Available *online* at : http://jnte.ft.unand.ac.id/

Jurnal Nasional Teknik Elektro

| ISSN (Print) 2302-2949 | ISSN (Online) 2407-7267 |

QCI Optimization to Minimize Latency and Enhance User Experience

Patria Adhistian [1], Priyo Wibowo [2]

[1] *Department of Industrial Engineering, Pamulang University, Jl. Surya Kencana No. 1 Pamulang, Tangerang Selatan, Indonesia*
[2] *National Research and Innovation Agency, Kawasan Puspiptek, Building 417, Setu, Tangerang Selatan, Indonesia*

ARTICLE INFORMATION

Received: January 30, 2024
Revised: June 1, 2024
Accepted: July 27, 2024
Available online: July 31, 2024

KEYWORDS

QoS, latency, throughput, DRX, Pre-allocation, PDCP, scheduling

CORRESPONDENCE

Phone: +62 (021) 7412566

E-mail: dosen01529@unpam.ac.id

A B S T R A C T

Limited QCIs (QoS Class Identifiers) restrict the handling different service types with varying quality requirements. This necessitates research on QoS management to minimize latency and improve user experience, particularly for real-time applications like video conferencing and online gaming. This paper proposes a combined optimization scheme targeting QCI 3 to reduce latency. The approach involves disabling DRX, optimizing pre-allocation, and reducing the PDCP discard timer. The optimization performance is studied by taking the case of an e-sport game that demands low network latency, affecting the quality of the players' experience. The optimization scheme was validated through functionality, resource allocation, and air interface latency tests conducted under actual e-sport gaming conditions. Network latency was measured every minute to evaluate the impact of optimization on esports games running under QCI 7, QCI 3, and optimized QCI 3. In addition, air interface latency for optimized QCI 3 under networks with poor coverage and very high-capacity networks was compared to latency under QCI 8 (basic), QCI 7, and regular QCI 3. The optimization strategy demonstrated a significant reduction in air interface latency, up to 19% improvement compared to non-optimized QCI 3. It has reduced air interface latency's maximum, minimum, and standard deviation values during gameplay. The strategy also ensured concurrent operation with multiple QCI values without compromising other application's throughput. The proposed optimization strategy effectively enhances the user experience by significantly reducing average latency and jitter.

INTRODUCTION

Most countries still use 4G LTE networks. In fact, according to a report by the GSMA [1], 4G coverage in low- and middle-income countries (LMICs) increased to 84% by the end of 2021. While 5G networks have been expanding, particularly in high-income countries [1], [2], 4G remains the most widely used mobile network technology globally. Despite the emergence of 5G technology, 4G LTE networks are still widely used and will continue to be used for the foreseeable future. Consequently, research on optimizing latency in 4G networks remains important to improve the network performance and user experience.

The LTE network uses QoS Class Identifiers (QCI) values to determine how to prioritize and handle the data packets. This ensures that the different classes of service receive the appropriate level of resources and support. ETSI TS 123.203 provides guidelines for QoS (Quality of Service) management, but the selection of an appropriate scheduling scheme for LTE networks is not standardized [3], [4]. Providers have the flexibility to configure and implement an algorithm that best suits the specific needs and concerns of the system. The document only specifies the QoS parameters and procedures used to manage the allocation and use of network resources, such as bandwidth and power. It also covers the QCI, Allocation and Retention Priority (ARP), Guaranteed Bit Rate (GBR), and Non-Guaranteed Bit Rate (NGBR). Each QCI value has a set of standard characteristics, including bit rate, packet loss, delay, jitter, reliability, and priority.

Even though QCI allows network providers to customize the quality of service levels according to application and user needs, QCI configuration requires a deep understanding of service characteristics and network requirements [5]. This complexity can be an obstacle, especially if not set up properly. A limited number of QCIs leads to limitations when handling different services with different quality requirements. Importantly, the effectiveness of QCIs varies greatly depending on how the network operator implements them.

Some algorithms have been proposed based on QoS provision to meet on-demand service requests. The DSA technique using Nash Bargaining Solution has been proposed to increase throughput in Long Term Evolution (LTE) networks with different types of base stations (BSs) [6]. The results show that the DSA technique has lower BER and higher throughput than the typical UL-LTE configuration. Radio Block usage ratio is introduced as one of the additional parameters for centralized and distributed radio resource allocation schemes to optimize the performance of both

https://doi.org/10.25077/jnte.v13n2.1193.2024

cell center and cell edge users along with the QoS provisions of the users [7]. A scheduling framework has been studied to address QoS requirements defined by 3GPP specifications for LTE networks. Simulation results show that the proposed scheduler fulfills the Packet Delay Budget for GBR bearers and performs efficiently for non-GBR bearers [8]. A scheduling algorithm that considers the priority of services in the QCI table, available resources, and traffic load is proposed using a scaling factor to convert priority values to the scheduling metric [9]. Enhanced Dynamic Scheduling (EDS) is proposed to rapidly adapt to changing conditions during an active VoLTE call and improve UL transmission delay and system capacity [10]. EDS can only be applied to voice packets since the packet time interval is constant for VoLTE. Two novel downlink LTE scheduling algorithms based on Reinforcement Learning (RL) and Q-learning techniques were also proposed to optimize spectrum utilization and fairness among primary and secondary users. However, it is less efficient than collaborative algorithms for licensed and unlicensed users [11]. A scalable priority-based resource allocation scheme aims to balance resource utilization and application priority support [12]. The proposed scheduling algorithm outperforms standard algorithms regarding resource sharing fairness, average resource utilization, QCI priority support, and delay budget violation. However, the algorithm aims to balance resources but may not be optimal in all scenarios.

Discontinuous Reception (DRX) can be used to optimize LTE system performance. DRX is a sleep mode mechanism commonly used in wireless communication systems to conserve battery power in user devices. However, it can also be utilized to enhance network responsiveness. The DRX mechanism works by dividing time into active and sleep periods. The length of the active and sleep periods is determined by the DRX configuration, which can be adjusted based on the network conditions and device requirements. By adjusting the DRX parameters, the network operator can balance the need for low-power consumption with the need for efficient data transfer and signaling in the network [13]–[20]. DRX can be configured differently for each QCI to optimize user devices' network performance and battery life. DRX can be configured to reduce the delay, increase the network's responsiveness, increase the data throughput, and reduce the number of signaling messages required.

In addition to DRX, pre-allocation optimization can also improve resource allocation and utilization. Due to the limited resources of wireless communication systems, resource allocation and scheduling through packet scheduling technology are crucial to improving network performance and ensuring fairness between users by maximizing network spectrum utilization and capacity based on multi-service QoS [21]. The pre-allocation mechanism, in terms of Virtualized Radio Resource Pre-Allocation, can guarantee QoS with the minimum requirements of radio resources [22]. The appropriate timing between uplink grant reception and the corresponding uplink data transmission can positively affect latency [23].

