

# Solving IPv4 Address depletion with MAP-T

Mapping of Address and Port – Translated

Xander Thuijs CCIE#6775 [xthuijs@cisco.com](mailto:xthuijs@cisco.com)

Distinguished Engineer IOS-XR/ASR9000/NCS5500/8000

Cisco Systems

# Executive Summary

## Problem

- Address the depletion of IPv4 allocation in many networks
- Cost reduction
- High performance translations (overhead of stateful xlate and lack (no) of logging of stateless)

## Analysis

- CGN is used
- NAT44(4) is currently defacto
- 1000 ports per subscriber desired\*
- Mapping/logging to trace pub v4 subscriber address plus port to user through syslog (or NF) requiring extensive dbases

## Reassessment

- V4 pub addr reuse
- Tracing and recording
- High performance lowest control plane overhead
- Logging is critical with desired “single shot” view all
- Using V6 natively for those services supporting reducing the pub v4/port

## Path Forward

- Ideally V6 based
- MAP-T on the CPE
- BNG v6 subscribers
- MAP-T border router to xlate v6->v4
- CPE will do nat44 based on port range assigned by MAP rules
- Logging/Trace through dhcp lease and MAP rule

# Handling the condition

Problem – Exhaustion of global IPv4 addresses.

Solution – Migrate to IPv6.

Obstacles –

- . Migration to IPv6 will happen slowly and in phases.
- . During this migration, we may have
  - . IPv6 only networks
  - . IPv4 only hosts and IPv4 only servers
- . So, we do need v4-v6 translation solutions during the migration in addition

