Shifting the Conversation from Speeds, Feeds, and WiFi 7 to a Multidimensional Model for Flawless Application Performance

A Technical Paper prepared for SCTE by

Adam Hotchkiss
VP, Product Solutions
Plume Design, Inc.
325 Lytton Ave, Palo Alto, CA 94301, USA
adam@plume.com

Table of Contents

Title	Page Number
1. Abstract	3
2. Introduction	3
3. Evolution of 'speeds and feeds' and their relevance	3
4. The Role of WiFi in modern QoE	3
5. Unpacking the Seven Dimensions of Full Stack Optimization QoE	4
6. Full Stack Optimization: A Revolutionary Approach	6
7. Conclusion	13
Abbreviations	14
Bibliography & References	14

List of Figures

Title	Page Number
Figure 1 - Full Stack Optimization dimensions	6
Figure 2 – Dynamic Deep Packet Inspection CPE architecture	7
Figure 3 - Households running real-time applications	8
Figure 4 - Streaming application market share	8
Figure 5 - Video conferencing application market share	9
Figure 6 - Network topology with adaptive WiFi functionality	10
Figure 7 - Internet latency measurements for real-time applications	11
Figure 8 - Prioritized packet delivery for low latency applications	11
Figure 9 - Latency improvement for different application types in congested WiFi network	12
Figure 10 - Measurements for Application Performance	12

List of Tables

Title	Page Number
Table 1 – Application session statistics	7

1. Abstract

Central to every 21st-century home are 'smart' applications that require robust, secure, and dependable internet to operate smoothly. These applications encompasses everything from 4K video streaming to advanced heating controls and video conferencing. With ever-increasing network complexity, households are progressively dependent on Communications Service Providers (CSPs) for expertly-managed connectivity solutions to ensure superior Quality of Experience (QoE) across floorplans, devices, and apps.

Our industry spends an astounding amount of time discussing WiFi 7's expansive speed and performance as a means to solve for evolving application use in the home, but that's the wrong conversation. WiFi 7 is only an enabling technology to easily connect and send data between devices. We must fundamentally shift the conversation around delivering flawless application experiences for every device in the consumer's home, regardless of WiFi standards.

This paper proposes a comprehensive QoE delivery approach down to the application level, harnessing the potential of Plume's Full Stack Optimization. It further offers technical guidelines for CSPs on the deployment of cloud-based learning algorithms. These algorithms, paired with precise Key Performance Indicator (KPI) metrics, leverage capabilities of any WiFi networking protocols, and open-source platforms to address the intricate challenges of ensuring optimal connectivity.

2. Introduction

In today's digital age, households are no longer just physical spaces but digital ecosystems. Real-time applications such as Netflix, FaceTime, Zoom, and various online gaming platforms have emerged as mainstays. With this transformation, there's an elevated demand for a new way to optimize and measure true consumer Quality of Experience (QoE). Historically, the answer lay in amplifying "speeds and feeds", but such strategies now exhibit diminishing returns in ensuring optimal user experience.

3. Evolution of 'speeds and feeds' and their relevance

For decades, when users encountered issues with their online experience, the primary solutions were straightforward: boost the internet bandwidth or transition to the next-gen WiFi technologies. This method predominantly produced satisfactory results; however, now it is faltering in plain sight. Recent observations indicate diminishing efficacy due to two critical shifts:

1. Internet bandwidth–the speed at which data enters a home–now often surpasses the bandwidth needs within the home's internal network
2. Modern applications have evolved. They no longer demand raw bandwidth, but require consistent, low-latency data streams for optimal QoE.

Consequently, we find ourselves witnessing a paradigm shift, where the emphasis migrates from solely amplifying aggregate capacity to ensuring a consistent, application-specific delivery on each device, regardless of its spatial position and environmental conditions in the home.

4. The Role of WiFi in modern QoE

As technology continually advances, newer iterations of WiFi naturally bring improvements. However, is there a direct and significant correlation between WiFi evolutions and QoE? While WiFi 7 has its advantages, it's critical to recognize that WiFi, in essence, is merely a medium for transmission of data between two points. WiFi specifications are centered around bandwidth and connectivity specifications, caring little about the individual applications that run on top.

For example, below is the typical WiFi technology benefit comparison chart:

WIFI TECHNOLOGY BENEFIT COMPARISON

	WiFi 4	WiFi 5	WiFi 6	WiFi 6E	WiFi 7
Year	2007	2013	2019	2020	2024
Protocol	802.11n	802.11ac	802.11ax	802.11ax	802.11be
Bands	2.4/5GHz	5 GHz only	2.4/5GHz	2.4/5/6*GHz	2.4/5/6 GHz
Peak Speed	600 Mb/s	6.9 Gb/s	6.9 Gb/s	6.9 Gb/s	46 Gb/s
Channel Widths	20/40 MHz	20/40/80/160 MHz	20/40/80/160 MHz	20/40/80/160 MHz	20/40/80/160 MHz/ 320 MHz
Security	WPA2	WPA2	WPA3	WPA3	WPA3
Key Features	4X4 MIMO LDPC Error Correction 64-QAM	8X8 MIMO 4X DL MU-MIMO Beam Forming[8] 256-QAM	8X8 MIMO 8X DL MU-MIMO Beam Forming OFDMA TWT 1024-QAM	8X8 MIMO 8X DL MU-MIMO Beam Forming OFDMA TWT *6 GHz added, USA ONly.	16X16 MIMO 16X DL MU-MIMO Multi-AP Multi-RU Puncturing Multi-Link Beam Forming OFDMA TWT 4096-QAM

