

ENOG73

SRv6 MUPのPoCを作った話

2022/03/25

古河ネットワークソリューション株式会社 技術企画開発部

林 秀明
藤原 達弥



ENOGとワタシ

- ENOG55 (SRv6回)がはじめてのENOGでした
 - 次は発表っ！と思ってはや2年……遂に発表できる機会を頂きました💕
 - あさまさん、ありがとうございます！



自己紹介

- 古河電工のルータ部門が2005年に分社化
- ハードウェア／ソフトウェアの製造・開発、ユーザサポートを自社で取り組み、ユーザの声が開発まで届く体制
- 技術企画開発部では日々提案される新技術に関して、技術の選定とPoC開発検証を通じて製品化に向けた課題抽出・解決をミッションとしています
- 今回はSRv6に関する事例を紹介したいと思います
- 今回の発表はdraft-mhkk-dmm-srv6mup-architecture-01に基づいた内容であり、最新のI-Dでは仕様が更新されています



SRv6の取り組み

諸事情により、内容非公開





なぜSRv6 MUP？

- モバイル市場で何か取り組めることはないか検討中に
うちのボスがMUPドラフトに興味津々
- 「JANOGまでに作ってみよう」と鶴の一声（2021年末）

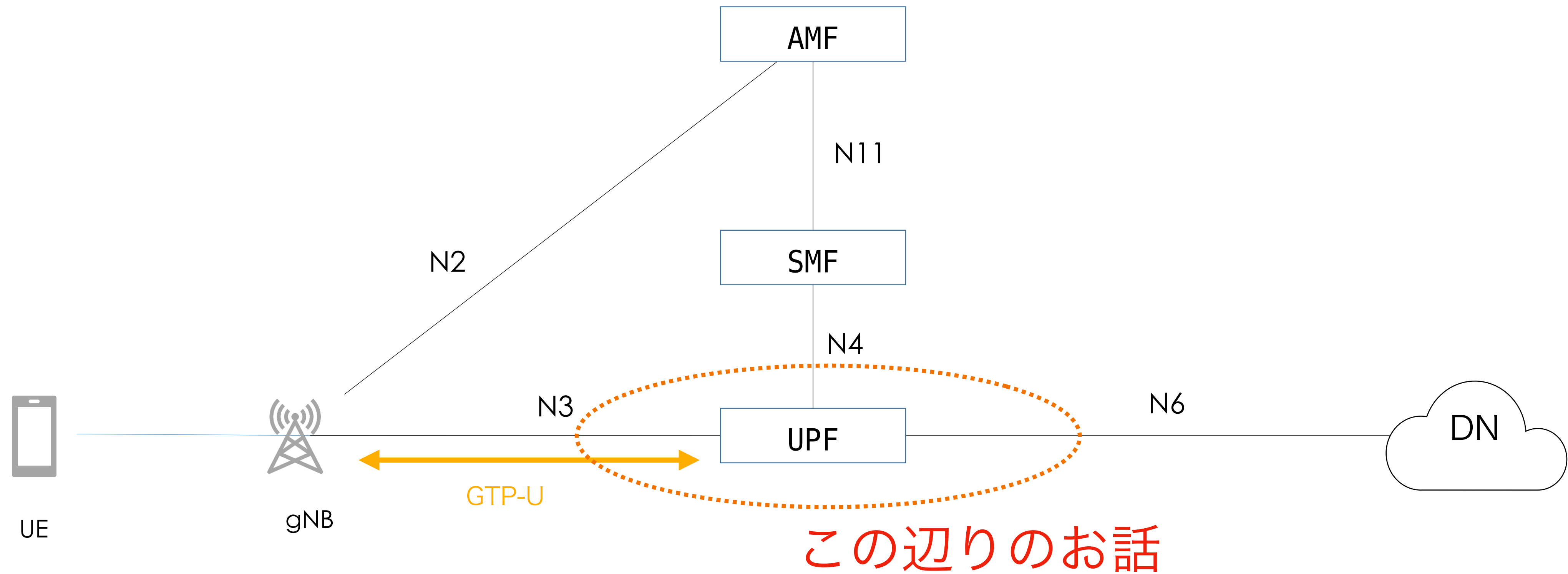


SRv6 MUP動作概要

draft-mhkk-dmm-srv6mup-architecture-01
(3/25 現在 draft は -03 に update)

