QoE Evaluation of Integrated Satellite-Terrestrial Network on a Real-World Testbed

1st Mahshid Noorani
Electrical and Computer Engineering
Institute for Systems Research
University of Maryland
College Park, MD, US
mnoorani@umd.edu

1st Asim Zoulkarni
Electrical and Computer Engineering
Institute for Systems Research
University of Maryland
College Park, MD, US
asimz@umd.edu

2nd John S. Baras
Electrical and Computer Engineering
Institute for Systems Research
University of Maryland
College Park, MD, US
baras@umd.edu

***Abstract*—Integrated Satellite Terrestrial Networks (ISTNs) play a critical role in achieving the key promises of 5G and 6G and providing broadband connectivity globally, especially where extending terrestrial infrastructure is cost prohibitive. Offloading part of the terrestrial traffic load to the satellite network is a cost-effective solution to avoid congestion and enhance the overall capacity of next-generation networks. However, such an integration introduces several design challenges with regard to managing the heterogeneity of the network, satisfying the latency and throughput requirements of different application types, and minimizing the overall cost of deployment. In this work, we set up a real-world ISTN testbed utilizing Hughes Network Systems' laboratory to evaluate the Quality of Experience (QoE) of several popular applications over a hybrid terrestrial-satellite broadband connection. The results indicate that the proposed ISTN's versatility and its ability to utilize intelligent path control improve QoE for common applications, even at modest levels of wireless bandwidth.**

***Index Terms*—Broadband, Satellite Internet, Terrestrial Wireless, 5G, Hybrid Networks, Real-Testbed, QoE**

I. Introduction

Abundant amount of information and services are available in today's broadband networks. The ongoing digital transformation requires effective utilization of these services and promises change in every aspect of work, life and society. Further, we are moving fast towards a networked immersed world, the metaverse. However, access to broadband is limited to 29% of the population world-wide as reported by OECD [1] – accepting great heterogeneity from country to country and geographic regions. The pandemic demonstrated that without widely available broadband access to these data, and the services they enable, we cannot have democratized education, healthcare, manufacturing, environmental monitoring, jobs, housing, food, and water resources. As we plan future communication infrastructures involving 5G, 6G, NextG, we must consider the implementation cost. It is clear that it is cost prohibitive to provide broadband to many areas of the world via fiber. Recent studies suggest that the current model of uniform pricing, which applies to customers of different needs, is likely to become unsustainable for rural broadband provision [2]. Broadband communications should be provided to all, in the same way we provide transportation roads to all, regardless of where they live.

Supported by Hughes Network Systems, LLC.

This observation leads us to consider hybrid communication technologies combining terrestrial and non-terrestrial means. Their cost-effective nature is an attractive element; however, we need to carefully investigate the requirements of typical applications in terms of latency and throughput requirements in order to be able to dynamically (slicing) provide high quality user experience. Various governments have recently introduced Infrastructure Bills providing funding to make Broadband Services available to most people [3]. However, the funding is not adequate and often the requirements for parameters such as latency and throughput are often excessive and do not represent experimental evidence of widely used applications (i.e., often require excessive response speed and bandwidth).

Towards achieving the key promises of ubiquitous high-performance broadband, it is essential to utilize the capacity of various types of communication networks (i.e., terrestrial, space, aerial) and supporting technologies, such as software defined networking, simultaneously, as opposed to the traditional standalone fashion. In fact, the recent advancement on SDN [4] and path control algorithms renders hybridization of networks an attractive approach. Satellites can deliver broadband highly cost effectively to rural and areas of low density, but the high delay communication may impact the Quality of Experience (QoE) of latency sensitive applications. 4G and 5G wireless coverage has become very extensive, but often narrow band connections are delivered in rural areas.

In this work, we systematically evaluate the QoE of several popular applications over satellite, terrestrial and hybrid terrestrial-satellite broadband connection. Analysis on latency sensitivity of popular applications, based on a soft clustering of applications as latency sensitive or insensitive shows that only roughly 15% percent of the downstream user traffic and 20% of the upstream user traffic is latency sensitive, therefore most of the traffic can be offloaded to satellites and this motivates the use of hybrid systems [5]. A well-designed hybrid system would intelligently integrate satellite and wireless networks to cost effectively deliver high performance.