The table[1] predominantly emphasizes parameters such as modulation rates, maximum bit rates, potential number of MIMO streams, and the spectrum utilized. While these elements are crucial for enhancing the bandwidth capacities between devices, they do not directly address the core KPIs of application QoE, which is central to the consumer experience. Furthermore, these specifications are largely theoretical; for instance, the prospect of households benefiting from a 16x16 MIMO and 46 Gbps throughput remains distant. Therefore, relying solely on WiFi technologies to address the growing application QoE challenge does not meet the comprehensive needs essential for delivering a Full Stack QoE experience.

5. Unpacking the Seven Dimensions of Full Stack Optimization QoE

To genuinely grasp and elevate Application Performance (i.e. QoE), it's imperative to understand and address its seven dimensions:

> **Space**: Network performance shouldn't be hostage to a home's design or size. It should consistently deliver across various architectural layouts.
>
> **Time**: It’s not just about speed, but also about unwavering performance irrespective of external temporal disruptions.

Load: As households contain an array of devices and applications, the network must dynamically adapt, recognizing and responding to myriad load scenarios.

Position: Device mobility within a home shouldn't compromise the QoE. Consistent, seamless application delivery is paramount.

Latency: In an era of real-time applications, latency becomes a critical KPI. Swift response times are non-negotiable.

Interoperability: A high QoE system must integrate seamlessly with multiple vendor hardware without compromising stability or performance.

Efficiency: Optimal utilization and allocation of resources underscore the efficacy of a QoE system.

Many current network systems predominantly addresses space and load requirements by integrating mesh extenders to enhance WiFi RF coverage. While effective in augmenting capacity in regions of the home distant from the primary gateway, this strategy falls short in delivering a fully optimized user experience. The performance of such systems remains largely determined and constrained by the type of WiFi technology deployed, and fails to capitalize on the capabilities of an advanced Full Stack Optimization framework.

To achieve a holistic system optimization, algorithmic adjustments are required to accommodate temporal variations induced by environmental factors, enable uninterrupted mobility across any position, and facilitate real-time application performance by reducing latency. Furthermore, Full Stack Optimization should be universally compatible across diverse Customer Premise Equipment (CPE) and System on Chip (SoC) vendors, incorporating a range of WiFi technologies and radio configurations.