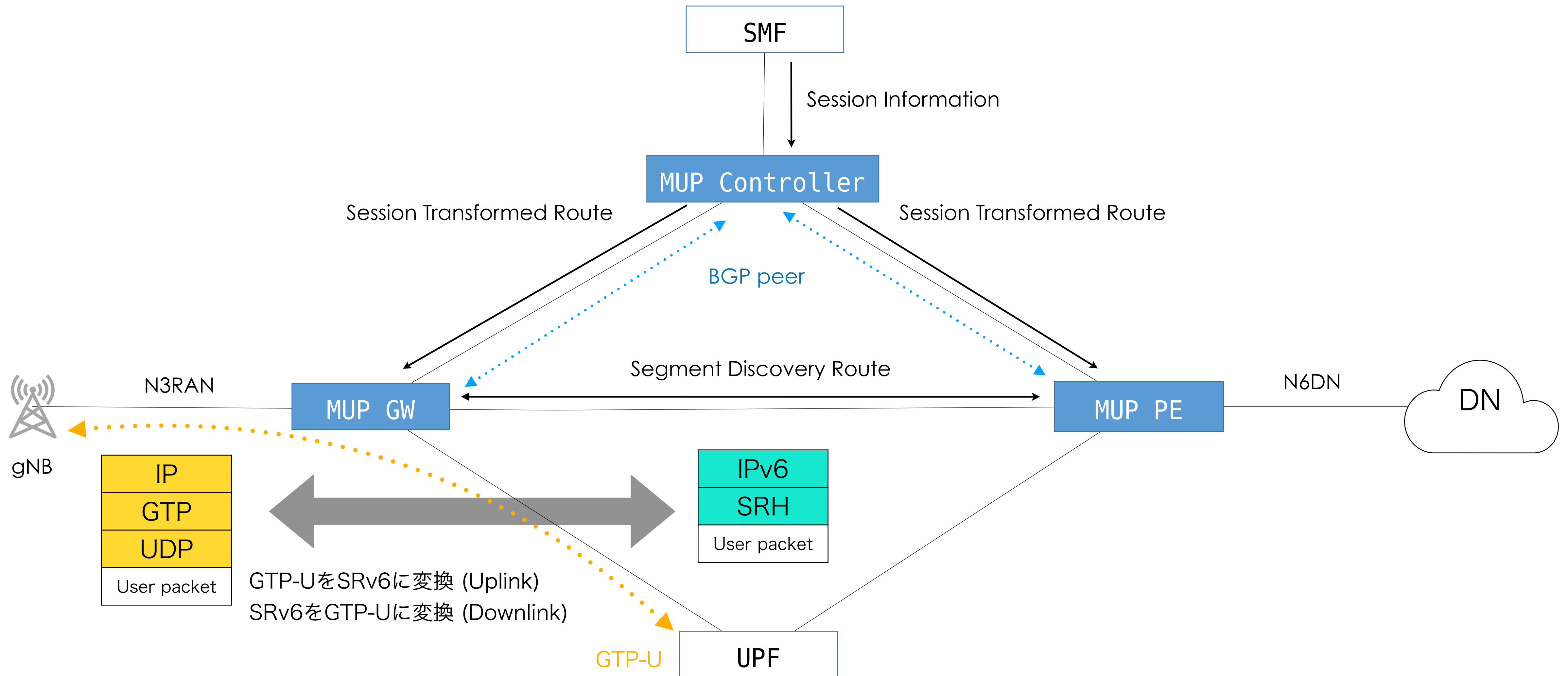


5GC Architecture





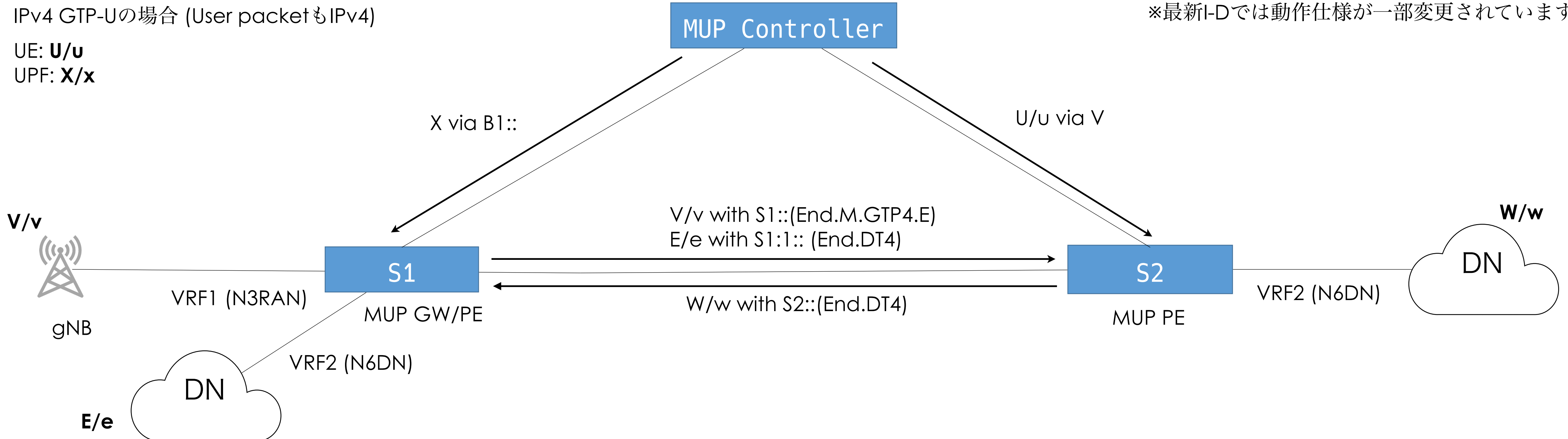
SRv6 MUP Architecture





BGPによる経路交換

※最新I-Dでは動作仕様が一部変更されています



S1のtable情報

経路表 (VRF1)	経路表 (VRF2)	SID table
X via B1::	W/w (SID S2::)	S1:: (End.M.GTP4.E)
V/v	E/e	S1:1:: (End.DT4)
		BSID B1:: (H.M.GTP4.D)
		SID list <S1:1::>

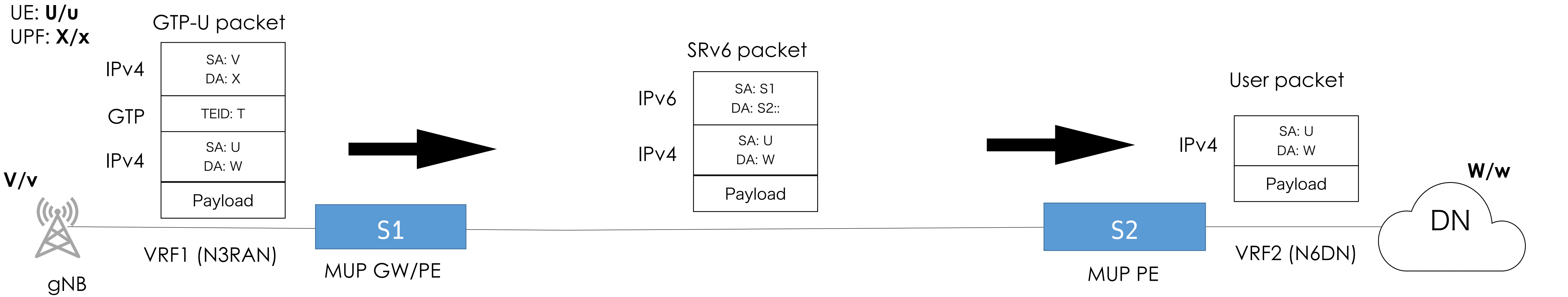
S2のtable情報

経路表 (VRF2)	SID table
W/w	S2:: (End.DT4)
U/u (SID S1::V:T)	
V/v (SID S1::)	
E/e (SID S1::1)	