Another scheme for optimizing LTE system performance is the management of the PDCP (Packet Data Convergence Protocol) buffer [24]. Dedicated buffers for each Radio Access Bearer are implemented at the PDCP layer in evolved Node B (eNB) [25]. These buffers prepare user data for transmission over the air interface and can become overwhelmed by the greedy nature of the Transmission Control Protocol (TCP). The standard 3GPP Discard Timer and conventional Active Queue Management (AQM) techniques are ineffective in managing the PDCP buffer. Therefore, two alternative mechanisms are proposed by [26] that are better suited for this environment. A novel buffer management scheme for PDCP buffers at the eNB was also proposed and tested using a set of simulations, demonstrating its effectiveness in achieving optimal PDCP buffer occupancy [27]. A cross-layer scheme that involves data duplication and integration at the PDCP sublayer of layer 2 for downlink operations can reduce the block error rate (BLER), improve throughput, and reduce latency in various scenarios [28]. Packet out-of-order in the PDCP layer, which can affect end-to-end throughput, can be minimized by using low-complexity traffic-splitting mechanisms [29].

Accurately identifying the type of traffic (e.g., VoIP, video streaming, web browsing) can be difficult, as packets may not always contain clear application-layer information [30]. Network traffic can be highly dynamic, with shifts in bandwidth usage, application mix, and user behavior over time. Different applications have varying latency and packet loss tolerance, which must be mapped to the appropriate QCI class. Keeping the QCI classification up-to-date and responsive to these changes can be a significant challenge for network operators [31]. Quality of Service (QoS) management in mobile networks needs to address these challenges, especially in traffic identification and ensuring proper QoS parameters are applied.

On the other hand, quality of experience (QoE) is the key to increasing customer satisfaction for internet service providers, which impacts revenue growth. Quality of Experience (QoE) is a measure of user perception of the quality of an Internet service or application that includes factors that contribute to satisfaction with the service [32], [33]. Certain QoS requirements must be met to achieve a high level of QoE. For example, real-time applications such as video conferencing and online gaming require low latency and jitter. Network performance, such as delay, latencies, and jitter, impact players' quality of experience [33], [34]**.** The response latency is a critical factor in user Quality of Service [35]. There is a correlation between latency and playability, with games noticeably getting worse as latency increases [36], [37]. Some cloud-based games are sensitive to latency, which degrades user performance [38].

In first-person shooter games, a high DPI setting is essential for faster cursor movement, while a high frame rate is also important for improving game smoothness. However, both are sensitive to network latency. Higher DPI and higher frame rate are more sensitive to network latency, as a small delay in receiving data can lead to uncontrollable cursor movement, a decreased frame rate, and the game not running smoothly. Small reductions in network latency can significantly improve player accuracy and score, enhancing the overall gaming experience [39]. Lower latency correlates with higher QoE (Quality of Experience), reducing frustration and increasing game responsiveness [39]. Optimizing latency is necessary to ensure responsive controls, accurate aiming, and smoother visuals, especially in competitive games.

This paper proposes a combined optimization scheme to address the challenges related to dynamic network traffic and improve user experience, especially latency and jitter. The case of e-sport gaming was chosen since the user experience is affected by network performance. Stability and low-latency network connections are essential for e-sport games. The strategy involves disabling DRX, optimizing pre-allocation, and reducing the PDPC discard timer to decrease average latency and jitter. The proposed strategy was validated through functionality and performance tests conducted under real e-sport gaming conditions, demonstrating its effectiveness. Air interface latency when playing a game using QCI 3 with optimization is measured and compared to QCI 8 as the baseline data service and QCI 7 and QCI 3, commonly used as dedicated bearers for online gaming.

METHODS

The response to changes in QCI parameters can be a challenge for network operators who must keep the QCI classification up-to-date with user application changes. Thus, this study was conducted by configuring different QoS levels using the QCI parameters. Specifically, QCI 8 was used as the default EPS Bearer for all data services, while QCI 7 was activated as a dedicated bearer during the user equipment (UE) play game. To alleviate network congestion during periods of heavy traffic, a dedicated bearer with QCI 3 was only activated for specific registered game applications. The QCI 3 bearer was automatically deactivated once the UE closed the gaming application. To optimize the network performance for gaming applications and improve the user experience, the following optimization strategies for QCI 3 were used:

To optimize network performance for gaming and enhance user experience, several optimization strategies were applied to QCI 3. First, the Discontinuous Reception (DRX) sleep mode mechanism was disabled for QCI 3, as DRX—while beneficial for conserving power—can introduce additional latency [13]–[20]. Disabling DRX eliminated latency caused by DRX cycles, particularly when the UE consistently received data during gameplay. The average latency of packet arrival during the DRX sleep period can be expressed as [40]:

$$L = \frac{1}{2}\left(\frac{1}{\lambda} + T_{ON}\right) + \frac{T_{OFF}}{2} - \frac{T_{ON}}{2}\left(1 - e^{-\lambda T_{ON}}\right) \quad (1)$$

where λ is the arrival rate, T_{ON} is the DRX on duration, and T_{OFF} is the DRX off duration.

The second optimization was pre-allocation tuning, which ensured timely resource allocation for uplink data transmission—an essential factor for maintaining low latency in real-time applications [7], [8], [10]. This was achieved through periodic uplink data scheduling while suppressing buffer status reports, thereby minimizing interference. The pre-allocation mechanism was applied exclusively to QCI 3 to reduce latency, as visualized in Figure 1.

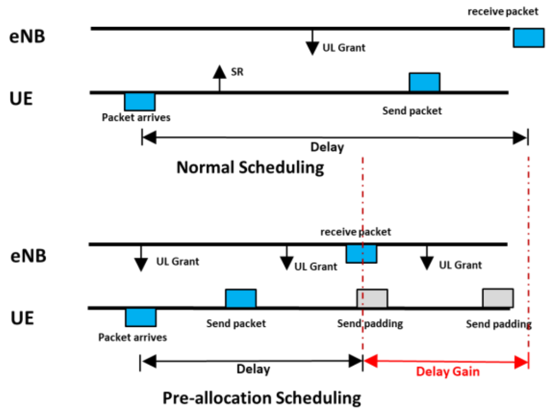


Figure 1. Pre-allocation Scheduling Optimization

Finally, the Packet Data Convergence Protocol (PDCP) discard timer was reduced to half of its default value to improve buffer efficiency and minimize latency [24], [27], [28]. The timer is triggered when a PDCP Protocol Data Unit (PDU) is sent from the sender to the receiver. If the PDU remains in the buffer beyond the timer duration, it is discarded by the eNB. This ensures that outdated or unacknowledged PDUs do not consume air interface resources, thereby supporting efficient data transmission—especially for small packets in gaming contexts.

