Identifying Persistent and Recurrent QoE Anomalies for DASH Streaming in the Cloud

Chen Wang*†, Hyong Kim*, Ricardo Morla†
*Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, USA
†INESC Porto and Faculty of Engineering, University of Porto, Porto, Portugal
chenw@cmu.edu, kim@ece.cmu.edu, ricardo.morla@fe.up.pt

***Abstract*—Quality of Experience (QoE) anomalies widely exist in all types of video services. As video services migrate to the Cloud, unique challenges occur to deploy video services in the Cloud environment. We study the QoE anomalies for users in a video service deployed in a production Cloud CDN. We use a QoE anomaly identification system, QRank, to identify anomalous systems. We consider Cloud CDN servers, Cloud CDN networks, transit networks, user access networks and different types of user devices. Our extensive experiments in production Cloud find several interesting insights about QoE anomalies of video streaming in the Cloud. 91.4% of QoE anomalies are detected on 15.32% of users. These users experience QoE anomalies persistently and recurrently. The Cloud servers and networks seldom cause QoE anomalies. More than 99.98% of QoE anomalies are identified in anomalous systems including the transit networks, the access networks and user devices. We infer that transit networks are the actual bottleneck systems for QoE anomalies in production Cloud. More than 95% of persistent and recurrent QoE anomalies are identified in less than 10 transit networks. We collect latency measurements to anomalous networks and the analysis indicates that the limited capacity in transit networks are the major cause of QoE anomalies. Resulting anomalies impair user QoEs persistently or recurrently. In order to provide good user QoE, the Cloud provider should identify transit networks that may become bottlenecks for high quality video streaming and appropriate peering with Internet Service Providers (ISPs) to bypass these bottlenecks[1].**

***Keywords*-QRank, QoE Anomalies, Transit networks**

I. Introduction

Quality of Experience (QoE) is important for video service. More and more video applications migrate to the Cloud to cache videos [1].

Existing Cloud platforms define service level agreements (SLAs) for different types of services. For example, Azure Cloud [2] defines the uptime percentage as an SLA for virtual machines. It refunds tenants when the percentage of a VM monthly uptime is below 95%. For Azure CDN service, Azure defines the percentage of HTTP transactions without error as its SLA. When the percentage of successful HTTP transactions is below 99.9%, it refunds users. These SLAs do not guarantee the QoE for users in video applications. Streaming videos from the Cloud involve many different systems. Videos are encoded in the Cloud virtual instances, cached in Cloud CDNs, transferred through transit networks, delivered through usersáccess networks, and played in usersdevices. Multiple stakeholders manage these systems. The Cloud providers manage virtual instances, CDN servers and Cloud CDN networks. Various Internet Service Providers (ISPs) manage different transit and access networks. Users manage their own devices. Anomalies in any systems can impair end user QoEs. Cloud providers need to find system bottlenecks causing QoE anomalies to improve their infrastructure provide quality video applications.

In this work, we analyze the QoE anomalies for DASH streaming [3] from a production Cloud CDN. We run a QoE anomaly identification system, QRank [4]. QRank uses real-time QoE measurements to identify the anomalous systems. QRank assumes that the system with users who experience lower QoEs is more likely cause QoE anomalies. QRank identifies anomalous system by ranking the QoEs in all systems in the video streaming. We consider Cloud CDN servers, Cloud CDN networks, transit networks, access networks and different types of user devices. We deploy 124 users worldwide in PlanetLab and Azure Cloud to run DASH video streaming sessions for 100 hours. QoE measurements from $124 \times 100 = 12400$ video sessions are collected. 9367 QoE anomalies with average length of 127.48 seconds are detected on 65 emulated users. Our extensive experiments in production Cloud find the following insights.

- Users experience QoE anomalies very differently. 1) **Recurrency**: 12.1% users experience QoE anomalies recurrently and experience 87.83% of QoE anomalies. 2) **Persistency**: 8.87% of users experience persistent QoE anomalies with duration over 15 minutes. 3) **Sparcity**: During 100 hoursvideo streaming, 41.1% of users experience only less than one QoE anomaly per hour and all QoE anomalies last less than 900 seconds. 47.58% of users do not experience QoE anomalies at all.
- According to QRank, access networks, transit networks and user devices incur more than 99.98% of QoE anomalies. Among those, 58.66% of QoE anomalies are identified in access networks, 38.14% in user devices, and 13.89% in transit networks.

[1]This work was supported in part by the FCT under Grant SFRH/BD/51150/2010 and an Azure Research grant provided by Microsoft Corporation.

- 95.38% of QoE anomalies in user devices and 97.23% in access networks are experienced by PlanetLab users. These users are emulated on 48 PlanetLab nodes that belong to 21 campus networks. We infer that PlanetLab nodes in those campus networks are capacity limited, causing QoE anomalies.
- QoE anomalies incurred in access networks and devices are due to PlanetLab conditions. We infer that transit networks could be the major cause of QoE anomalies for real world video application in the Cloud. For all QoE anomalies identified in transit networks, more than 95% of QoE anomalies are identified in only 10 transit ISPs.

These results have an important implication. In order to provide good user QoE, the Cloud provider should identify transit networks that may become bottlenecks for high quality video streaming and appropriate peering with ISPs to bypass these bottlenecks.

II. Data Collection

A. Video Streaming Testbed in the Cloud

We set up a video server in Azure Cloud [5] as the origin video server. We cache the video content in Azure CDN. We run DASH (dynamic adaptive streaming over HTTP) video streaming in 100 PlanetLab servers [6] and 24 Azure virtual machines around the world. Each user request a video every hour and each video streaming session last around 50 minutes. We run DASH video streaming for 100 hours and collect QoE measurements from 12400 video sessions.

B. QoE Measurement

We use a chunk based QoE model in [7] to measure end user QoE in run time for DASH. DASH is prevalent in most commercial VoD systems, such as YouTube, Netflix, Hulu and Amazon Prime.