SRv6 MUP動作説明 (UpLink)

IPv4 GTP-Uの場合 (User packetもIPv4)



※最新I-Dでは動作仕様が一部変更されています

- S1の動作
- ① X (UPF)宛ての packets を受信すると、経路表 (VRF1) に従い、nexthopがB1::でBSIDなので、SID tableを確認
 - ② B1::はH.M.GTP4.DなのでGTP-Uのヘッダをはずして、BSIDと紐づいたSID listのSID S1:1::が所属するVRF2の経路表を確認
 - ③ innerのDA Wに基づいてSRv6の packets にして(H.Encap.Red)転送

S1のtable情報 (必要なもののみ)

経路表 (VRF1)	経路表 (VRF2)	SID table
X via B1::	W/w (SID S2::)	S1:1:: (End.DT4)
		BSID B1:: (H.M.GTP4.D)
		SID list <S1:1::>

- S2の動作
- ① DAのS2::はEnd.DT4なのでSRv6のヘッダをはずして packets を転送

S2のtable情報 (必要なもののみ)

経路表 (VRF2)	SID table
W/w	S2:: (End.DT4)



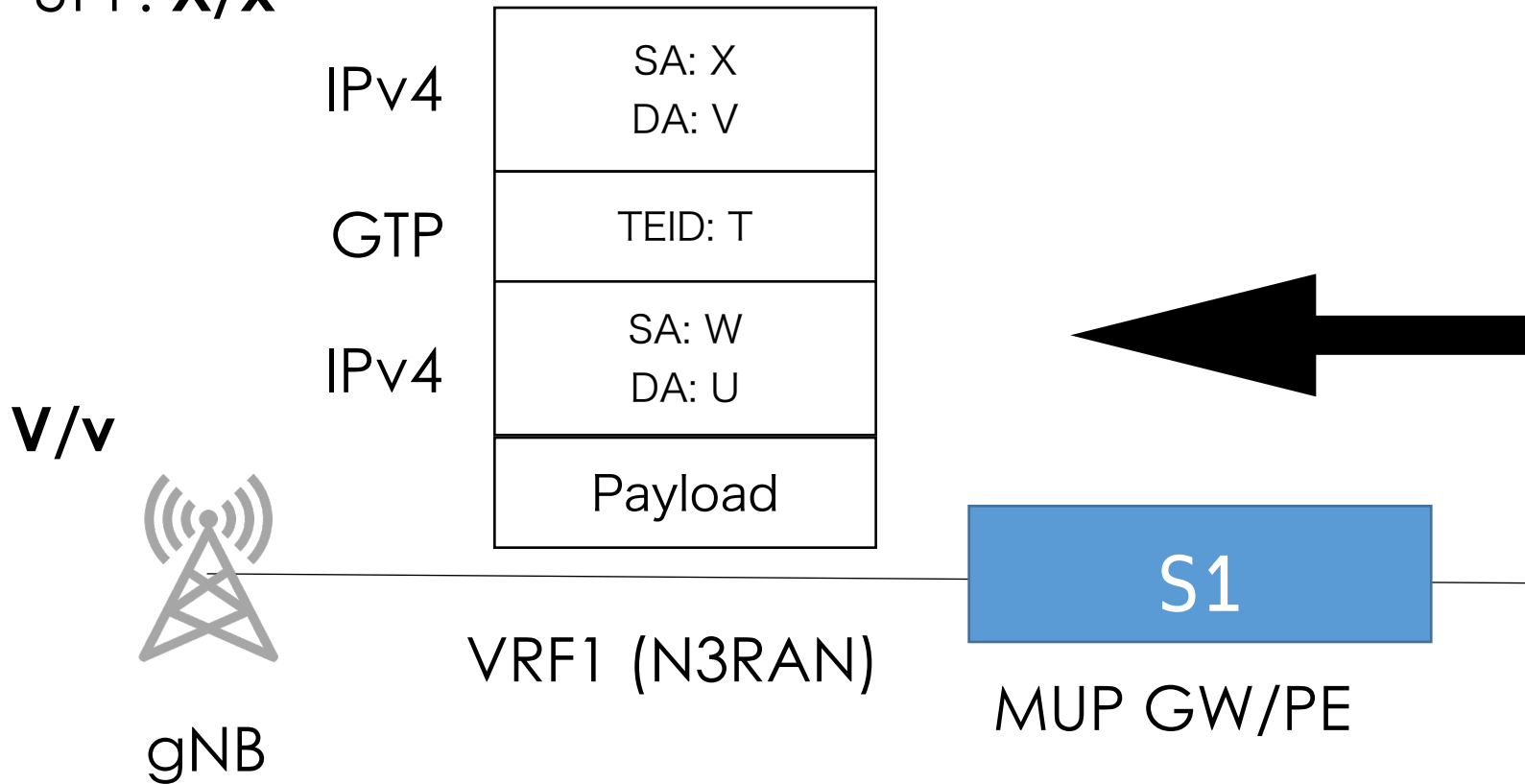
SRv6 MUP動作説明 (DownLink)

※最新I-Dでは動作仕様が一部変更されています

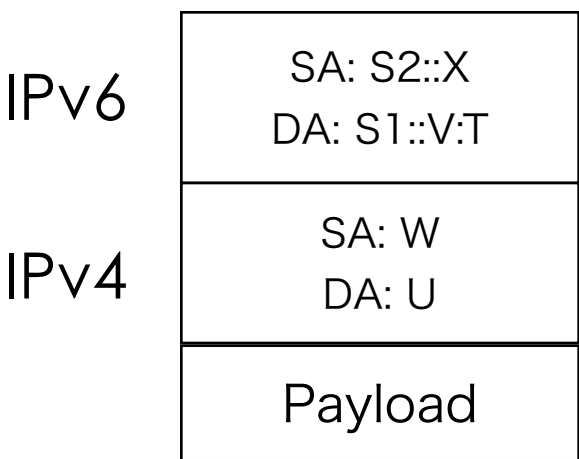
IPv4 GTP-Uの場合 (User packetもIPv4)