The testing of QCI 3's performance and functionality with optimization was divided into functionality, resource allocation, and air interface latency tests. The real-time e-sport game, “Player Unknown Battle Ground” (PUBG), was chosen because it has several video quality options related to the DPI (Dots Per Inch) and frame rate level. It is an important part of the first-person shooter gaming experience [39]. The DPI dan frame rate level refers to the visual fidelity and smoothness of the game graphics, which is essential for competitive games like PUBG. Players can adjust the display quality and frame rate settings for the best gaming experience. DPI and frame rate are interconnected factors that determine how cursor or crosshair movement is translated on the screen. Higher DPI settings result in faster cursor movement, while higher frame rate settings improve the smoothness of the game. High DPI levels and high frame rates were chosen in the test.

To evaluate the impact of varying Quality of Service (QoS) levels on mobile gaming performance, this study employed multiple Quality of Service Class Identifier (QCI) configurations. The test scenario involved configuring QCI 8 as the default Evolved Packet System (EPS) Bearer for all data traffic, activating QCI 7 as a dedicated bearer during gaming sessions, and using QCI 3 as a congestion management bearer during periods of high network traffic. A cellular network testbed was developed to emulate realistic LTE conditions and support comprehensive experimentation.

The experimental setup consisted of four key components: the network environment, QoS configuration, network monitoring tools, and performance data metrics. The network environment featured an LTE eNodeB as the base station, which supported dynamic QCI assignments, while the User Equipment (UE) devices were used to simulate real user behavior and traffic patterns. The QoS configuration followed a structured scheme

https://doi.org/10.25077/jnte.v13n2.1193.2024

where QCI 8 was assigned as the default bearer, QCI 7 was activated to prioritize traffic during active gaming sessions, and QCI 3 was selectively enabled for registered gaming applications during heavy traffic periods to alleviate network congestion.

To monitor network behavior and performance in real time, Wireshark and iPerf were employed as the primary monitoring tools. These tools captured detailed performance data during gaming sessions using the PUBG application under scenarios that closely simulated real-world usage. Log data was collected from the UE and network-side equipment to analyze behavior across the different QCI configurations.

Performance evaluation focused on two principal metrics: latency and throughput. Latency was measured by calculating the round-trip time (RTT) of data packets transmitted by the UE, while throughput was assessed based on the actual data transfer rate observed during gameplay. These metrics were gathered and analyzed, particularly under simulated heavy traffic conditions, to assess how each QCI level affected the quality of the gaming experience.

This study's general procedures for measurement and data analysis followed a systematic approach to assess the performance of various QCI configurations under different network conditions. Initially, baseline measurements were obtained using a default network setup with QCI 8 as the standard bearer. During this phase, fundamental performance metrics were recorded without activating any specialized QoS configurations, serving as a reference point for subsequent comparisons.

In the next phase, QCI 7 was activated as a dedicated bearer during gaming sessions to prioritize gaming traffic. Performance data—latency, throughput, jitter, and packet loss—was collected and analyzed while the User Equipment (UE) actively engaged in gameplay. This phase provided insight into the network behavior when enhanced QoS was explicitly applied for gaming.

Then, QCI 3 was activated to simulate conditions of heavy network traffic. This bearer was selectively applied to specific registered gaming applications, such as PUBG, during simulated congestion scenarios. Performance data was gathered under these high-load conditions to evaluate how well QCI 3 managed traffic. After the gaming application was closed, the QCI 3 bearer was automatically deactivated, and additional performance measurements were taken to observe any post-deactivation effects on network behavior.

Finally, a comprehensive comparative analysis was conducted by evaluating the collected performance metrics across the three configurations—QCI 8, QCI 7, and QCI 3. This analysis focused on assessing the individual and relative impacts of each QCI level on gaming performance, particularly during peak usage periods. The comparison highlighted how tailored QoS configurations can significantly enhance user experience in latency-sensitive applications like mobile gaming.

Functionality Test

The Respond to Registered Games test proved that QCI 3 would only be activated by specific registered games. The functionality of the dedicated bearer QCI 3 was also evaluated by measuring the throughput of nine different activities running sequentially. This method is used to simulate the dynamics of network traffic while observing the response and identifying the network type. The test process started by deactivating flight mode, running the Speedtest application, running the YouTube application, running the PUBG game application, opening Speedtest again, returning to PUBG, returning to YouTube, and then returning to PUBG again.

Resource Allocation Test

Measuring the throughput while playing an e-sport game was used to measure resource allocation and occupation. These measurements were used to verify whether the optimization scheme was working as intended and to ensure the optimization did not negatively impact other network traffic. The dedicated bearer QCI 3 was activated during the PUBG game application running. Peak and average throughput were measured during each session of the PUBG game. Three SIM cards were used in this study, including one supporting Dual Bearer QCI 3 GBR DL 20Mbps/UL 10Mbps, one supporting Dual Bearer QCI 3 GBR DL 100Kbps/UL 50Kbps, and one using the Default Bearer QCI 7 non-GBR. DL/UL PRB utilization and DL/UL cell throughput during gaming were monitored with live performance monitoring with a one-minute granularity counter.

Air Interface Latency Test

Air interface latency describes the delay that occurs over the mobile network. It directly reflects the perceived user experience, especially for real-time applications such as esports games. While QCI prioritizes traffic, it does not directly control latency. Measuring air interface latency after optimization can show whether the changes achieve the goal of reducing latency for prioritized traffic. Measuring air interface latency also validates whether disabling DRX and reducing timers effectively results in faster data transmission for prioritized traffic. This is important to ensure the effectiveness of the optimization scheme.

To study the effect of optimization, latency measurements were performed every minute on networks dedicated to playing the e-sport game, namely QCI 7, QCI 3, and QCI 3 with optimization. Air interface latency when playing games for QCI 3 with optimization was also tested in a bad coverage network (RSRP -100dBm, SINR -5 dB) and very high-capacity network (PRB 98%, 40 active users) and compared with QCI 8 (baseline), QCI 7, QCI 3, and QCI 3 with optimization.

RESULTS AND DISCUSSION

QoS parameters such as latency and jitter are correlated with QoE [33], [39], particularly in gaming and multimedia applications. Latency and bandwidth consumption in Cloud Gaming can cause a deterioration in Quality of Experience (QoE) for users [41]. The current study, which shows improving QoE through optimized QCI parameters, highlights the critical role of QoS management in user satisfaction. A combined optimization scheme targeting QCI 3 to reduce latency involves disabling DRX, optimizing pre-allocation, and reducing the PDCP discard timer. Disabling DRX reduced service latency by eliminating the delay caused by the mechanism [13]–[20]. Packet delay and latency are reduced by modifying the PDCP dump timer, which leads to more efficient

PDCP buffer management. [24], [27], [28]. The pre-allocation mechanism ensures timely resource allocation for uplink data transmission so it can maintain low latency in real-time applications [7], [8], [10].

