Routing and Packet Scheduling For Virtualized Disaggregate Functions in 5G O-RAN Fronthaul

Cheng Cheng Zhang, Kim Khoa Nguyen, Mohamed Cheriet
École de technologie supérieure, Université du Québec,
Montreal, Canada
Email: {chengcheng.zhang.1}@ens.etsmtl.ca, {kim-khoa.nguyen, mohamed.cheriet}@etsmtl.ca

GLOBECOM 2021 - 2021 IEEE Global Communications Conference | 978-1-7281-8104-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/GLOBECOM46510.2021.9685802

***Abstract*—Open Radio Access Network (O-RAN) is an innovative RAN architecture designed to revolutionize 5G and beyond mobile networks. O-RAN virtualizes the fronthaul network functions into O-RAN Centralized Unit (O-CU), O-RAN Distributed Unit (O-DU) and O-RAN Radio Unit (O-RU). Unfortunately, no standard data communication mechanism has been defined for the communication between these elements. Therefore, O-DUs may not work efficiently in O-DU pool, limiting the RAN performance. This paper investigates an optimized solution for routing and packet scheduling, allowing multiple O-DU pools to communicate with their O-RUs meeting the requirements of different 5G classes of service. We propose an O-DU pool architecture and formulate the problem of optimal routing and packet scheduling to forward Orthogonal Frequency-Division Multiplexing (OFDM) symbols over the optimal routes and map UDP packet sizes to fragment OFDM symbols. Numerical results show our solution can select the optimal routes and packet sizes to carry requested traffic. Moreover, in the multiple O-DU pools coexisting, we use the Dynamic Programming (DP) algorithm to find out the optimal global solution and a greedy algorithm to approximate the solution in near real-time.**

***Index Terms*—O-RAN, 5G, Routing, Packet scheduling.**

I. Introduction

The Fourth Generation (4G) RAN has Option 8 [1] to decouple their functions in Remote Radio Unit (RRU) and Baseband Unit (BBU). Unfortunately, the decoupled option still lacks the flexibility to adapt to dynamic services. O-RAN alliance, based on the 3rd Generation Partnership Project (3GPP), conceives its disaggregated components using the Virtualised Network Function (VNF) [2]. In Fig. 1 of the disaggregated 5G ORAN architecture, O-RAN takes 3GPP defined Option 2 and Option 7 [1] to decouple the O-CU, O-DU and O-RU. In the O-CU, Radio Resource Control (RRC) layer manages radio bearers, while Service Data Adaptation Protocol (SDAP) layer handles the mapping between Quality of Service (QoS) flows and data radio bearers established in Packet Data Convergence Protocol (PDCP) layer. In O-DU, Radio Link Control (RLC) layer, Medium Access Control (MAC) layer and High Physical (High PHY) Layer provide the sorting of QoS flows to apply services guaranteed by the selection of transport formats to map to physical resources blocks. In O-RU, the Low Physical (Low PHY) layer cooperated with the Radio Frequency (RF) layer to generate the 5G NR signals. The F1 [2] from O-RAN is the interface between O-CU and O-DU and the enhanced Common Public Radio Interface (eCPRI) protocol [2] is an interface for control data and user data between O-DUs and O-RUs.

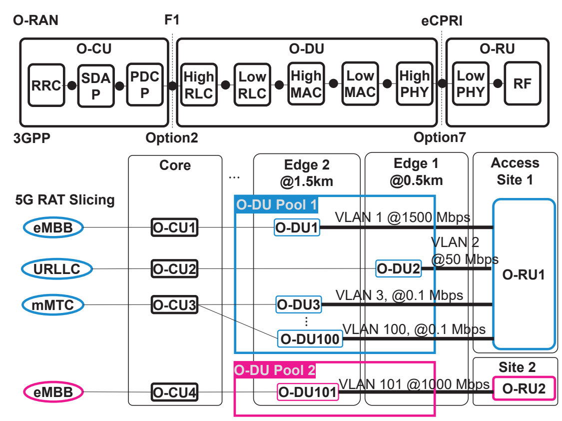

Fig. 1. O-RAN architecture and an example of O-DU placement for three types of 5G services in the 5G RAT slicing