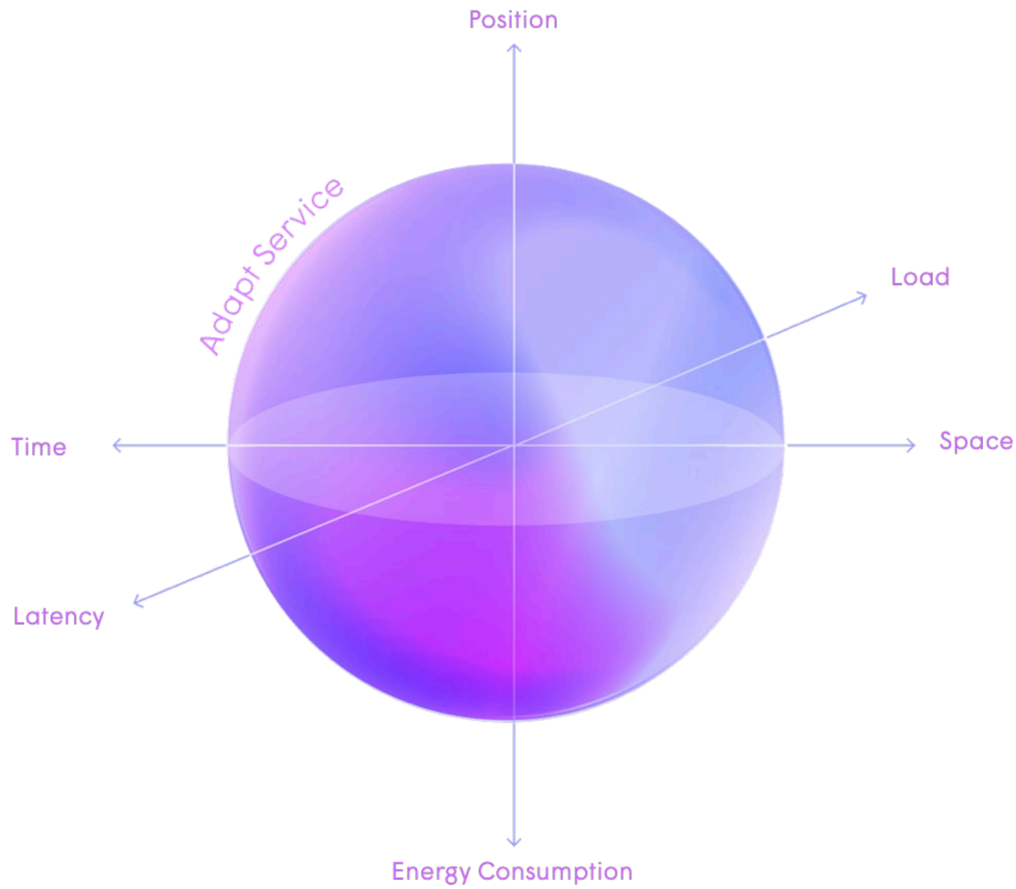

Figure 1 - Full Stack Optimization dimensions

6. Full Stack Optimization: A Revolutionary Approach

Full Stack Optimization embodies a comprehensive strategy for enhancing QoE, requiring:

1. Visibility across the stack: Merely relying on WiFi statistics is myopic. By employing Dynamic Deep Packet Inspection (DDPI), networks gain detailed insight into the devices, applications, and people who use them [2]. This granular understanding, when fused with WiFi metrics, empowers cloud-based algorithms to proactively adapt and optimize the network landscape.

Plume utilizes OpenSync in the CPE to perform DDPI functions to identify applications in real time on the network [3]. Packets from each IP traffic flow may be requested by any number of CPE user space applications for processing according to the desired functions, which may include cybersecurity protection, parental controls or application optimization functions. Applications are identified by inspecting anywhere from 1 to 12 packets in each application session. After identification, application sessions may be passed through without change, blocked or modified by manipulating the IP headers of the traffic flow according to a policy applied from the cloud. The traffic flow packet processing task is offloaded from the CPU by training the SoC hardware acceleration engine to take over the handling of the packets when all user space packet inspection demands are satisfied. After application identification is complete the application is monitored and scored for QoE delivery in the network.

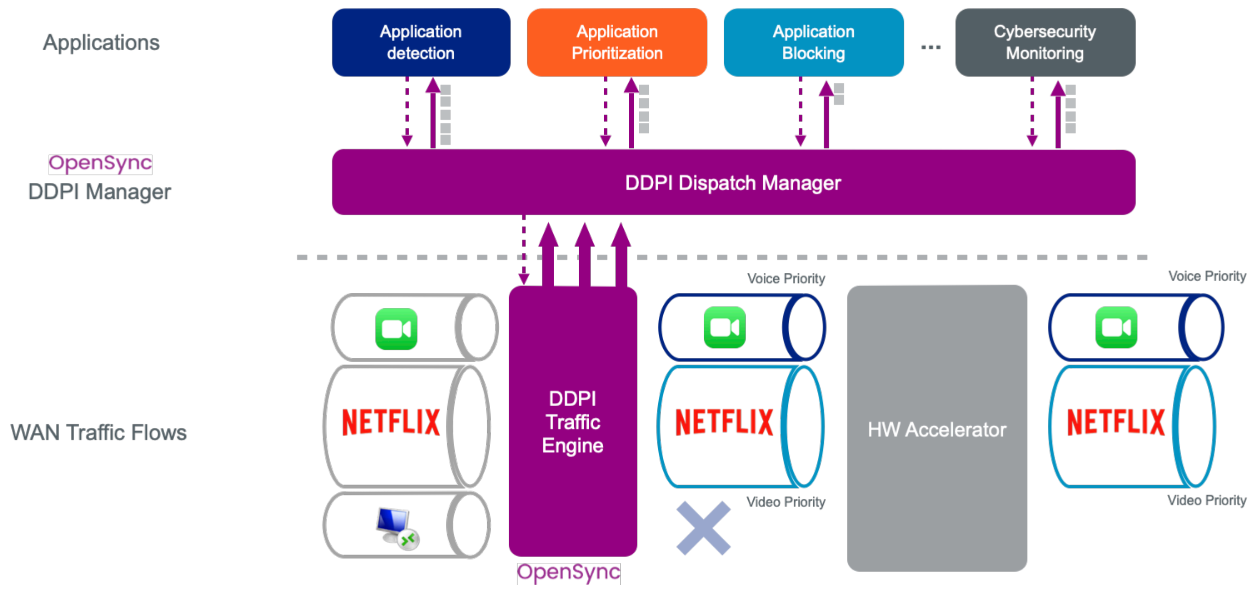

Figure 2 – Dynamic Deep Packet Inspection CPE architecture

To quantify the shift in the home to real-time applications we measured applications running in Plume managed households over a 30 day period in July and August 2023. Application information was detected and measured by using DDPI in OpenSync enabled CPEs for households with one or more APs and subscribed to Plume's HomePass service. The results show:

Table 1 – Application session statistics

Application Statistic	Value
Average number of application sessions	863 per HH/day
Average number of real-time applications sessions	64.7 per HH/day
Percentage of total applications that are real-time	78%

While the number of real-time application sessions ran in a location is only 7.5% of the total sessions, the real-time applications make up 78% of the applications in the home. To provide more clarity, we measured application usage on US based Plume customer homes over a 30-day period in July and August 2023 to better understand the importance of optimizing around real-time applications.

