P4-enabled Trusted WLAN Access Network with Edge Computing Support

Ze-Yu Jin
Dept. of CS
NYCU, Taiwan

Wai-Lok Shum
Dept. of CS
NYCU, Taiwan

Li-Hsing Yen
Dept. of CS
NYCU, Taiwan

Chien-Chao Tseng
Dept. of CS
NYCU, Taiwan

ABSTRACT

The 3rd Generation Partnership Project (3GPP) proposed a new non-3GPP access network, the Trusted WLAN Access Network (TWAN), in Release 16 of the 5th Generation Mobile Communication Network (5G). TWAN enables devices without 5G capability to access 5G networks via WLAN, allowing operators to serve non-5G devices in the market through Wi-Fi services. This paper divides TWAN into a control plane and a user plane, which are handled by an SDN controller and P4 switches, respectively. This setup, named P4-TWAN, is further integrated with User Plane Function (UPF) to eliminate unnecessary tunneling. Experimental results show that P4-TWAN processed user traffic at line rate while the integration with UPF benefited from the elimination of unneeded tunneling.

CCS CONCEPTS

• Networks~Network services~Programmable networks

KEYWORDS

5G Access Network, Non-3GPP Access, P4, Edge Computingt

1 Introduction

The 3rd Generation Partnership Project (3GPP) proposed the Non-3GPP Interworking Function (N3IWF) to allow 3GPP-capable devices to access 5G core (5GC) networks through non-3GPP and untrusted access networks such as wireless local area networks (WLANs). For non-3GPP but trusted access networks (such as operator-controlled Wi-Fi), 3GPP introduced the Trusted WLAN Access Network (TWAN) [1] in 5G NR Release 16. TWAN comprises the Trusted WLAN Access Point (TWAP) and the Trusted WLAN Interworking Function (TWIF). TWAP is a Wi-Fi node with 802.1X authentication capability, and TWIF is responsible for processing control messages from TWAP and non-5G-capable over WLAN (N5CW) devices. Notably, TWAP forwards user traffic and handles registration messages on behalf of N5CW devices (Fig. 1).

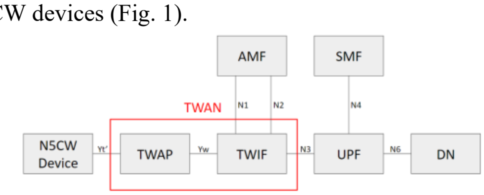

Figure 1: TWAN Architecture [1]

To enable low-latency communication, 5G networks incorporate edge computing by deploying Edge User Plane Functions (edge UPF) and edge servers near access networks [2]. In practical deployments, edge service providers may set up one-to-one instead of one-to-many access models to provide isolated ultra-low latency connections to users [3] [4]. However, such a setup lacks the flexibility to precisely provide communication resources to exactly match the user demands. Furthermore, access nodes (where TWAN resides) and gateways (where edge UPF resides) could be physically close in an edge computing environment, yet they need to create GPRS Tunneling Protocol User Plane (GTP-U) [5] tunnels. This configuration consumes extra bandwidth, considering that IP networking is not needed between them. A possible enhancement is to integrate the functionalities of TWAN and UPF in a single device to eliminate the tunneling. Nevertheless, the integration should also allow the device to act as a standalone TWAN to make it a one-size-fits-all solution.

This paper showcases a design and implementation of TWAN, namely, P4-TWAN, for future 5GC networks. P4-TWAN follows the principle of Control and User Plane Separation (CUPS) by using software-defined networking (SDN) controllers for handling control messages and Programming Protocol-Independent Packet Processors (P4) [6] switches for hardware-accelerated user traffic processing. It enables line-rate processing of N5CW traffic. We also present an implementation and optimization scheme to combine the P4-TWAN and the UPF functionalities in a single device. The integrated device, named P4-TWAN-UPF, can operate as a standalone TWAN and an integrated TWAN-UPF device at the same time. We conducted experiments to test the performance of our implementations. The results confirmed that P4-TWAN maintains line rate during GTP-U handling and that P4-TWAN-UPF maintains line rate processing for bidirectional traffic.

The rest of this paper is organized as follows. Sec. 2 reviews related work. Sec. 3 presents the design and the implementation of P4-TWAN-UPF. Sec. 4 reports the experimental results. The last section concludes this paper.