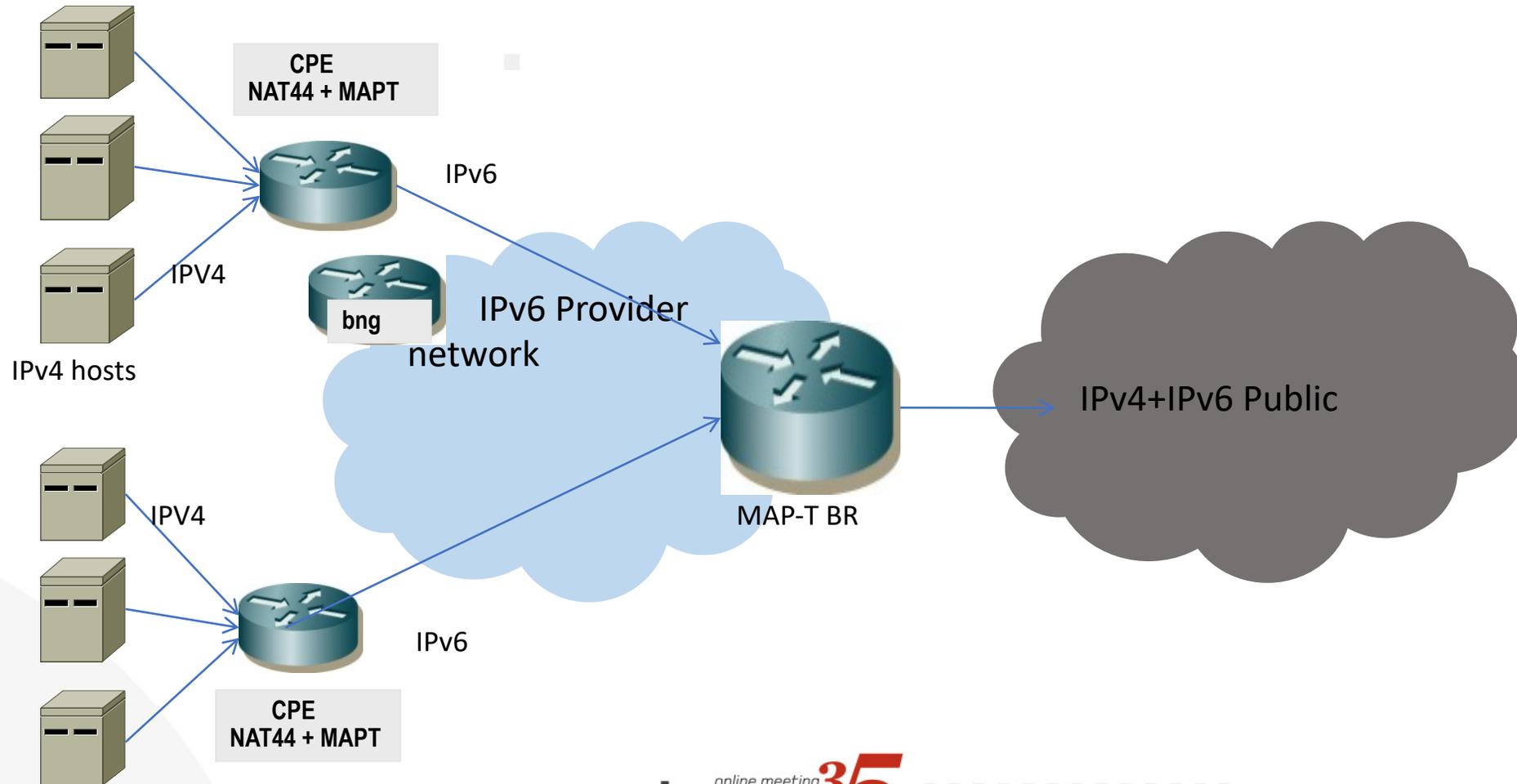
# What is MAP-T

- Originally intended as a v6 transition technology for providers with a v6 core only requiring to transport v4
- Provides for embedding v4 addresses as part of a v6 (prefix) with port awareness
- Provides for address translation
  - This functionality put MAP-T on the map 😊
- DHCP provides mapping rules to CPE to perform **xlate** and **encap**
  1. V6 address of the border gateway
  2. Public v4 to use
  3. Port range allowed for this subscriber
  4. Of course, the subscriber v6 prefix
- Is stateless: MAP border router, provider does not need to carry all individual translations
  - Reducing logging requirements hence storage as needed in nat44(4)
- HW Implementation generally through a policy based routing (PBR) infrastructure

# MAP-T components

- Mapping Rule
  - Construct of how to map a IPv4 address to an IPv6 prefix
    - The IPv4 prefix is a prefix used for the “IPv4 NAT Pool” or shared IPv4 address(s)
  - The IPv6 prefix is the aggregate of the prefixes assigned to individual subscribers
  - Used by MAP-T CE to determine its IPv4 address and optionally the layer-4 port allocation
- Allows a **MAP-T BR (Border Router)** to support IPv4/IPv6 stateless translation
  - Border router can sit anywhere in the network
- **Base Mapping Rule (BMR)**
  - The mapping rule(s) used by a MAP-T CPE given a list of mapping rules
- **Default Mapping Rule (DMR) or BR Prefix**
  - An IPv6 prefix used to embed IPv4 destination addresses (RFC 6052)
  - Similar to IPv4 default route
    - Allows for natural redundancy through anycast gateway!