The remainder of this paper is organized as follows. In Sec-

2770-7679/22/$31.00 ©2022 IEEE
DOI 10.1109/FNWF55208.2022.00106

tion II, the testbed set-up is illustrated. Section III elaborates on the application selection and evaluation process. We discuss the experiment set-up, parameters, and methodology in Section IV, followed by results in Section V, and conclusion and future work in Section VI.

II. Testbed

We set up a testbed on Hughes Network Systems' lab to study cost effective ISTNs that meet the expectation of end-users in terms of QoE. We construct a hybrid network architecture involving a terrestrial wireless and a satellite component, which is designed to meet the 100 Mbps and 20 Mbps downlink and uplink throughput requirements, respectively, matching the most recent proposed US broadband definition. This is achieved by utilizing the already existing terrestrial wireless infrastructure and augmenting the total throughput provided to the customer with the introduction of a high-throughput satellite link. This idea is realized by advanced path steering algorithms that classify and intelligently steer packets between the high latency and low latency paths (Hughes Active Technologies) and is particularly attractive for rural and sparsely populated areas, where the existing wireless infrastructure is insufficient, yet with the introduction of a versatile high throughput satellite link, sustainable broadband connectivity to populations that are adversely affected by the digital divide can be provided. The hybrid architecture addresses delay-related issues by maintaining the high-throughput satellite link for the routing of latency insensitive traffic, while allowing the use of the wireless channel for the more throughput-wise limited latency sensitive traffic.

The ISTN block diagram is depicted in Figure 1. Equipment used to set-up the testbed are described in Table I. We used two computing devices to simulate the wireless and satellite networks and used the APC equipment for active switching between the two networks. This set-up provided us the ability to set up and precisely control network parameters such as throughput and latency during the QoE evaluation tests. The switch is used for the definition of three virtual local area networks (VLANs) linked to the Hughes local area network, the cellular network and simulated network separately. The Wi-Fi router is connected to the simulated network and provides for that access to wireless devices.

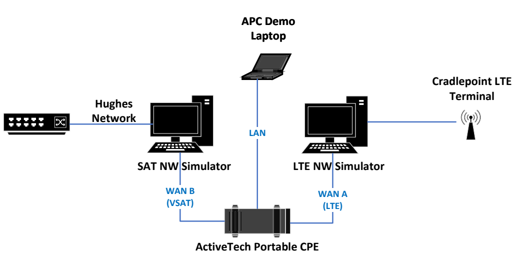

Fig. 1: ISTN Testbed simplified block diagram: The terrestrial wireless simulator is used for throughput parametrization.

Equipment	Definition
Qotom-Q555G6	ActiveTechnologies Portable Customer Premises (APC) equipment is an industrial PC with multiple Ethernet ports running portable version of Hughes ActiveTechnologies software. Also supports TCP satellite acceleration.
Dell OptiPlex 3010 SFF	PC with three 1 Gbps Ethernet interfaces running the wireless WTB network simulator
HP Pavilion s5-1024	PC with four 1 Gbps Ethernet interfaces running the satellite WTB network simulator
Cisco Switch SG250-26	26-Port Gigabit Smart Switch
TL-SG1005P TP-Link Switch	5-Port Gigabit Desktop Switch with 4-Port PoE
Cradlepoint IBR350LPE	4G LTE Cellular Router

TABLE I: Equipment for the testbed setup.

III. Application Selection & Evaluation

In order to decide which applications we select for our evaluation, we study the most prominent traffic generating ones, globally. The Sandvine Internet Phenomena reports [6]–[8], provide information on the applications and the percentage of the total internet traffic they generate, for both downstream and upstream directions. We consider the top ten application categories i.e., video streaming, web browsing, gaming, social, file sharing, marketplace, in decreasing order of traffic share, and observe that the four first categories account for approximately 80% of the total traffic. Additionally, during the COVID-19 pandemic, remote work and learning lead to the widespread use of video conferencing applications, which we study separately from video streaming. Hence, in this work we proceed with the analysis of and experimenting on our testbed with, popular applications selected from the video streaming, video conferencing, and web-browsing application categories.