Functionality

When the flight mode of the UE was disabled, the authentication process ran and activated the default EPS bearer (QCI 7). After the UE opened the PUBG games application, the dedicated bearer with QCI 3 was activated. The dedicated bearer with QCI 3 was deactivated when UE closed the game application. Figure 2 shows that the QCI 3 dedicated bearer is only used when the UE detects a PUBG IP already registered in the library whitelist.

In the functionality test, multiple bearers with multiple QCI values were activated in parallel. Multiple bearer tests were conducted to ensure that QCI 3 with optimization can work together with other active QCI. The dedicated bearer QCI 3's functionality test used 9 different activities that ran sequentially, as shown in Figure 3. The QCI 3 dedicated bearer was activated when UE opened the PUBG game. It is only triggered when UE accesses a specific registered IP (PUBG Server IP with specific DPI). When the UE transferred to other service packets not registered as DPI, it used the QCI 7 default bearer in parallel with QCI 3. PUBG packages continued to flow even though the dedicated bearer QCI 3 and default bearer QCI 7 were active. When the UE switched to HOME and opened Speedtest during play PUBG, it occupied the QCI 7 default bearer. It can be seen that there is no throughput limitation, as packages can still reach 40 Mbps for download and 20 Mbps for upload. The same is also observed when switching to HOME and opening YouTube. YouTube packages flowed using QCI 7, and there is no throughput limitation, as it can still reach 40 Mbps for download and 20 Mbps for upload. The dedicated bearer QCI 3 activation did not impact the throughput of the Speedtest and YouTube applications, which use the default bearer QCI 7. The dedicated bearer QCI 3 only impacted PUBG games.

Resource Allocation

Stability and low-latency network connections are important in online applications, but fairness in using network resources must also be guaranteed. Measurement of the throughput while playing an e-sport game is used to measure resource allocation and resource occupation. A dedicated bearer with QCI 3 was active during the PUBG game application running. The resource throughput consumed while playing the PUBG game is shown in Figure 4, and the peak and average throughput during each session of the PUBG game are shown in Table 1.

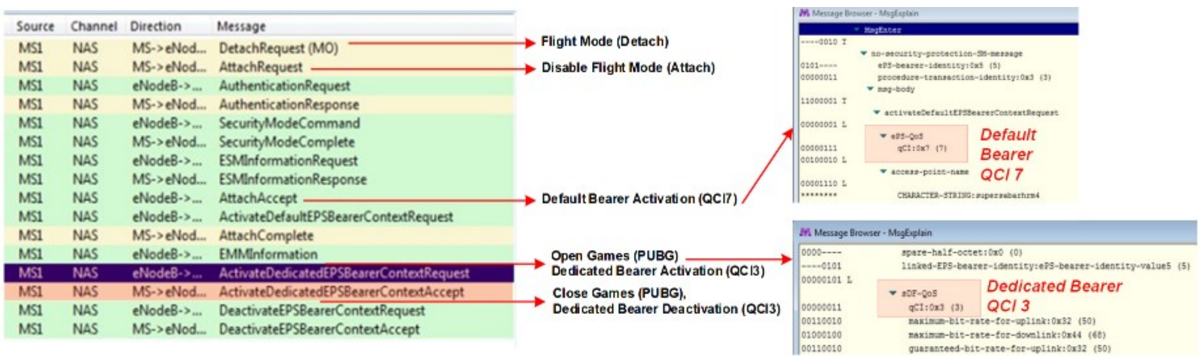


Figure 2. Respond QCI 3 to Registered Application (PUBG)

