in any of the above category are classified as "abnormal" connections.

We list in Table I the relative part of each of the above mentioned type of termination, found in the captured traffic.

TABLE I. TRAFFIC BY TYPE OF TERMINATION

	Packets ($\times 10^6$)				Traffic (GB)				Connections ($\times 10^3$)	
	Down		*up*		*down*		*up*			
Normal	44	*43%*	29	*42%*	52	*45%*	5.4	*35%*	967	*26%*
Eligible	6.8	*7%*	4.3	*6%*	8.9	*8%*	0.8	*5%*	137	*3.7%*
Interrupted	13	*13%*	7.4	*11%*	18	*16%*	0.8	*5%*	20	*0.6%*
Abnormal	17	*17%*	12	*17%*	19	*17%*	3.5	*23%*	329	*9.0%*
Unfinished	8.9	*9%*	6.2	*9%*	11	*9%*	1.2	*8%*	358	*9.7%*
Non TCP	11	*11%*	10	*15%*	5.5	*5%*	3.5	*23%*	1831	*51%*
Total	101		69		115		15		3643	

III. CONNECTION-LEVEL ANALYSIS

A. Global analysis

We consider in this section the difference in the performance between the normal and the interrupted connections. We first compare the mean throughput of the connections classified according to their type of termination. The *effective throughput* is obtained from the total amount of data in the connection divided by the connection duration, reduced by silence periods larger than five seconds. The *instantaneous throughput* is calculated by counting the number

of bits transferred during one measured Round Trip Time (RTT), as reported by the number of bytes in the acknowledgement. We observe from Table II that the upward throughputs are slightly lower for the interrupted connections. However, the downward throughputs that are expected to influence the user satisfaction in a stronger way are surprisingly much larger for the interrupted connections.

TABLE II.   UP AND DOWN THROUGHPUTS (IN KBIT/S) ACCORDING TO THE TERMINATION TYPE

Type of termination	Mean Throughput		Effective Throughput		Instantaneous Throughput	
	Up	*Down*	*Up*	*Down*	*Up*	*Down*
Normal	24	67	28	90	1612	431
Interrupted	22	515	24	536	1006	1158

We posit that several hypotheses may explain this surprising result.

H1) A first hypothesis is that the throughput of the interrupted connections is good, but that other performance criteria bother the customers.

H2) A second hypothesis is that the termination behavior of the TCP connections depends on the application and its implementation on the client and/or on the server side. Some applications may have higher throughputs and for some reasons higher interruption rates, which may impact these global measurements. We consider this hypothesis in the next subsection. Similarly, some customers or some servers with high throughputs may be subjected to more TCP interruptions.

H3) A third hypothesis is that, during high throughput transfers, the users may be more sensitive to the performance degradation. We look at this hypothesis in section IV.