C. QoE Anomaly

Users expect an acceptable QoE from their VoD service [8]. The VoD providers can determine the acceptable QoE value q_0 by studying the user engagement [9]. We define QoE anomaly as any fault or anomaly that degrades end user QoE below q_0 . Chunk QoE values can be computed at run time from the bitrate of a chunk and the freezing time. QoE anomaly can be detected at run time if QoE is monitored on the video player. We choose $q_0 = 2$ for the rest of the paper. QoE lower than 2 indicates that the user is streaming the video at the two lowest bitrates or is freezing.

Anomalies usually affect QoE for more than one chunk. If two chunk QoEs below q_0 are detected within a time interval, we attribute them to the same anomaly. We assume that two chunks with QoE value below q_0 within an interval of N chunks are caused by the same anomaly. The duration of the QoE anomaly is determined from the reception of the first to the last chunk with QoE below q_0. The video chunk

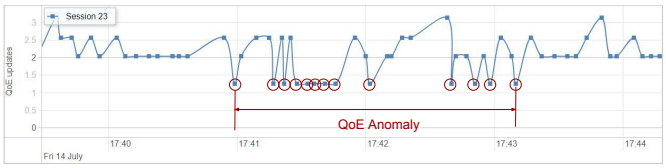


Figure 1. An example of QoE anomaly

in our experiment is a 5-second video segment. We choose $N = 12$ and the minimum interval to separate two distinct QoE anomalies is $12 \times 5 = 60$ seconds. Figure 1 shows QoEs monitored on one video streaming session and a QoE anomaly spanning 2 minutes and 12 seconds. We further classify QoE anomalies into three categories:

- **Severe QoE anomaly**: In the period of a QoE anomaly, the percentage of chunks with QoE values below q_0 is equal or greater than 70%.
- **Medium QoE anomaly**: In the period of a QoE anomaly, the percentage of chunks with QoE values below q_0 is less than 70% and equal to or greater than 20%.
- **Light QoE anomaly**: In the period of a QoE anomaly, the percentage of chunks with QoE values below q_0 is less than 20%.

III. QRank System and Monitoring System

A. QRank

QRank is a QoE anomaly identification system for VoD service[2]. It identifies the bottleneck system causing QoE anomalies. Specifically, QRank detects QoE anomalies based on chunk QoE values, discovers the underlying network topology and systems used for video streaming via traceroute measurements, and identifies anomalous systems by ranking the QoEs in all systems involved in the video streaming. Here is an example of all systems involved in a video streaming session. We use a desktop in Carnegie Mellon University campus network to stream a video cached in Microsoft Azure CDN. We probe the CDN server several times and discover the underlying network topology for video service as shown in Figure 2. The video traffic is delivered from the CDN server to the user through routers in multiple networks: the Cloud network, several different transit networks managed by different Internet Service Providers (ISPs), and the campus network. QRank considers the server, the Cloud network, the transit networks, the access network and the user devices as possible anomalous systems. We analyze all QoE anomalies identified from our extensive experiments in production Cloud environments.

B. Geographical distributed monitoring system

We use a distributed monitoring system to validate QRank and to analyze the root causes of QoE anomalies. We deploy

[2]The details of QRank system can be found in [4]. In this paper, we only discuss the results identified by QRank for Azure CDN.

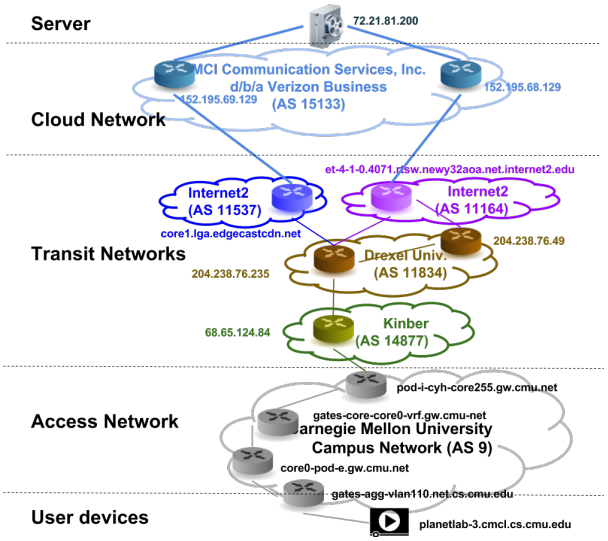


Figure 2. Systems involved in video streaming from the Cloud CDN

124 monitoring agents in 100 PlanetLab nodes and 24 Azure servers around the world. These agents monitor the network performance by probing the routers and CDN servers. The performance of a network is measured by latencies from an agent to a router in the network. Only the closest agent to the network and the server is used to probe. The "closest" agent is chosen based on its geographical distance to the network and the server. We estimate the network link latency by measuring the latencies to both adjoining routers of the link. The link latency is the subtraction of two. If the response time from a further router is lower than a closer router, we assume the link latency is zero. All latency measurements are collected every 30 seconds and $traceroute$ measurements are collected every 10 minutes.

IV. Descriptive Statistics of QoE Anomalies

A. Prevalence of QoE anomalies among users

During 100 hours of video streaming from 124 emulated users in PlanetLab and Azure, QRank detects totally 9440 QoE anomalies. 65 users out of 124 experience QoE anomalies. We count the number of QoE anomalies per user in Figure 3. Results show that a small number of users experience a huge number of QoE anomalies while most users have no QoE anomalies or few QoE anomalies in two days. The top 10 users account for 8049 QoE anomalies in total of 9440 QoE anomalies (85.26%). Most QoE anomalies are severe (4485 out of 9440) and medium (4861 out of 9440) anomalies. Among all QoE anomalies, more than 99% of QoE anomalies are severe and medium QoE anomalies. It indicates that during anomaly period, users have more than 20% chunks with QoE values less than q_0. They probably stream videos at two lowest bitrates all the time.

