



Janog 50th SRv6 MUP: スケーラブルな実装と課題

July 13th 2022

Tetsuya Murakami

Arcus Company Evolution



Company

Founded 2016
Headquarters:
San Jose, CA
100+ employees

Target Customers

Fortune 100
Data Center, Cloud, & CDN
Telco/Service Providers

Board Investors

- **Guru Chahal:** Lightspeed Venture Partners
- **Steve Herrod:** General Catalyst
- **Chris Rust:** Clear Ventures
- **Ankur Prakash :** Liberty Global Ventures

Advisors

- Sumeet Arora: SVP, Thoughtspot, ex-SVP/GM Service Provider, Cisco
- Roy Chestnutt: ex EVP & Chief Strategy Officer, Verizon
- Yun Freund: SVP, Equinix, Advisory Board, Google Cloud
- Amarjit Gill: Investor & Entrepreneur
- Bask Iyer: CEO Baskmind, ex CIO VMWare, Dell, Juniper, Honeywell
- Krish Prabhu: ex CTO, AT&T
- Dave Ward: CEO, PacketFabric, ex-CTO Cisco, Fellow Cisco & Juniper



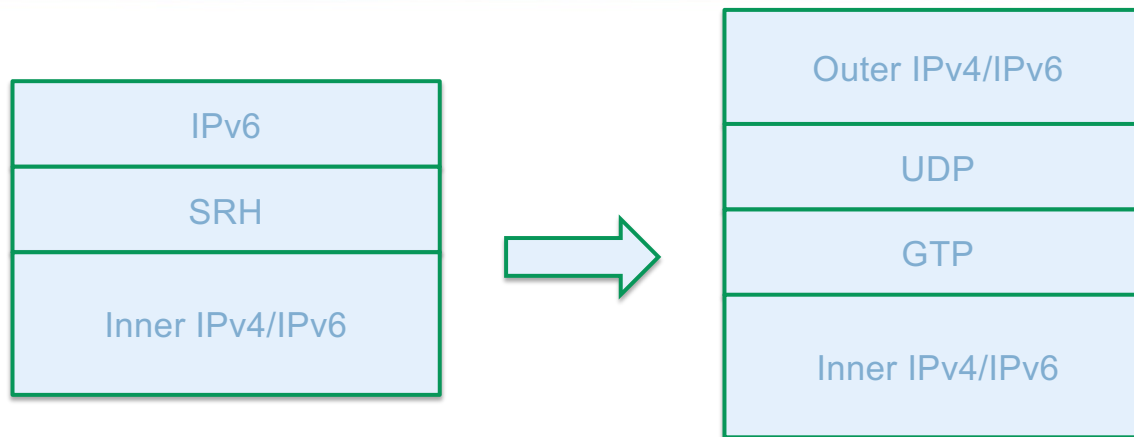
Chairman and CEO

- **Shekar Ayyar**
VMware Executive 15+ years

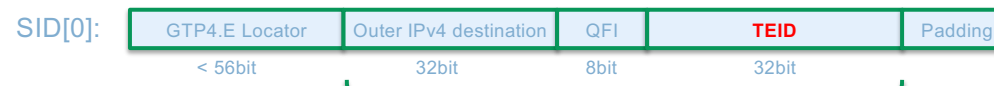
Tech Founder and CTO

- **Keyur Patel**
Cisco
Distinguished Engineer
Routing, VPNs, Networking Protocols
Co-Chair IETF BGP & Sidrops WG

SRv6 MUP: SID Format (GTP4/6.E)



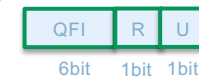
GTP4.E case)



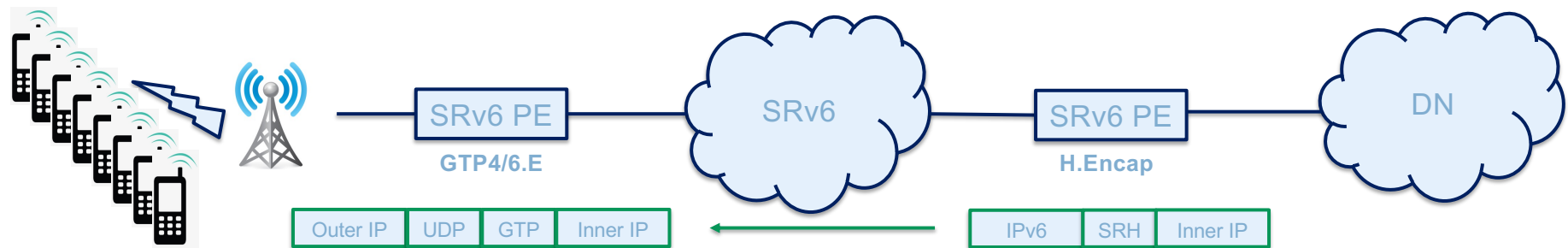
GTP6.E case)