We consider now the influence of the end-to-end packet losses on the TCP interruptions. More precisely, we consider a rough estimate of the packets lost by the TCP connections: the number of out-of-sequence packets divided by the total number of data packets. Figure 2 shows the complementary distribution function of the upward and downward loss rates of TCP connections according to their type of termination. It shows the proportion of connections with a loss rate higher than a given value. We remark that the interrupted connections observe much more packet losses: 17% of interrupted downward connections observed loss vs. 1.6% of normal ones. Interrupted connections also observe higher loss rates than normal connections as the curves are shifted to the right. We however note that most of the interrupted connections (83%) do not observe any loss. So, a high loss rate may explain TCP interruptions for some, but certainly not in most cases.
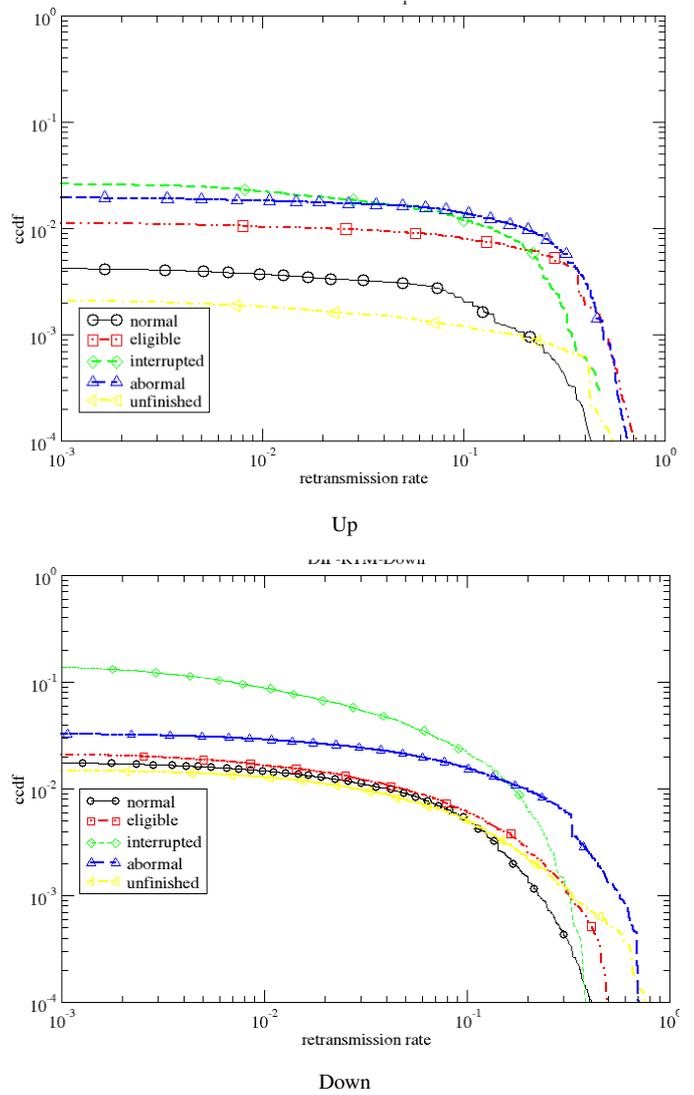

Up

Down

Figure 2.   Mean loss rate of TCP associations

Figure 3 presents the local and the distant round-trip times for TCP connections, according to their type of termination. We define the local (resp. distant) round-trip time as the delay seen by the probe on downward (resp. upward) connections between a data packet and its acknowledgment. Retransmissions are not taken into account. We remark in Figure 3 that the local round-trip times of the interrupted connections are a bit larger, as their distributions are shifted to the right. The distant round-trip times of interrupted and normal connections appear quite similar. The fact that the main difference between the interrupted and normal connections is on the local loop might indicate that a possible cause of interruptions could be local congestion due to the presence of multiple parallel TCP connections on the access link of the customer.
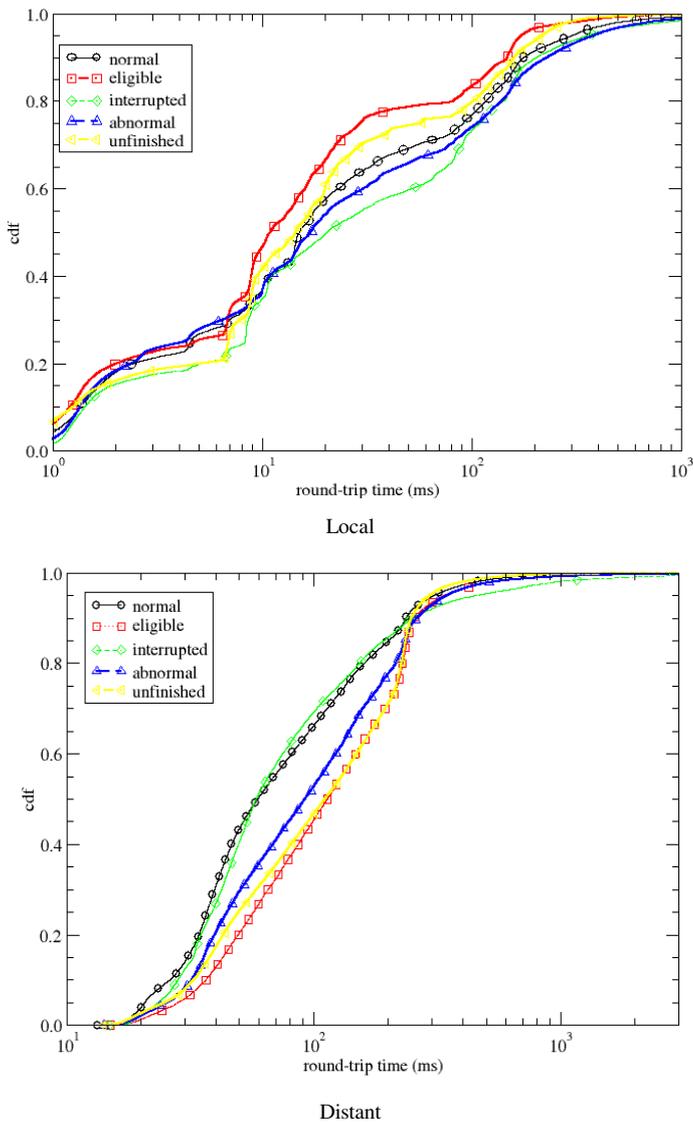
Local