2 Related Work

The UPF integrated with P4-TWAN in this paper is a follow-up of Chen's work [7], which designed and implemented a 5G UPF using SDN and P4 technologies. The functions of the UPF include Packet Detection Rule (PDR), Forwarding Action Rule (FAR), and QoS Enforcement Rule (QER). The overall architecture comprises three parts: the 5G control plane, the SDN control plane, and the SDN user plane. The former two parts correspond to the PFCP Agent and UPF-c App, respectively, on an Open Networking Operating System (ONOS) [8] controller, and the last part corresponds to the UPF-u pipeline on a P4 switch. The PFCP Agent processes PFCP messages between the UPF and the 5G control plane. After processing, the UPF-c converts the processing results into Flow Rules, which are then sent to the P4 switch through the ONOS controller. The UPF-u implements PDR, FAR, and QER tables to support these three UPF functions.

Aghdai et al. [9]presented a P4-based 4G edge gateway to handle tunneled traffic between eNodeB and EPC. If the gateway detects packets from eNodeB that should be directed to edge servers, it decapsulates the packets before forwarding them to the edge servers, where a Load Balancer distributes workload across servers. The architecture is similar to 5G's Uplink Classifier UPF and Anchor UPF, but it does not eliminate bandwidth overhead demanded by GTP-U tunneling.

SD-Fabric [10] uses SDN and P4 to build data center networks with integrated UPF functionality. It virtualizes the network into a single UPF to support low-latency, high-throughput edge computing. All P4 switches in the user plane share the same pipeline, avoiding redeployment for function switching, while the control plane uses separate applications for different functions. Our work references this design when integrating TWAN and UPF. However, like [9], SD-Fabric still incurs extra bandwidth for GTP-U tunneling.

In short, current research on mobile network edge computing does not cut off redundant tunneling. This paper addresses this gap by integrating P4-TWAN with UPF.

3 Design and Implementation of P4-TWAN-UPF

Figure 2 shows the architecture of our design. TWAP is a Wi-Fi node capable of performing 802.1X authentication. N5CW devices can communicate with TWAP using Wi-Fi but are unable to perform standard registration with 5GC. Nevertheless, they can perform EAP-AKA' authentication procedure with TWAP. TWIF is the core of TWAN. We followed the principle of CUPS to divide the functionality of TWIF into TWIF Control Plane (TWIF-CP) and TWIF User Plane (TWIF-UP). We implemented TWIF-CP on an ONOS controller, implemented TWIF-UP on P4 switches, and integrated TWIF-UP and UPF-UP on the same P4 switch using a combined pipeline.

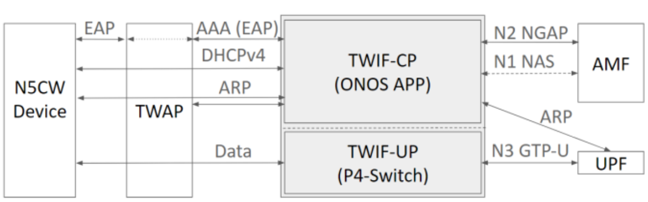

Figure 2: TWIF CUPS Architecture

3.1 TWIF Control Plane (TWIF-CP)

TWIF-CP needs to communicate with the Access and Mobility Management Function (AMF) in 5GC. Refer to Figure 3 for the architecture of TWIF-CP, which comprises five key components. The *N2 module* enables communication with AMF, handling NGAP encoding/decoding, and SCTP transport. Besides N2 signaling messages, TWIF-CP needs to handle two types of control messages from TWAP and N5CW devices. One type is from WLAN (such as AAA, ARP, and DHCP) for authentication and IP address allocation. For this part, we design corresponding In-band Network Control Modules: *RADIUS/DHCP Agents* for protocol encoding/decoding and an *ARP Proxy* for ARP handling. The other type is N1 NAS (Non-Access Stratum) between a UE and AMF. TWIF-CP should process N1 NAS signaling messages on behalf of the authenticated N5CW devices. To this end, TWIF-CP should keep track of the status (such as registered, authenticated, or PDU session created) and identities (such as MAC address in WLAN and Subscription Permanent Identifier (SUPI) in 5G) of each N5CW device. Our design is to create a *Relay UE* instance for each attached N5CW device and let the Relay UE process N1 NAS messages, maintain UE information, and encode/decode N1 messages for the N5CW device.