UE: **U/u**
UPF: **X/x**

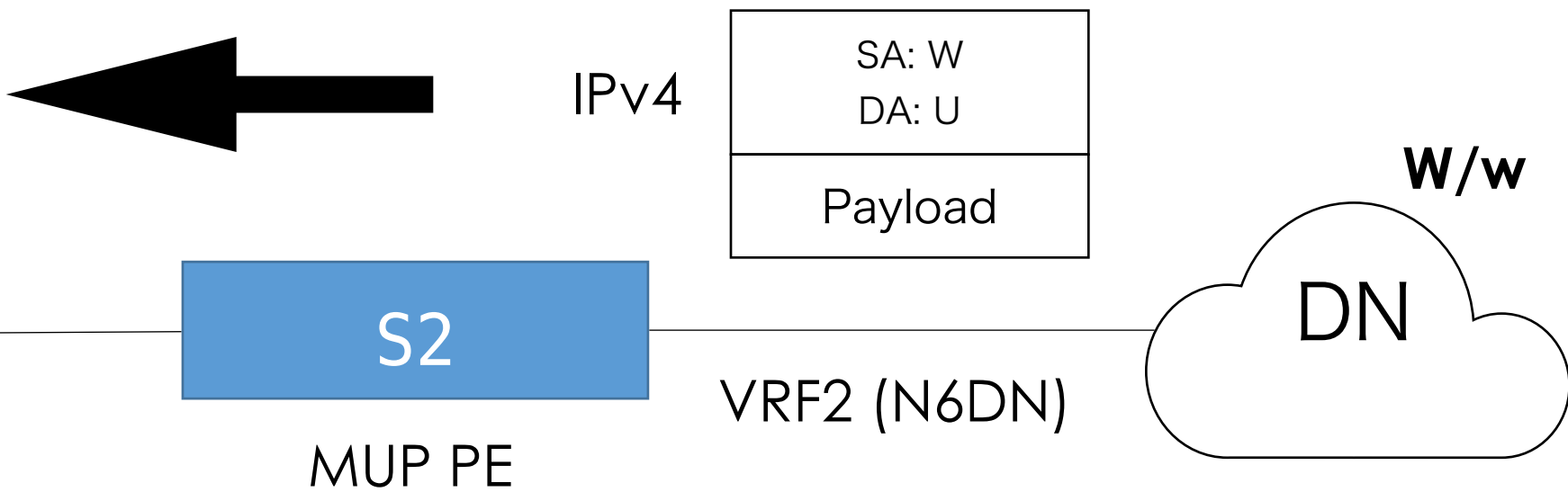
GTP-U packet



SRv6 packet



User packet



S1の動作

- ① DAのS1::V:T (S1::)はEnd.M.GTP4.EなのでSRv6のヘッダをはずして、SIDに埋め込まれたGTP-UパケットのSA(X), DA(V), TEID(T)を元にGTP-Uパケットを作成し転送

S2の動作

- ① U宛てのパケットを受信すると経路表 (VRF2)を見て、SRv6のパケットにして(H.Encap.Red)転送 (SIDにはGTP-UのDA (V)とTEID (T)を、SRv6パケットのSAにGTP-UのSA (X)を埋め込む)

S1のtable情報 (必要なもののみ)

経路表 (VRF1)	SID table
V/v	S1:: (End.M.GTP4.E)

S2のtable情報 (必要なもののみ)

経路表 (VRF2)
U/u (SID S1::V:T)

作ってみよう



実装ターゲット

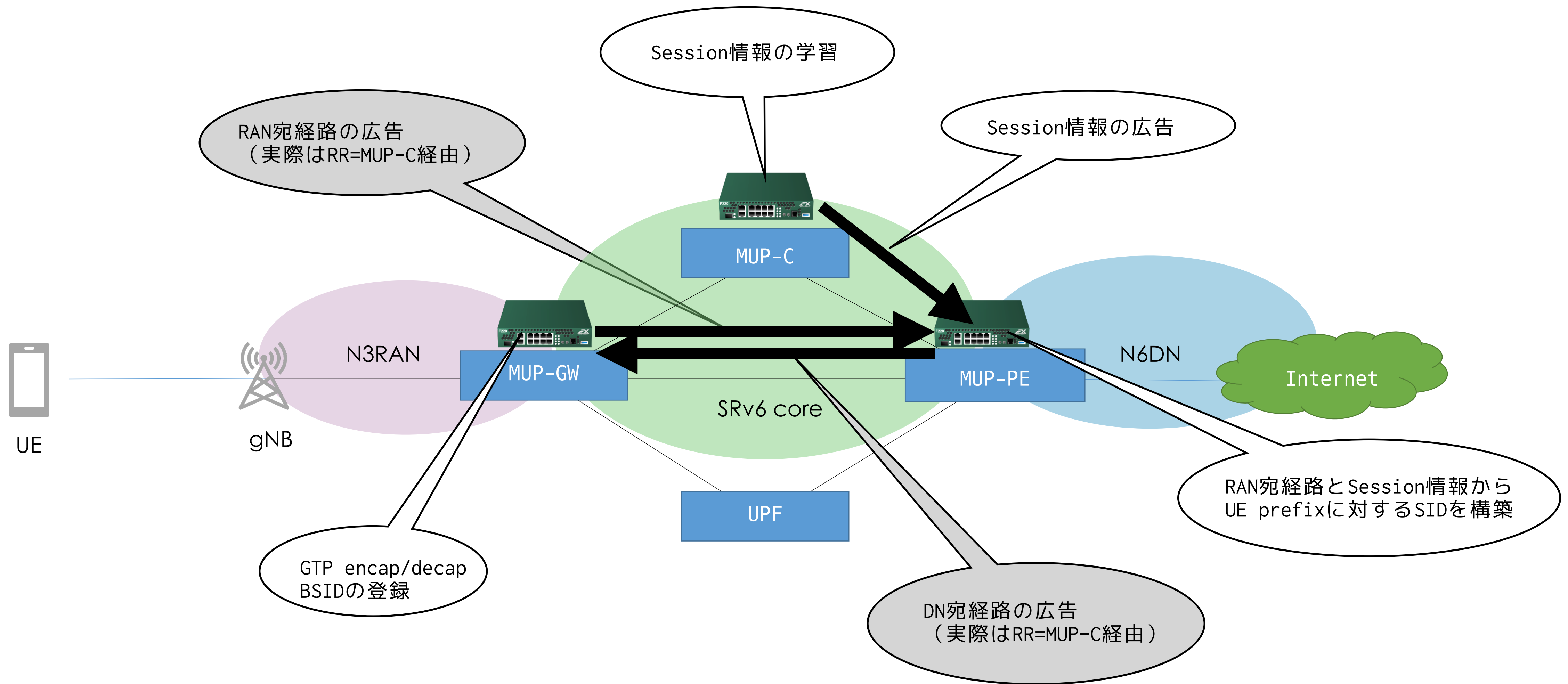
- SRv6対応のFITElnet FX201(キャリア向け)、FITElnet F220-EX(小型CPE)に実装

