Discussions on Integrating AN and UPF

draft-zzhang-dmm-mup-evolution-02

Jeffrey Zhang, Keyur Patel, Luis Contreras, Kashif Islam
IETF 115, London
Basic Idea

• In 5G, UPFs are more and more distributed close to gNB-CU (AN)
  • For MEC, private networks, or local Internet peering
  • Could be co-located – with direct link in between or even running on the same server
  • Distributed UPF means distributed DNs – implemented as VPNs
• In 6G, what if AN and UPF are integrated into a single NF (ANUP)?
  • A flattened, routing/switching-based architecture
    • ANUP is a router/switch with wireless/wired connections
      • Foundation of Internet
    • 3GPP/wireless technologies responsible for wireless access
      • Mobility Management, UE authentication/authorization, …
    • IETF/wireline technologies for the rest
UE IP/Ethernet traffic on top of radio stack
UE IP/Ethernet traffic on top of radio stack

UE IP/Ethernet traffic in GTP-U tunnel
UE IP/Ethernet traffic on top of radio stack

UE IP/Ethernet traffic in GTP-U tunnel

relay

Local DN

Transport

Central DN

Local DN

VRF1

VRF2

PE1

PE2

PE3

VRF1

VRF2
UE1

DU1

CU1

UPF1

UE IP/Ethernet traffic on top of radio stack relay UE IP/Ethernet traffic in GTP-U tunnel

Local DN

PE1

VRF1

VRF2

Local DN

Central DN

Transport

Local DN

PE2

VRF1

VRF2

PE3

VRF1

VRF2

UE2

DU2

CU2

UPF2
UE IP/Ethernet traffic on top of radio stack

UE IP/Ethernet traffic in GTP-U tunnel

Local DN

Central DN

Transport
UE IP/Ethernet traffic on top of radio stack

UE IP/Ethernet traffic in GTP-U tunnel

 relay

Transport

Central DN

Local DN

Local DN
Disclaimer

• The work needs to be done in 3GPP
• We’re discussing here only to first socialize the idea among IETF/wireline-friendly people
• Only if we get enough support among mobile operators may we bring it to 3GPP for further work
Advantages

• Simplified signaling and optimized data plane
• Simplified and flattened architecture unified for wireline/wireless
• Many 5G special features/procedures are not needed anymore or can be greatly simplified
  • MEC
  • 5MBS
  • LAN-type services
  • …
New Discussion Points

• Reduced signaling
• Can’t you use ULCL I-UPF to do local breakout already?
• What if I have to use NAT with overlapping private address spaces in different regions?
• Aren’t you increasing the burden on the AN?
• Aren’t you making the AN more complicated?
• Aren’t you going against the microservice architecture that breaks up functionalities?
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 needed for:
    • Initial setup
      • Periodical refreshes to maintain the state

• Since no tunnel is used with a router/switch ANUP:
  • No need to setup and maintain the state
  • Signaling only needs to tell ANUP which DN a PDU session belongs to
    • It’s a new signaling, but for 6G a lot of change will happen anyway
NAT with Overlapping Private Address Spaces

• An operator may have to use overlapping private address spaces for different regions

• With central UPF:
  • One UPF for each region with a NAT GW attached to each UPF
  • UE $\leftarrow\rightarrow$ Internet traffic goes through the NAT GW
  • Intra-region UE $\leftarrow\rightarrow$ UE traffic does not need NAT
  • Inter-region UE $\leftarrow\rightarrow$ UE traffic goes through two NAT GWs

• With distributed UPF/ANUP:
  • One hub DN router for each region with NAT GW attached to the hub router
  • The rest is the same as the central UPF case
  • No difference whether ANUP is used or not
What about ULCL I-UPF for MEC

• 5G ULCL I-UPF does allow local breakout for MEC purpose even with central UPF deployment
• It is a special and extra “user plane function”
  • Sits in the path of N3 tunnels and exams *all* N3 traffic
    • Intercepts local breakout traffic based on filtering rules signaled from SMF
  • AN puts on GTP-U encapsulation, then all traffic gets peeked through the encapsulation and some traffic gets decapsulated by a nearby ULCL I-UPF
    • A waste avoided by distributed UPF/ANUP
• Target scenario of ANUP is distributed UPF co-located with AN
  • The ANUP or the distributed (anchoring) UPF replaces the special ULCL-UPF
  • ANUP can be used together with central UPF
    • Some sessions tunneled to the central UPF
    • Some sessions terminated and traffic routed locally
• In other words, no need for the special ULCL I-UPF anymore
Increased Burden on the AN?

• All necessary UPF functionalities need to be done either on previously separate UPF or on the integrated ANUP
  • All in the same edge DC
  • Note that ANUP is to integrate *co-located* AN and UPF
• In case of cloud-native implementation, it does not matter if the load is on AN, UPF, or ANUP
  • It’s actually easier for capacity planning in case of ANUP
    • Previously one needs to decide how much resource to allocate to UPFs
      • Besides the compute resources (and PDU sessions) allocated to ANs
    • Now one only needs to decide how much compute resources and which sessions are assigned to which ANUPs
Increased Complexity on AN?

• UPFs and ANs effectively are hub and spoke PEs for the DN VPN
  • ANs are spokes and UPFs are hubs
  • GTP-U tunnels correspond to <base label, VPN label> label stacks
    • TEIDs correspond to per-prefix labels
  • No VRFs on ANs because there is no need for IP lookup
    • UL traffic is always sent to the hub (UPF)
    • DL traffic is relayed to UEs based on TEID
      • Just like that pre-prefix labels removes the need for IP lookup for to-CE traffic

• What is added in ANUP is the routing in VRFs
  • For local routing purpose
  • Routing functions are becoming commoditized
    • especially with server-based implementations
Isn’t This against Microservices Architecture?

• Per https://microservices.io/:

  Microservices is an architectural style that structures an application as a collection of services that are:
  
  Highly maintainable and testable; Loosely coupled; Independently deployable;
  Organized around business capabilities; Owned by a small team

  The microservice architecture enables the rapid, frequent and reliable delivery of large, complex applications

• Microservices is about decomposing complex applications

• ANUP is about integrating forwarding functions for high performance data forwarding
Keep the Discussion Going!

• Really appreciate the comments
• Could you discuss this with your (3GPP/wireless) colleagues?
• Will continue to discuss with more mobile architects and update the draft accordingly