Distant

Figure 3. Distribution of round trip times

From the global analysis presented in this section and considering some criteria like local round-trip times or packet loss, we observe that interrupted connections globally observe lower performance than normal connections. But we also notice more surprising results: the interrupted connections have usually larger throughput, at different timescales, and even packet losses and round-trip times are not sufficient to explain the TCP interruptions. We have then considered many hypotheses to explain these surprising results. The first possible explanation (**H1**) should be rejected as the most interrupted connections do not observe packet losses, and the distributions of the local and the distant round-trip times of interrupted and normal TCP connections are not dramatically different. A first conclusion, which we draw from the above findings, is that the global performance metrics are not sufficient to explain why the TCP connections are interrupted. We will check the other hypotheses in the following sections.

B. Analysis per application

We consider more precisely in this subsection the differences between the main applications, especially their interruption rate and their performance. Connections are classified in applications by an Orange-internal DPI (Deep Packet Inspection) tool. This tool has been compared with Tstat in [7]. The list of identified applications detected by this tool is shown in Table III.

TABLE III.   APPLICATIONS CLASSES

Class	Application / Protocol
WEB	HTTP and HTTPS browsing
P2P	eDonkey, eMule obfuscated, Bittorrent, Gnutella, Kazaa, SoulSeek, Filetopia, Others
MAIL	SMTP, POP3, IMAP, IMAPs, POP3s, HTTP Mail
CHAT	MSN, IRC, Jabber, Yahoo Msn, HTTP Chat
OTHERS	NBS, Ms-ds, Epmap, Attacks, Nntp, LDAP, Microsoft SQL, Oracle SQL, mySQL
DOWNLOAD	FTP data, FTP control, HTTP file transfer
GAMES	Blizzard Battlenet, Quake II/III, Counter Strike, HTTP Games
STREAMING	MS Media Server, Real Player, iTunes, Quick Time
UNKNOWN	others

We mainly observe in our traffic traces the eight types of applications listed in Table IV. Some connections are not recognized and therefore, denoted as "unknown". The relative part of each application, according to the number of connections, of up and down packets and bytes, and the interruption rate are given in Table IV. We notice that these various metrics display rather different proportions. For example, Streaming represents only 1.6% of connections, but 42% of bytes downward. Download represents 0.4% of connections but 14% of bytes downward. Conversely, recognized P2P represents (similarly to the unknown connections) 31% of connections and only 10% of the bytes downward. Hence, the Streaming and the Download connections are much larger in size than the connections of the other applications.

TABLE IV.   TRAFFIC PER APPLICATION

application	packets				Traffic				connect.	interrupt
	down		up		down		up			
	$\times 10^6$	%	$\times 10^6$	%	GB	%	GB	%	$\times 10^3$	%
Streaming	34	33	18	25	49	42	1.1	6.9	57	9.8
Web	27	27	20	29	31	27	3.6	24	1191	0.7
P2P	15	15	14	20	12	10	7.1	47	1120	0.3
Download	11	11	5.9	8.5	16	14	0.3	2.1	15	4.7
Unknown	7.1	7.0	6.8	9.8	3.7	3.3	2.0	13	925	0.1
Games	2.8	2.7	1.9	2.7	2.2	1.9	0.1	0.9	28	0.1
Mail	1.5	1.5	1.3	1.9	1.3	1.1	0.7	4.9	29	0.5
VOIP	0.6	0.6	0.6	0.8	0.3	0.3	0.1	0.6	25	1.6
Chat	0.3	0.3	0.2	0.3	0.2	0.2	0.1	0.6	9	14.7
total	100.7		68.8		114.9		15.2		3643	0.6

We remark in the last column that the interruption rate is highly correlated with the application: three applications have much larger interruption rates than the other ones: Chat (15%), Streaming (10%) and Download (5%). However, in general, the mean interruption rate for all the TCP connections is 0.6%, much smaller than the interruption rate of the individual applications: Chat, Streaming and Download.