The TWIF *Manager*, as the core, coordinates internal operations (message conversion between the two types of control messages, resource allocation, and network topology management), manages Relay UE instances, and stores network configuration data provided by administrators. Finally, the Flow Handler controls TWIF-UP by establishing and updating flow rules on the TWIF-UP.

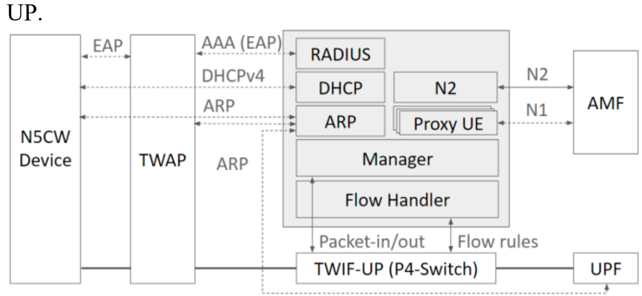

Figure 3: TWIF-CP Architecture

3.2 TWIF User Plane (TWIF-UP)

TWIF-UP processes uplink and downlink user traffic on P4 switches. When receiving an uplink packet from N5CW devices, TWIF encapsulates it with an outer GTP-U header before forwarding it to the UPF. On the other hand, when downlink traffic passes through TWIF, TWIF removes the outer header and then delivers it to N5CW devices. TWIF must validate the TEID (Tunnel Endpoint Identifier) and the destination IP address in downlink packets, and modify the MAC address before the packets exit to ensure delivery to the correct N5CW device.

Our prototype uses the Tofino 1 P4 Model with Ingress and Egress pipelines, each featuring a Parser and Deparser. The TWIF-UP pipeline is structured into five table groups (Figure 4).

Table 1: Experimental Environment

Device Type	CPU/Model	Memory/Port	Network/System
N5CW Device	Intel i7-10700	16 GB	Wi-Fi 6 AX201/Ubuntu 18.04.5
TWAP	Edgecore EAP 101	-	OpenWRT 21.02
P4-Switch (TWIF-UPF)	Inventec P4-Switch	10 Gbps	P4 SDE 9.3.0/Combined Pipeline
Server (Hosting VM)	Intel Xeon E5-2620 v4	128 GB	3×Intel 82599ES/Proxmox 7.1
VM	8 vCPU	8 GB	PCIE Passthrough/Ubuntu 20.04.5
SDN Controller	-	-	ONOS 2.6
5G Core Network	-	-	free5GC

In traffic processing, tables have specific roles. The L2 Table targets TWAP traffic by MAC and sets egress. PKT-OUT filters CPU packets, removes content, and defines egress. INC detects in-band messages (RADIUS, DHCP, ARP) and routes to CPU; PKT-IN adds metadata. Uplink PDR identifies N5CW traffic by Source IP and assigns FAR IDs; FAR applies GTP-U encapsulation and configures egress. Downlink PDR detects traffic by Inner Source IP and TEID, strips GTP-U headers, assigns FAR IDs; FAR adjusts MAC and egress. Key headers enable UPF-UP integration and efficient handling.

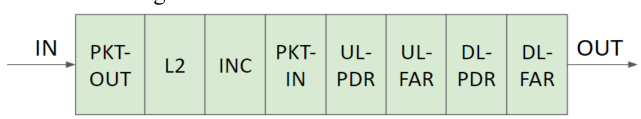

Figure 4: TWIF-UP Pipeline Architecture

3.3 Integration With UPF

Three key requirements underpin the integration of TWIF-UP and UPF-UP on the same P4 switch. First, the new architecture must retain the ability to operate as a standalone TWIF, ensuring backward compatibility. Second, the integration should be invisible to the 5G control plane; that is, the 5G control plane should continue to perceive TWAN and the integrated UPF as separate entities. Third, the design must support future scalability. To meet these criteria, the paper adopts a cascading approach rather than a direct merge, combining TWIF-UP and UPF-UP into a unified P4 pipeline (Figure 5). This method preserves the packet classification and processing capabilities of both components, enabling seamless operation

