Mobile User Plane Evolution

Distributed/Lite/Integrated User Plane Function (UPF)

ChiNOG10, 10/6/2022

Jeffrey Zhang, Juniper Networks

Agenda

- Mobile User Plane Background
- RAN Centralization and CN Distribution
- Alternative User Plane Implementation/Deployment
- One Step Further in 6G
- Summary

MCN & User Plane

- A Mobile Communication Network (MCN) includes RAN and CN
  - Radio Access Network (RAN) and Core Network (CN)
  - Connected over a transport network
- Planes in an MCN
  - User Plane: data plane that carries mobile user traffic
    - Spans from User Equipment (UE, or mobile device) to a User Plane Function (UPF) in CN
  - Control Plane: to set up resources needed for mobile communication
  - Management Plane & Synchronization Plane

RAN & Centralization Trend

- RAN: A network of radio access components that terminate the air interface from UEs
  - NodeB → eNB → gNB
- D-RAN: Distributed and self-sufficient standalone NodeB/eNB
- C-RAN: Distributed gNB-DUs separated from centralized gNB-CUs
  - DU: Distributed Unit
  - CU: Centralized Unit
- D-RAN → C-RAN

CN & Distribution Trend

- CN: the brain of an MCN; to enable and implement mobile services
  - With control plane and user plane components
  - Not a topological concept (anymore)
- CUPS: **C**ontrol-plane **U**ser **P**lane **S**eparation
  - Allows control plane and user plane components to scale independently
  - Allows the components to deploy independently
    - Centralized control plane components (SMF/AMF and others in 5G)
    - Distributed user plane components (UPF in 5G)
  - A CN concept that has been extended to RAN as well
    - gNB-CU → gNB-CU-CP + gNB-CU-UP
- UPFs are being distributed close to gNB-CU
  - For MEC or optimized local traffic (between UEs or to/from local DN or Internet peering)
    - MEC: Multi-access Edge Computing
  - Distributed UPFs meet centralized gNB-CU
    - Collocated in Edge DC (or even Far Edge DC) – maybe running on the same server

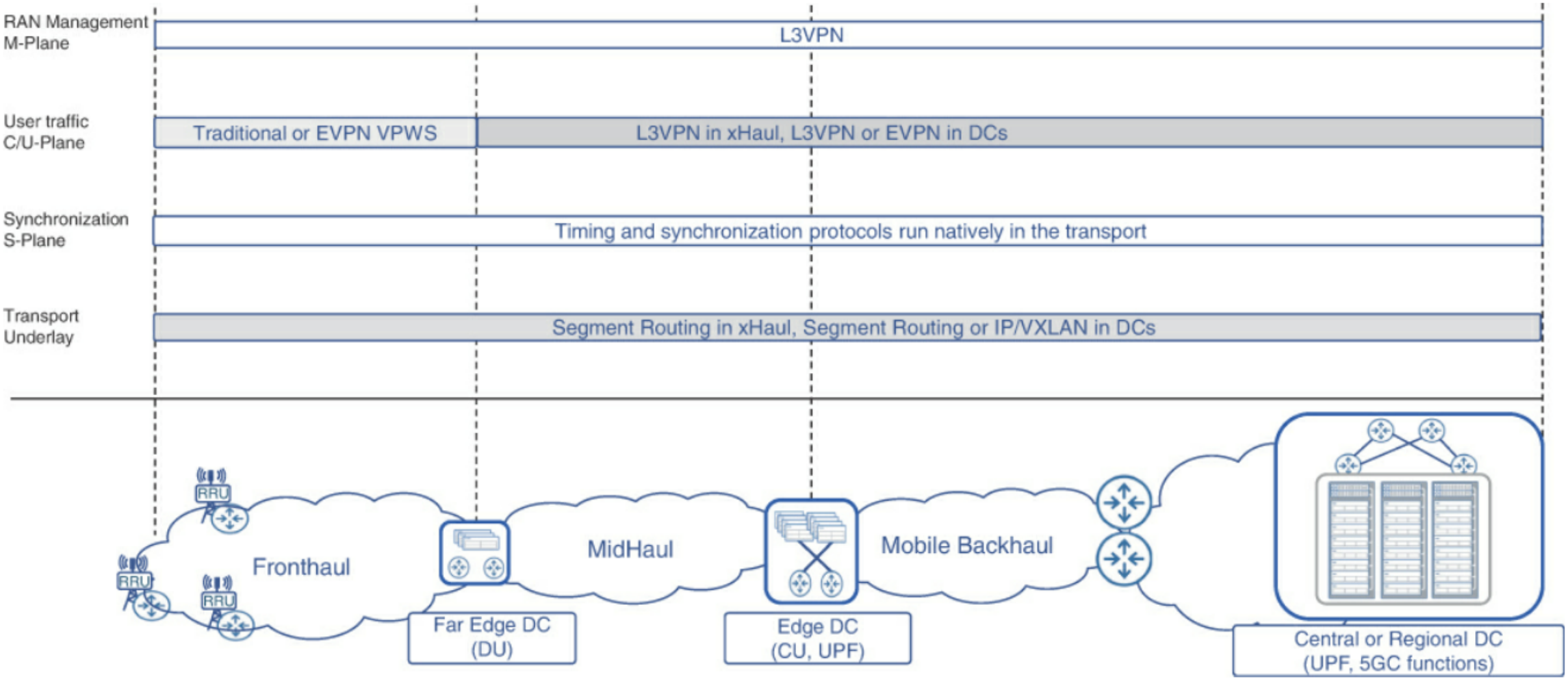


Figure 8-14 from the book “Network Architect’s Guide to 5G” (ISBN-13: 978-0-13-737684-1)