*) QFI format

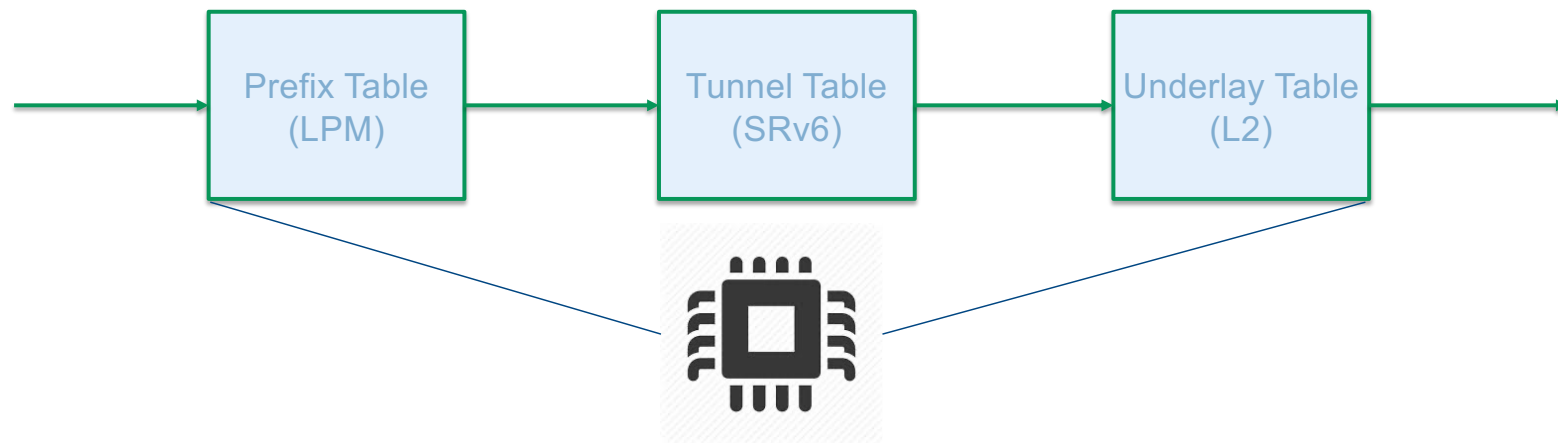


SRv6 MUP: Packet forwarding for Downlink



- UE毎にユニークなTEIDが割り当てられる。
- UE毎にIngress PEが管理するPrefix SIDが異なる。
- UEが増えると、PEが管理するPrefix SIDが増える。

SRv6 MUP: ハードウェア構成



- Prefix table: L3 routing table
- **Tunnel Table: SRv6 encap info** (SID, Source, etc)
- Underlay Table: Layer2 nexthop info

SRv6 MUP: Many UE devices....



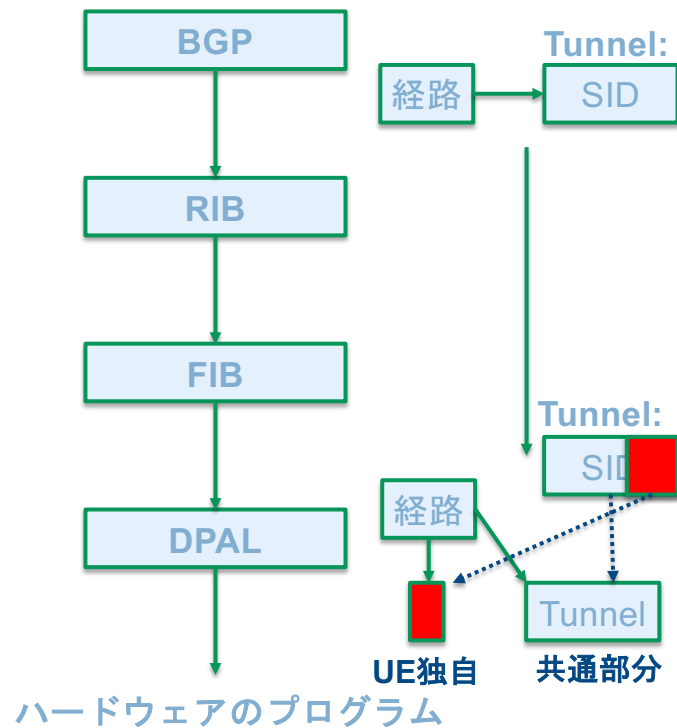
- UE毎にユニークなSIDが割り当てられる。
- Tunnel EntryにてSRv6 Ecapに必要な情報(SID, Source addressなど)を持つ。
- UEが増えると、、、
 - SIDが増える。
 - Tunnel Entryが増える。
 - Tunnel Tableは内部TCAMで保持（スケール限定的）
 - Prefix Tableは外部TCAMで保持（スケール可能）

SRv6 MUP: Many UE devices....