We notice that some users experience QoE anomalies with an average period longer than 30 minutes. We denote anomalies lasting longer than 900 seconds (i.e. 15 minutes)

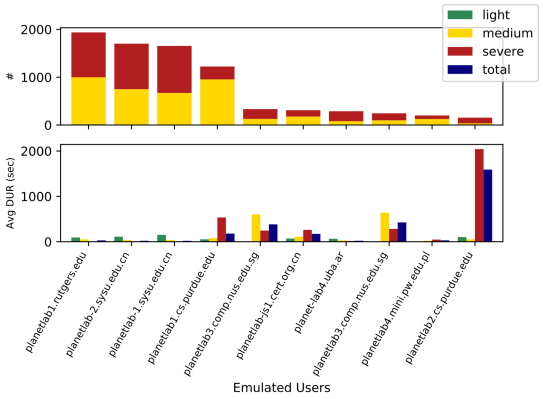


Figure 3. Count and the average duration of QoE anomalies per user (Show top 10 users with maximum number of QoE anomaly)

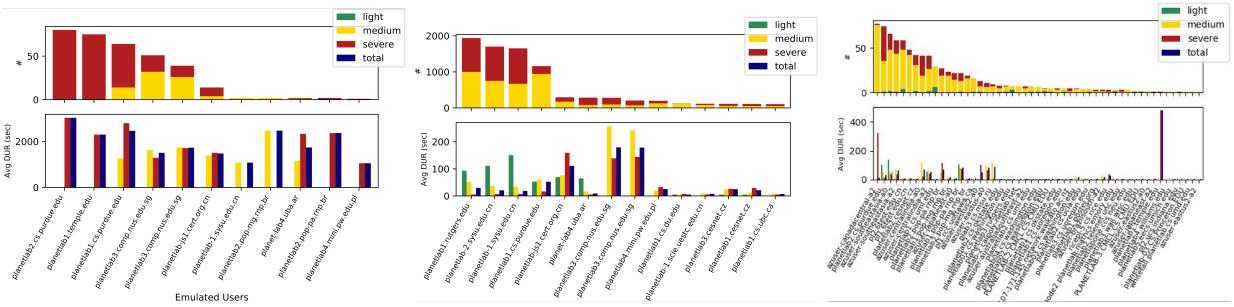
as *persistent* QoE anomalies. QRank detects that there are 11 users experiencing 332 *persistent* QoE anomalies. The top 6 users experience more than 97% (323 in 332) *persistent* QoE anomalies. All of them are PlanetLab users. We list the number and the average duration of all persistent QoE anomalies they experience in Figure 4 (a). Almost all persistent QoE anomalies are severe and medium QoE anomalies. There are users experiencing short QoE anomalies that occur frequently. These anomalies last less than 15 minutes but occur *recurrently*, namely on average occurring more than once per hour. We denote these as recurrent QoE anomalies. All users with *recurrent* QoE anomalies are shown in Figure 4 (b). Among 65 users with QoE anomalies, there are 14 users experiencing *recurrent* QoE anomalies. The top 4 users with most *recurrent* anomalies get 77.8% (6453 out of 8292) of all *recurrent* QoE anomalies. All other QoE anomalies are denoted as *occasional* QoE anomalies, which last less than 900 seconds and on average occur less than once per hour. Figure 4 (c) shows that the distribution of *occasional* QoE anomalies among users is in a long-tail shape. Overall, 47 users only have *occasional* QoE anomalies and 4 users experience both *occasional* and *persistent* QoE anomalies. Among all 9440 QoE anomalies, there are 332 persistent QoE anomalies, 8292 recurrent QoE anomalies and 812 occasional QoE anomalies. Around 91.4% of QoE anomalies are *persistent* and *recurrent*. These anomalies occur on 19 users. (15 users with *recurrent* anomalies plus 11 users with *persistent* anomalies minus 7 users with both anomalies as shown in Table I.) From above results, we show that *persistent* and *recurrent* QoE anomalies are not prevalent and only occur on few users. *Occasional* QoE anomalies are prevalent among users and its distribution over users follows a long-tail distribution. We later show that these occasional QoE anomalies are mostly caused by occasional traffic congestion in different networks.

B. Types of anomalous systems

QRank identifies QoE anomalies in user devices/home networks, access networks, transit networks, cloud networks

Table I
OF USERS WITH QOE ANOMALIES

Type of QoE anomalies on emulated users	# of users	Emulated User Examples
Total # of emulated users	124	N/A
Emulated users got QoE anomalies	65	N/A
Emulated users get ***occasional*** QoE anomalies	51	azuser-canadacentral-a2, planetlab2.cs.okstate.edu
Emulated users get ***recurrent*** QoE anomalies	15	planetlab1.rutgers.edu, planetlab-2.sysu.edu.cn
Emulated users get ***persistent*** QoE anomalies	10	planetlab2.cs.purdue.edu, planetlab1.temple.edu



(a) Persistent QoE anomalies (b) Recurrent QoE anomalies (c) Occasional QoE anomalies

Figure 4. Count and the average duration of QoE anomalies over users

and CDN servers. When a QoE anomaly is detected on a video session, all systems in the video streaming are analyzed. QRank identify the system with the lowest average QoE value during the anomaly period as the anomalous system. We count the number and the average duration of QoE anomalies identified in different types of anomalous systems in Table II . We observe that 62.36% (5887 out of 9440) QoE anomalies identify network systems as the anomalous systems. Among those anomalies, 94.58% (5568 out of 5887) QoE anomalies identify access networks as their anomalous systems. 23.32% (1373 out of 5568) QoE anomalies identify transit networks as anomalous systems. Only 1 QoE anomaly identify Cloud networks as anomalous systems. QoE anomalies identified in access and transit networks usually last long. Anomalies caused by access networks on average last 139.22 seconds. Anomalies caused by transit networks on average last 185.992 seconds. Client devices/home networks are identified as anomalous systems for 3573 QoE anomalies (37.85%) and these anomalies on average last 87.153 seconds. We then detail anomalies in each type of systems.