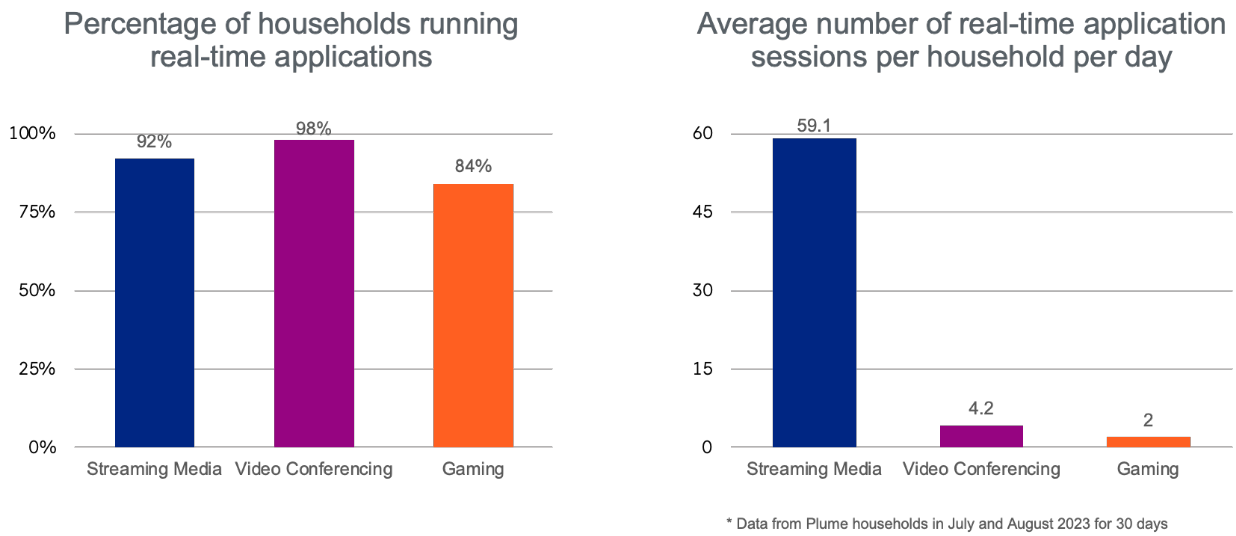

Figure 3 - Households running real-time applications

Additionally, the percentage of households running real-time applications is overwhelmingly high as shown above, with video conferencing being run at least once in 98% of the homes and gaming present in 84% of the homes over a 30-day period of monitored Plume subscribers.

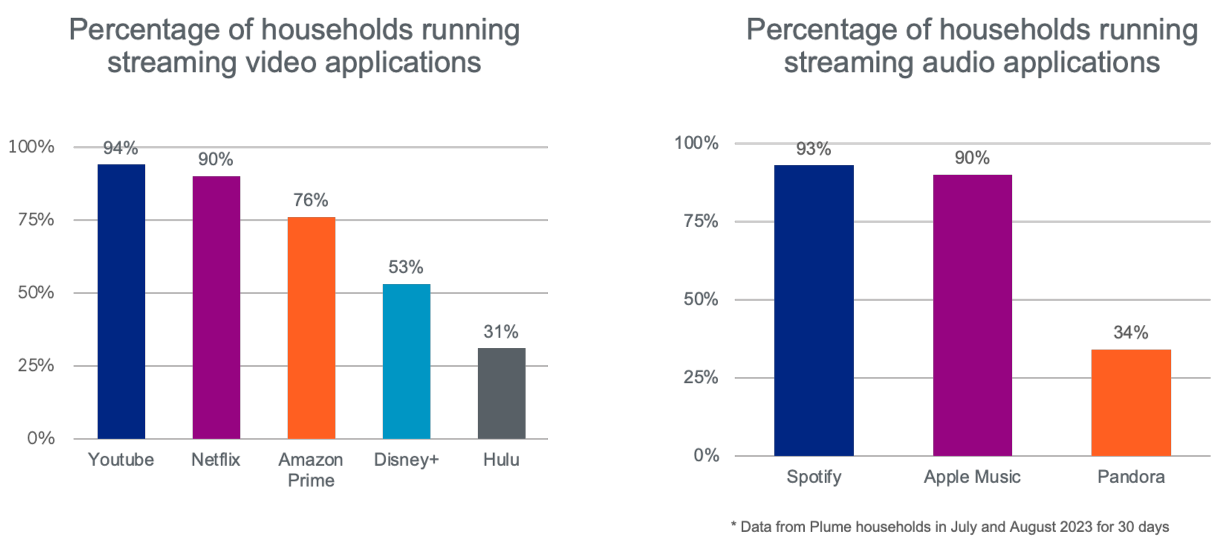

Figure 4 - Streaming application market share

Streaming Media services are dominated by YouTube, Netflix, Spotify, and Apple Music, each with at least one stream played in 90% of households or higher over the 30-day period.