Furthermore, to more accurately capture user experience, we study these applications in granular phases specific for each application category, as shown in Table II, so that functionalities like searching, loading, runtime, for video streaming are evaluated separately with regards to their resulting QoE.

In each application, we set variables such as QUIC (Quick UDP Internet Connection) enabling, video codec, video resolution, duration of browsing, browsing phrase, and noise suppression setting constant to ensure consistency of the experiments that result in reliable and comparable outcomes in terms of QoE. For video conferencing, we set the video resolution to 720p to provide the minimum high definition requirement using YouTube application. The resolution for the Netflix application is set to high (four quality options are provided i.e., low, medium, high, auto). We set the video codec to vp9, which is one of the most common codecs and is the codec that Netflix application uses as well; this selection allows us to stay consistent across applications in the video streaming category. The QUIC encrypted transport layer network protocol was enabled. This protocol is designed to make HTTP traffic more secure, efficient, and faster. We conducted experiments with both QUIC enabled and disabled and observed a higher quality of experience while QUIC was enabled.

Application Group	Application	Phases
Video Streaming	*YouTube* *Netflix*	• Searching • Loading • Runtime • Fast-Forward
Video Conferencing	*Zoom* *Microsoft Teams*	• Joining • Video off • Video on • Video on & Screen sharing
Browsing	*Chrome* *Firefox*	• Submit a query (seacrh) • Check 2 first results • Next Page

TABLE II: Selected applications are identified in the table above. QoE descriptions in plain language were then defined for each application, each mapped to a mean opinion score value (see Section IV-A).

IV. Experiment

In this section, we evaluate the minimum throughput requirements of commonly used internet applications that preserves appropriate QoE, for QoE-based optimal resource provisioning. More precisely, we quantify the QoE based on the mean opinion score (MOS), obtained from participants in a survey we conducted.

A. Mean Opinion Score

A well-established metric for the quantification of the user quality of experience is the Mean Opinion Score (MOS), also used as the ground-truth reference for research on objective quality modeling [9]. It is the mean of the values on a predefined scale that subjects assign to their opinion of the performance of the network system, originally introduced either for evaluation of conversation or for listening to spoken material [10]. Subjects provide their input score according to a description of what each value in a range from one to five corresponds to, in terms of explicitly described experiences; an example of the MOS we defined for each application and their phases is shown in Table III. Table III provides the opinion score definition that is used for the three phases of the browsing experience. We study each application category at a phase granularity so that different functionalities (of potentially different latency requirements) are split into different phases and we expand the MOS terminology to reflect the experience of different applications and their discrete phases, as shown for video streaming and conferencing in Figures 2, 3, respectively.

B. Survey

To evaluate the performance of the hybrid network and compare the performance of the five network settings mentioned in Table IV, we performed a survey on ten participants. The baseline experiments correspond to the current standard for broadband in the US i.e., 25 Mbps and 3 Mbps in terms of throughput for the downlink and uplink channels, respectively, while the proposed standard requires 100 Mbps and 20 Mbps for the downlink and uplink channels, respectively.

The survey participants are subjected to watching randomized series of segmented videos and are requested to score

Fig. 2: This figure provides the scoring definitions for each phase of a benchmark video streaming session.

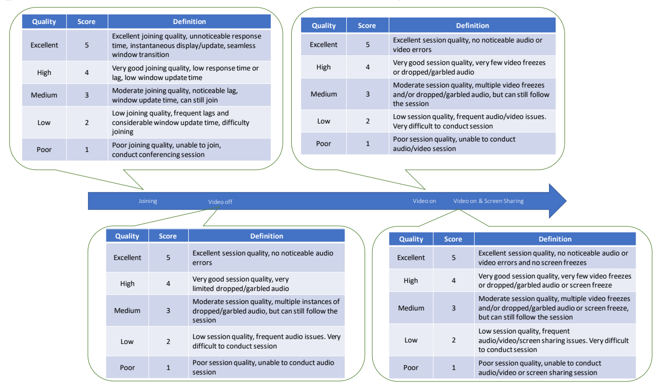

Fig. 3: Different opinion score definition tables are provided for each application and its distinct phases. In the figure above, the scoring definitions for each phase of a video conferencing session are provided.