C. QoE anomalies identified in access networks

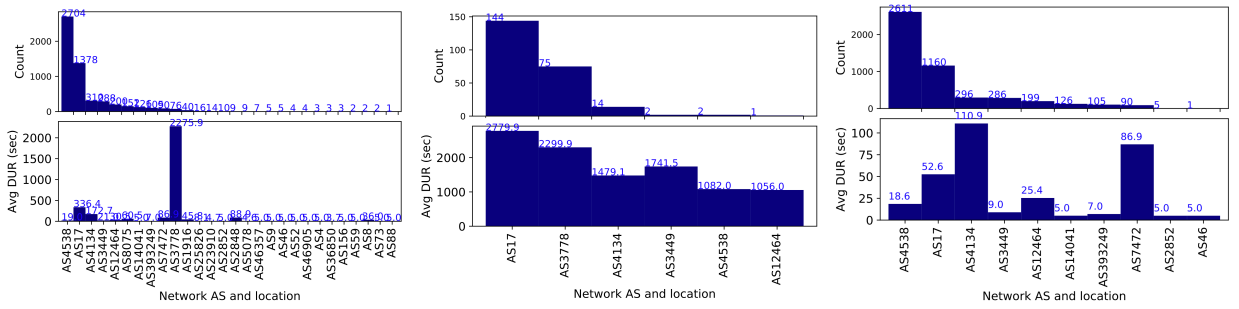
In Figure 5, we count the number and the average duration of QoE anomalies caused by access networks. The top three networks totally incur 78.8% (4392 out of 5568) of all anomalies in this category. These are AS4538 with Name "China Education and Research Network Center", AS17 with name "Purdue University" and AS4134 with name "No.31,Jin-rong Street". We study users connecting through these networks and we find that AS4538 is the access network of "planetlab-1.sysu.edu.cn" and "planetlab-2.sysu.edu.cn". AS17 is the access network for "planetlab1.cs.purdue.edu", "planetlab2.cs.purdue.edu". AS4134 is the access network for user "planetlab-js1.cert.org.cn". These users get many anomalies as shown in Figure 3. As shown in Figure 5 (b), AS17 and AS4134 also incur persistent QoE anomalies. There is another network, AS 3778 with name "Temple University", incurring persistent QoE anomalies that last more than 38 minutes (2299.9 seconds) on average. Among total 238 persistent anomalies caused by access networks, the top 3 networks incur 233 persistent anomalies. As shown in Figure 5 (c), most recurrent anomalies also occur in few networks. The top 3 networks totally incur 83.36% (4067 out of) of all recurrent anomalies identified in access networks. The top 6 networks incur total 95.88% (4678 out of 4879). These recurrent anomalies last from 9 seconds to 2 minutes (110.9 seconds) on average. There are also networks that only cause QoE anomalies occasionally, such as AS46357 (California Polytechnic State University) and AS 88(Princeton University).

D. QoE anomalies identified in transit networks

In Figure 6, we study QoE anomalies caused by transit networks. There are totally 38 anomalous transit networks and they totally cause 1373 QoE anomalies. AS262589 (INTERNEXA Brasil Operadora de Telecomunicacoes S.A) is identified to cause the most QoE anomalies. The distribution of QoE anomalies among transit networks has a long tail. The top 10 transit networks incur more than 82.45% (1132 out of 1373) of all anomalies in this category. We notice that AS 7922 (Comcast Cable Communications, LLC) and AS 3491 (PCCW Global, Inc.) cause 94.68% of persistent QoE anomalies (89 out of 94). In addition to AS 3491, there are 4 other transit networks that cause many recurrent QoE anomalies. They are AS19037 (AMX Argentina S.A.), AS262195 (Transamerican Telecomunication S.A.), AS262589 (INTERNEXA Brasil Operadora de Telecomunicacoes S.A), and AS6762 (TELECOM ITALIA SPARKLE S.p.A). AS3491 incurs recurrent QoE anomalies

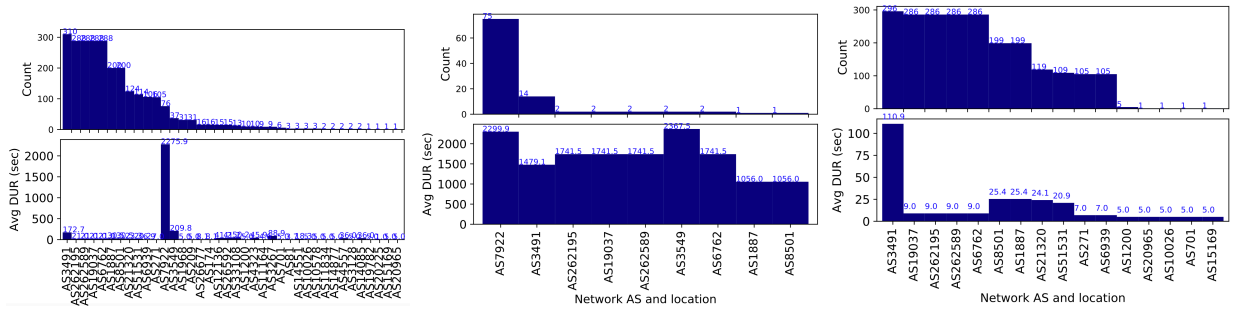
Table II
QOE ANOMALY STATISTICS PER ANOMALOUS SYSTEM TYPES

Type of anomalous systems	Networks				Client	Server
Anomalous system	Cloud Network	Transit Network	Access Network	All Networks	User devices/Home Network	Servers
Count of anomalies	1	1373	5568	5887	3573	0
Mean Duration (sec)	5.0	185.992	139.22	133.97	87.153	N/A



(a) All QoE anomalies (b) Persistent QoE anomalies (c) Recurrent QoE anomalies

Figure 5. The count and the average duration of QoE anomalies identified in access networks



(a) All QoE anomalies (b) Persistent QoE anomalies (c) Recurrent QoE anomalies

Figure 6. The count and the average duration of QoE anomalies identified in transit networks managed by different ISPs.

lasting longer than 100 seconds on average. The other 3 transit networks incur recurrent QoE anomalies lasting less than 10 seconds on average.