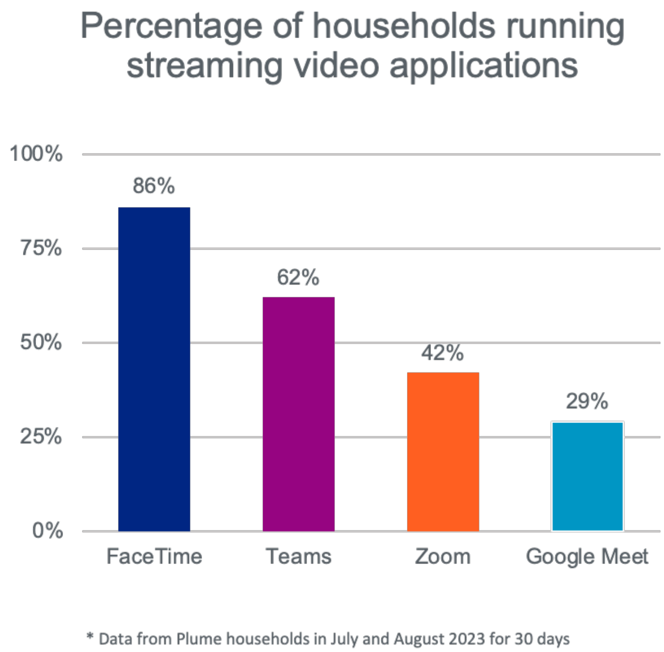

Figure 5 - Video conferencing application market share

Video Conferencing was dominated by personal devices using FaceTime, but followed by more business-related video conferencing services such as Microsoft Teams, Zoom, and Google Meet.

Based on the widespread and continued growth of real-time applications in households, high consumer QoE may benefit from giving these applications special consideration- particularly through the implementation of low-latency delivery. The efficacy of commonly utilized real-time applications, as gauged by their performance and ensuing QoE, is intrinsically tied to the latency and its consistency [4] [5].

2. Resource Manipulation and Allocation: To effectively accommodate a range of application needs, networks must extend their capabilities beyond mere WiFi utilization. In the current landscape, numerous applications demonstrate heightened sensitivity to latency, making it imperative for networks to offer pathways with minimized delay. Essential to achieving optimal application performance is the network's agility in identifying, classifying, and prioritizing applications according to their distinct requirements. Predictive algorithms, empowered by cloud-based AI and ML models, are instrumental in facilitating this dynamic adaptation.

The foundation of network optimization is a high-bandwidth and low-congestion WiFi network [6]. WiFi bandwidth allocation is computed for each access point, device, and expected applications running in the home based on previous network usage by the household. Learning models predict the needed WiFi capacity and compute the best use of spectrum, topology, and channel bandwidths to meet the application demand.