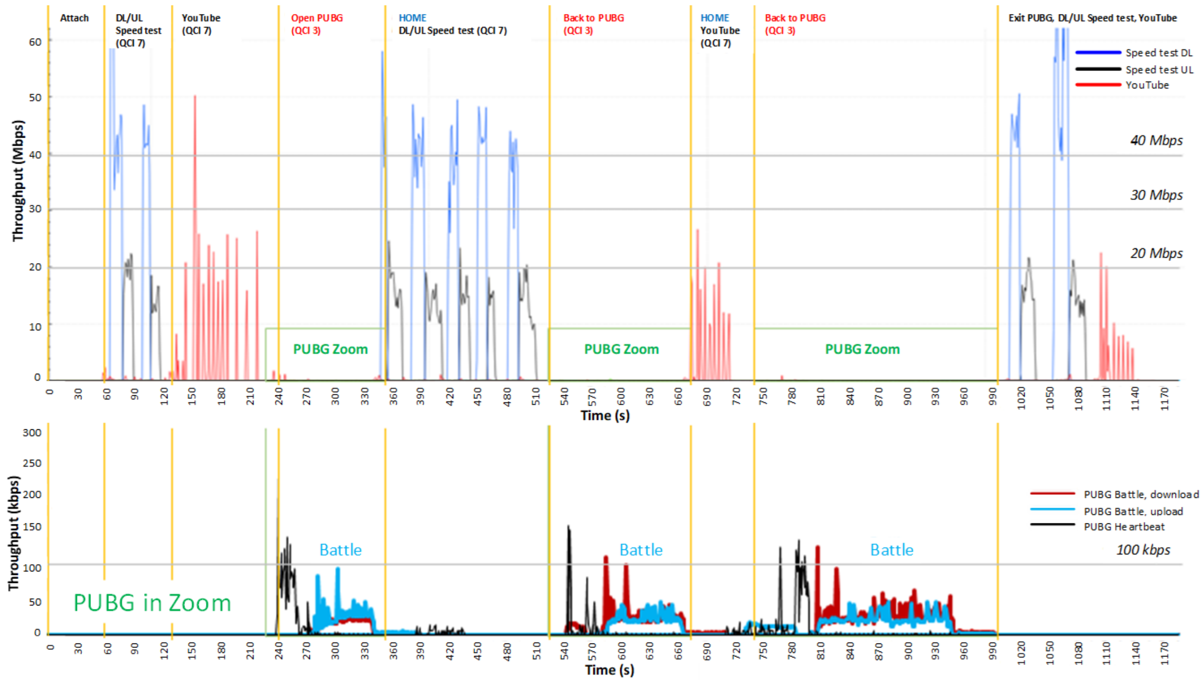


Figure 3. Dedicated Bearer QCI 3 Functionality Tests

https://doi.org/10.25077/jnte.v13n2.1193.2024

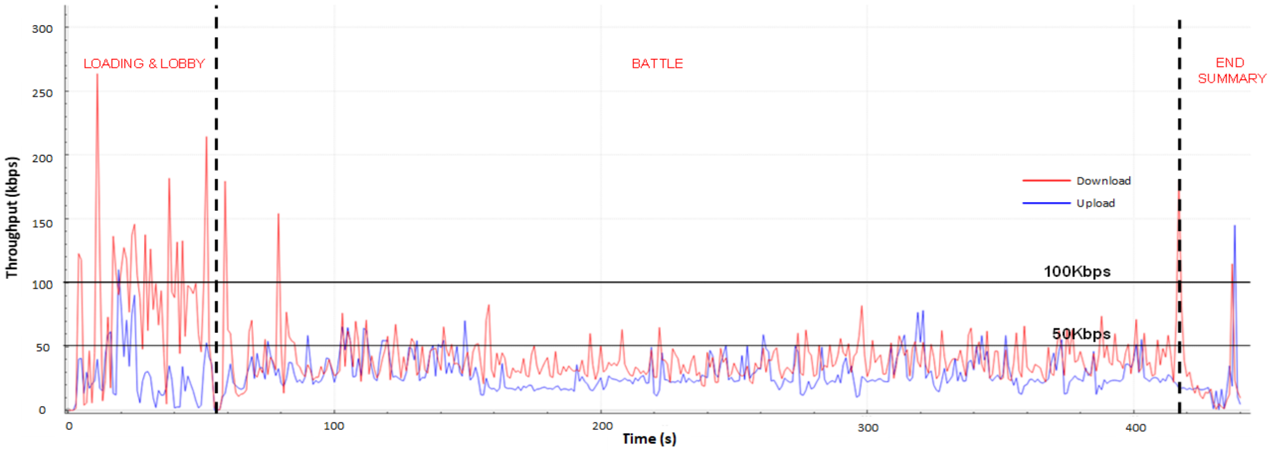


Figure 4. Throughput During Play PUBG Game

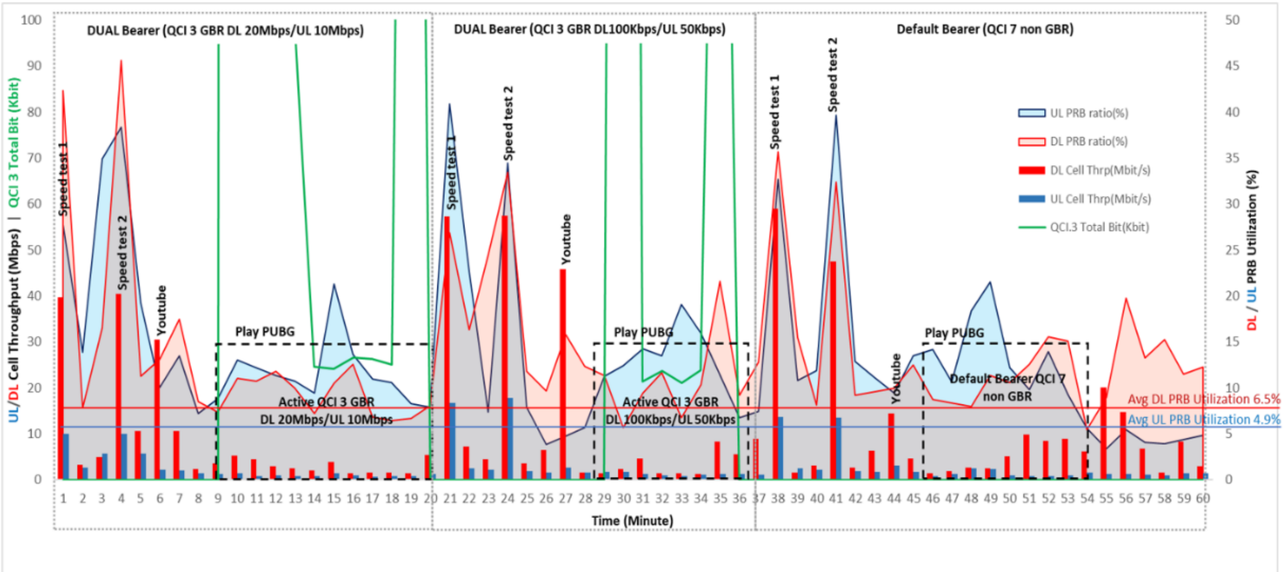


Figure 5. RAN Resource Occupation

Table 1. Throughput During Playing The PUBG Game

Game Session	Download (kbps)		Upload (kbps)	
	Peak	Average	Peak	Average
Loading	250	76.11	120	25.8
Lobby	220.2	39.3	28.9	28.9
Battle	82.8	39.3	76.1	28.9
Heartbeat	1.09	0.2	1.09	0.2
End Summary	122.8	29.2	151.1	20.3

Table 1 shows that the average download throughput in the lobby session, battle session, and end summary session are 76.1 kbps, 39.3 kbps, and 29.2 kbps, respectively. Meanwhile, the average upload throughput for the lobby, battle, and end summary sessions are 25.8 kbps, 28.9 kbps, and 20.3 kbps, respectively. The average throughput during the lobby session is the highest, while during the battle session, which requires the highest stability, less throughput is needed.

The PUBG game session that impacts the gaming experience is the battle session. The average throughput for this session is 39.3 kbps and 28.9 kbps for download and upload, respectively. Activating QCI 3 on a specific game DPI can guarantee the gaming experience during the battle session. A dedicated bearer with GBR DL 100 kbps and UL 50 kbps is sufficient for the PUBG game. Updating game software/applications and downloading inside games generally requires large resources. To reduce the low experience due to this process, it is proposed to use QCI 3 Maximum Bit Rate (MBR) for the uplink and downlink, which is the same as the existing default bearer AMBR (Aggregate Maximum Bit Rate). This parameter is used to ensure that the network can provide a consistent level of service for different types of traffic and to prevent one type of traffic from consuming too much bandwidth and impacting the performance of other traffic.

RAN (Radio Access Network) resource occupation measurement ensures that QCI 3 won't reserve network resources with defined Guaranteed Bit Rate UL/DL 10 Mbps/20 Mbps or 50 kbps/100 kbps. RAN resource occupation, as shown in Figure 5, indicates the resource efficiency for GBR vs. non-GBR applications. It demonstrates how guaranteed bit rates (GBR) ensure a consistent quality of service by reserving necessary resources for specific activities (playing PUBG). Using QCI 3, GBR ensures that the resources are allocated efficiently, preventing overuse or underuse. Optimization impacts the lower PRB utilization for QCI 3. GBR shows efficient resource usage without reserving excessive bandwidth, aligning with the optimization strategy to improve user experience without affecting overall network performance. The consistent throughput for QCI 3, even under different GBR settings, correlates with reduced latency and jitter, as shown in the study. This is essential for real-time applications like gaming. The ability to switch between different activities (Speed test, YouTube, PUBG) without significant spikes or drops in PRB utilization indicates a well-optimized network capable of handling dynamic traffic efficiently. Network Load Handling: the graph also shows the network's performance under high-capacity conditions (e.g., multiple Speed tests and YouTube sessions). The PRB utilization lines and throughput areas indicate the network's ability to maintain performance even when heavily loaded, supporting the effectiveness of the optimization scheme.