E. QoE anomalies identified in devices

Around 37.85% (3573 out of 9440) QoE anomalies identify user devices as anomalous systems. In Figure 7, we show the count and the average duration of all anomalies identified in devices. In our experiments, we use PlanetLab nodes and Azure servers to emulate users. We notice that PlanetLab nodes totally cause more than 95% (3408 out of 3573) of QoE anomalies in this category. The Azure servers only incur less than 5% of QoE anomalies. In real world, the client-side anomalies can be caused by users' devices, such as TV, phone, pad, or home network devices, such as routers or modems. Among all anomalies caused by devices, there are 92 persistent QoE anomalies and 3194 recurrent QoE anomalies. All of them are identified in PlanetLab nodes. We infer that these PlanetLab nodes have outbound capacity limit.

V. ROOT CAUSE ANALYSIS FOR QOE ANOMALIES

A. Root Cause Analysis for Persistent QoE Anomalies

Transit and access networks in our experiments cause most persistent QoE anomalies. We study network measurements during two persistent QoE anomalies. We find that latencies to anomalous networks also increase and fluctuates during anomalies.

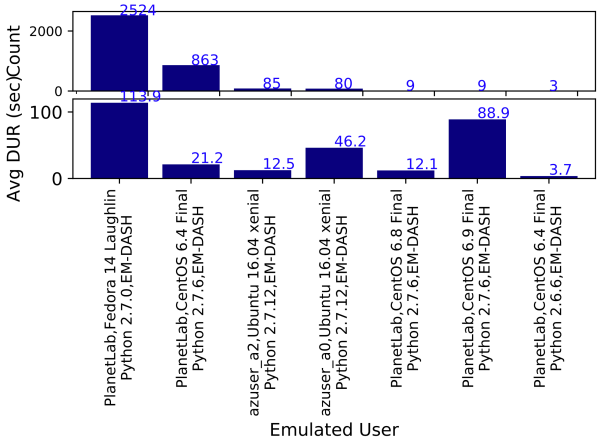


Figure 7. Count and the average duration of QoE anomalies per user (Show top 10 users with maximum number of QoE anomaly)

B. Persistent QoE anomaly identified in access network

In Figure 8, we show an example of persistent QoE anomaly on user "planetlab2.cs.purdue.edu" identified in access network AS17 (Purdue University). Figure 8 (a)

draws the located suspect nodes that are exclusively on the routes of users with QoE anomalies. Figure 8 (b) highlights the access network that is identified by QRank as the anomalous system. QRank compares aggregated QoEs among all networks involved and identify the one with the lowest aggregated QoE (0.835) as the anomalous system. We then show the probed latencies to all routers that are on the streaming path in Figure 8 (c). It shows that the latencies from the user to the routers in the network AS 17 fluctuate strongly. AS17 is the access network. The latencies to those routers are expected to be much lower than routers in other networks. However, we observe that the latencies to the router "128.10.127.251" in AS17 is similar to or just slightly lower than the latencies to the server "72.21.81.200". The latencies to "128.10.127.251" can increase up to longer than 80 milliseconds. The maximum latency to the router is even longer than the maximum latency to the server. It indicates that the end-to-end latency to the server is mainly attributed by the latency in access network AS17. We then estimate the link latencies from $traceroute$ data. We draw the estimated latencies for all links on the user's path in Figure 8 (d). We observe that the latencies for links in AS17 are on average larger and fluctuate stronger. Both users through AS17 network experience persistent QoE anomalies. It indicates the users choose the lowest bitrate during the anomaly. It is reasonable to infer that there is not enough capacity in network AS17, which cause the QoE anomalies.

C. Persistent QoE anomaly identified in transit network

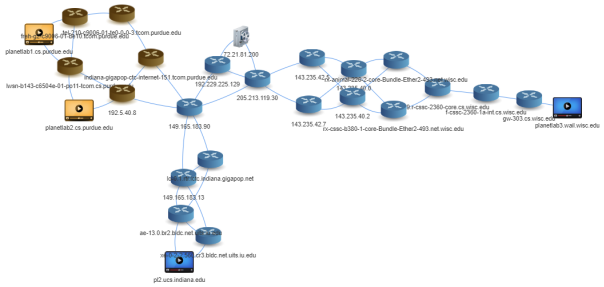
In Figure 9, we show the latencies measured during an example persistent QoE anomaly identified in transit network AS7922 (Comcast Cable Communications, LLC). The QoE anomaly last 2193 seconds on emulated user "planetlab1.temple.edu". We probe all routers on the user's path to its cache server. The latencies to all routers from the user is shown in Figure 9 (a). The monitoring agent on the user probes routers with 10 pings every minute. The latency is estimated by the mean of round trip times of 10 pings. It shows that the latencies from user to routers in AS7922 increases frequently. These routers include "69.241.67.106", "50.207.243.129" and "69.241.67.186". We also probe the cache server with $traceroute$ measurement every 10 minutes. We estimate link latencies from $traceroute$ responses between adjoining hops. The link latencies during the QoE anomaly period is shown in in Figure 9 (b). It shows that the latencies on link between router "50.207.243.129" and the router "69.139.192.169" are on average higher than other links. It can be verified in Figure 9 (a) as well. The probed latencies to router "69.139.192.169" is low. The maximum latency to "69.139.192.169" is 2.5259 ms. However, the latency to "50.207.243.129" is on average higher (4.7534 ms) with stronger fluctuations. The maximum latency to 50.207.243.129 is 33.4833 ms. This can be caused by dynamic queue length on this router. It also indicates that the traffic through the transit network is bursty.