Table V provides for each application, the mean performance metrics for the normal and the interrupted TCP connections. A common observation to almost all of the applications is that, the mean downward throughput of the interrupted connections is usually larger than that of normal connections. The other results are more specific to each application.

1) Chat
The interrupted chat connections have higher upward and downward throughputs, lower local and distant round-trip times, and higher loss rates. So, the loss rate per application is the sole performance criteria which may explain an interruption of the TCP connections. As not all the interrupted chat connections observe packet loss, and as the interruption rate of chat connections is much higher than for the other applications, we should assume some other reasons behind the interrupted chat connections. Maybe chat user usually close the window once they are done with the chatting, thereby interrupting a TCP connection in the end.

2) Download
We further notice that the upward throughputs of interrupted download connections are 30% below those of the normal ones. In contrast, the downward throughputs, which should expectedly be more important for the users, are higher in the case of interrupted connections. Therefore, the throughput degradation should not be the primary reason for the interruption. On the other hand, we observe that the local round-trip times of the interrupted connections are much larger. Although their mean distant round-trip time is relatively a bit smaller, their combined sum — the end-to-end round-trip time — is however 70% larger. The loss rates of the interrupted download connections in both directions are also slightly higher than those of the normal download connections.

3) Games
Similarly, the downward throughputs of interrupted Games connections are larger than those of the normal ones, but the ratio is largely below the one for upward throughputs which may be a significant criterion for this application. Both up and down loss rates are slightly higher for interrupted Games connections, but local and distant round-trip times are much larger. Hence, interruption of Games connection seems then to be correlated with large end-to-end round-trip times, low upward throughputs and high up and downward loss rates.

4) Mail
All average up and down throughputs of the interrupted Mail connections are larger than those of the normal Mail connections. The mean round-trip times are slightly larger, with the loss rates in both directions almost doubled. Hence, the interruptions of Mail connections appear to be correlated with the increasing end-to-end round-trip times and loss rates.

5) P2P
The conclusions for P2P are similar to those of Mail, with the supplementary observation of considerably low instantaneous upward throughput of the interrupted connections.

6) Streaming
For streaming transfers, all the average up- and downward throughputs of interrupted connections are larger than those of the normal connections. The local mean round-trip time of the interrupted connections is slightly higher, while the distant one is significantly smaller as compared to the normal connections. The upward loss rate is exactly similar while the downward loss rate is slightly higher. The interruption of streaming can not be clearly correlated with the network performance. There appear to be two possible reasons behind this observation. First, interruptions could be due to the implementation of streaming clients. We observed in [11] that one of the client-side web browser interrupts a TCP connection as soon as the request for the video container is made, henceforth, initiating a new TCP connection to play the video. Second, users could interrupt the connections, which may be for example due to the lack of their interest in the content of the video.

7) VOIP
Round-trip times of the interrupted VOIP connections are a bit (20%) larger, but loss rates are smaller, and most of the average throughputs are larger, except the upward instantaneous throughput. Hence, the VOIP users appear to be more sensitive to the round-trip times.

TABLE V.   MEAN VALUES OF TRAFFIC AND PERFORMANCE METRICS OF CONNECTIONS ACCORDING TO THE APPLICATION AND THE TERMINATION OF THE CONNECTION.

application	termination	Mean throughput (kbits/s)		Round trip time (ms)		Packet loss rate (%)	
		down	*up*	*distant*	*local*	*up*	*down*
Chat	*normal*	15	6	125	106	2	7
	interrupted	68	26	113	54	7	10
Download	*normal*	376	21	86	79	4	3
	interrupted	538	15	69	213	4	4
Games	*normal*	55	18	126	88	8	3
	interrupted	326	8	957	296	14	5
Mail	*normal*	56	25	21	56	2	3
	interrupted	144	37	26	70	3	6
P2P	*normal*	4	4	271	168	4	4
	interrupted	12	8	572	410	7	7
Streaming	*normal*	257	13	77	92	3	1
	interrupted	1501	30	29	95	3	2
VOIP	*normal*	23	9	134	92	7	6
	interrupted	97	21	160	114	3	5
Web	*normal*	72	30	34	95	5	6
	interrupted	180	26	50	133	5	5

8) Web
Finally, downward throughputs of the interrupted Web connections are larger, but upward throughputs are slightly smaller than those of the normal connections. Both up and downward loss rates appear to be almost similar. However, the

round-trip times are significantly larger for the interrupted connections. Consequently, the round-trip times seem to be the most important user-perceived performance criterion for this kind of interactive application, as it influences directly the response time of a webpage.