3GPP has defined Radio Access Technology (RAT) slicing in the Fifth Generation (5G). Three main categories of services in 5G are shown in Fig. 1. Ultra-Reliable Low Latency Communications (URLLC) service [3] requires the O-RAN to minimize the latency and consider the sufficient bandwidth to transfer the traffic . The enhanced Mobile Broadband (eMBB) service [3] demands super large bandwidth and excellent quality for the application, e.g. 8K Resolution. The massive Machine Type Communication (mMTC) service [3] relies on the massive connection in the limited resources in the fronthaul network. Each service can be composed of O-RAN virtual elements placed at different locations. The architecture also allows coexistence of multiple vendors as shown in the example in Fig. 1. We assume there are two access sites to support a cell site in a region, and O-CUs, which aggregate and control the resource and the mapping with 5G RAT slicing, are in the fixed location in the core. O-RU 1 running on the Site 1 from a vendor in blue color requires O-DU pool 1 created in Edge1 and Edge2 with different distances to maintain three services simultaneously to access 5G RAT slicing. O-RU 1 is linked to the edges through VLANs with different bandwidth.

978-1-7281-8104-2/21/$31.00 ©2021 IEEE

The eMBB service which requires a superfast broadband but not extremely low latency, has to place the O-DU 1 in Edge 2 at the distance of 1.5 km from O-RU 1. The O-DU 1 associates VLAN 1 with 1500 Mbps high bandwidth. In the Edge 2, the mMTC service requires massive connections from O-DU 3 to O-DU 100 but low bandwidth of 0.1 Mbps per each to route the traffic from the Internet of Things (IoT). On the other hand, the URLLC service requires O-DU 2 to be placed on Edge 1 at the distance of 0.5 km to get low latency through VLAN2 with 50 Mbps bandwidth. Another vendor in pink color provides O-RU 2 on Site 2 to connect with the O-DU pool 2 to get only the eMBB service via O-DU 101 through the VLAN101 at 1000 Mbps bandwidth.

The aforementioned example suggests a challenging issue of routing and scheduling of an adaptive Ethernet fronthaul for different 5G services in multiple RAT slicing, which has not been discussed in prior work [4] [5] [6]. It is a high-complexity problem because an O-RU may support multiple slices simultaneously to provide different 5G services to subscribers. This requires multiple O-DU pools to be launched at different locations and connected to the O-RU via different rates. In our proposed architecture in the Fig. 1, we design an O-DU pool, as a group of O-DUs connected to the same O-RU. Each O-DU in the same pool provides different network characteristics based on the service requirement in the RAT slicing. We propose then a solution to route traffic from the O-RU to O-DU at an optimal physical distance and with an optimal bandwidth to meet RAT slicing requirements. Our proposed solution is different from the conventional VLAN that is mainly used to separate traffic in the Ethernet.

We combine the routing problem with the packet scheduling problem to provide an integral solution. The scheduling problem determines the optimal packet size and rate on the VLAN to carry 5G OFDM symbols. It is also a challenging problem in O-RAN because 3GPP has defined multiple waveform parameters, which results in various sizes of OFDM symbols according to 5G numerology parameter setup. However, the eCPRI protocol which links O-RU to O-DU is based on UDP protocol with a limited payload size of 1550 Bytes. Therefore, the O-DU must fragment the OFDM symbol into smaller pieces to fit in the UDP payload. Although O-RAN architectures allow the co-existence of multiple O-DU pools and O-RUs, it has not defined any OFDM symbols scheduling mechanism for mixed numerology parameters used on O-DU. Therefore, selecting an optimal UDP loading rate to fragment OFDM symbols to UDP payload is required to enable multiple O-DU pools working efficiently together in an O-DU pool. Fig. 2 presents an example of this problem. The O-DU sends OFDM symbols to its O-RU over an Ethernet-UDP/IP network. An OFDM symbol in each slot in numerology 0 has $66.76\mu s$ length. After the Control Plane (C-Plane) data arrived in the O-RU that schedules to plot the upcoming OFDM data symbol over its array antenna to do beamforming, O-DU starts fragment its OFDM symbol into a dedicated length to fit the UDP payload size in the User Plane (U-Plane), which will be carried over the encapsulated in IP over the Ethernet network.