諸事情により、内容非公開



※FITElnet は古河電気工業株式会社の登録商標です

SRv6 MUPを動かすためにやること





MUP-GW:BSIDの登録

- BSID H.M.GTP4.DをStatic Local SIDとして登録し、SID listと経路とCLIで紐付ける

```
F220-MUP-GW# show segment-routing srv6 sid
```

SID	Function	Context	Owner	State
2001:1:1::	H.M.GTP4.D		NSM	InUse

```
F220-MUP-GW# show ip route vrf VRF1 10.0.0.127/32
```

```
Routing entry for 10.0.0.127/32
Known via "static", distance 1, metric 0, best, redistributed
Encapsulation Information:
  Tunnel Type: SRv6
  Tunnel IF: Tunnel1 (Data: 0xf6a01ab8)
  Tunnel ID: 229
  Tunnel Endpoint: 2001:1:1:: (System VRF-ID: 0)

2001:1:1::, Tunnel1 (Tunnel-ID:229), RD 65000:1, System VRF-ID 1, NHD LINK Tunnel1 (2), refcnt 1
```

I-D -02 以降では BSID を使わない方向へ



MUP-GW:Segment discovery route の広告

- Function End.M.GTP4.Eを割り当ててMUP-PEに広告

```
F220-MUP-GW# show ip bgp vpnv4 all
```

Network	Next Hop	Metric	LocPrf	Weight	Path
Route Distinguisher: 65000:1 (VRF1)					
*> 192.168.2.25/32	0.0.0.0			32768	i
*>i192.168.30.2/32	192.168.2.25		100	0	i
Route Distinguisher: 65000:2 (VRF2)					
*>i0.0.0.0/0	2001:db8::2	0	100	0	?
Route Distinguisher: 100.0.0.3:1					
*>i192.168.30.2/32	192.168.2.25		100	0	i

```
F220-MUP-C# show ip bgp vpnv4 all 192.168.2.25/32
```

```
Route Distinguisher: 65000:1
BGP routing table entry for 192.168.2.25/32
  Advertised to non peer-group peers:
    2001:db8::2
  Local, (Received from a RR-client)
    2001:db8::1 (metric 20) from 2001:db8::1 (100.0.0.1)
      Origin IGP, localpref 100, valid, internal, best
      Extended Community: RT:65000:1
      Original RD:65000:1
      BGP Prefix-SID: SRv6 L3VPN 2001:1:: (L:16.16, F:16.0, T:16.32) End.M.GTP4.E
      Local Label: no label
      Remote Label: 1120
      Path Identifier (Remote/Local): /0
      Last update: Fri Jan 21 18:23:54 2022
```




MUP-C:セッション情報の学習

- SMFからMUP-Cにセッション情報を送ることができれば良いが、SMFはイジれない……。どのようにして学習させるか？
- CLI/REST-APIで設定することに

```
POST /api/:api-version/routing/srv6mup/sessions
{
  "ue-prefix" : "192.168.30.2/32" ,
  "gnb-address" : 192.168.2.25,
  "teid" : 16777480
}
```

- TEIDはどうやって知り得る？ パケットを覗き見して設定♡



MUP-C:セッション情報の広告

- どのようにしてMUP-PEに伝えるか？

UE prefix	VPNv4/VPNv6 NLRI	RTによるimportが容易
gNB address	MP Next-hop	
TEID	Color Extended community	単にColorが使いたかっただけ。 QFIなども必要なので、Tunnel Encapsulation attributeが吉

```
F220-MUP-C# show ip bgp vpnv4 all
Network          Next Hop          Metric    LocPrf Weight Path
Route Distinguisher: 65000:1
*>i192.168.2.25/32 2001:db8::1      100      0 i
Route Distinguisher: 65000:2
*>i0.0.0.0/0      2001:db8::2      0        100    0 ?
Route Distinguisher: 100.0.0.3:1
*> 192.168.30.2/32 192.168.2.25     100    32768 i
```



MUP-PE:UE prefixのSIDを構築

- Segment discovery route とセッション情報からSIDを構築

F220-MUP-PE# **show ip bgp vpnv4 all**

Network	Next Hop	Metric	LocPrf	Weight	Path
Route Distinguisher: 65000:1					
*>i192.168.2.25/32	2001:db8::1		100	0	i
Route Distinguisher: 65000:2 (VRF2)					
*> 0.0.0.0/0	10.5.0.4	0		32768	?
*>i192.168.2.25/32	2001:db8::1		100	0	i
*>i192.168.30.2/32	192.168.2.25		100	0	i
Route Distinguisher: 100.0.0.3:1					
*>i192.168.30.2/32	192.168.2.25		100	0	i