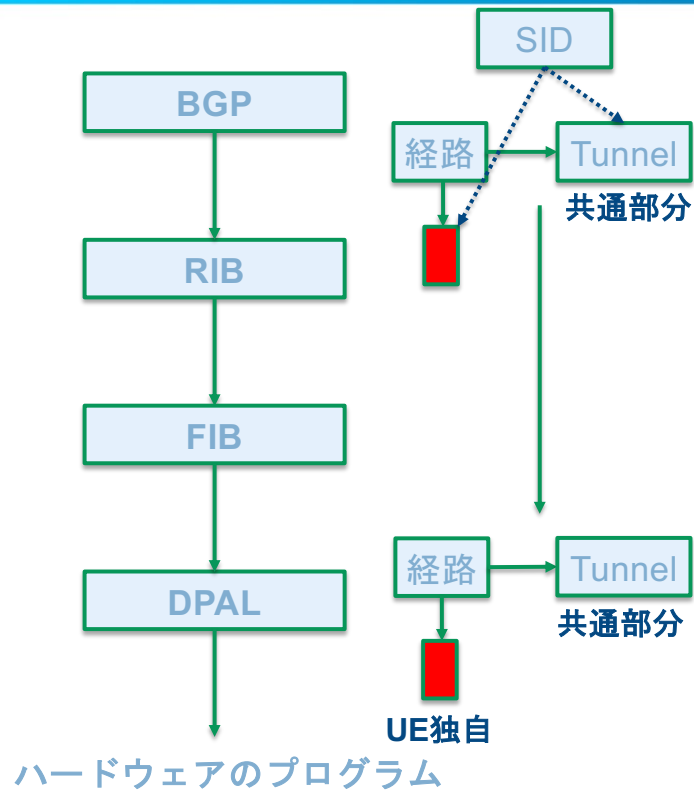
- **Tunnel Table (内部TCAM)**
 - GTP4/6.Eの共通部分(Locator)のみを保持
- **Prefix Table (外部TCAM)**
 - L3 Routing entry(UE経路)を保持し、GTP4/6.Eの非共通部分(UE毎にユニーク部分)を保持
 - GTP4.E => 72 bit
 - GTP6.E => 40 bit

SRv6 MUP: モジュール構成



- 誰がLocator(共通部分)とUE部分(非共通部分)を管理するか？
- Option1)
 - ハードウェアに近い人が管理する(DPAL)。
 - BGP/RIB/FIBはスケール環境、非スケール環境で共通にできる。
 - DPALでの管理は複雑。。。
 - SIDは、Tunnel情報として登録される。DPALで共通部分だけでtunnelを生成することはできるが、その時UE経路がない場合もあり(Tunnel生成後にUE経路を生成する必要があるため)、UE部分をどうやって経路と結びつけるか。。。。
 - UE経路が特定のTunnelを示すindexを持ち、BGP/RIB/FIBでのTunnel Indexが膨大になる(UE毎に異なるTunnel)。
 - RIB/FIB/BGPとDPALでは、異なるTunnel生成スキームとなり、装置全体に対して一貫性がなく、デバッグが難しくなる。

SRv6 MUP: モジュール構成



- 誰がLocator(共通部分)とUE部分(非共通部分)を管理するか？
- Option2)
 - コントロールプレーン(BGP/RIB/FIB)もハードウェアの構成を意識する。
 - BGP/RIB/FIBもハードウェアの構成によって、Tunnelの生成、UE経路の生成を変える。
 - BGP/RIB/FIB/DPALが共通のTunnel生成スキームを使用するため、デバッグが容易
 - BGP/RIB/FIBも共通部分でのみTunnelを生成し、UE部分は経路の属性として渡すため、DPALでの処理が容易。
 - BGP/RIB/FIBが生成するTunnelを指す、indexが少なくて済み、同じindexをBGP/RIB/FIB/DPALが使用できる。

SRv6 MUP: どちらがいい？



- **コントロールプレーンは、Platform Independentであるべき？**
 - 装置毎にコントロールプレーンが動作を変える必要はない。
 - エンジニア観点からは理想的だし、望ましい。
 - ただ、実装の観点からは、
 - DPALが複雑化する
 - Tunnel生成スキームがモジュール単位で異なる。
 - Tunnel管理の複雑化
 - メモリの使用用増加
 - デバッグビリティの複雑化
 - etc...
- **あまり拘らず、コントロールプレーンがデータプレーンを助けてもいい**

SRv6 MUP: 今後の課題



- Platform independentは、大切だけど、、、
 - 状況に応じて、Platformに合わせたコントロールプレーンのデザインにすることは必要。
- コントロールプレーン、データプレーンのデザイン・実装
 - 両方の知識や経験がないと、大変になることも多い。
- 運用者、SW開発者、HW開発者が開発当初から密な連携が必要