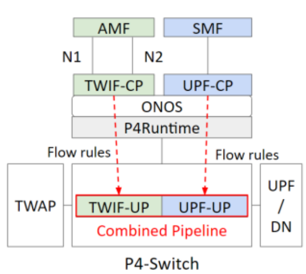

Figure 5: Combined Architecture

In implementing the UPF functionality within the P4-TWAN framework, we reused the existing UPF-CP and UPF-UP components from P4-UPF [7] without modification. In the Combined Pipeline design, uplink traffic flows sequentially through the TWIF Uplink section and then the UPF, while downlink traffic traverses the UPF before reaching the TWIF Downlink section. To achieve this, the paper extracts UPF-UP's traffic-processing tables (including Packet Detection Rules [PDR], Forwarding Action Rules [FAR], and QoS Enforcement Rules [QER]) and inserts them between the TWIF Uplink and Downlink stages, as depicted in Figure 6. Finally, to optimize resource utilization, functionally redundant tables, such as PKT-IN/OUT, are merged across TWIF-UP and UPF-UP, streamlining the overall pipeline structure while maintaining performance.

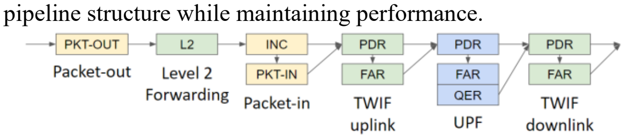

Figure 6: Combined Pipeline Architecture

4 Performance Evaluation

When conducting the performance evaluation of P4-TWAN-UPF, we utilized free5GC [11] as the 5G core network and two traffic generators to emulate N5CW devices and UPF/Data Networks for testing the user plane performance. Figure 7 shows the network topology and Table 1 details hardware/software configurations.

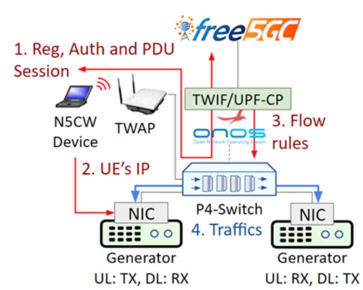

Figure 7: Testbed and Setup Procedures

4.1 TWIF-UP GTP-U Tests

This test is to assess whether TWIF-UP in the Combined Pipeline can maintain line rate while correctly processing GTP-U. Uplink traffic was generated using iPerf [12], while downlink traffic employed Scapy [13]. The original packet size was 1450 Bytes, expanding to 1486 Bytes after GTP-U encapsulation (i.e., a 2.5% increase). The test comprised three groups, each lasting 12 seconds, with data statistics collected over the middle 10 seconds. The three test groups are as follows. Uplink traffic at 95% and 100%

bandwidth evaluated TWIF-UP's GTP-U processing under partial load and full-load line rate maintenance, while 100% downlink traffic assessed its GTP-U processing and line rate retention.

For each test group, we measured the traffic rates at the ingress and egress ports, respectively, on the P4 switch. Figure 8 shows that, after passing through TWIF-UP, the uplink traffic rate in Group 1 increased by 0.24 Gbps while the downlink traffic rate in Group 3decreased by 0.24 Gbps. For a 10 Gbps link, this aligned with the expected 2.5% discrepancy, confirming correct GTP-U processing. Results from Groups 2 and 3 further demonstrated that TWIF-UP maintains line rates on GTP-U handling.

4.2 TWIF-UPF-UP Tests

We first address the issues of GTP-U tunneling within the decoupled access gateway architecture. Traffic traversing between an access network (AN) and UPF is encapsulated in GTP-U, which requires higher bandwidth than the UE-AN and UPF-DN segments. Assume identical bandwidth limits for UE-AN, AN-UPF, and UPF-DN segments. If UE-AN traffic reaches the bandwidth limit, traffic encapsulation will cause packet loss and reduced end-to-end throughput. This reduction is inversely proportional to the original packet size, i.e., smaller packets increase the overhead ratio of GTP-U headers.

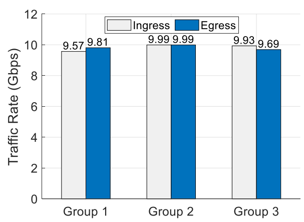

Figure 8: TWIF-UP Functional and Line Rate Test Result

We investigated whether the integrated TWIF-UPF-UP architecture avoided throughput degradation under full traffic load. We tested four different payload sizes (800, 1000, 1200, 1400 Bytes), each lasting 12 seconds with 10-second data collection. iPerf generated all traffic, and results measured RX/TX rates (packets/second) on the generator NICs.