F220-MUP-PE# **show ip bgp vpnv4 all 192.168.2.25/32**

Route Distinguisher: 65000:1
BGP routing table entry for 192.168.2.25/32
Advertised to non peer-group peers:
2001:db8::2
Local, (Received from a RR-client)
2001:db8::1 (metric 20) from 2001:db8::1 (100.0.0.1)
Origin IGP, localpref 100, valid, internal, best
Extended Community: RT:65000:1
Original RD:65000:1
BGP Prefix-SID: SRv6 L3VPN 2001:1:: (L:16.16, F:16.0, T:16.32) End.M.GTP4.E
Local Label: no label
Remote Label: 1120
Path Identifier (Remote/Local): /0
Last update: Fri Jan 21 18:23:54 2022

Segment Discovery Route

F220-MUP-PE# **show ip bgp vpnv4 all 192.168.30.2**

Route Distinguisher: 65000:2 (VRF2)
BGP routing table entry for 192.168.30.2/32
Not advertised to any peer
Local
192.168.2.25 from 2001:db8::3 (100.0.0.3)
Origin IGP, localpref 100, valid, internal, best, installed
Extended Community: RT::100.0.0.3:1 Color:16777480
Original RD:100.0.0.3:1
BGP Prefix-SID: SRv6 L3VPN 2001:1:46:c0a8:219:1:1:800 (L:16.16, F:16.72, T:0.0) End.M.GTP4.E
Local Label: no label
Remote Label: Exp-Null
Path Identifier (Remote/Local): /0
Last update: Fri Jan 21 18:39:07 2022

Route Distinguisher: 100.0.0.3:1
BGP routing table entry for 192.168.30.2/32
Not advertised to any peer
Local
192.168.2.25 from 2001:db8::3 (100.0.0.3)
Origin IGP, localpref 100, valid, internal, best
Extended Community: RT::100.0.0.3:1 Color:16777480
Original RD:100.0.0.3:1
Local Label: no label
Remote Label: Exp-Null
Path Identifier (Remote/Local): /0
Last update: Fri Jan 21 18:39:07 2022

Session Transformed Route

動作を検証してみる



検証環境

- 偶然、社内にPrivate LTE環境が => これで検証してみよう

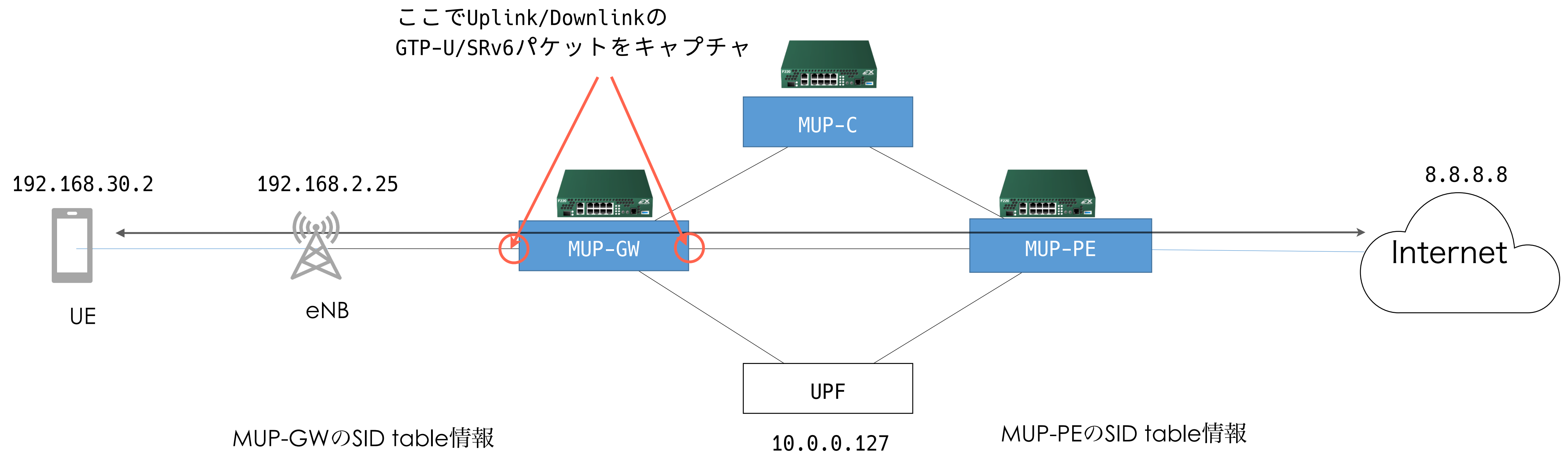
諸事情により、内容非公開





構成図と試験内容

- UEから8.8.8.8向けにpingで疎通を確認



MUP-GWのSID table情報

SID table
2001:1:46:: (End.M.GTP4.E)
2001:db8:0:1:1:: (End.DT4)
BSID 2001:1:1:: (H.M.GTP4.D)
SID list <2001:db8:0:1:1::>

MUP-PEのSID table情報

SID table
2001:db8:0:2:4b:: (End.DT4)



Uplink GTP-U

> Frame 5: 134 bytes on wire (1072 bits), 134 bytes captured (1072 bits) on interface en4, id 0
> Ethernet II, Src: [REDACTED], Dst: Furukawa_f0:5a:6d (00:80:bd:f0:5a:6d)
> Internet Protocol Version 4, Src: 192.168.2.25, Dst: 10.0.0.127
> User Datagram Protocol, Src Port: 2152, Dst Port: 2152
✓ GPRS Tunneling Protocol
 > Flags: 0x30
 Message Type: T-PDU (0xff)
 Length: 84
 TEID: 0x01000237 (16777783)
> Internet Protocol Version 4, Src: 192.168.30.2, Dst: 8.8.8.8
> Internet Control Message Protocol

Src: 192.168.2.25 (eNB)

Dst: 10.0.0.127 (UPF)