D. Root Cause Analysis for Recurrent QoE Anomalies

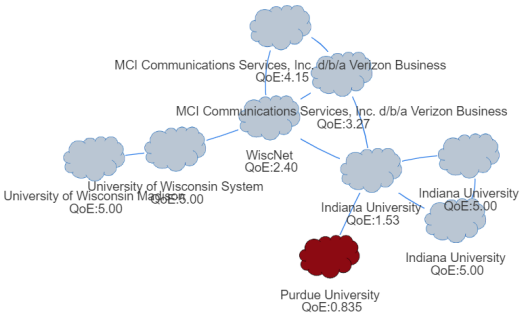
1) Recurrent QoE anomalies identified in access network: Figure 10 (a) shows recurrent QoE anomalies on user "planetlab-1.sysu.edu.cn" and "planetlab-2.sysu.edu.cn" who both access Internet through the network AS4538. The QRank identifies AS4538 as the anomalous network. We probe all routers in the streaming paths from a monitoring agent in the closest Azure region. In Figure 10(b). It shows that the latencies to the routers in AS4538 increase above 200 ms frequently. We also highlight the anomaly periods in shaded area in Figure 10 (b). Many QoE anomalies co-occur with those latency peaks. From Figure 10 (a), we see that both emulated users have QoE just above 2 even in non-shaded area. It can be inferred that there is not sufficient capacity for them to stream higher bitrates. When the latencies to routers in AS4538 increases, the packets going through the network have longer delays, resulting QoE below 2. We also study the latencies to various routers in AS4538 through a 2-day period. The latencies show strong fluctuations. We find that the latency peaks (above 200 ms and increases up to 250 ms) occur recurrently with an average interval of 83.83 seconds.

E. Recurrent QoE anomalies identified in transit network

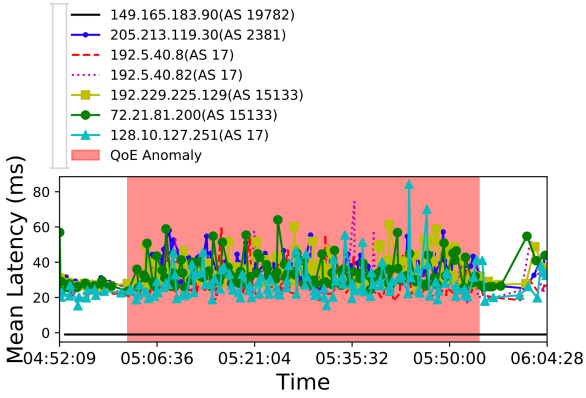
In Figure 6 (c), we show the count of recurrent QoE anomalies in various transit networks. The top anomalous transit networks are AS 3491, AS19037, AS262195, AS262589, etc. We pick up a QoE anomaly on a user streaming through AS262589 to show possible anomaly root causes. The recurrent QoE anomaly is on user "planet-lab4.uba.ar". The anomaly occurs very often during the streaming. We only draw QoE curve in a 10 minute time window as shown in Figure 11 (a). QRank identifies the anomaly in the transit network AS262589. We let a monitoring agent on the user probing all routers on its streaming path. Figure 11 (b) shows the latencies to all routers. The latencies to all routers seem very steady. There is one router "177.84.161.134" in AS262589 with constant long latencies (around 150 ms). The probed latencies to the router are even longer than the probed latencies to the server ("192.16.48.200" in AS15133). It is either due to the long queue length on the router or the low priority of $Ping$ traffic on the router. Long queue indicates that the capacity of the router is barely adequate to handle the traffic. We also observe that user "planet-lab4.uba.ar" overall has a medium QoE with chunk QoE values from 2 to 3. The QoE frequently drops below 2 but the QoE anomalies usually last less than 10 seconds. We infer that the recurrent QoE anomalies are related to the recurrent increases of traffic in AS262589. The capacity in AS262589 might be insufficient



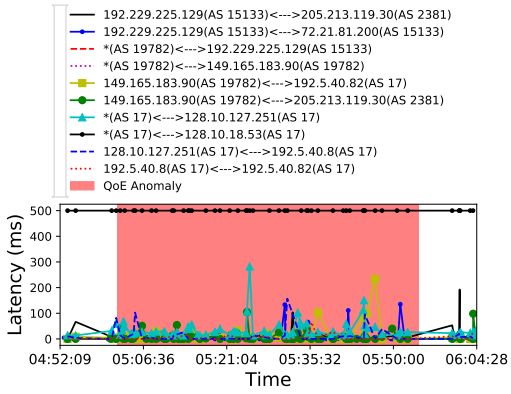
(a) Localization of suspect nodes



(b) Anomalous Systems identified by QRank

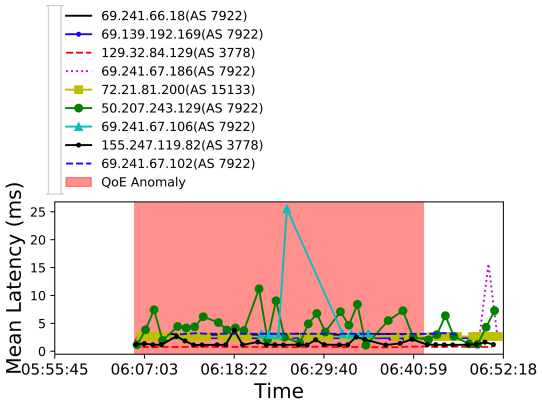


(c) Latencies to all routers involved in the video streaming

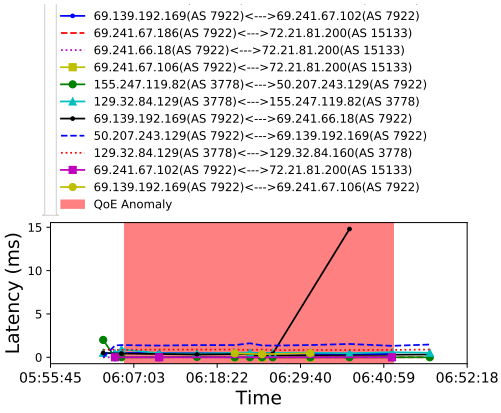


(d) Inferred latencies on all links (Traceroute)