Looks Familiar …

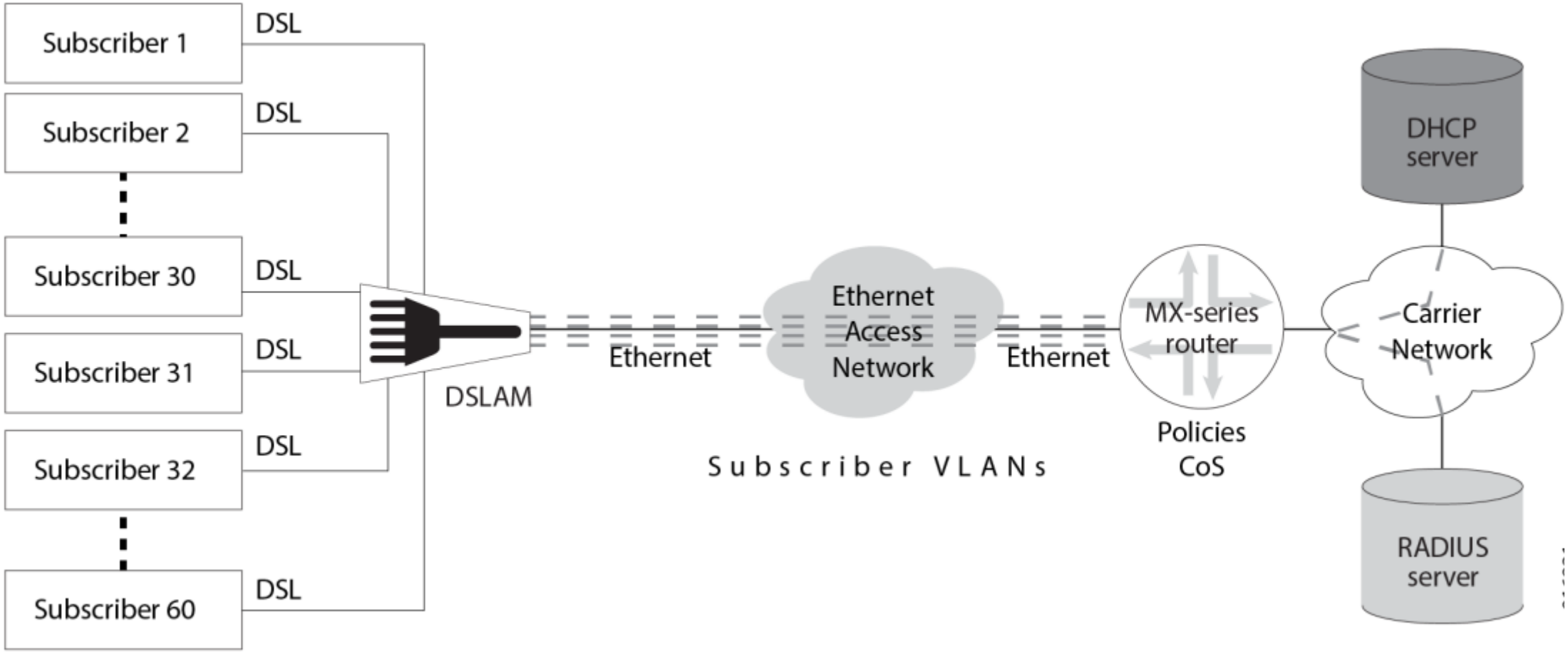


Figure 2 of Juniper's "UNDERSTANDING SUBSCRIBER MANAGEMENT AND BNG"