TEID: 16777783 (Uplink)

Src: 192.168.30.2 (UE)

Dst: 8.8.8.8 (Google)



Uplink SRv6

> Frame 9: 138 bytes on wire (1104 bits), 138 bytes captured (1104 bits) on interface en4, id 0
> Ethernet II, Src: Furukawa_f0:5a:6d (00:80:bd:f0:5a:6d), Dst: Furukawa_4c:b1:8b (00:80:bd:4c:b1:8b)
> Internet Protocol Version 6, Src: 2001:db8::1, Dst: 2001:db8:0:2:4b::
> Internet Protocol Version 4, Src: 192.168.30.2, Dst: 8.8.8.8
> Internet Control Message Protocol

Src: 2001:db8::1 (MUP-GWのSRv6 Encap時のsrc address)

Dst: 2001:db8:0:2:4b:: (MUP-PEのEnd.DT4のSID)



A dark gray, rounded rectangular shape covers the majority of the page, likely a redaction or placeholder. The background shows a grid pattern with various colored cells (purple, pink, red, white) and some faint text, including the number '430' in the top right corner.

Src: 2001:db8:0:a00:7f::2

UPFのアドレス情報(10.0.0.127)が埋め込まれている

Dst: 2001:1:46:c0a8:219:1:1:800

MUP-GWのLocator

Function
(End.M.GTP4.E)

eNBのアドレス情報
(192.168.2.25)

QFI, TEID情報
(QFI: 0, TEID (Downlink): 16777480)

```

0                                     1                                     2                                     3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-----+-----+-----+-----+-----+-----+-----+-----+
|   QFI       |R|U|                               PDU Session ID      |
+-+-+-+-+-----+-----+-----+-----+-----+-----+-----+-----+
|PDU Sess(cont')|
+-+-+-+-+-----+-----+-----+-----+-----+-----+-----+-----+

```




Downlink GTP-U

> Frame 6: 134 bytes on wire (1072 bits), 134 bytes captured (1072 bits) on interface en4, id 0
> Ethernet II, Src: Furukawa_f0:5a:6d (00:80:bd:f0:5a:6d), Dst: [REDACTED]
> Internet Protocol Version 4, Src: 10.0.0.127, Dst: 192.168.2.25
> User Datagram Protocol, Src Port: 34835, Dst Port: 2152
▼ GPRS Tunneling Protocol
 > Flags: 0x30
 Message Type: T-PDU (0xff)
 Length: 84
 TEID: 0x01000108 (16777480)
> Internet Protocol Version 4, Src: 8.8.8.8, Dst: 192.168.30.2
> Internet Control Message Protocol

Src: 10.0.0.127 (UPF)

Dst: 192.168.2.25 (eNB)

TEID: 16777480 (Downlink)