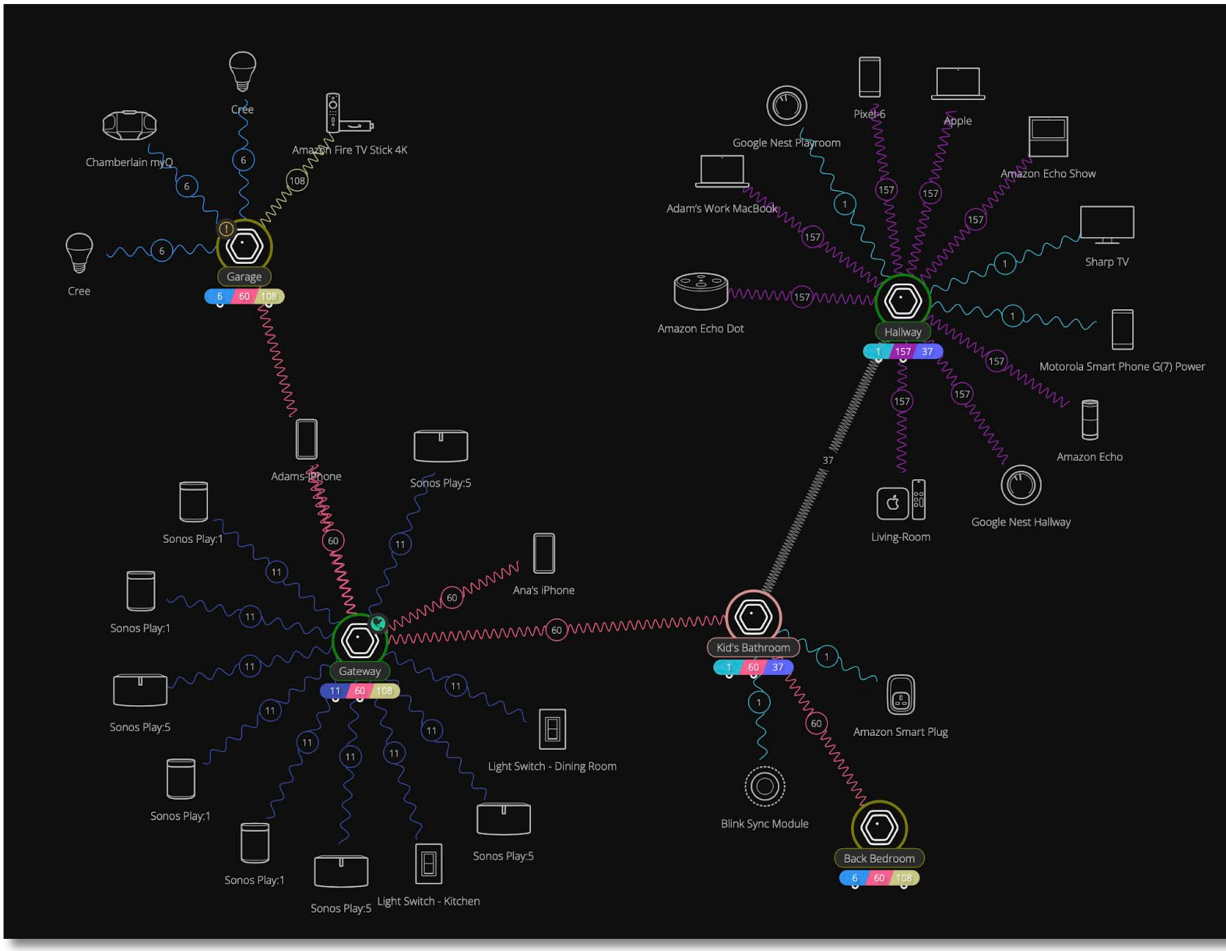

Figure 6 - Network topology with adaptive WiFi functionality

The figure above shows an example of a typical home network topology utilizing 7 different WiFi channels in 2.4GHz, 5GHz and 6GHz with WiFi 5, WiFi 6, and WiFi 6E technologies utilized. Use of different WiFi channels provides the highest amount of capacity with lower latency.

Apart from robust WiFi connectivity, latency optimization plays a pivotal role in delivering superior QoE for real-time applications. For instance, a latency exceeding 100 milliseconds from the end device to the internet server can significantly compromise the QoE in gaming, especially for cloud-based services. Likewise, real-time video conferencing experiences degradation with latency surpassing 200 milliseconds [4,5].

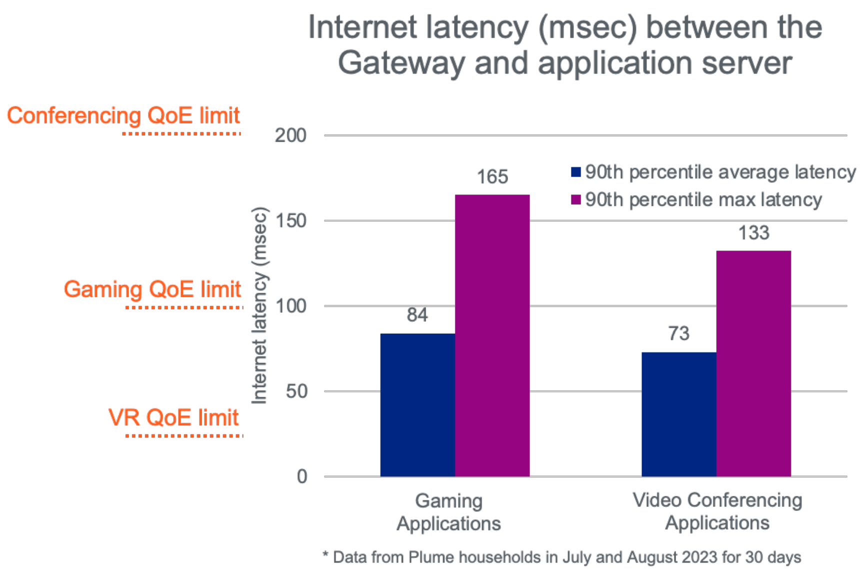

Figure 7 - Internet latency measurements for real-time applications

Measured latency between the household gateway and the internet server already measured maximum latency of 165 msec and 133 msec for gaming and video conferencing applications respectively, leaving little-to-negative latency margin in the home.

Utilizing packet prioritization for designated latency-sensitive applications can significantly mitigate latency issues during periods of temporary WiFi network congestion. Enhanced latency performance is achieved when packets from these identified applications receive precedence over those from non-latency-sensitive sources at the queuing points of each CPE interface. There are several queuing mechanisms available, such as WiFi WMM queues, OFDMA schedulers, and MLO with WiFi 7 [7]. In scenarios where the bandwidth on the WAN interface is restricted, as seen in certain DSL, DOCSIS, and Fixed Wireless Access (FWA) connections, a distinct queue can be set up for latency-sensitive traffic. This ensures that extensive buffered data does not delay the processing of latency-sensitive packets [8].