Quality	Score	Description
Excellent	5	Excellent searching quality, unnoticeable response time, instantaneous suggestion display/update, seamless page transition
High	4	Very good searching quality, low response time or lag, low suggestion and page update time
Medium	3	Moderate searching quality, noticeable lag, suggestion and page update time, can still search
Low	2	Low searching quality, frequent lags and considerable suggestion and page update time, difficulty searching
Poor	1	Poor searching quality, unable to submit query, conduct searching session

TABLE III: Opinion score definition guide for browsing application evaluation, provided to survey participants.

the internal segments comprising the use of an application, referred to as phases (see Table II), based on score definitions as indicated in Table III. For each phase, different functionalities of the same application can be evaluated separately as to how they perform over high or low latency networks. For this purpose, we focus on applications used by a common household representative of video streaming applications such as, YouTube and Netflix, online collaboration tools such as, Zoom and Microsoft Teams, and web-browsers such as, Chrome and Firefox. The participants evaluate the QoE of each phase in experiments repeated under different network parameters as indicated in Table IV.

DL/UL (Mbps)	**5/1**	**3/1**	**2/1**	**25/3**	**100/20**
Wireline	✗	✗	✗	✗	✓
Satellite	✓	✓	✓	✓	✗
Hybrid (Wireless)	✓	✓	✓	✗	✗

TABLE IV: For hybrid simulations, the satellite variable is bound to 3/1 (determined by the first round of survey results based on satellite only), while the (terrestrial) wireless throughput is a free variable. The wireline 100/20 experiment serves as a baseline. The X/Y abbreviation is used to denote downlink and uplink throughput, X,Y, in Mbps, respectively.

V. Results

After obtaining the MOS for each of the various phases of each application, we proceed with the estimation of the satellite and wireless throughput requirements for satellite and hybrid terrestrial-satellite networks. We run our experiment for multiple combinations of downlink and uplink throughput values ranging from 1 to 5 Mbps. This range is chosen based on selected applications vendors' information, as well as tests to estimate such parameters that range in acceptable QoE (e.g. Figure 4). For example, video conferencing applications such as Zoom have throughput requirements ranging from 600 kbps (down,up) to 3.0/3.8 Mbps (down/up) [11].

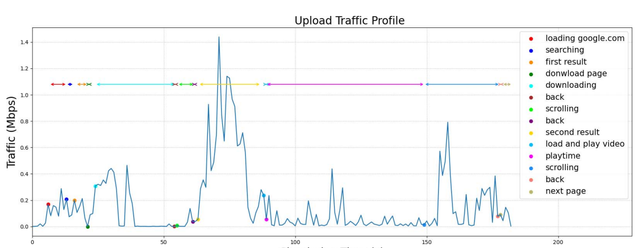

Fig. 4: Uplink traffic load over terrestrial network for browsing application on Google search engine [12].

In the Tables V,VI,VII, we present our findings on QoE over hybrid terrestrial-satellite network for the three application categories of video streaming, video conferencing, and web browsing, respectively. The evaluation of the throughput requirements for each phase is obtained from the minimum downlink capacity parameter that yielded at least a *medium* MOS. The results for each application were obtained in a way similar to that of the YouTube video streaming application results, summarized in Table V.

We study the satellite connection for 3 Mbps downlink and 1 Mbps uplink speeds (abbreviated as 3/1) which yield at least *medium* mean opinion score for all phases of video streaming and hybrid satellite-terrestrial with throughput parameters 3/1-2/1, respectively, which yield *high* experience for latency sensitive phases (i.e., searching). Subsequently, we evaluate the hybrid network with varying wireless downlink throughput parameters and observe at least a *high* mean opinion score for all phases of video streaming. The QoE we observed for the two streaming applications studied is similar, though the way each handles resolution quality in low throughput differs. YouTube preserves the specified resolution (i.e., 720p) and the video frame rate may decrease as a result, while Netflix degrades the resolution to avoid frame rate decreases; the latter is sometimes perceived as a better experience.