A Physical Resource Block (PRB) divides the available bandwidth into smaller units, enabling the network to allocate resources efficiently to different users and services. Depending on network conditions, the number of PRBs available for downlink and uplink transmission may vary. During the gameplay of PUBG using three different SIM cards, the downlink/uplink Physical Resource Block Utilization and the downlink/uplink cell throughput were almost the same. This indicates that the DL Guaranteed Bit Rate (GBR) of 20Mbps does not reserve 20Mbps of cell throughput. It was observed that DL PRB utilization is much higher compared to dual bearer (QCI 3 GBR DL 100Kbps/UL 50Kbps) or default bearer (QCI 7 non-GBR). This implies that more data is transmitted from the network to the user's devices than dual and default bearer, but it does not reserve cell throughput. GBR QCI only consumes the required resources based on the throughput requirement and does not reserve the maximum with GBR value.

Air interface Latency

Air interface latency is the delay when transmitting data wirelessly over a cellular network. This latency can significantly impact the performance of applications that require real-time communication, such as e-sport gaming. The Radio Access Network provides mobile devices with wireless access to the network. It is closest to the end user, consisting of base stations (eNB), base station controllers (BSCs), and mobile switching centers (MSCs). RAN is also responsible for providing quality of service (QoS) mechanisms and implementing the RAN feature (QCI optimization).

The comparison of latency between QCI 7, QCI 3, and QCI 3 with optimization is shown in Figure 6. The average latency of QCI 3 with optimization is lower compared to QCI 3 and QCI 7. The default QCI 3 (without optimization) can reduce average latency by 47% more than QCI 7, while QCI 3 with optimization can reduce 19.7% more than QCI 3 or 66.7% more than QCI 7. This lower average latency is important in improving the gaming experience for users. Lower latency can prevent input lag, less accuracy, game freeze, and difficulty interaction in a real-time multiplayer game like PUBG.

The latency standard deviation for QCI 3 with optimization is lower than the default QCI 3 and QCI 7. The standard deviation of latency can be used as an indicator of jitter. When the standard deviation of the latency is low, it suggests that the delay between packets is more consistent, which means there is less jitter in the network. The jitter of QCI 3 with optimization is reduced by 140% compared to default QCI 3 or 596% compared to QCI 7. This indicates that QCI 3 with optimization effectively reduces average latency and the latency standard deviation (jitter).

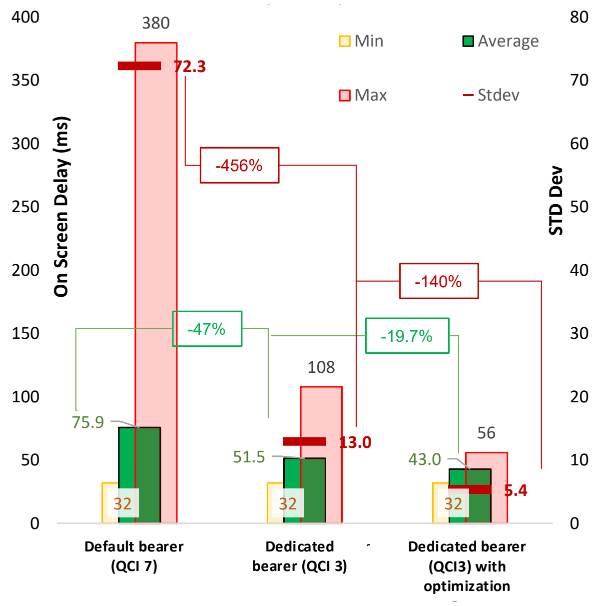


Figure 6. Latency Comparison Between QCI 7, QCI 3, and QCI 3 After Optimization

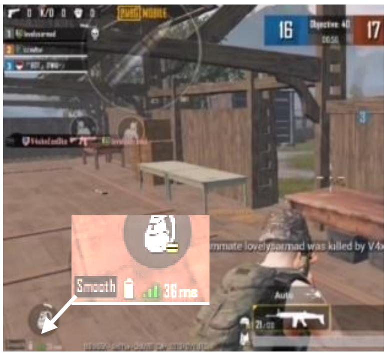
The air interface latency of QCI 3 with optimization was also tested under a bad coverage network (RSRP< -100dBm, SINR < 0dB) and a very high-capacity network (PRB > 97%, Active User = 40). The latency was compared for QCI 8 (baseline), QCI 7, QCI 3, and QCI 3 with optimization, and the results are displayed in Figure 7. Additionally, the average air interface latency experienced by users while playing the popular game PUBG was compared, as shown in Table 2.

Table 2. Throughput While Playing The PUBG Game

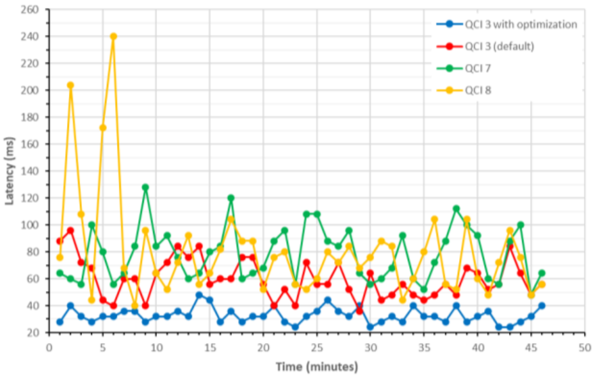
Air interface latency (ms)	QCI 8	QCI 7	QCI 3	QCI 3 with optimization
Average	67.9	67.2	57.4	46.2
Minimum	40	36	36	24
Maximum	240	160	96	84
Stdev	24.7	22.5	11.7	10.7

As shown in Table 2, the average air interface latency during PUBG gameplay was 67.9 ms for QCI 8, 67.2 ms for QCI 7, 57.4 ms for QCI 3, and 46.2 ms for QCI 3 with optimization. QCI 3 with optimization was found to reduce the average latency by up to 32% compared to QCI 8, 31% compared to QCI 7, and 19% compared to QCI 3 without optimization. Furthermore, QCI 3 with optimization reduced the maximum, minimum, and standard deviation values of air interface latency while playing games. Optimization strategy can improve user experience by reducing average latency and latency variations (jitter).

https://doi.org/10.25077/jnte.v13n2.1193.2024



(a)



(b)

Figure 7. Air Interface Latency While Playing PUBG Game (a) on The UE Screen and (b) Comparison Between QCIs

However, this work only focuses on QCI 3 optimization for e-sport games, which may not be generalized to other applications or network conditions. The performance evaluation of the optimization was performed under specific conditions of poor coverage and high-capacity networks only, which may not reflect all real-world scenarios. Although QCI 3 does not reserve network resources, its impact on network performance when multiple applications run simultaneously has not been explored extensively. These findings can be further advanced with future research on QoS management and latency optimization, potentially leading to improved user experience in real-time applications. Further research is needed to enrich the impact of optimization on a wider range of services and network environments.