And This …

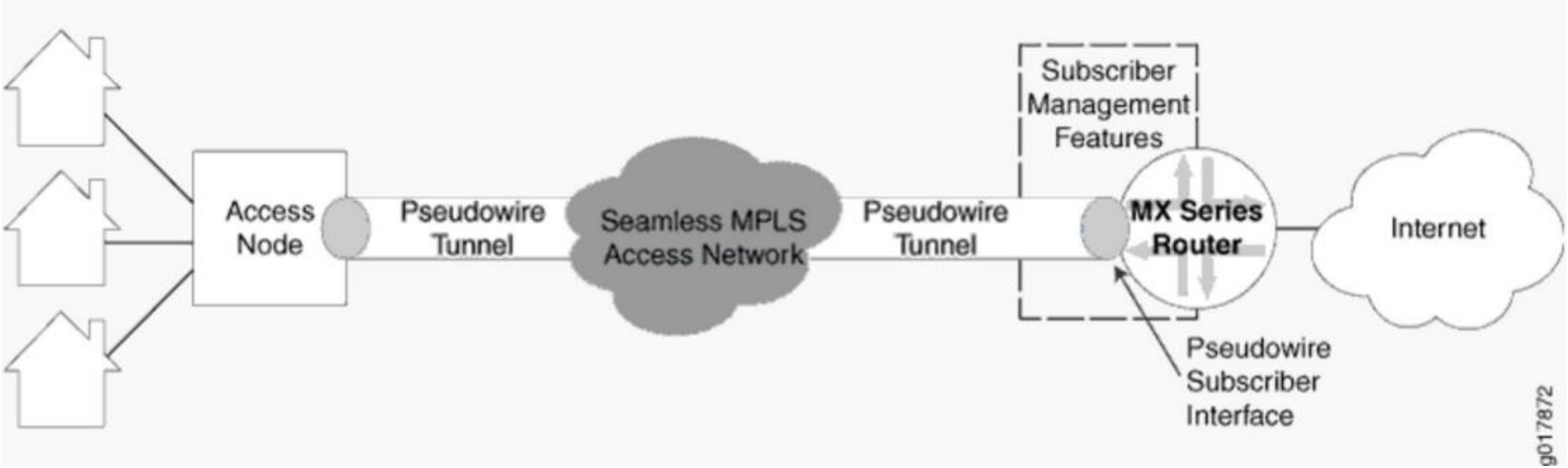
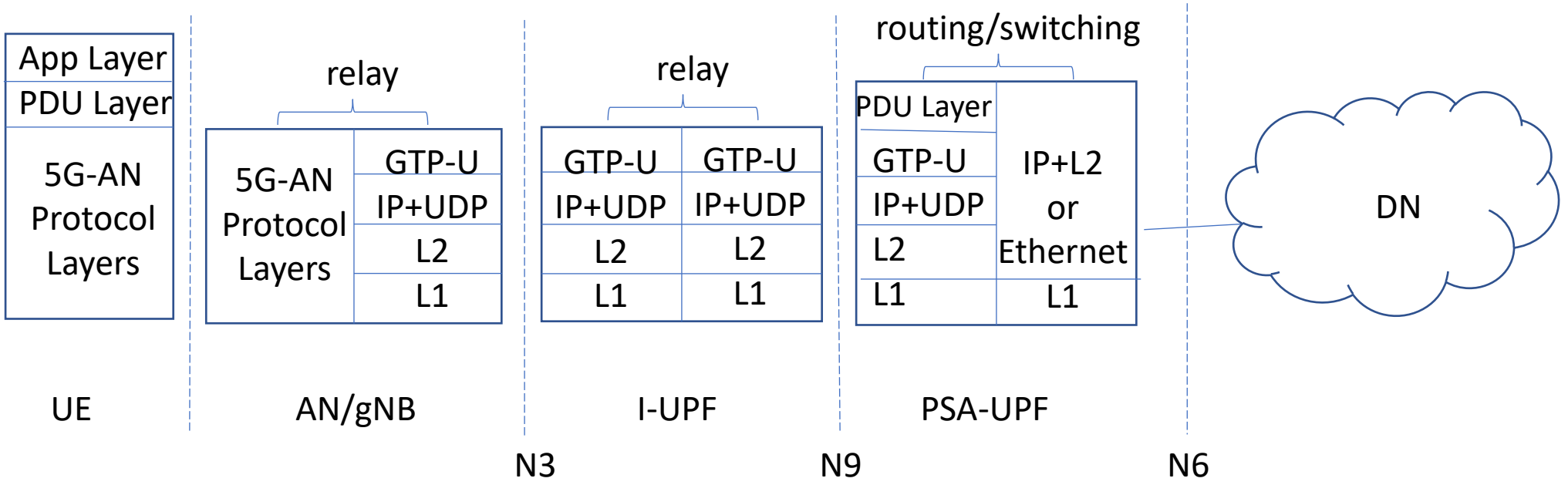


Figure 4 of Juniper's "UNDERSTANDING SUBSCRIBER MANAGEMENT AND BNG"

- eNB/gNB are like the DSLAM or Access Node in the above pictures
- UPF are like the BNG (Broadband Network Gateway)

User Plane Components

- UPF: the CN component of the User Plane
  - Like a BNG router; route/switch to/from Data Network (DN) via N6 interface
  - Traditional UPFs are implemented with PDR/FAR rules from N4 signaling
    - PDR: Packet Description Rule FAR: Forwarding Action Rule
    - Functionality-wise they route/switch traffic based on IP/Ethernet header
- gNB: the RAN component of the 5G User Plane
  - Terminate air interface
  - Relay IP/Ethernet traffic between UE and UPF
    - Over GTP-U tunnels (N3 interface/tunnel)
    - Each tunnel is for a session; identified by <gNB, gNB-TEID, UPF, UPF-TEID>
      - TEID: Tunnel endpoint ID (like an MPLS label)
- N3 tunneling is over an IP transport that is itself an IPVPN over transport infrastructure