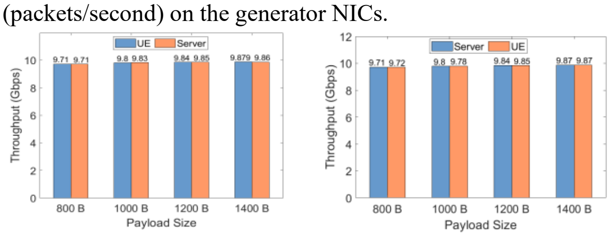

Figure 9: TWIF-UPF-UP Throughput and Line Rate: Uplink (Left) and Downlink (Right)

Figure 9 shows no significant end-to-end throughput degradation in TWIF-UPF-UP for different packet sizes in both uplink and downlink traffic. The results demonstrated the benefit of eliminating GTP-U encapsulation and also confirmed the ability of TWIF-UPF-UP to maintain line rate processing for bidirectional traffic.

5 Conclusion

Our paper proposes an SDN- and P4-based TWAN architecture that separates TWAN into control and user planes, enhancing control message processing efficiency and enabling line-rate user traffic handling. Furthermore, by integrating UPF functionality, the integrated TWAN-UPF allows edge computing traffic to be directly forwarded to edge servers, eliminating the need for intermediate devices in traditional designs and reducing extra computational resource consumption.

REFERENCES

[1] 3GPP. 2022. System architecture for the 5G System (5GS) (Release 16) 3GPP TS 23.501 V16.13.0. https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/16.13.00_60/ts_123501v161300p.pdf (accessed June 15, 2025).

[2] N. Makondo, H. I. Kobo, T. E. Mathonsi, D. Du Plessis, T. M. Makhosa and L. Mamushiane. 2024. An Efficient Architecture for Latency Optimisation in 5G Using Edge Computing for uRLLC Use Cases. In *2024 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems* (icABCD). IEEE, Port Louis, Mauritius, pp. 1-7.

[3] E. Goshi, H. Harkous, S. Ahvar, R. Pries, F. Mehmeti and W. Kellerer. 2024. Joint UPF and Edge Applications Placement and Routing in 5G & Beyond. In *2024 IEEE 21st International Conference on Mobile Ad-Hoc and Smart Systems* (MASS). IEEE, Seoul, Korea, Republic of, pp. 446-452.

[4] Leyva-Pupo, I., Santoyo-González, A., & Cervelló-Pastor, C. 2019. A Framework for the Joint Placement of Edge Service Infrastructure and User Plane Functions for 5G. *Sensors*, 19(18), 3975.

[5] 3GPP. 2021. General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U) (Release 16) 3GPP TS 29.281 V16.2.0. https://www.etsi.org/deliver/etsi_TS/129200_129299/129281/16.02.00_60/ts_129281v160200p.pdf (accessed June 1, 2025).

[6] Pat Bosshart, Dan Daly, Glen Gibb, Martin Izzard, Nick McKeown, Jennifer Rexford, Cole Schlesinger, Dan Talayco, Amin Vahdat, George Varghese, and David Walker. 2014. P4: programming protocol-independent packet processors. S*IGCOMM Comput. Commun*, (44)3, 87–95.

[7] Chen, Bo-Fan. 2021. P4 UPF-NS: P4-Enabled 5G User Plane Function with Network Slicing and Bandwidth Management. https://hdl.handle.net/11296/v365ej (accessed June 12, 2025).

[8] ONF. 2014. Open Networking Operating System (ONOS). https://opennetworking.org/onos/ (accessed June 2, 2025).

[9] A. Aghdai, M. Huang, D. Dai, Y. Xu and J. Chao. 2018. Transparent Edge Gateway for Mobile Networks. In *IEEE 26th International Conference on Network Protocols* (ICNP). IEEE, Cambridge, UK, pp. 412-417.

[10] ONF. 2021. SD-Fabric. https://docs.sd-fabric.org/master/index.html# (accessed June 1, 2025).

[11] free5GC. https://free5gc.org/ (accessed June 21, 2025).

[12] iPerf - The ultimate speed test tool for TCP, UDP and SCTP. https://iperf.fr/ (accessed June 21, 2025).

[13] Scapy Project. https://scapy.net/ (accessed June 21, 2025).

Acknowledgment

This work was supported in part by the National Science and Technology Council, Taiwan, under grants NSTC 113-2218-E-011-009, 113-2221-E-A49-179, 114-2218-E-011-003, and 114-2221-E-A49-009.