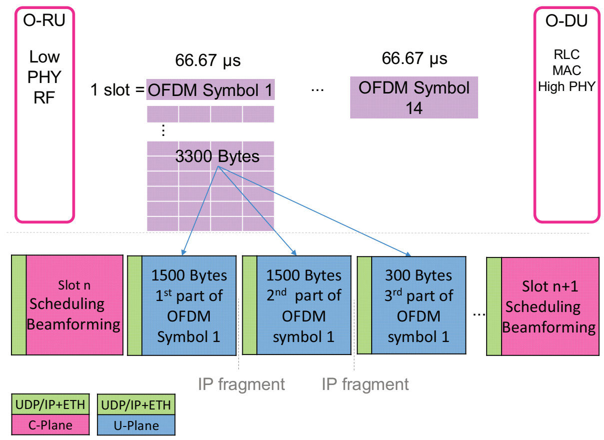

Fig. 2. An example for an O-DU to fragment the data in an OFDM symbol into 3 UDP packets carried in the capsule of UDP/IP over Ethernet protocol.

From our example shown in Fig. 2, the first OFDM symbol has 3300 bytes of binary data, while each UDP packet has a payload of a maximum of 1550 bytes. If O-DU schedules UDP packet size as 1500 bytes, the third UDP packet takes 300 bytes. It is a bad scheduling, because it wastes the utilization of the bandwidth that causes delay. If O-DU schedules in small packet size, it is also bad scheduling and well explained in Fig. 5 because of the delay. The problem becomes even more challenging if we inverse the objective function to place O-DUs that are initiated, launched and deleted dynamically. It scales the volume of connections and takes the migration from one location to another depending on the demand of subscribers' behaviors in the RAT network slices.

Prior works dealing with O-RAN integration are still limited at the architectural level and ignore the operational complexity. In particular, the dynamic establishment of a dedicated link, routing multiple O-DU pools, and scheduling UDP packets to minimize response time from the O-DU pool to O-RUs have not been addressed so far. In this paper, we investigate a practical implementation of integration by taking advantage of O-RAN fragmentation that is a function in the O-RAN to fragment OFDM symbols into a small UDP packet size. We propose a new optimal routing and a packet scheduling method to forward data and a greedy algorithm to approximate the optimal solution.

We summarize the contributions of this work as follows.

- We propose a new architecture of O-RAN to support 5G services in the 5G RAT slicing.
- We propose an optimal routing model and a packet scheduling model to route and encapsulate UDP packets according to the arrival symbol rate on the links between the O-DU pool and O-RUs to meet 5G requirements of minimal costs.
- We design a dynamic programming (DP) algorithm to solve the joint optimization problem of routing and packet scheduling. In addition, we design a greedy algorithm to

approximate optimal results in nearly real-time.

II. System Description

We propose our architecture of O-RAN in Fig. 3. In our proposed system, we group all O-DUs bound with the same O-RU as an O-DU instance. All the O-DU pools are running on the O-DU Pool, a computing resource. The physical resource of O-DU pool can be very flexible to deploy comprehensive 5G computing resources, including Cloud core network, edge computing resource or even on the access computing resource, etc. The location of the O-DU depends on the service requirement. For example, an O-DU may be established on the access region on the site may support the low latency service requirement, while an O-DU in edge resources is a candidate location to support the superfast broadband service. As per O-RAN specification from the technical specification, the O-DU shall support Ethernet II as Layer 2, including the tagging aspects , untagged Ethernet, single tagged (802.1Q) Ethernet, and dual-tagged (802.1ad) Ethernet [7]. The Ethernet for the O-DU is mandatory, but the choice of tagging leaves for operators to decide. [7]. O-DU shall also support IPv4 as Layer 3 that is mandatory in the O-DU while using IPv4 [7]. When IPv4 is used, Cooperative Transport Interface (CTI) and data messages are encapsulated in UDP packets [7]. CTI is an interface in O-RAN to support the resource allocation in the transport networks to transfer User Plane (U-Plane) and Control Plane (C-Plane) traffics between O-DU and O-RU [7]. C-Plane and U-Plane are protocols used for transferring control signals and user data respectively.