CONCLUSIONS

The optimization strategy of QCI 3 has greatly improved user experience by reducing average latency and jitter. Performance and functionality tests in the case of an e-sport game have shown that QCI 3 optimizations can be enabled in parallel with multiple QCI values without affecting the throughput of other applications. QCI 3 does not reserve network resources and only uses the necessary resources based on scheduling throughput requirements without reserving the maximum GBR value. Under bad coverage and a very high-capacity network, QCI 3 with optimization was found to reduce the average latency by up to 32% compared to QCI 8, 31% compared to QCI 7, and 19% compared to QCI 3 without optimization. The optimization of QCI 3 has reduced the maximum, minimum, and standard deviation values of air interface latency during gameplay. Thus, the optimization can significantly enhance the user experience by reducing average latency and latency variations (jitter). However, it is necessary to explore the effect of optimization on other applications.

ACKNOWLEDGMENT

The authors thank PT XL Axiata for its technical resources and the Institute for Research and Community Service, Pamulang University, for their support.

REFERENCES

[1] A. Delaporte and K. Bahia, "The State of Mobile Internet Connectivity 2021," *GSMA Intell.*, pp. 1–66, 2021.

[2] K. Bahia and A. Delaporte, "The State of Mobile Internet Connectivity 2022," *GSMA Intell.*, pp. 1–61, 2022.

[3] 3GPP TS 123.203, "Technical Specification, Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture," vol. 15.4.0. ETSI, p. 262, 2018.

[4] 3GPP TS 123.203, "Technical Specification, Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture," vol. 17.2.0. ETSI, p. 272, 2022.

[5] E. Zeydan, J. Mangues-Bafalluy, O. Dedeoglu, and Y. Turk, "Performance comparison of QoS deployment strategies for cellular network services," *IEEE Access*, vol. 8, pp. 176073–176088, 2020, doi: 10.1109/ACCESS.2020.3027626.

[6] N. I. Othman, A. F. Ismail, K. Badron, W. Hashim, and S. Pinardi, "Improvements for Uplink Long Term Evolution (Ul-Lte) in Heterogeneous Network Using Dynamic Spectrum Allocation Technique," *Int. J. Comput. Networks Commun.*, vol. 14, no. 6, pp. 21–35, 2022, doi: 10.5121/ijcnc.2022.14602.

[7] R. Gatti and Shivashankar, "Improved resource allocation scheme for optimizing the performance of cell-edge users in LTE-A system," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 1, pp. 811–819, Jan. 2021, doi: 10.1007/S12652-020-02084-x.

[8] P. Ameigeiras *et al.*, "3GPP QoS-based scheduling framework for LTE," *EURASIP J. Wirel. Commun. Netw.*, vol. 78, no. 1, pp. 1–14, Dec. 2016, doi: 10.1186/S13638-016-0565-9.

[9] E. K. Hong, J. Y. Baek, Y. O. Jang, J. H. Na, and K. S. Kim, "QoS-guaranteed scheduling for small cell networks," *ICT Express*, vol. 4, no. 4, pp. 175–180, Jan. 2018, doi: 10.1016/J.ICTE.2017.11.017.

[10] A. Gizik, O. A. Sensoy, and E. Masazade, "Enhanced Dynamic Scheduling for Uplink Latency Reduction in Broadband VoLTE Systems," in *55th-Asilomar Conference on Signals, Systems and Computers*, 2021, p. (pp. 13-17). doi: 10.1109/IEEECONF53345.2021.9723407.

[11] N. N. Sirhan and M. Martinez-Ramon, "Cognitive Radio Resource Scheduling Using Multi-Agent Q-Learning for Lte," *Int. J. Comput. Networks Commun.*, vol. 14, no. 2, pp. 77–95, 2022, doi: 10.5121/ijcnc.2022.14205.

[12] U. Singh *et al.*, "Scalable priority-based resource allocation

scheme for M2M communication in LTE/LTE-A network," *Comput. Electr. Eng.*, vol. 103, p. 108321, 2022, doi: https://doi.org/10.1016/j.compeleceng.2022.108321.

[13] A. T. Koc, S. C. Jha, R. Vannithamby, and M. Torlak, "Device power saving and latency optimization in LTE-a networks through DRX configuration," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 5, pp. 2614–2625, 2014, doi: 10.1109/TWC.2014.031914.131298.

[14] F. Moradi, E. Fitzgerald, M. Pioro, and B. Landfeldt, "Flexible DRX optimization for LTE and 5G," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 607–621, Jan. 2020, doi: 10.1109/TVT.2019.2952251.

[15] K. Wang, X. Li, H. Ji, and X. Du, "Modeling and Optimizing the LTE Discontinuous Reception Mechanism Under Self-Similar Traffic," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5595–5610, Jul. 2016, doi: 10.1109/TVT.2014.2328232.

[16] I. Baig, S. Chaudhuri, and D. Das, "Downlink interference control through adaptive DRX in a carrier aggregation enabled LTE-Advanced heterogeneous networks," in *2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Jun. 2017, p. (pp.1-6). doi: 10.1109/ANTS.2016.7947786.

[17] S. Khwandah, B. Mohamad, J. Cosmas, P. Lazaridis, and Z. Zaharis, "Video performance using adaptive DRX switching for LTE-A," in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jul. 2017, p. (pp. 1-5). doi: 10.1109/BMSB.2017.7986128.

[18] J. M. Liang, J. J. Chen, H. H. Cheng, and Y. C. Tseng, "An energy-efficient sleep scheduling with QoS consideration in 3GPP LTE-advanced networks for internet of things," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 3, no. 1, pp. 13–22, 2013, doi: 10.1109/JETCAS.2013.2243631.

[19] J. M. Liang, J. J. Chen, P. C. Hsieh, and Y. C. Tseng, "Two-Phase Multicast DRX Scheduling for 3GPP LTE-Advanced Networks," *IEEE Trans. Mob. Comput.*, vol. 15, no. 7, pp. 1839–1849, Jul. 2016, doi: 10.1109/TMC.2015.2480090.

[20] O. Ergul, O. Yilmaz, A. T. Koc, and O. B. Akan, "DRX and QoS-aware energy-efficient uplink scheduling for long term evolution," in *2013 IEEE Global Communications Conference (GLOBECOM)*, Jul. 2014, p. (pp. 4644-4649). doi: 10.1109/GLOCOMW.2013.6855684.

[21] Y. Wang *et al.*, "Uplink Resource Scheduling Strategy in Wireless Networks," in *2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology (CCET)*, Aug. 2020, p. (pp. 294-298). doi: 10.1109/CCET50901.2020.9213145.