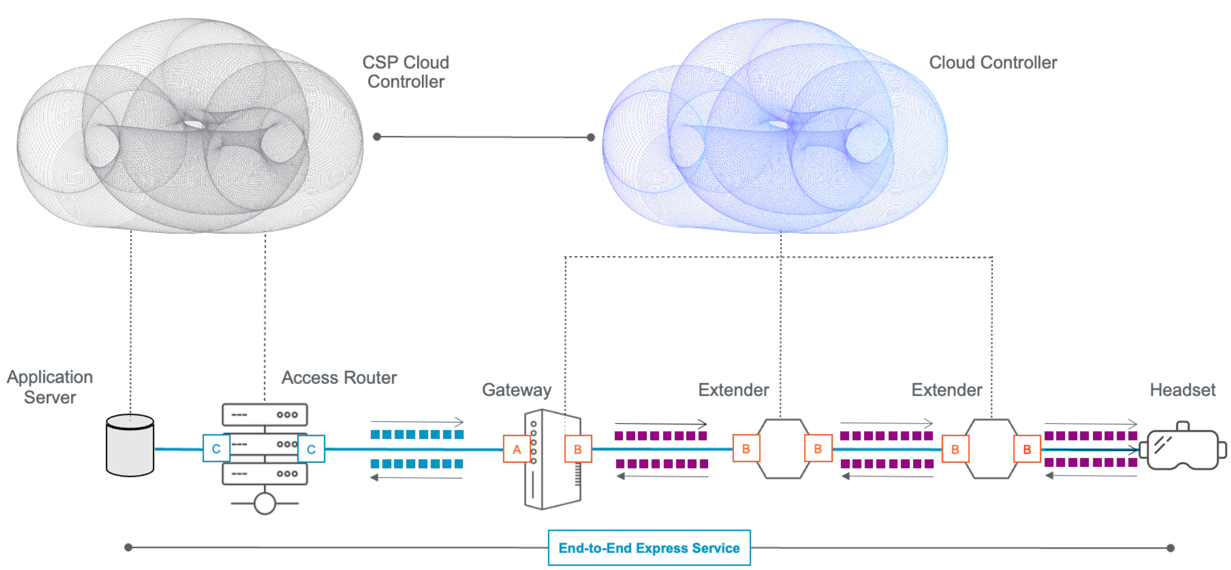

Figure 8 - Prioritized packet delivery for low latency applications

As demonstrated and measured in the figure below of packet latency versus different levels of prioritization, it is not adequate to simply throw more bandwidth at the problem to solve consumer expectations. When application session prioritization is implemented, the maximum in-home latency for prioritized application can be reduced by more than 60% in the presence of high congestion compared to best effort transmission. Full Stack Optimization must continuously provide active latency reduction to match the requirements of each individual application session for each consumer household.

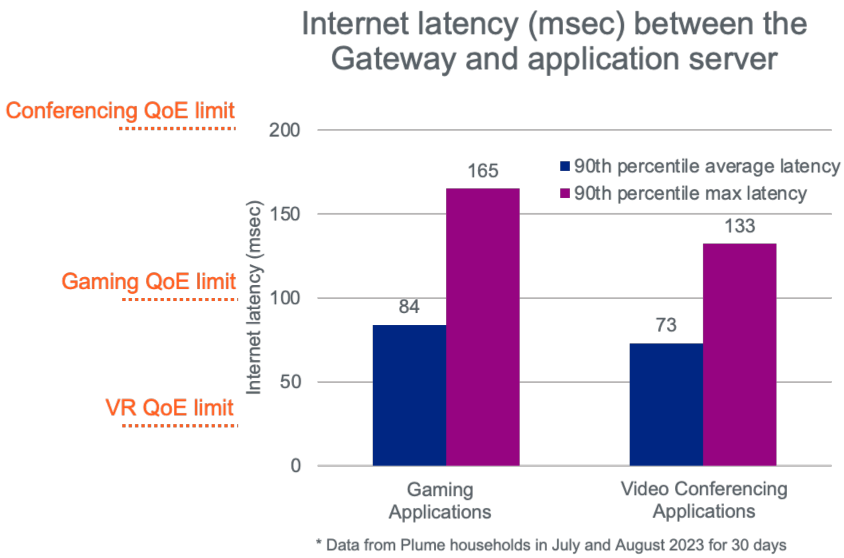

Figure 9 - Latency improvement for different application types in congested WiFi network

3. Monitoring and Assessment: Beyond generic metrics like speed tests, it's essential to delve deeper. Real-time monitoring of every application's performance, encompassing parameters like latency, throughput, and consistency, is crucial. This comprehensive monitoring translates into actionable insights, shaping continual refinement strategies.