In our proposed architecture, we take the single tag VLAN to design our transport network on IPv4 over Ethernet to setup Point to Point (P2P) connections between O-DUs in an O-DU pool and an O-RU. Fig. 3 has also shown an example to integrate services from two vendors into our proposed O-RAN architecture design, e.g. the vendor1 marked in pink color has an O-DU pool 1 and O-RU 1, while vendor2 which marked in purple color has O-DU pool 2 and O-RU 2. We articulated two problems to enable the coexistence of multiple O-DU pools to transfer packets to O-RUs in the pool. The O-DU pool 1 has reserved three VLANs, and each VLAN that represents a point-to-point connection from the O-RU to the RAT slicing via an O-RU has a different VLAN ID, physical distance, and bandwidth to transfer their packets, while O-DU pool 2 reserved two VLANs as shown in the Fig. 3. O-DU pools can manage the O-DUs inside to support different RAT slicing for the O-RU by selecting VLANs to forward packets. We assumed O-RUs have their installation in fixed locations. The accumulated delays in each O-DU pool vary because the number of VLAN is the limited resource in the pool. Therefore, the routing problem is arranging the number of VLAN, bandwidth, and physical distance optimally to increase the number of O-DU pools.

Fig. 4 explains the second problem we proposed about the scheduling mechanism and the delay model for the UDP packet sent from O-DU pool to the O-RU over the IPv4 over the Ethernet network. The O-DU pool at the fronthaul interface

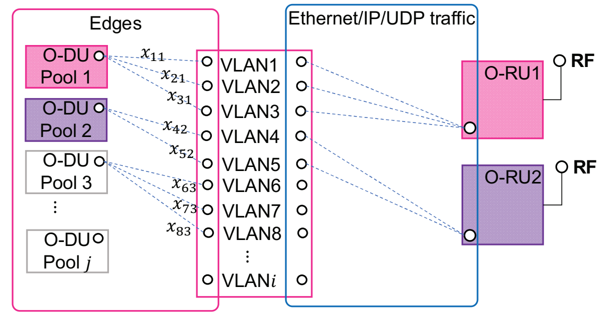

Fig. 3. Connection map between O-DU pools with VLANs in the O-DU pool.

uses eCPRI messages to transfer C-Plane control and U-Plane data messages to O-RU in sequentially accumulated delays. We assume at the n slot, and $N-1$ modulated OFDM symbols. Those symbols require the same amount of C-Plane control messages to plot over the phase array antenna. Furthermore, we only consider an O-RU dedicated to serving for a single O-DU pool without retransmission.

After a network transmission delay, all control messages arrived at O-RU fronthaul interface. In the C-Plane, O-RU needs a delay of Tadv_cp_dl in the downlink to translate the control messages' parameters. Then O-DU pool initiates its first OFDM symbol ready to send over the U-Plane. U-Plane uses IP over the Ethernet network to transfer data, while the OFDM modulated user data is carried over the UDP protocol encapsulated in the UDP/IP protocol. As we have articulated the problem in section II, the UDP size limitation may not contain OFDM symbols' full size. So, O-DU instant at the fronthaul interface may fragment the OFDM symbol into multiple small packet sizes to fit the UDP payload. The example is shown in the Fig. 4, an O-DU pool fragmented the first OFDM symbol into four pieces to load into four UDP payloads, which requires four to be transferred over the Ethernet network, while the second OFDM symbol requires two IP packets. The problem is finding out the load rate to fill up bits from an OFDM symbol into the UDP payload to control the UDP packets' number.

Two problems have to be addressed, namely the routing problem and the packet scheduling problem. The routing problem is associating the optimal routes from an O-DU pool to a VLAN link dedicated by this VLAN ID and bandwidth, and the UDP packet scheduling problem is to minimize system response time among O-DU pools in the O-DU pool [8]. We solved the problems to determine the globally optimal set of the combination among routing, packet sizes and bandwidth to assign to the O-DU pool. We also apply the greedy algorithm to approximate the routing, packet sizes, and bandwidth by an optimal local value implemented by the first candidate fit principle because of the solution's complexity.