[22] W. Chen, X. Xu, C. Yuan, J. Liu, and X. Tao, "Virtualized Radio Resource Pre-Allocation for QoS Based Resource Efficiency in Mobile Networks," in *2017 IEEE Global Communications Conference*, Jan. 2017, p. (pp. 1-6). doi: 10.1109/GLOCOM.2017.8254655.

[23] N. Patriciello, S. Lagen, L. Giupponi, and B. Bojovic, "The impact of NR scheduling timings on end-to-end delay for uplink traffic," in *2019 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2019, p. (pp. 1-6). doi: 10.1109/GLOBECOM38437.2019.9013231.

[24] U. Toseef, C. Goerg, T. Weerawardane, and A. Timm-Giel, "Performance comparison of PDCP buffer management schemes in LTE system," in *IFIP Wireless Days*, Oct. 2011, p. (pp. 1-3). doi: 10.1109/WD.2011.6098217.

[25] 3GPP TS 36.323, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification," vol. 17.2.0, no. 17. p. 58, 2023.

[26] U. Toseef, T. Weerawardane, A. Timm-Giel, and C. Görg, "LTE system performance optimization by discard timer based PDCP buffer management," in *8th International Conference on High-capacity Optical Networks and Emerging Technologies*, Dec. 2011, p. (pp. 116-121). doi: 10.1109/HONET.2011.6149800.

[27] P. Szilagyi and C. Vulkan, "Efficient LTE PDCP buffer management," in *2015 IEEE International Conference on Communications (ICC)*, Sep. 2015, p. (pp 5928-5934). doi: 10.1109/ICC.2015.7249267.

[28] A. M. Ahmed, A. Patel, and M. Z. A. Khan, "Reliability Enhancement by PDCP Duplication and Combining for Next Generation Networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, Apr. 2021, p. (pp.1-5). doi: 10.1109/VTC2021-Spring51267.2021.9449079.

[29] J. Sun, S. Zhang, S. Xu, and S. Cao, "High Throughput and Low Complexity Traffic Splitting Mechanism for 5G Non-Stand Alone Dual Connectivity Transmission," *IEEE Access*, vol. 9, pp. 65162–65172, 2021, doi: 10.1109/ACCESS.2021.3076301.

[30] M. Finsterbusch, C. Richter, E. Rocha, J.-A. Muller, and K. Hanssgen, "A Survey of Payload-Based Traffic Classification Approaches," *IEEE Commun. Surv. Tutorials*, vol. 16, no. 2, pp. 1135–1156, 2014, doi: 10.1109/SURV.2013.100613.00161.

[31] A. Azab, M. Khasawneh, S. Alrabaee, K.-K. R. Choo, and M. Sarsour, "Network traffic classification: Techniques, datasets, and challenges," *Digit. Commun. Networks*, 2022, doi: https://doi.org/10.1016/j.dcan.2022.09.009.

[32] M. Seufert, S. Wassermann, and P. Casas, "Considering User Behavior in the Quality of Experience Cycle: Towards Proactive QoE-Aware Traffic Management," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1145–1148, 2019, doi: 10.1109/LCOMM.2019.2914038.

[33] F. Metzger *et al.*, "An Introduction to Online Video Game QoS and QoE Influencing Factors," *IEEE Commun. Surv. Tutorials*, vol. 24, no. 3, pp. 1894–1925, Sep. 2022, doi: 10.1109/COMST.2022.3177251.

[34] S. Flinck Lindstrom, M. Wetterberg, and N. Carlsson, "Cloud Gaming: A QoE Study of Fast-paced Single-player and Multiplayer Gaming," in *2020 IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC)*, Dec. 2020, p. (pp. 34-45). doi: 10.1109/UCC48980.2020.00023.

[35] Y. Lin and H. Shen, "CloudFog: Leveraging Fog to Extend Cloud Gaming for Thin-Client MMOG with High Quality of Service," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 2, pp. 431–445, 2017, doi: 10.1109/TPDS.2016.2563428.

[36] S. Vlahovic, M. Suznjevic, and L. Skorin-Kapov, "The impact of network latency on gaming QoE for an FPS VR game," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, Jun. 2019, p. (pp.1-3). doi: 10.1109/QOMEX.2019.8743193.

[37] P. B. Beskow, K. H. Vik, P. Halvorsen, and C. Griwodz, "The partial migration of game state and dynamic server selection to reduce latency," *Multimed. Tools Appl.*, vol. 45, pp. 83–107, Oct. 2009, doi: 10.1007/s11042-009-0287-7.

 https://doi.org/10.25077/jnte.v13n2.1193.2024

[38] M. Claypool and D. Finkel, “The effects of latency on player performance in cloud-based games,” in *2014 13th Annual Workshop on Network and Systems Support for Games*, 2014, p. (pp. 1-6). doi: 10.1109/NetGames.2014.7008964.

[39] S. Liu, M. Claypool, A. Kuwahara, J. Scovell, and J. Sherman, “The Effects of Network Latency on Competitive First-Person Shooter Game Players,” in *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*, 2021, pp. 151–156. doi: 10.1109/QoMEX51781.2021.9465419.

[40] Y. Liu, M. Huynh, A. Mangla, and D. Ghosal, “Performance analysis of adjustable discontinuous reception (DRX) mechanism in LTE network,” in *2014 23rd Wireless and Optical Communication Conference (WOCC)*, 2014, pp. 1–6. doi: 10.1109/WOCC.2014.6839948.

[41] J. R. Ky, B. Mathieu, A. Lahmadi, and R. Boutaba, “ML Models for Detecting QoE Degradation in Low-Latency Applications: A Cloud-Gaming Case Study,” *IEEE Trans. Netw. Serv. Manag.*, vol. 20, no. 3, pp. 2295–2308, 2023, doi: 10.1109/TNSM.2023.3293806.

AUTHOR(S) BIOGRAPHY

Patria Adhistian

He received a B.Eng in Informatics from Telkom University in Bandung, Indonesia, in 2007 and a Magister Management from Diponegoro University in Semarang, Indonesia in 2010. He has over 15 years of experience in IT and Telecom as a professional with in-depth knowledge in VAS, Data Service, Data Core network areas, Project Delivery Services, Core Networks, Wirelines, and Wireless Technology. He began his career as an engineer in a telecommunication company in 2008 and as a lecturer in the Department of Industrial Engineering at Pamulang University in 2016.

Priyo Wibowo

He received a B.Sci. degree in Physics from Yogyakarta State University, Yogyakarta, Indonesia 2004, and a master's degree in Electrical Engineering, specializing in Telecommunications Engineering from the University of Indonesia, Depok, Indonesia. He has experience working as the Technical Manager at the Electromagnetic Laboratory of the Indonesian Institute of Science. He is currently a researcher at the National Research and Innovation Agency. He is also a visiting lecturer at the Pamulang University. His research interests include applied electromagnetics, electromagnetic compatibility, testing technology, instrumentation, telecommunications, and antennas.