Summarizing the results listed in Table V, the download and the Streaming connections have the highest throughputs and also high interruption rates. This explains the surprising observation that we discussed in the previous section. Another obvious conclusion that further confirms the findings of the previous works is that the sensitivity of the user to the end-to-end performance actually depends on the application.

IV. TCP CONNECTION BREAKDOWN

The objective of the analysis presented in this section is to get a look closer to the user perception. Instead of considering connection-level metrics assessed on the whole duration of TCP connections, we consider transfer-level metrics for the successive alternate exchanges between the end-hosts. Those exchanges are usually due to the requests from the local client followed by answers from the distant server. We have indeed noticed in [1] that roughly 90% of the connections observe an alternate exchange of groups of packets (blocks, trains…) from one end of the connection and then from the other. We consider only this kind of connections in the rest of this section. For these connections, we propose in [1] to differentiate on each side of the transfer the delay to prepare the data from the delay to transfer it.
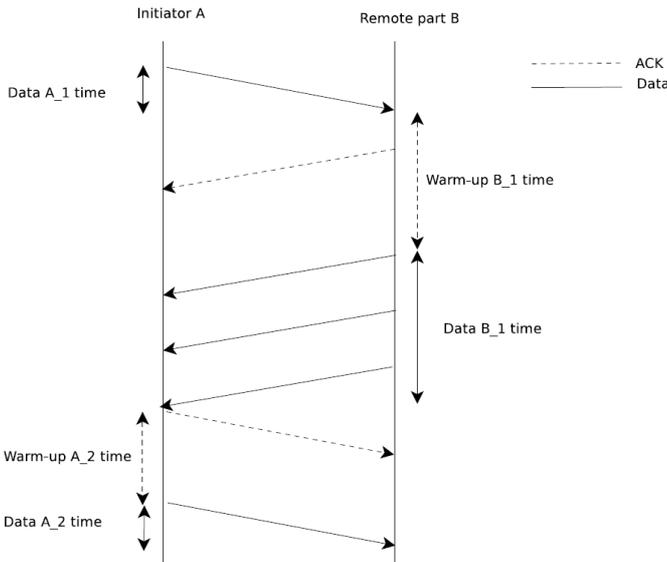

Figure 4.   Breakdown of TCP connections in alternate exchanges (from [1])

More precisely, we distinguish:

- The think-time, or warm-up, which is the silence between the last data packet from the previous train in the other direction and the sending of the first data packet of the next train. On the server side, this silence

may be used to prepare the data to send, to request information from other servers, etc. On the user side, it may be used to read the information received and think about a next eventual request. The server-side warm-up times are referred to as the distant warm-up times in this paper.

- The transmission delay, during which packets are on-line, which elapse from the date of departure of the first packet of the train to the date of arrival of the last packet of the train.

Assuming that the user should not be bothered by its own think-time, Table VI considers only the distant warm-up time. In particular, it presents the mean and the standard deviation, of all the distant warm-up times. It also lists the mean of the ratio per connection of the last to the mean of all distant warm-up times. We observe from this table that the distant warm-up of the interrupted connections is usually not significantly larger than the distant warm-up of the normal connections. It is 10% larger for Web and Mail, 40% larger for P2P, but it is slightly smaller for Games and VOIP, and much more for Download, Streaming and Chat. We observe similarly that the variations of the warm-up delays are also much smaller for the interrupted connections, except for P2P for which they are 50% larger. On the third part of the table, we notice that the last distant warm-up of the interrupted connections is roughly larger than for the normal connections for P2P and Games (roughly twice), and especially for Mail (more than five times).

TABLE VI.   MEAN, STANDARD DEVIATION AND RATIO OF LAST TO MEAN DISTANT WARM-UPS ACCORDING TO THE TERMINATION.