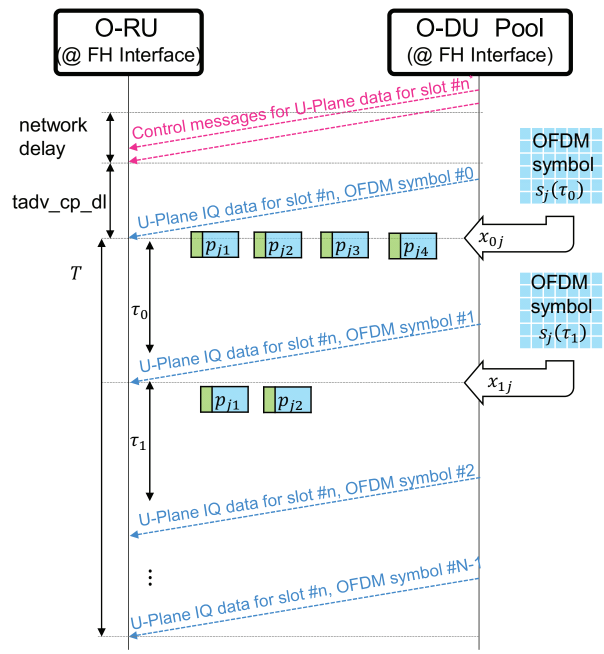

Fig. 4. The UDP packet mapping in the delay between O-DU pool and O-RU.

TABLE I
LIST OF SYMBOLS

x_{ij}	is the decision variable which denotes the connection between O-DU and VLAN of links.
B_j	denotes the total bandwidth of the j O-DU pool.
T_j	is the time required to transmit $s_j(\tau)$ OFDM symbols for the j O-DU pool.
K_j	is the maximum number of VLANs that j O-DU pool can associate.
$B(\cdot)$	denotes the bandwidth assigned to links associated to the j O-DU pool.
$F(\cdot)$	is the loading rate from OFDM symbols to UDP packets.
$s_j(\tau)$	is the symbol size of the j O-DU pool at slot τ.
p_{jk}	is the packet size of k packet generated by the j O-DU pool.
G_j^*	is the minimum response time in j O-DU pool.
W_j	is a weighting parameter for the j th O-DU pool.
α	is a scale parameter for the heuristic function.
n	denotes the total number of O-DU pools in the system.

III. Problem Formulation

A. O-DU Pool Routing Model

Fig. 3 presents a mapping for the connections between O-DU pools and VLANs created in the O-DU pool. There is i VLANs and j O-DU pools. We define x_{ij}, which is a decision variable, indicates that the O-DU pool 1 associates its links of VLANs when $x_{ij} = 1$; otherwise when $x_{ij} = 0$ [9].

$$x_{ij} = \begin{cases} 1 & , connection \\ 0 & , otherwise \end{cases} \tag{1}$$

For example, in Fig. 3, VLAN1, VLAN2 and VLAN3 associated with the O-DU pool 1, and we denote them x_{11}, x_{21}, and x_{31} respectively.

We assume that A VLAN can only be associated to an O-DU pool, such as

$$\sum_{\forall j} x_{ij} = 1 \tag{2}$$

Furthermore, the connection constraint that an O-DU pool has to be associated with at least one VLAN to route the traffic and less than K_j, the maximum number of VLANs for the j O-DU pool.

$$1 \leq \sum_{\forall i} x_{ij} \leq K_j \tag{3}$$

B. O-DU and O-RU Packet Scheduling Model

We formulate the problem by a set of fragmented UDP packets. The k th packet $p_{jk} \in \mathbb{Z}^+$ of the j O-DU pool such that $0 \leq p_{jk} \leq 1550 Bytes$, which is fragmented from an OFDM symbol $s_j(\tau)$ during the slot τ of the j O-DU pool.

$$\sum_k p_{jk} \geq s_j(\tau),\ \forall j \tag{4}$$

$F(\cdot)$ is a function to map the bandwidth of a dedicated link x_{ij}.

$$F(x_{ij}) = \frac{x_{ij} \sum_k p_{jk}}{\tau} \tag{5}$$

By $M/M/1$ queueing theory, the expected waiting time for 14 slots of τ of OFDM symbols should be less than T.