PDU Session: a 5G UE-UPF user plane connectivity for a specific DN

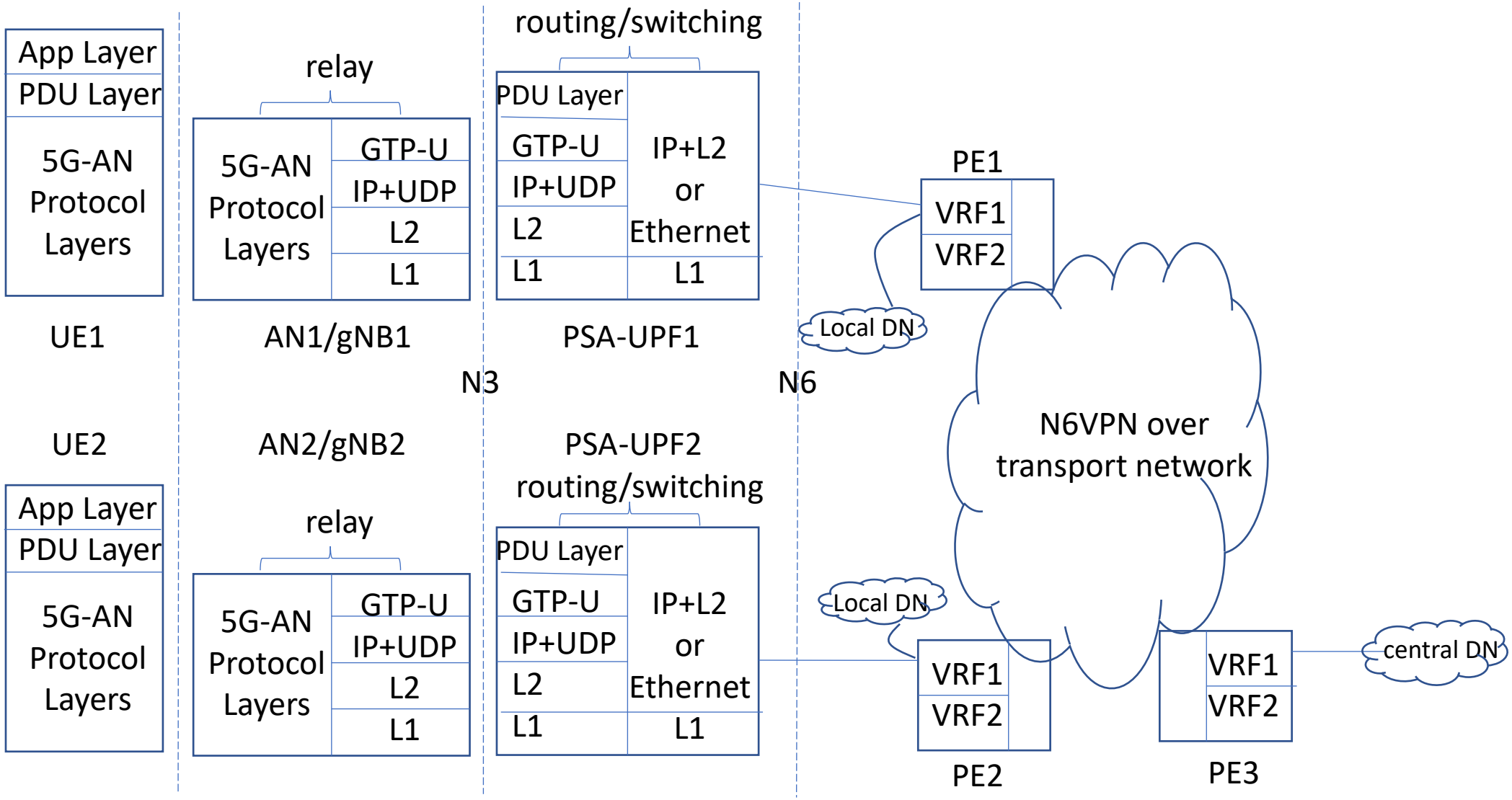
PSA-UPF: PDU Session Anchor UPF

I-UPF: Intermediate UPF

Recap: 5G User Plane Background

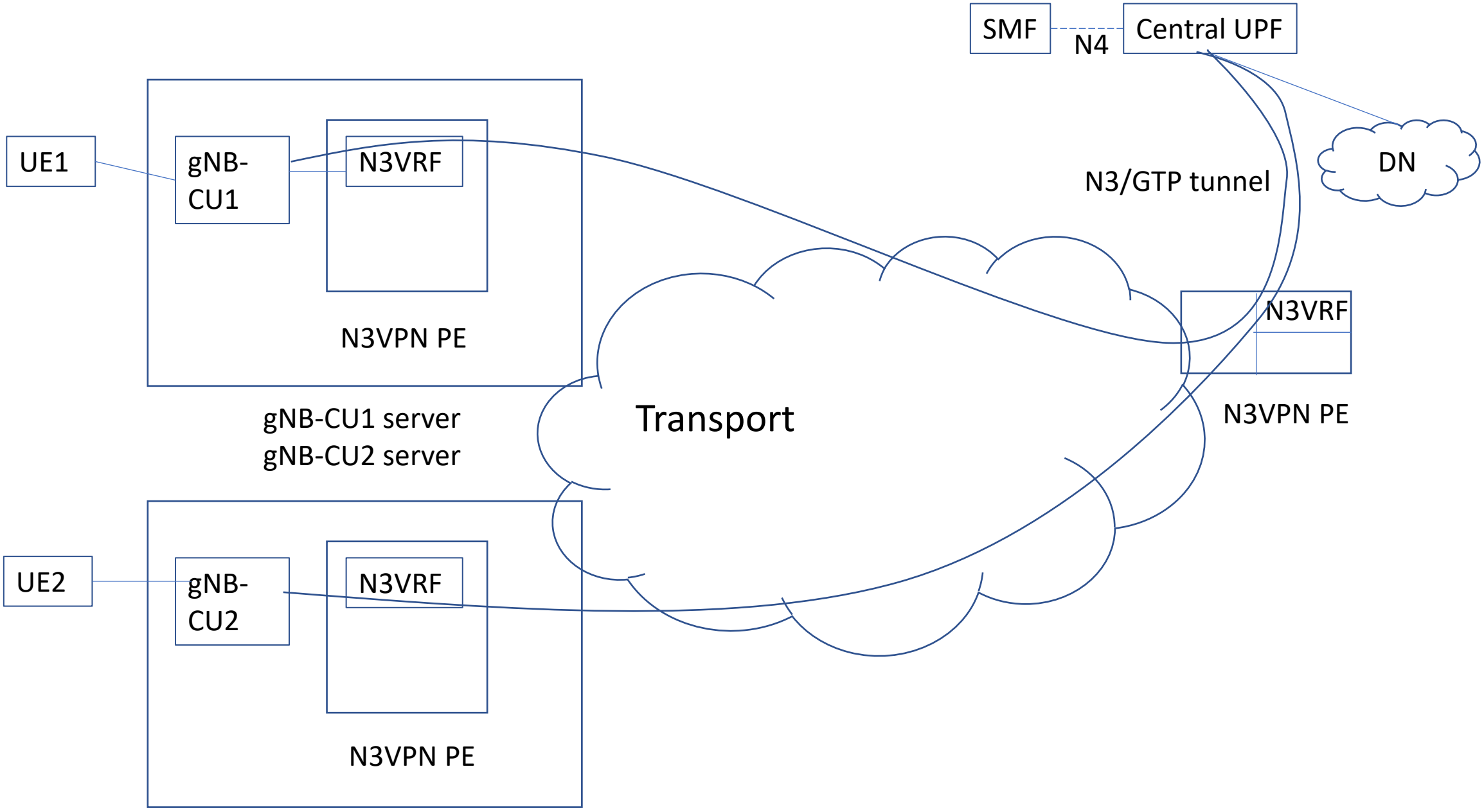
- Traditionally, UPFs are centrally deployed
  - This has the advantage of session continuity when UEs move around
  - They also scale up with dedicated hardware platform
- For MEC, UPFs are distributed close to gNB-CU (an Access Node or AN)
  - This means distributed DNs as well (implemented as VPN - N6VPN)
- For C-RAN, gNBs are being centralized
  - Leads to co-located UPF and AN in Edge DC
    - With direct/short link in between or even running on the same server
- N3 tunnel between ANs and UPF are through an IP transport network
  - That is a VPN (N3VPN) over a converged transport infrastructure
  - N3VPN PEs are close to ANs/UPFs (which are N3VPN CEs)