Distant warm-up	mean (seconds)		standard deviation (seconds)		last / mean	
	normal	interrupted	normal	interrupted	normal	Interrupted
Web	0.7	0.7	4.2	2.0	3.6	2.7
P2P	0.5	0.7	0.8	1.2	5.6	11.0
Download	1.0	0.3	2.6	1.0	0.8	0.7
Mail	0.6	0.7	1.9	1.8	2.6	14.5
Games	0.3	0.2	1.6	0.5	1.2	3.0
Streaming	1.3	0.3	6.1	0.8	3.2	1.3
Chat	4.1	0.4	10.7	0.7	11.1	1.1
VOIP	0.8	0.6	2.7	1.1	19.8	1.2

We observe similar differences between the normal and the interrupted connections when looking more rigorously at the distributions of the warm-ups instead of simply looking at the mean and the standard deviations. Unfortunately, due to the lack of space, we cannot include all the distributions in this paper. As examples, we display only the distributions for the upward and the downward transfer durations of Web and Games connections in Figure 5.
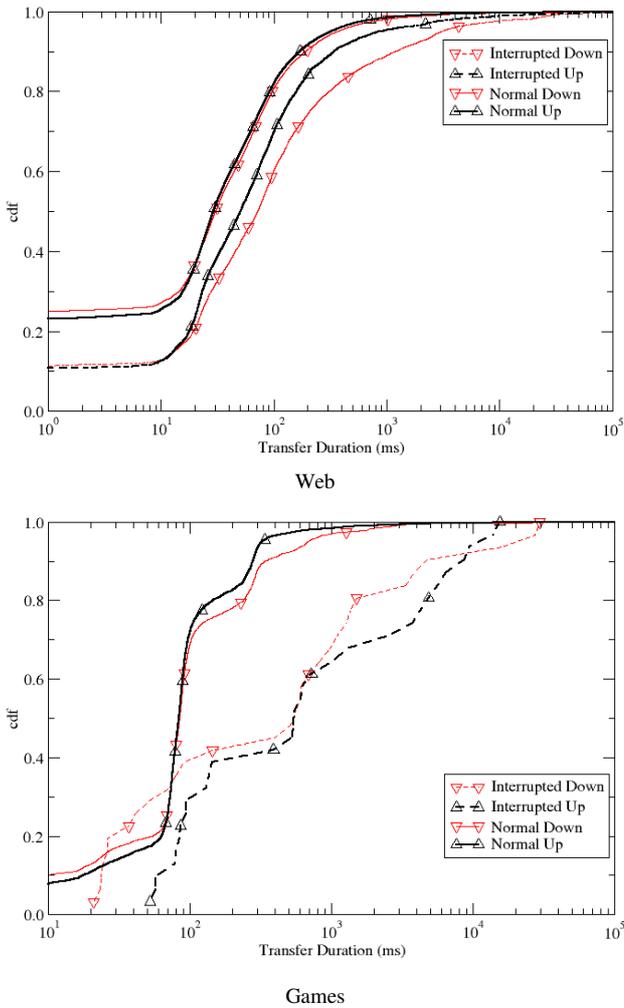
Web

Games

Figure 5.   Compared distributions transfer durations of normal and interrupted Web and Games connections

TABLE VII.   MEAN, STANDARD DEVIATION AND RATIO OF LAST TO MEAN TRANSFER DURATIONS ACCORDING TO THE TERMINATION.

Downward transfer durations	mean (seconds)		standard deviation (seconds)		last / mean	
	normal	*interrupted*	*normal*	*interrupted*	*normal*	*interrupted*
Web	0.3	1.7	0.7	3.4	4.2	10.0
P2P	1.0	6.1	2.8	12.2	3.3	18.0
Download	1.2	21.7	2.8	31.3	10.3	75.1
Mail	0.1	2.7	0.4	6.9	1.2	1.1
Games	0.2	2.3	0.7	4.1	6.0	55.5
Streaming	1.2	5.4	2.1	7.7	12.3	100.7
Chat	3.0	0.3	8.0	0.3	8.3	1.0
VOIP	1.0	1.1	2.7	8.2	3.5	1.9
Upward transfer durations	**mean (seconds)**		**standard deviation (seconds)**		**last / mean**	
	normal	*interrupted*	*normal*	*interrupted*	*normal*	*interrupted*
Web	0.2	1.2	0.5	2.2	3.4	10.5
P2P	1.2	2.9	3.4	6.0	3.7	4.4
Download	2.5	18.9	5.0	27.0	16.8	57.7
Mail	0.3	0.2	0.8	0.6	1.7	8.7
Games	0.3	6.0	0.9	11.7	5.1	164.3
Streaming	4.0	9.3	7.0	13.8	20.6	161.6
Chat	1.7	0.2	5.7	0.3	10.4	2.8
VOIP	1.2	0.3	3.6	0.4	3.9	3.2