PoC実装してみた感想

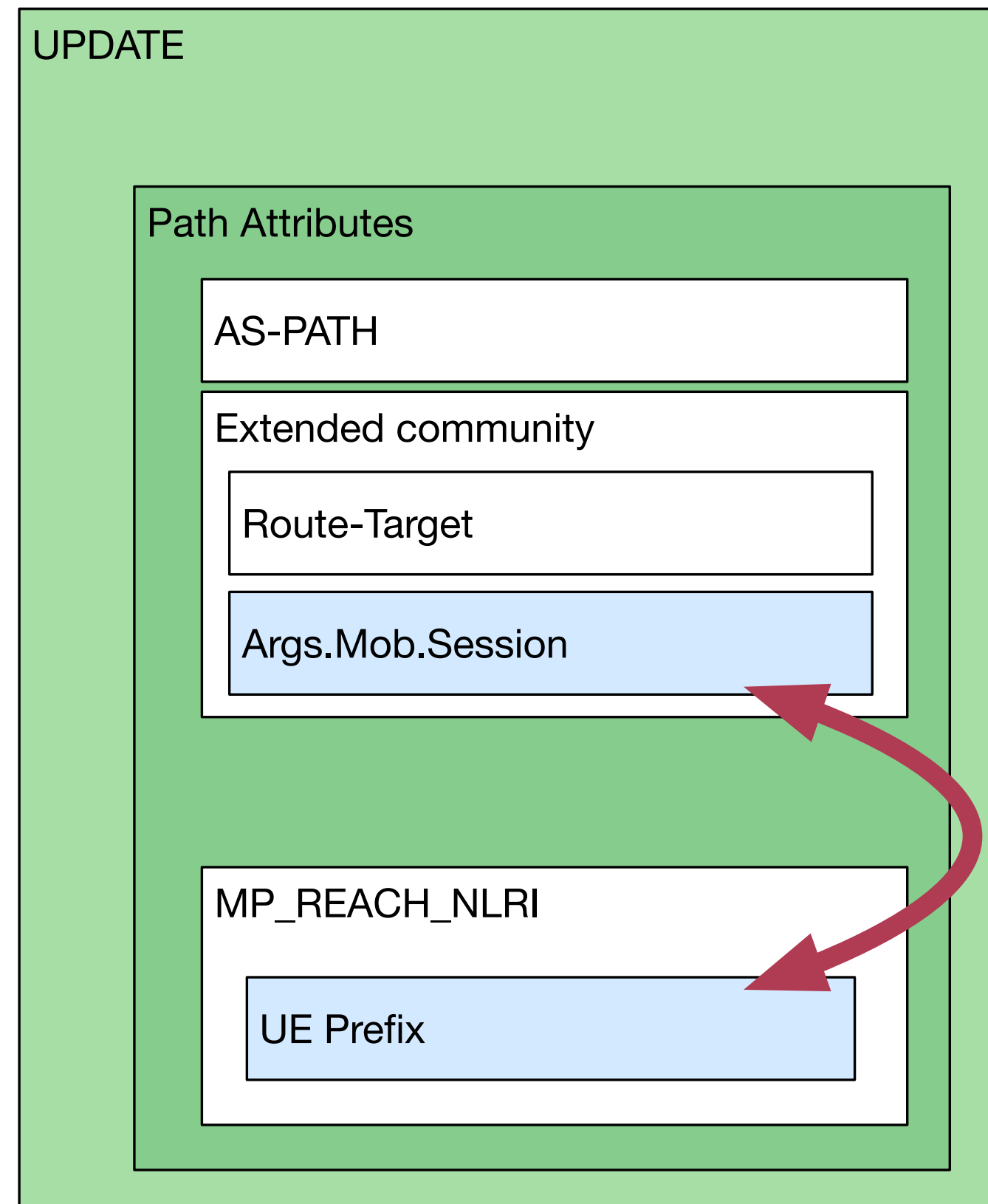


感じたこと

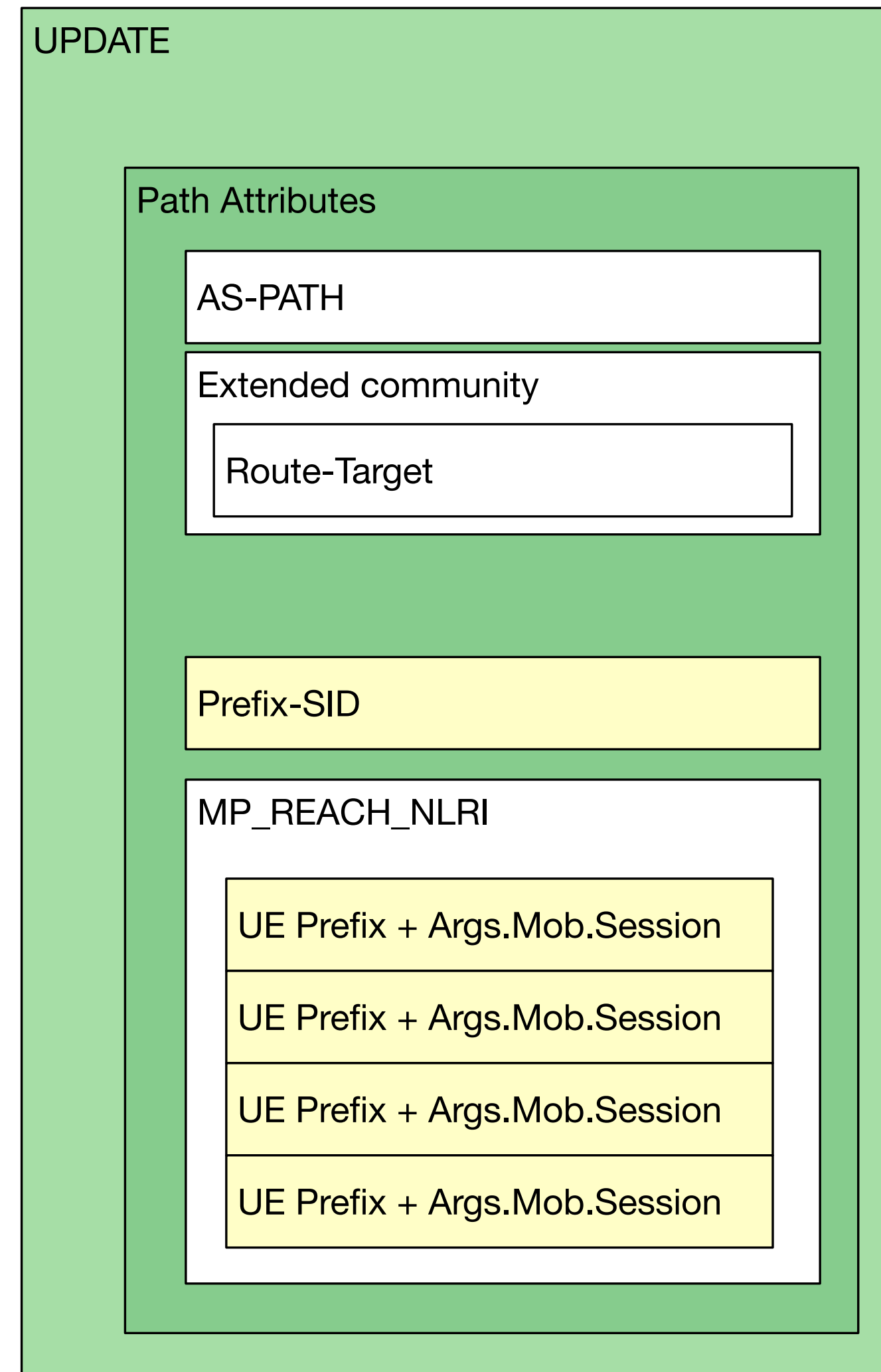
- TEIDの情報をattributeに入れると、MultiPrefixで広告できない
 - NLRIにArgs.Mob.Sessionを入れて新しいアドレスファミリを定義した方がよい
- Segment discovery routeが持つLocator情報はMUP-Cがセッション情報のPrefix-SIDとして広告した方がMUP-PEの受信経路数が減る
 - MUP-PEはSID構築のためだけにSegment discovery routeを学習している



MUP-CからのUPDATE



一つのUPDATEメッセージで複数の UE Prefixを送信できない
大量の UE Prefixを広告するとき収束性能に効いてくる



NLRIに含めることで
複数の UE Prefix を送信可能



- draft-mpmz-bess-mup-safi-00が`submit

- MUP用のSAFIを新たに用意し、NLRIを定義

Type 1 Session Transformed route (ST1)

[illegible]



今後の取り組み

- draft-mpmz-bess-mup-safiによるBGP-MUP実装
- ハンドオーバー動作確認
- IPv6(UE prefix/gNB address)
- 5Gへの対応確認



まとめ

- SRv6 MUP実装について紹介しました
- SRv6 MUPが既存モバイル網の構成を変更することなく、トラフィックをSRv6網へ誘導できることが実証できました。
- 実際に手を動かしてみて、セッション情報のより良い広告方法が検討できました。
 - I-Dへのフィードバックを行っていきたいと思います
- 実装募集。相互接続実験しましょう