$$\frac{1}{B_j} \sum_i \sum_j \frac{F(x_{ij})}{B(x_{ij}) - F(x_{ij})} \leq T \tag{6}$$

C. Objective Functions

1) eMBB mapping problem: eMBB service demands super-fast bandwidth and the excellent quality for the application. The objective function is to maximize the total bandwidth.

$$\begin{aligned} \text{Max.} \quad & \sum_{i,j} B(x_{ij}) \\ \text{S.t.} \quad & (2)(3)(6) \end{aligned} \tag{7}$$

2) mMTC mapping problem: The mMTC requires a very large number of connections with limited resources in the fronthaul network. The objective is to maximize the number of connections.

$$\begin{aligned} \text{Max.} \quad & \sum_{i,j} x_{ij} \\ \text{S.t.} \quad & (2)(3) \end{aligned} \tag{8}$$

3) URLLC service mapping problem: URLLC service in the RAT slicing requires the minimal latency. Therefore, the objective function is to minimize the delay of each O-DU pools.

$$\begin{aligned} \underset{x_{ij}}{\text{Min.}} \quad & \frac{x_{ij} \sum_k p_{jk} - s_j}{B(x_{ij})} \\ \text{S.t.} \quad & (2)(3)(4)(5)(6) \end{aligned} \tag{9}$$

IV. Algorithmic Solutions

To solve three optimization problems (7) , (8), and (9), we design two algorithms: a greedy and a DP algorithm.

Algorithm 1 Greedy Algorithm

Input data: Number of O-DU pools
Output data: Minimum Response Time
Function `generateMatrix()`**:**
    Matrix = oduInstNum $\bigotimes$ vlanBWSet $\bigotimes$ udpSizeSet $\bigotimes$ distance
    return Matrix
Function `checkConstrains(`$Matrix$`)`**:**
    for *each num in Matrix.oduInstNum* **do**
        for *each vlanBW and updSize and distance in Matrix* **do**
            Check constrains
            Fragment the OFDM symbol
            Mapping with UDP payload
            Mapping with VLAN bandwidth
            for *Select VLAN* **do**
                Objective function to calculate delay time
            end
        end
        return *Response time matrix*
    end
Function `Main`**:**
    `/* First fit greedy method */`
    Matrix = `generateMatrix();`
    resultMatrix = `checkConstrains(`*Matrix*`);`
    for *item in resultMatrix* **do**
        if *item.currentRespTime* $\leq$ *item.previousRespTime* **then**
            return minimal first-fit response time
            break;
        end
    end
return

Algorithm 2 Dynamic Programming Algorithm

Input data: Bandwidth, Connection, Packet size
Output data: Minimum Response Time
1. Initializing $\alpha \in \mathbb{R}$ s.t. $\alpha \in (0, 1)$.
2. Iterating the recursive function until it converges,

$$V_{n+1} = V_n + \alpha(G_n^* - V_n)$$

where

$$V_n \equiv \frac{\sum_{j=1}^{n-1} W_j G_j^*}{\sum_{j=1}^{n-1} W_j} \text{ for } \forall n \leq 2, \text{ and } W_j \in (0,1)$$

and assign to G_j^* by searching the optimal value of the objective function for each O-DU pool j for URLLC service.

$$G_j^* \equiv \frac{x_{ij}^* \sum_k p_{jk}^* - s_j}{B^*(x_{ij})}$$

Algorithm 1 is a greedy method to search in a four-dimensional matrix data structure generated in generateMatrix function. We use the tensor product of vectors in each axis of the matrix from VLANs, bandwidth, UDP packet size to distance. As the number of O-DU pools increases, the matrix becomes more complicated in combinations. The checkConstrains function has a for-loop to iterate the O-DU pools among all edge resources. By checking constraints, we iterate every item in the matrix. If all constraints are satisfied, we batch the data and use the objective function to calculate the response time based on the URLLC service. It is a first-fit greedy algorithm. The algorithm searches the data set that we batched from the checkContrains function. It starts from the beginning of the dataset to compare the current minimum response time and the previous minimum response time. We design the greedy algorithm in the main function to find the approximately minimum delay. The greedy algorithm stops until it finds the first-fit minimum value from the data set.