Figure 8. A persistent QoE anomaly identified in access network



(a) Mean latencies from a user to routers on the session's path

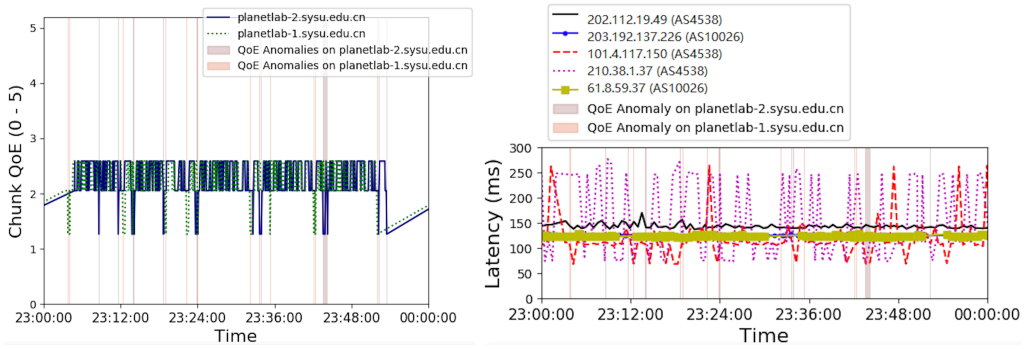


(b) Inferred latencies on all links (Traceroute)

Figure 9. A persistent QoE anomaly identified in transit network

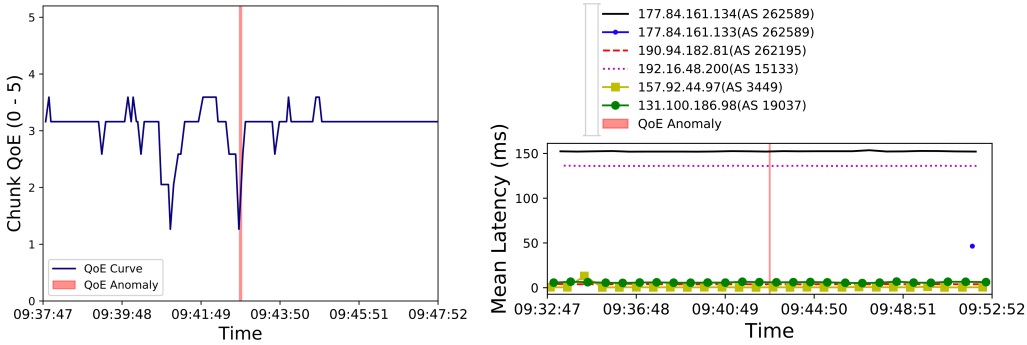
to provide higher QoE, so a slight increase in traffic can decrease the user QoE below 2.

1) Recurrent QoE anomalies identified in devices: From Figure 4 (b), we notice that the user with the most recurrent QoE anomalies is "planetlab1.rutgers.edu". The anomalous system for the anomaly is identified in the device. The device is a PlanetLab node installed with Fedora 14 Laughlin OS. It runs our emulation code of DASH player in the environment of Python 2.7.0. In Figure 12, QRank identifies the anomaly in the client as it is the only node exclusively on the anomalous user's path. There are many QoE anomalies like this. We find that all persistent and recurrent QoE anomalies in this category are caused by PlanetLab nodes. We infer that these PlanetLab nodes may have outbound capacity limit. There are only 4.9% QoE anomalies identified in Azure A0 and A2 instances. These are all occasional QoE anomalies. Azure A0 and A2 instances are the most economical virtual machines that share physical resources with other tenants. Their performance can degrade occasionally under interference.



(a) QoEs on anomalous users (b) Latencies from an agent to routers in suspect networks

Figure 10. Recurrent QoE anomalies identified in access network



(a) QoEs on the user with recurrent QoE anomaly (b) Latencies probed from the user to all routers on useŕs path

Figure 11. Recurrent QoE anomalies identified in transit network

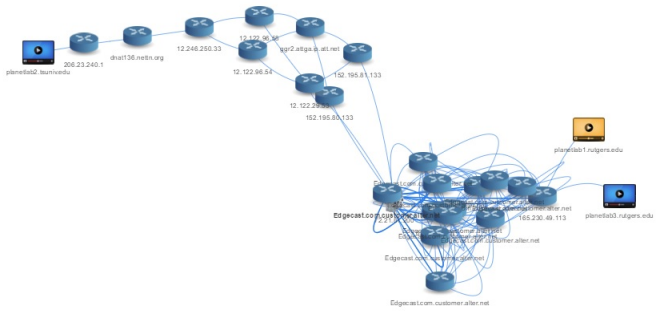


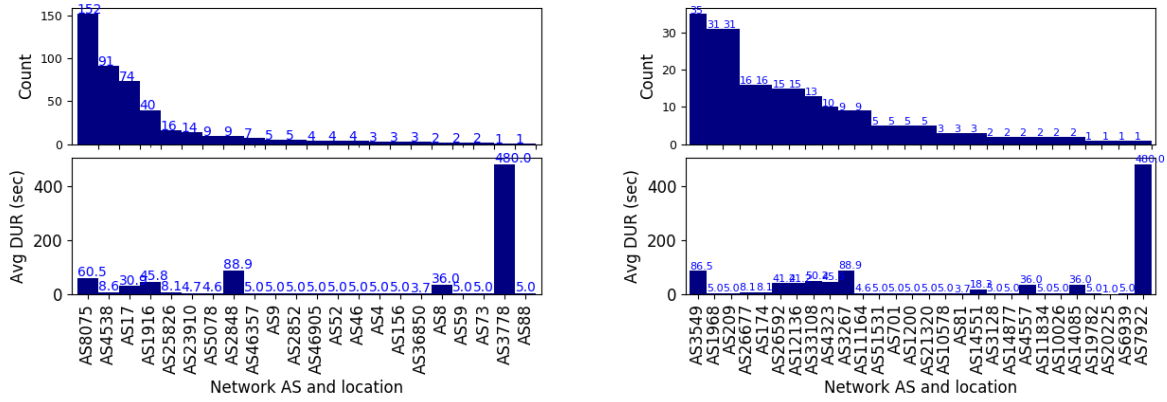
Figure 12. QoE anomaly identified in device