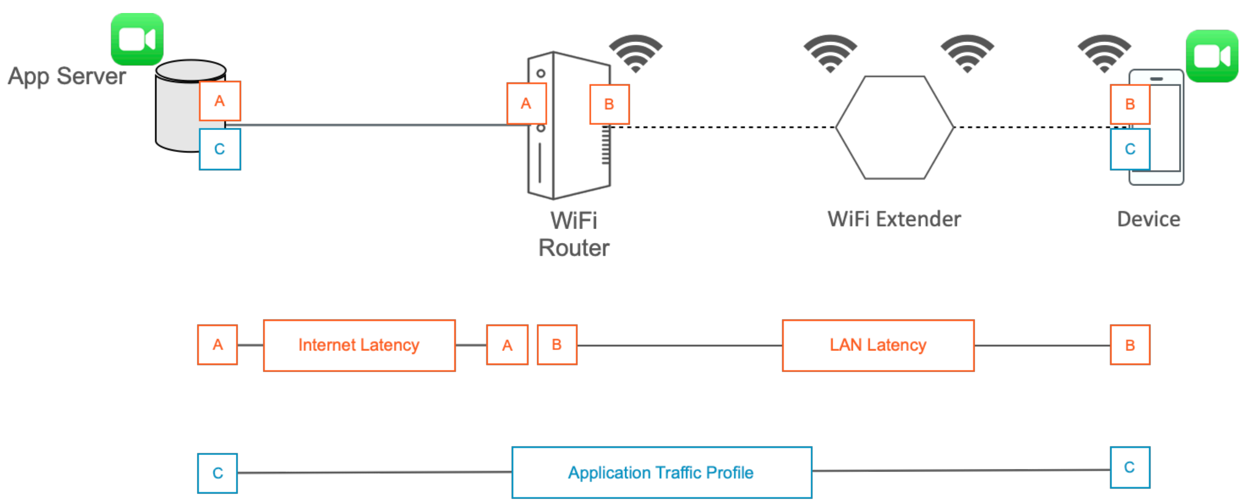

Figure 10 - Measurements for Application Performance

In the assessment of Application Performance, an inclusive evaluation is required, encompassing both latency and throughput measurements. Latency metrics are measured from both the Wide Area Network

(WAN) segment of the internet and the Local Area Network (LAN) within premises. This dual-segment assessment is pivotal in pinpointing the origins of excessive latency that may impede optimal Application Performance. Application throughput is determined from the internet application server directly to the consumer device, facilitated by real-time surveillance of the packets traverse the WiFi router.

The culmination of these metrics - throughput, latency, and their consistency - produces an Application Performance score tailored to the distinct necessities of the specific application. For example, Voice Over Internet Protocol (VOIP) calls are predominantly weighted on latency and the uniformity in packet delivery, notwithstanding their lower throughput. In contrast, high-definition video streaming places a substantial emphasis on the unwavering delivery of high throughput and can tolerate a considerable degree of latency. Finally, cloud gaming applications require a combination of both ultra-low latency control packet feeds, and low latency high bandwidth video streams simultaneously in different directions. Application Performance scores may be used by CSP for call center troubleshooting, subscriber self-help, and network level KPIs to measure the delivery of application performance across different subscriber types and equipment configurations.

7. Conclusion

The burgeoning prevalence of real-time applications across households emphasizes the urgency for a measurement standard that's both elevated and consistent. The days of singular strategies, like bandwidth enhancement or WiFi upgrades, are waning. Engineers and network architects must embrace a multi-dimensional approach. The seven-dimensional model, integrated within the Full Stack Optimization feature set, charts a promising path toward a future where each application's performance is seamless, user-centric, and consistent.

Abbreviations

QoE	Quality of Experience
DDPI	Dynamic Deep Packet Inspection
LAN	Local Area Network
CSPs	Communication Service Providers
MHz	Megahertz
GHz	Gigahertz
Mb/s	Megabit per second
MIMO	Multiple-Input Multiple-Output
AP	Access Point
AI	Artificial Intelligence
ML	Machine Learning
FWA	Fixed Wireless Access
WAN	Wireless Area Network
OFDMA	Orthogonal Frequency-Division Multiplexing
MLO	Multi-Link Operation
WMM	Wireless Mobility Management
VOIP	Voice Over Internet Protocol

Bibliography & References

[1] https://www.truecable.com/blogs/cable-academy/wifi-7-vs-wifi-6-whats-the-difference

[2] Tomasz Bujlow, Valentín Carela-Español, Pere Barlet-Ros, Independent Comparison of Popular DPI Tools for Traffic Classification, http://tomasz.bujlow.com/publications/2014_journal_elsevier_comnet_independent_comparison.pdf

[3] OpenSync 3.0 Overview, EUB-020-013-00 OpenSync 30 Overview.pdf, Flow Service Manager – FSM 53-55

[4] Normoyle, A., Guerrero, G., & Jörg, S., Player Perception of Delays and Jitter in Character Responsiveness, proceedings of the ACM Symposium on Applied Perception (SAP 2014), 117-124

[5] Eric Baccei, Alexander Clopper, Taylan Sel, An Evaluation of Zoom and Microsoft Teams Video Conferencing Software with Network Packet Loss and Latency, Worcester Polytechnic Institute

[6] Adam Hotchkiss, Bill McFarland, The Plume Advantage: Putting the Adapt in Adaptive WiFi, https://content.plume.com/v/the-plume-advantage-putting-the-adapt-in-adaptive-wifi

[7] Monica Paolini, Senza Fili, Wi-Fi 7: The next generation in the evolution of Wi-Fi, https://6ghz.info/wp-content/uploads/2022/07/SenzaFili-IntelTB-Wi-Fi7-The-next-generation-in-the-evolution-of-Wi-Fi.pdf

[8] Amir Mirzaeinnia, Mehdi Mirzaeinia, and Abdelmounaam Rezgui, Latency and Throughput Optimization in Modern Networks: A Comprehensive Survey, September 2020