Table VII shows the mean, the standard deviation of the transfer delays of up and downward trains, and the ratio of the last transfer delay to the mean transfer delay. We first remark that for all the applications, the warm-up delays and the transfer delays are found in the same order of magnitude, namely from a few hundreds of milliseconds to few seconds. The transfer durations in both directions are also found within the same magnitude. We clearly notice in Table VII that for all the applications (except of VOIP and Chat), the interrupted connections observe much larger and much more varying transfer durations in both directions than the normal connections. Their last transfer duration is also many times larger, which points at some kind of starvation problem.

Finally, we define the response time as the delay between the time of the first packet of a request sent by a user and the last packet of the answer to this request, received on the client-side. The response time is similarly defined as the sum of the up- and downward transfer delays and of the distant warm-up delay. Table VIII presents the mean, the standard deviation and the ratio of the last to the mean response times, according to the termination and to the application. We notice in Table VIII that the interrupted connections of all the applications except of chat and VOIP, experience larger response times as compared the normal ones. We observe that for these applications, the part of response time due to the network delays as well as the ratio of the last response time to the mean are also larger. We observe different results considering the standard deviations of the response time. In this case, we remark that VOIP interrupted connections observe larger variations than normal connections. On the contrary, the variations of the response times of Web and Streaming interrupted connections are not significantly different from their normal connections. These observations may indicate that Web and Streaming users are more bothered by the large response times, while VOIP users are more disturbed by the varying response times. For most of the applications, the users appear to be very sensitive to the response times. Hence, the interrupted connections in these cases can be taken as an indication of the longer response times. It reflects that the users are more critical about the response times of the requested content. They become impatient and click away, if the transfers take slightly longer, as mentioned in [14]. They spend a constant time on the network, expecting the content to be delivered during this time [21]. They abandon the transfers, if it takes more than their expected time.

TABLE VIII.   MEAN, STANDARD DEVIATION AND RATIO OF LAST TO MEAN RESPONSE TIMES ACCORDING TO THE TERMINATION.

Response time	mean (seconds)		standard deviation (seconds)		last / mean	
	normal	*interrupted*	*normal*	*interrupted*	*normal*	*interrupted*
Web	1.3	3.6	4.3	4.0	3.7	8.7
P2P	2.7	9.7	2.9	12.2	3.9	13.4
Download	4.7	40.9	3.9	31.3	11.6	66.5
Mail	1.1	3.6	1.9	7.1	2.1	4.2
Games	0.8	8.5	1.8	4.2	4.0	130.0
Streaming	6.4	15.0	6.5	7.8	15.6	136.6
Chat	8.8	0.8	13.3	0.7	10.0	1.5
VOIP	3.0	2.1	3.8	8.3	7.8	1.9

V. Conclusions

The objective of this study was to check if the interruption of TCP connections is actually correlated with the bad network performance. Our main conclusion is that this is actually true. Moreover, a second and the important conclusion is that the breakdown of TCP connections in alternate packet trains is critical in detecting this correlation for the most applications. We have indeed observed that the usual connection-level performance metrics of the interrupted connections are not very different, and sometimes better than those of the normal connections. However, considering separately the main applications and especially the request-level performance metrics show much better correlations between the interruption rate and the network QoS, and it also explains some paradoxical results observed earlier. A third conclusion is that diverse applications have different sensitivity to the various performance criteria.

Complementary studies try to point out the root causes [20][1] of the performance problems disturbing users according to these various applications, in order to propose some ways to reduce their impact and maximize the customer experience. The influence of the various performance criteria on the QoE considered in this document could also be more precisely characterized to define the utility functions, which may further be used to design and optimize the network and the services, taking into account the customer experience. The results observed should also be validated by comparison with similar traffic analyses in other conditions, such as others hours-of-day or in the context of wireless access networks.