2) Root Cause Analysis for Occasional QoE Anomalies: Occasional QoE anomalies are widely identified in various types of networks. In the above sections, we show that more than 95% of persistent and recurrent QoE anomalies are identified in 3 access and 2 transit networks. The occasional QoE anomalies are identified in more networks. The occasional QoE anomalies distribute over various anomalous networks following long tail distributions as shown in Figure 13. There are no special patterns observed in latencies to these networks. As these anomalies usually last very short period and do not recur, they might be caused by occasional bursty traffic in these networks. There is only one occasional QoE anomaly identified in the Cloud network. It is on user "planetlab1.cesnet.cz". We study the anomaly identification result from QRank. We find that all systems involved the video streaming are identified anomalous. As QRank identifies anomalous system purely based on users' QoE, its accuracy is poor when there is no other users using the same network. As the only QoE anomaly identified in Cloud network is due to the insufficient accuracy of QRank system, we believe the Cloud network seldomly causes QoE anomalies. The accuracy of QRank is detailed in [4].

VI. Related Work

A. QoE anomaly analysis

Existing studies collect and analyze the QoE measurement from YouTube [10], large-scale live video streaming events [11], and Internet streaming services [12]. Pedro et al [10] study the correlation between the server changes and the QoE relevant degradations. They infer that the root causes behind QoE degradations are linked to Google CDNś server selection strategies. Their measurement data are from one ISP at single location and their conclusion may not be true for users worldwide. Juncheng et al[12] collect QoE measurement data worldwide from 379 video service providers. They cluster QoE anomalies over the space of client/session attributes, including the CDN, the client AS and the connectivity type. However, they ignore many other systems involved in the video streaming, such as transit networks. A client can receive videos from the same CDN through different transit ISPs that would give completely



(a) Access networks

(b) Transit networks

Figure 13. Occasional QoE anomalies identified in various types of networks

different QoEs. [11] analyzes Quality of Experience for a live streaming event in North America and finds lower engagements for users with QoE impairments.

B. QoE anomaly localization and diagnosis

QWatch system [7], locates nodes that are exclusively on the routes of users with QoE anomalies as suspect nodes. However, node-level localization provides little insights about the systems. It is also less accurate when the load-balancing networks introduce dynamic routing. [13] diagnoses QoE anomalies for video streaming on mobile devices. They correlate QoE anomalies with anomalies detected in network/device system measurements. However, detailed network and system measurements are usually not available to the VoD providers. We use QRank [4] to identify the QoE anomalies in production environment. QRank identifies QoE anomalies at system level without network measurements.

VII. Conclusion

As video services starts migrating to the Cloud, video service providers are wondering whether the Cloud can provide good Quality of Experience (QoE) for their users. In this paper, we emulated 124 users around the world to perform DASH video streaming from Microsoft Azure Cloud CDN to measure the performance of the Cloud CDN in terms of user QoE. We collected QoE anomalies from 12400 video sessions and identified the anomalous systems that cause those QoE anomalies. Interestingly, the Cloud CDN does not incur any QoE anomalies. Instead, transit networks, access networks and devices are major causes of QoE anomalies. Besides, more than 91.4% QoE anomalies are experienced by only 15.32% users and these users experience QoE anomalies either persistently or recurrently. 2 transit networks and 3 access networks incur more than 95% of all persistent QoE. 6 access networks and 10 transit networks incur more than 95% of all recurrent QoE anomalies. If capacity in these anomalous networks can increase, more than 95% QoE anomalies would be prevented. We conclude that to provide good QoE for video services, the Cloud provider should work with access/transit ISPs to increase the capacity of end-to-end connections.

References

[1] Y. Wu, C. Wu, B. Li, X. Qiu, and F. C. Lau, "Cloudmedia: When cloud on demand meets video on demand," in *ICDCS*. IEEE, 2011.

[2] "Azure SLA," https://azure.microsoft.com/en-us/support/legal/sla/, 2017.

[3] T. Stockhammer, "Dynamic adaptive streaming over http: standards and design principles," in *MMSys*. ACM, 2011.

[4] C. Wang, "QoE Based Management and Control for Large-Scale VoD System in the Cloud – chapter 5," Ph.D. dissertation, Carnegie Mellon University, 2017.

[5] "Microsoft azure," https://azure.microsoft.com/en-us/, 2017.

[6] B. Chun, D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. Bowman, "Planetlab: an overlay testbed for broad-coverage services," *SIGCOMM Review*, 2003.

[7] C. Wang, H. Kim, and R. Morla, "QWatch: Detecting and Locating QoE Anomaly for VoD in the Cloud," in *CloudCom*. IEEE, 2016.

[8] Y. Zhao, H. Jiang, K. Zhou, Z. Huang, and P. Huang, "Meeting service level agreement cost-effectively for video-on-demand applications in the cloud," in *INFOCOM*. IEEE, 2014.

[9] E. Marilly, O. Martinot, H. Papini, and D. Goderis, "Service level agreements: a main challenge for next generation networks," in *ECUMN*. IEEE, 2002.

[10] P. Casas, A. D'Alconzo, P. Fiadino, A. Bär, A. Finamore, and T. Zseby, "When youtube does not work: Analysis of qoe-relevant degradation in google cdn traffic," *TNSM*, 2014.

[11] A. Ahmed, Z. Shafiq, and A. Khakpour, "QoE analysis of a large-scale live video streaming event," in *SIGMETRICS*. ACM, 2016.

[12] J. Jiang, V. Sekar, I. Stoica, and H. Zhang, "Shedding light on the structure of internet video quality problems in the wild," in *CoNext*. ACM, 2013.

[13] G. Dimopoulos, I. Leontiadis, P. Barlet-Ros, K. Papagiannaki, and P. Steenkiste, "Identifying the root cause of video streaming issues on mobile devices," in *CoNext*. ACM, 2015.