Throughput (Mbps)				**Quality of Experience** (MOS)			
Terrestrial		**Satellite**					
DL	***UL***	***DL***	***UL***	***Searching***	***Loading***	***FF***	***Runtime***
-	-	25	3	4	5	5	5
-	-	3	1	3.5	4	3.5	4
2	1	3	1	4.5	4	4.5	4
3	1	3	1	4.5	4	4.5	4
5	1	3	1	4.5	4	4.5	5
100	20	-	-	5	5	5	5

TABLE V: The data in the table above quantify the improvement of the internet over satellite after the introduction of a wireless link of limited throughput for **video streaming** (YouTube). *DL*: Downlink, *UL*: Uplink, *FF*: Fast-Forward.

Throughput (Mbps)				**Quality of Experience** (MOS)			
Terrestrial		**Satellite**					
DL	***UL***	***DL***	***UL***	***Join.***	***Vid. off***	***Vid. on***	***Screen sh.***
-	-	25	3	4	4.5	4.5	4
-	-	3	1	4	4	3.5	3
2	1	3	1	4	5	4	4.5
3	1	3	1	5	5	4.5	4.5
5	1	3	1	5	5	5	5
100	20	-	-	5	5	5	5

TABLE VI: quantification of the improvement of the internet over satellite after the introduction of a wireless link of limited throughput for **video conferencing** (Microsoft Teams).

Throughput (Mbps)				**Quality of Experience** (MOS)		
Terrestrial		**Satellite**				
DL	***UL***	***DL***	***UL***	***Search***	***Check 2 Res.***	***Next Page***
-	-	25	3	4	4	5
-	-	3	1	4	3	3
2	1	3	1	4	4	3
3	1	3	1	5	4	4
5	1	3	1	5	5	5
100	20	-	-	5	5	5

TABLE VII: Quantification of the improvement of the internet over satellite after the introduction of a wireless link of limited throughput for **web browsing** (Google Chrome).

For video conferencing, we observe that the 2/1-3/1 hybrid terrestrial-satellite throughput parameters yield *high* to *excellent* experience for all phases. For the video-off phase, where only audio is exchanged, its quality is reported at least *high*, with the one observed negative factor being delay. Delay remains consistently high in satellite-only networks, while on hybrid terrestrial-satellite networks audio can be transmitted through the low latency link, improving user experience. With the addition of video traffic in the next phase, the user experience is preserved with the introduction of the terrestrial link while requiring modest throughput for either transport i.e., 2 and 3 Mbps for the terrestrial and satellite downlink channels, respectively. In order for the QoE to remain acceptable during the video-on phase, the video and audio need to remain synchronized regardless of which transport each follows. This requires that the path steering algorithm either routes both audio and video flows through the same transport or compensates for the delay difference of the two transports to avoid lip sync issues and preserve *high* QoE. During the final phase, one user shares their screen and new screen sharing traffic is introduced, while the video traffic is reduced as a result of the video compression for the transmission of just the necessary video thumbnail users view of each other in applications such as Microsoft Teams or Zoom. High user experience for this phase is preserved by the same throughput parameters. The experience for the joining phase, where the user waits to be admitted to the call, is either *high* or *excellent* across all our experiments.

For web browsing applications, a combination of terrestrial downlink throughput of 3 Mbps with a satellite downlink throughput of 3 Mbps results in *excellent* searching experience, with regards to live suggestion update and perceived page loading time as the user submits a search query, scrolls through the suggested results and triggers loading of small size thumbnails, that nonetheless need to be displayed quickly for high QoE. Similarly with the video streaming applications not involving video uploading, web browsing is not an uplink intensive activity and the respective throughput of 1 Mbps is sufficient for acceptable user experience.

We conclude that even with modest wireless throughput requirements, the system produces high user QoE and the downlink and uplink terrestrial and satellite wireless throughput required is no more than 2/1 and 3/1, respectively for video streaming and conferencing, and 3/1 for the terrestrial and satellite links for high user QoE on web browsing.