Algorithm 2 is a DP method that iterates a recursive Bellman equation until it converges. α is a scaling parameter to control the step of the heuristic function $G_n^* - V_n$. We define G_j^* as the minimum response time by searching the optimal connections x_{ij}^*, packet sizes p_{jk}^* and bandwidths $B^*(x_{ij})$ which are assigned to each of O-DUs inside the O-DU pool j. the j th O-DU pool . V_n is a normalized weighted sum of the minimum response time G_j^* from 1 to the $n-1$, where W_j is a weighting parameter of a real number between open interval of 0 and 1.

V. Numerical Results

A. Settings

In this simulation, we simulated the O-DU pool to apply 8 O-DU pools. Each O-DU pool contains at least one or more VNF of O-DU, requiring a VLAN to route the UDP packet over a distance and bandwidth in the requirement from 5G RAT slicing. So it is a challenging problem to simulate because we need to construct a four-dimensional matrix of distance, VLANs with different bandwidths associated with each O-DU pool, UDP packet size and O-DU pools. For simplicity, we set that Maximum VLAN numbers in the O-DU pool are less and equal to 50. One O-DU pool can associate a maximum of 3 VLANs that are assigned with different bandwidths from a set of 50 Mbps, 100 Mbps and 150 Mbps. The distance is less and equal to the minimum required distance. The total bandwidth upper limit in the O-DU pool is less and equal to 1 Gbps physical port. An OFDM symbol can be fragmented into a maximum number of 12 packets. The O-DU pool has four loading rates for the UDP payload of 12400 bits per millisecond, 11000, 10000, and 9000 bits per millisecond.

To validate Algorithm 1, we set up a single O-DU pool associated with three VLANs of the assigned bandwidth 50 Mbps, 100 Mbps, and 150 Mbps at distances 50m, 1000 m 5000 m. O-DU pool allows the O-DUs to classify their fragmented packets over three VLAN to forward the UDP packets. The VLAN classification principle has the mechanism to map the number of fragmented packets from the OFDM

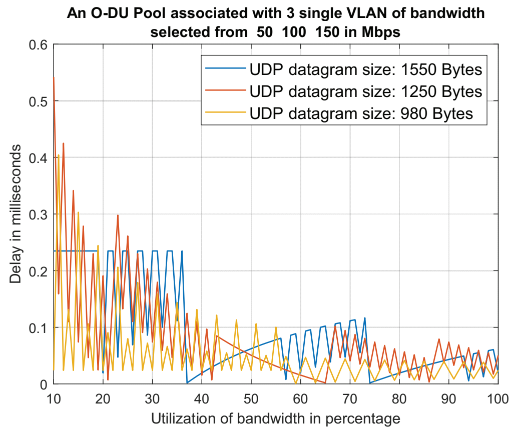

Fig. 5. An O-RU instance generates UDP size of 1550 Bytes, 1375 Bytes and 1250 Bytes over the selective three VLANs assigned with 50 100 or 150 Mbps.

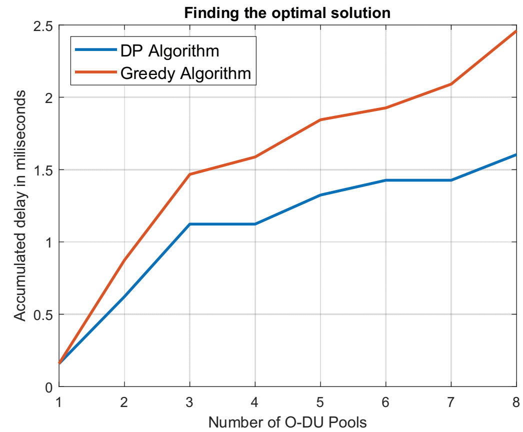

Fig. 6. Comparing the approximation of the greedy algorithm with the optimal values of DP algorithm in the accumulative response time regarding the incremental number of O-DU pools.