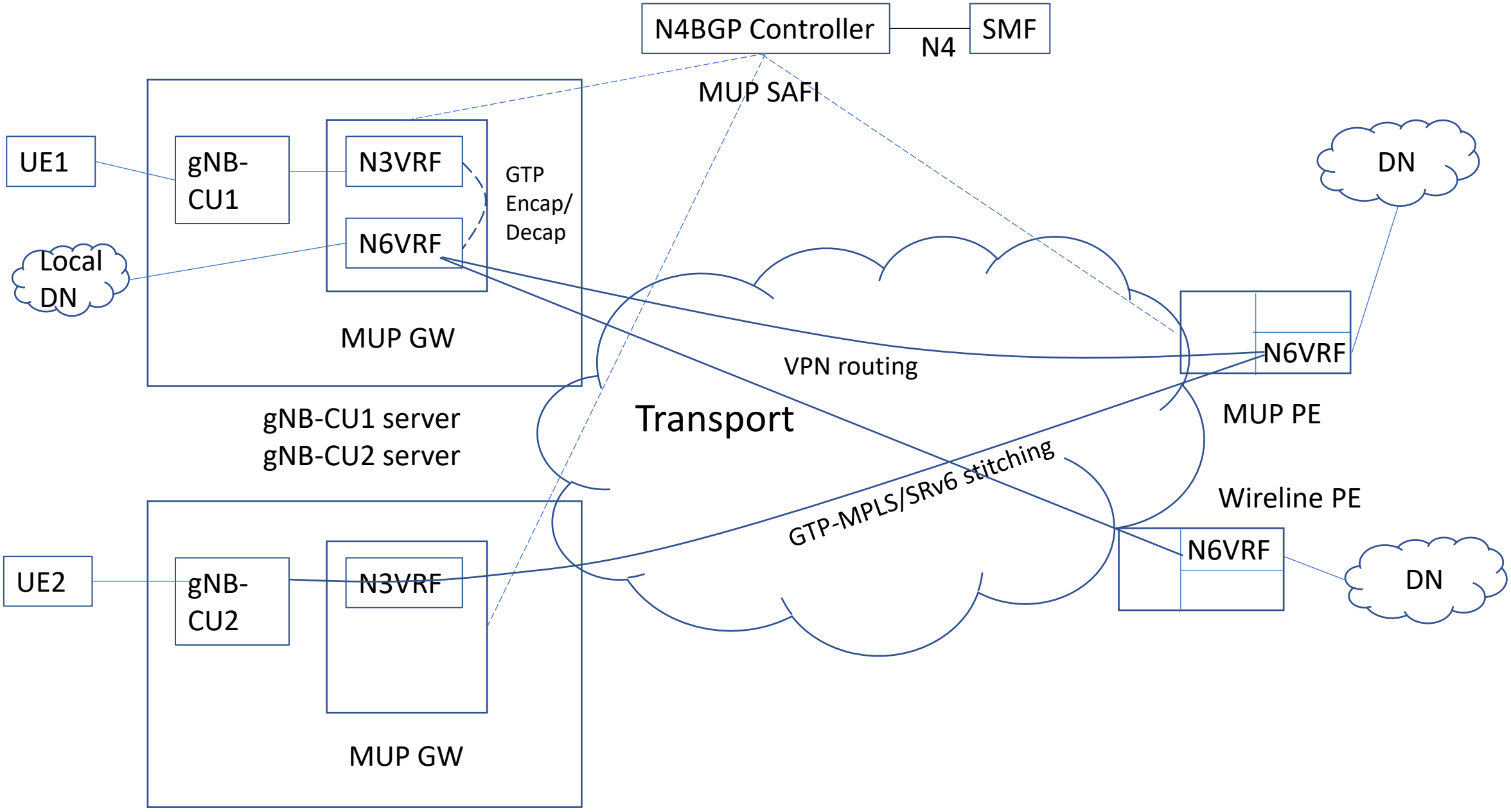
Distributed UPF and DN



Alternative Implementation/Deployment

- Implement distributed UPF as a router/switch
  - Based on N4 or BGP signaling translated from N4
  - Trim off some functionalities that are not needed
    - E.g., billing, LI for some deployment scenarios
    - Referred to as UPF-Lite
- Integrate N3VPN/N6VPN PE function into it
  - Referred to as MUP GW in SRv6 MUP Architecture
    - An SRv6 specific, router based, and partial implementation of Distributed UPF
    - The concept is actually not SRv6 specific


SMF
N4
Central UPF
DN
N3/GTP tunnel
UE1
gNB-
CU1
N3VRF
N3VPN PE
gNB-CU1 server
gNB-CU2 server
Transport
N3VRF
N3VPN PE
UE2
gNB-
CU2
N3VRF
N3VPN PE


N4BGP Controller
N4
SMF
MUP SAFI
UE1
gNB-CU1
N3VRF
N6VRF
GTP Encap/ Decap
Local DN
MUP GW
gNB-CU1 server
gNB-CU2 server
DN
VPN routing
Transport
N6VRF
MUP PE
GTP-MPLS/SRv6 stitching
Wireline PE
N6VRF
DN
UE2
gNB-CU2
N3VRF
MUP GW

An Alternative View

- The collection of distributed <N4BGP Controller, MUP GWs, MUP PE> appears to the SMF/gNBs as a single/central UPF
  - No change of 3GPP architecture/signaling
  - An alternative to distributed “traditional” UPFs
    - Mobility handled by UE host routes
- This is actually SR-agnostic
  - Works equally well with (SR-)MPLS
- This is so far just for partial UPF functions
  - For complete set of UPF functions, either extend this UPF-lite or just deploy traditional but distributed UPFs