References

[1] "Revisiting the performance of short TCP transfers", A. Hafsaoui, D. Collange, G. Urvoy-Keller, *Networking 2009, Aachen (Germany),* May 2009.

[2] "User Patience and the Web: a hands-on investigation", D. Rossi, C. Casetti, M. Mellia, *IEEE Globecom 2003, San Francisco (USA),* December 2003.

[3] "Understanding end-user Perception of Network Problems", J.S. Miller, A. Mondal, R. Potharaju, P. Dinda, A. Kuzmanovic, *ACM Sigcomm Workshop on Measurements Up the STack (W-MUST), Toronto (Canada),* August 2011.

[4] "Vienna Surfing - Assessing Mobile Broadband Quality in the Field", R. Schatz, S. Egger, *ACM Sigcomm Workshop on Measurements Up the STack (W-MUST), Toronto (Canada),* August 2011.

[5] "Performance of Networked Applications: The Challenges in Capturing the User's Perception", D. Joumblatt, R. Teixeira, J. Chandrashekar, N. Taft, *ACM Sigcomm Workshop on Measurements Up the STack (W-MUST), Toronto (Canada),* August 2011.

[6] "HomeMaestro: Distributed monitoring and diagnosis of performance anomalies in home networks", T. Karagiannis, C. Gkantsidis, P. Key, E. Athanasopoulos, E. Raftopoulos, *Microsoft Technical Report MSR-TR-2008-161*, October 2008.

[7] "Revealing the unknown ADSL traffic using statistical methods**",** M. Pietrzyk, G. Urvoy-Keller, *Traffic Monitoring and Analysis Workshop (TMA), Aachen (Germany),* April 2009.

[8] "Mean Opinion Score (MOS) Terminology", *ITU-T Recommendation P.800.1*, July 2006.

[9] "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs", *ITU-T Recommendation P.862*, February 2001.

[10] "Objective perceptual multimedia video quality measurement in the presence of a full reference", *ITU-T Recommendation J.247*, August 2008.

[11] "Classification of TCP connection termination behaviors for mobile Web", J. Shaikh, M. Fiedler, P. Arlos, T. Minhas, D. Collange, *IEEE Globecom Workshop on Smart Communication Protocols and Algorithms (SCPA),* December 2011.

[12] "An analysis of TCP reset behaviour on the Internet", M. Arlitt, C. Williamson, ACM Sigcomm Computer Communication Review (CCR), 2005.

[13] "Vocabulary for performance and quality of service", *ITU-T Recommendation P. 10/G. 100 (incl. Amendment 2),* 2008.

[14] "For Impatient Web Users, an Eye Blink Is Just Too Long to Wait", S. Lohr, New York Times, March 2012.

[15] "Multimedia Quality Assessment Standards in ITU-T SG12", P. Coverdale, S. Möller, A. Raake, A. Takahashi, *IEEE Signal Processing Magazine, Issue 6, Volume 28, pp. 91-97,* 2011.

[16] "Youtube Everywhere: Impact of Device and Infrastructure Synergies on User Experience", A. Finamore, M. Mellia, M. Munafo, R. Torres, S. R. Rao, *Technical Report 418, Purdue University,* May 2011.

[17] "Youtube Characterization: A View From the Edge", P. Gill, M. Arlitt, Z. Li, A. Mahanti. *In IMC '07, pp. 15-28,* 2007.

[18] "Can Internet Video-on-Demand be Profitable?", C. Huang, J. Li, K. W. Ross. *In SIGCOMM '07, pp. 133-144,* 2007.

[19] "Network Characteristics of Video Streaming Traffic", A. Rao, Y. Lim, C. Barakat, A. Legout, D. Towsley, W. Dabbous. *In ACM CoNEXT '11,* 2011.

[20] "A Root Cause Analysis Toolkit for TCP", M. Siekkinen, G. Urvoy-Keller, Ernst W. Biersack, D. Collange, Computer Networks, vol. 52 N° 9, pp. 1846-1858, June 2008.

[21] "Quality of Experience from User and Network Perspectives", J. Shaikh, M. Fiedler, D. Collange, Annals of Telecommunication, vol. 65, number 1-2, Feb. 2010.