symbol with the UDP loading rate that controls the UDP payload size for τ. We set up the $\tau = 1ms$. The default UDP payload size is 1550 Bytes, which is 12400 Bits. If a VLAN route of 50 Mbps is selected, the O-DU pool may transmit maximal four packets. The bandwidth 100 Mbps may allow a maximum of eight packets, while 150 Mbps for 12 packets. For example, during the τ period, the O-DU pool loads the UDP packet with a maximum rate of 12400 Bits per millisecond. For three packets (e.g. 7800 Bits, 9800 Bits, 11520 Bits) during the τ, O-DU pool selects the 50 Mbps route of VLAN to forward. The delay time for the O-DU pool is 0.4 milliseconds. The simulation of the VLAN classification mechanism in Fig. 5 for three O-DU dwelling in the single O-DU pool at 3 UDP loading rate can fill 1550 Bytes of payload, 1250 Byte and 980 Bytes of payload size. Fig. 5 interprets as the increase of the utilization of the VLAN bandwidth, the delay varies from three UDP payload sizes.

B. Results

The blue line in the Fig. 6 is drawn by the numerical result that finds out the optimal global value by DP algorithm in the matrix of O-DU pools, UDP packet size, and bandwidth for the associated VLAN. By comparing the solution's baseline, we implemented a greedy algorithm that approximates the optimal value by the first fit principle. From the pattern we learned from the single O-DU pool, we can iterate the four-dimensional matrix in sequential order, the greedy algorithm stops until it reaches its first minimum value of accumulated delay among 8 O-DU pools. The red line is drawn by the numerical result from the greedy algorithm. It interprets that the greedy algorithm would well approximate the solution in the lightweight and fast computing resources.

As the interpretation from Fig. 6, the average difference of between DP algorithm and greedy algorithm is 1.1 ms, as O-DU pools increased from 1 to 8. The system has the minimum 0.1586 ms difference in the accumulated response time when single O-DU pool is launched. And the error of the approximation from greedy method to DP method is increasing and up to 1.604 ms, while 8 O-DU pools are launched to operate in the coexisting manner by sharing the VLANs, bandwidths in the system. However, compared with the complexity $\mathcal{O}(n^2 \log n)$ of DP method, our greedy algorithm has complexity of $\mathcal{O}(\log n)$ in the first-fit principle. Since the complexity of searching reduced by n^2 time, greedy algorithm is a good approximation method for our system.

VI. Conclusion

This paper has presented problems to optimize routing and scheduling between O-DU pools and O-RUs to meet 5G requirements of minimal costs. Through simulation results of the multiple O-DU pools coexisting in the O-DU pool, we use the greedy algorithm to compare the difference between the optimal value and to approximate optimal value achieved by DP algorithm. In the future, we want to implement Reinforcement Learning algorithms for our system and extend our applications to involve 5G beamforming.

References

[1] 3GPP, " 5G; NG-RAN; Architecture description," 3GPP TS 38.401 version 16.2.0 Release 16.

[2] O-RAN ALLIANCE, "O-RAN Architecture Description," Specification Version v1.0 - February 2020.

[3] 3GPP, " Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Release description," 3GPP TR 21.915 version 15.0.0 Release 15

[4] S.K. Singh et al., "The Evolution of Radio Access Network Towards Open-RAN: Challenges and Opportunities," IEEE WCNCW, June 2020.

[5] P.K. Taksande et al., "Open5G: A Software-Defined Networking Protocol for 5G Multi-RAT Wireless Networks," IEEE WCNCW, June 2020.

[6] A. Umesh et al "Overview of ORAN Fronthaul Specification," NTT Docomo Technical Journal, Vol.20, No.1, Jul. 2019.

[7] O-RAN ALLIANCE, "O-RAN Fronthaul Working Group 4 Cooperative Transport Interface Transport Control Plane Specification," Specification Version 1.0 - April 2020.

[8] A. Atallah et al., "Routing and Scheduling of Time-Triggered Traffic in Time-Sensitive Networks," IEEE Transactions on Industrial Informatics, vol.16, no. 7, pp.4525-4534, 2020.

[9] J. Wu et al., "Adaptive Flow Assignment and Packet Scheduling for Delay-Constrained Traffic Over Heterogeneous Wireless Networks," IEEE Transactions on Vehicular Technology, vol. 65, No. 10, pp. 8781-8787, 2016.