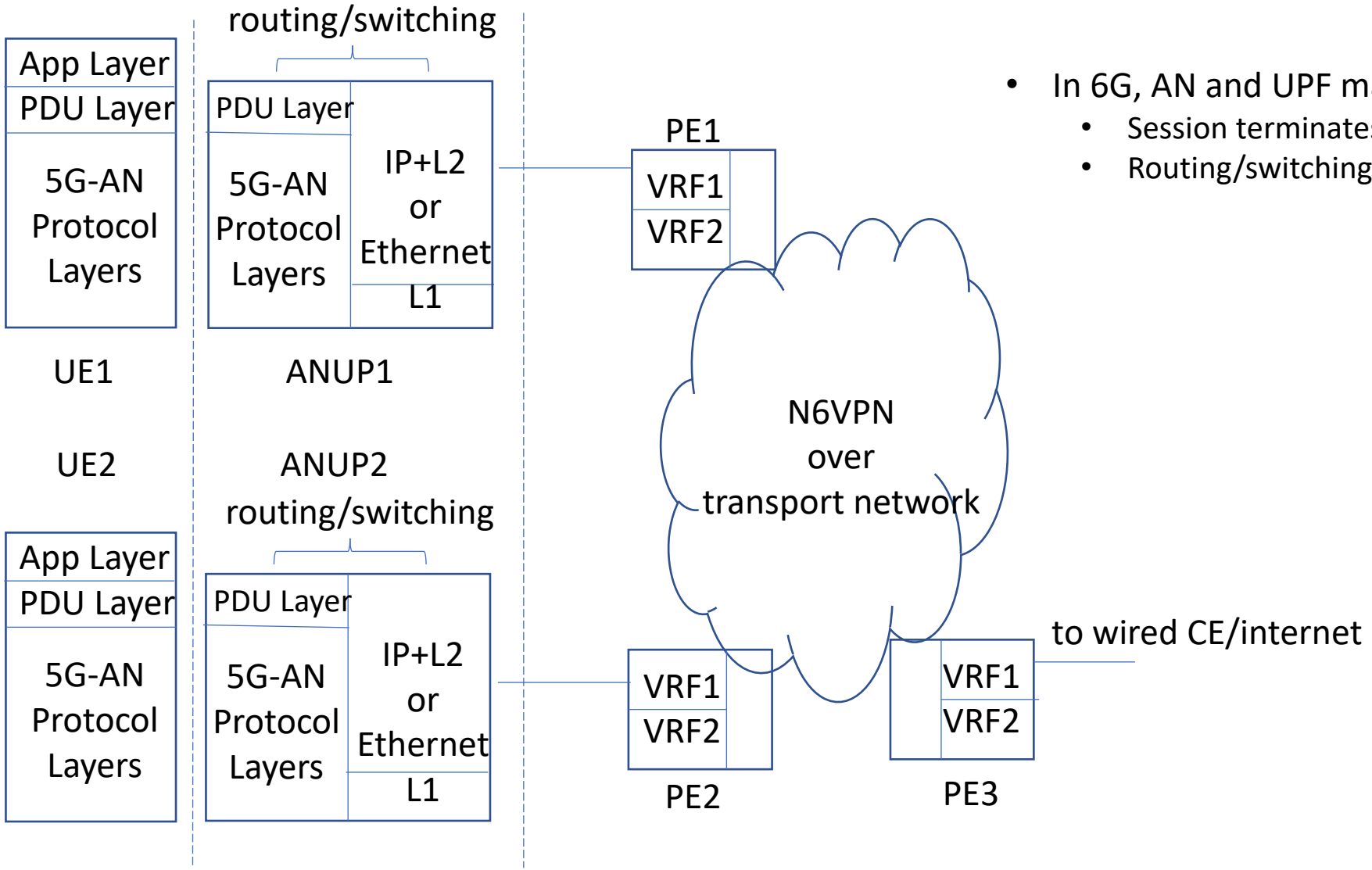
Motivation for UPF-lite

- Alternative to traditional hardware/vendor-tied central UPF
  - Distribution, scale-out
  - Disaggregation, virtualization
- Optimal traffic handling
  - Distributed local routing of UE-UE and UE-DN (Internet/DC) traffic
    - Reduced failure domain
  - Less overhead even when backhauling to a hub PE is needed
    - Previously: GTU-U header + N3VPN header
    - UPF-lite: N6VPN header
- Works for all scenarios but especially good for Fixed Wireless Access (FWA)
  - Not only for MEC
  - FWA dose not require mobility support and only prefix routes are needed
  - FWA does not require rich billing functionalities

One Step Further in 6G?

- What if, in 6G, AN and UPF are integrated into a single NF (ANUP)?
  - Optionally with N6VPN PE function built in
- A flattened, routing/switching-based architecture
  - ANUP is a router/switch with wireless/wired connections
    - Optionally with N6VPN PE functionality
  - 3GPP/wireless technologies responsible for wireless access
    - Mobility Management, UE authentication/authorization, …
    - Just like that IEEE technologies are for Ethernet connection
      - And Wi-Fi technologies for Wi-Fi connection to a Wi-Fi router
  - IETF/wireline technologies for the rest
- Consistent for wireline/wireless

Integrated AN/UP in 6G?



- In 6G, AN and UPF may be integrated into ANUP
  - Session terminates at ANUP
  - Routing/switching at ANUP

Not Really A Drastic Change

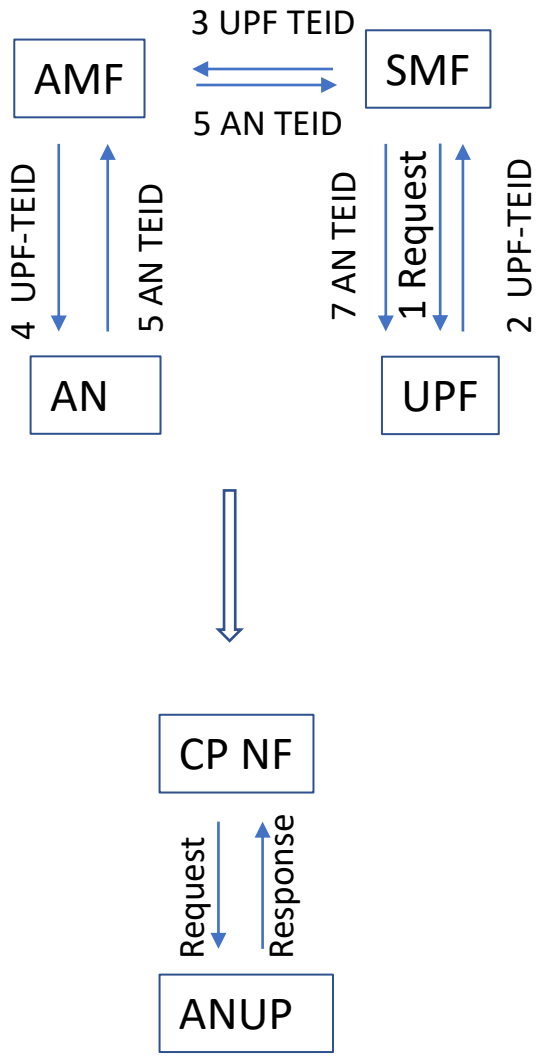
- Previously, ANs and UPF form a hub-and-spoke VPN
  - UPF is the hub PE and ANs are the spoke PEs
  - GTP-U tunnel corresponds to VPN tunnel (base tunnel label + VPN label)
  - No VRF on ANs because:
    - All UpLink (UL) traffic goes to hub so no IP lookup needed for UL traffic
    - “per-UE” TEID on AN side avoids IP lookup on AN for DownLink (DL) traffic
      - Just like “per-FEC” (vs. per-VRF) VPN label or option-B stitching
  - Of course, centralized N2/N4 (vs. distributed BGP) signaling is used
- We’re now just adding VRF on the ANs
  - For MEC purpose (or for optimal UE-UE traffic)
  - Of course, no longer restricted to GTP-U tunneling