VI. Conclusion & Future Work

A key aspect of providing broadband services to maximum number of people at an affordable price is the detailed understanding and quantification of the response latency and throughput requirements of commonly used Internet applications, specifically the low-latency throughput requirements. In this work, we evaluated the viability of a hybrid GEO-LTE/5G system to support the traffic demand of typical users with acceptable and competitive user QoE by deploying a real-world ISTN testbed. We conclude that throughput parameters of 2/1-3/1 (downlink/uplink in Mbps), for hybrid terrestrial-satellite network respectively, yield high QoE for video streaming and conferencing applications, which are the most demanding in terms of throughput, while an increase of the previous low-latency terrestrial downlink throughput parameter by 1 Mbps increases QoE from medium to high for web browsing applications. In future work, we plan to add other applications, including gaming, to our study as well as to incorporate the results of a large-scale survey.

Acknowledgment

We thank Sampath Ramaswami and Marjan Nowzadi from Hughes Network Systems for useful discussions and insights that greatly assisted the research, and for providing their suggestions and comments on the manuscript; we thank Doug Dillon, Shubham Sharma, Wesley Chen, and Randy Beazell from Hughes Network Systems for providing assistance with the testbed. This work was supported by Hughes Network Systems, LLC.

References

[1] Broadband Portal - OECD. https://www.oecd.org/sti/broadband/broadband-statistics/. Accessed 25 Aug. 2022.

[2] R. Layton, and P. Potgieter. "Rural Broadband and the Unrecovered Cost of Streaming Video Entertainment". International Telecommunications Society (ITS), 2021. ideas.repec.org, https://ideas.repec.org/p/zbw/itsb21/238035.html.

[3] "Infrastructure Investment and Jobs Act Implementation." Federal Register, 10 Jan. 2022, https://www.federalregister.gov/documents/2022/01/10/2022-00221/infrastructure-investment-and-jobs-act-implementation.

[4] N. Torkzaban, A. Gholami, J. S. Baras and C. Papagianni, "Joint Satellite Gateway Placement and Routing for Integrated Satellite- Terrestrial Networks," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-6

[5] A. Gholami, N. Torkzaban, P. Nikolaidis, A. Zoulkarni, J. S. Baras, "Classification of Application Traffic Demand and Flows based on Latency Sensitivity", Hybrid Networks Center (HyNet), University of Maryland, College Park.

[6] Sandvine, "The Global Internet Phenomena Report" Sandvine, Plano, TX, USA, 2019. September, 2019. [Online]. Available: https://www.sandvine.com/global-internet-phenomena-report-2019

[7] Sandvine, "The Global Internet Phenomena Report, COVID-19 Spotlight" Sandvine, Plano, TX, USA, 2020. May, 2020. [Online]. Available: https://www.sandvine.com/covid-internet-spotlight-report

[8] Sandvine, "2021 Mobile Internet Phenomena Report" Sandvine, Plano, TX, USA, 2021. May 2021. [Online]. Available: https://www.sandvine.com/download-mobile-internet-phenomena-report-2021

[9] J. Xu, L. Xing, A. Perkis and Y. Jiang, "On the Properties of Mean Opinion Scores for Quality of Experience Management," 2011 IEEE International Symposium on Multimedia, 2011, pp. 500-505, doi: 10.1109/ISM.2011.88.

[10] ITU-T, "Recommendation P.800.1 – Mean Opinion Score (MOS) terminology," July 2006.

[11] Zoom Video Communications, Inc., "Zoom system requirements: Windows, macOS, Linux", May 2022. [Online]. Available: https://support.zoom.us/hc/en-us/articles/201362023-Zoom-system-requirements-Windows-macOS-Linux

[12] N. Torkzaban A. Zoulkarni, P. Nikolaidis, A. Gholami, M. Noorani and J. S. Baras, "Integrated Satellite Terrestrial Networks: A GEO-based Resource Provisioning Use-case," 2022 First ACM/IEEE Next G Summit, Johns Hopkins University Applied Physics Lab, Laurel, Maryland, USA 2022.