Advantages

- Simplified, flattened architecture unified for wireline/wireless
  - Simplified signaling
  - Optimized data plane
- Many 5G special features/procedures are not needed anymore or can be greatly simplified
  - MEC
  - 5MBS
  - LAN-type services
  - …
- Integrate when desired/feasible, separate when you have to
  - There are still scenarios for separate AN and UPF

Simplified Signaling

- In 5G, N3 tunneling is used between separate AN and UPF
  - Even if they're co-located
  - Multi-step N2/N4/N11 signaling involved
- Since no tunnel is used with a router/switch ANUP:
  - Signaling only needs to tell ANUP which DN a PDU session belongs to
  - It's new signaling, but 6G will have a lot of changes anyway



Optimized Data Plane

- Direct/short/internal AN-UPF connection is removed
- GTP-U encap/decap is removed
  - Better throughput/performance
  - Reduced latency

Multicast

- ANUP is a router/switch
  - With wireless connections to UEs and wired connections to DNs
- Multicast DL traffic arrives on ANUP via whatever DN multicast means
  - Ingress Replication, PIM, BIER, P2MP, whatever
    - IETF/wireline technology
  - Then delivered to attached UEs via P2P/P2MP radio bearers
    - 3GPP/wireless technology
- Multicast UL traffic arrives on ANUP and then
  - Delivered to other ANUPs and DN routers via whatever DN multicast means
  - If needed, also delivered to locally attached UEs via P2P/P2MP bearers

MEC

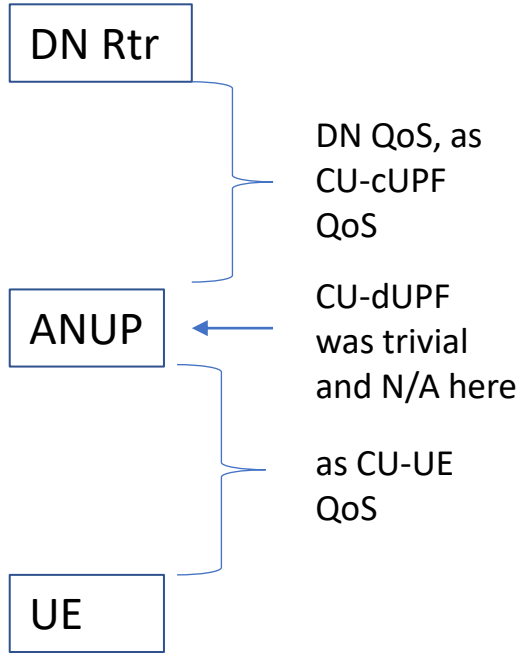
- ANUP is a router/switch
  - With wireless connections to UEs and wired connections to DNs
- No special 3GPP-specific MEC procedure needed at all
  - ANUP routes among DN and UEs directly
  - The DN is a local DN with local MEC resources
  - Application traffic routed to/from local/central resources transparently
    - Identifying/locating resources (via DNS or other means) is a generic function
    - It does not matter if it is wireline or wireless, MEC or not

LAN-type Services

- ANUP is a router/switch
  - With wireless connections to UEs and wired connections to DNs
- No special 3GPP-specific LAN-type service procedure needed at all
  - LAN-type services are IP/E-VPN reinvented anyway
  - ANUP routes among DN and UEs directly
    - The DN and relevant UEs of the same LAN-type service
  - When a LAN-type service span across multiple ANUPs:
    - The DN (most likely implemented as IP/E-VPN) connects them together, or,
    - ANUPs can have IP/E-VPN PE functionality built-in
- Seamless wireline/wireless integration
  - Including EVPN all-active multi-homing via wireline and wireless
  - LAN-type services are reinvented IP/E-VPN anyway

QoS

- The ANUP-UE QoS is still like CU-UE QoS
- The QoS previously between CU and co-located UPF is trivial anyway and with integrated ANUP it is N/A
- The QoS previously related to N3 tunneling (w/o co-location) is now QoS in DN
  - Previously, N3-related QoS is realized through the transport infrastructure
  - Now it is DN (VPN) QoS realized through the same transport infrastructure
- The QoS parameters signaled to ANUP will be used for:
  - QoS between ANUP and UE, and,
  - QoS between DN routers and ANUP
    - DN routing signaling could be enhanced – e.g., the UE routes advertised into DN could carry QoS information so that DL traffic will be subject to the desired QoS handling



Will 3GPP Accept Integrated ANUP?

- It seems a natural evolution
  - To people familiar/friendly with IETF/wireline technologies
  - But a big paradigm shift on 3GPP/wireless side
- But the work is to be done in 3GPP
- Trying to get support from mobile operators
  - Socializing the idea first among their IETF/wireline people on mobile side
    - https://datatracker.ietf.org/doc/draft-zzhang-dmm-mup-evolution/
    - Appreciate your feedback and support! zzhang@juniper.net
  - Exploring prototype implementation for demo
- Will bring to 3GPP if we get enough support
  - The work is on 3GPP not IETF side

Summary

- With 5G, traditional central UPFs can be transparently replaced with a collection of <N4BGP controller, MUP GW, MUP PE>
  - Removal of N3VPN; Integration of N6VPN PE function; local routing
- With 5G, gNBs are being centralized while UPFs are being distributed
  - They meet and co-locate in the edge or even far edge DCs
- In 6G, the co-located gNB/AN and UPF function may be integrated into a single function (ANUP)
  - A router/switch with wireless/wired connections
  - Simplified architecture, optimized signaling and data plane
  - Operator’s feedback/support appreciated!