# MAP-T A High-level View



# MAP-T Translation

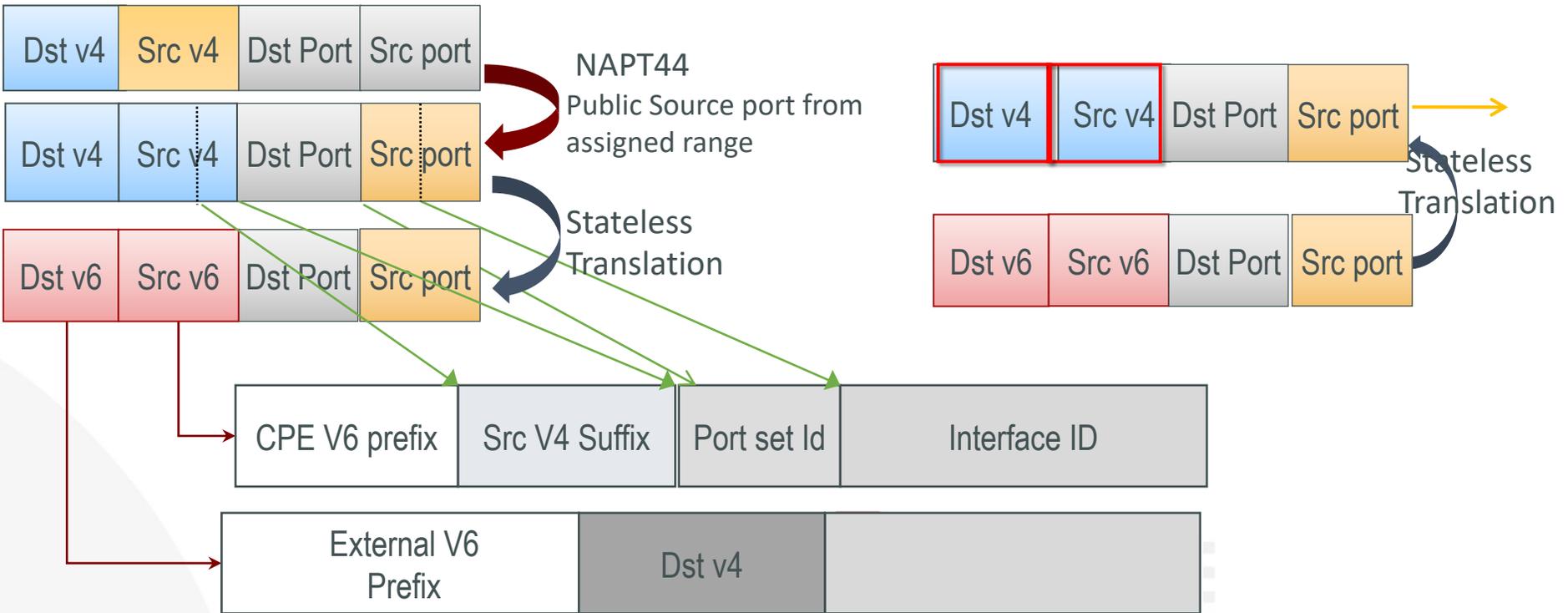
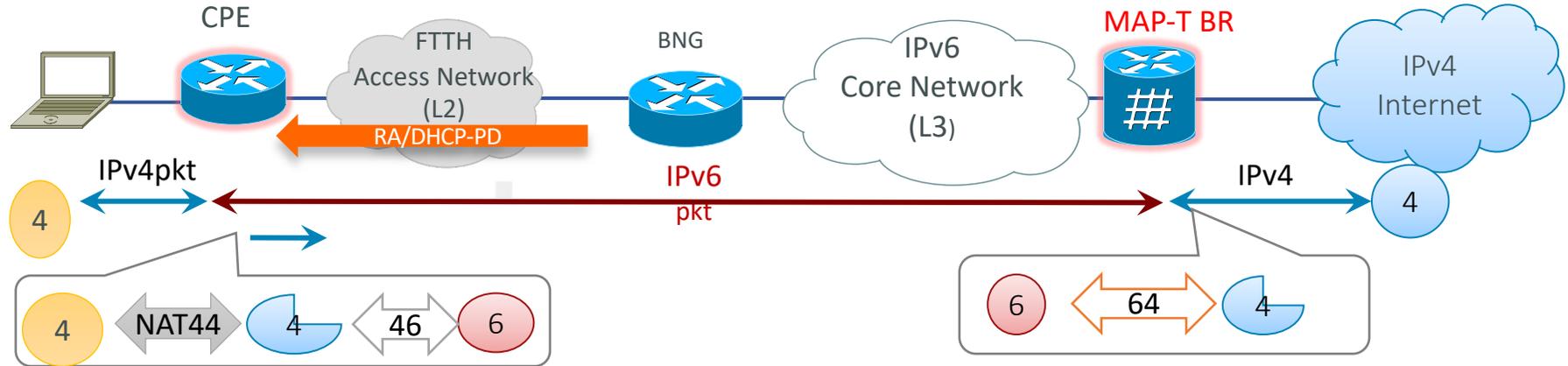
- Stateless double translation – once in CPE, once in SP BR
- CPE will still do NAT44 (like before).
- CPEs will share public IPv4 address.
- Each CPE assigned unique source port range for NAT44
  - After NAT44, source port can identify the CPE v6 prefix
  - x bits of source port is “CPE id” or “port set id”

Example - 7 bits for port-set-id – 128 CPEs with 512 ports each



16 bit source port

# MAP-T Translation

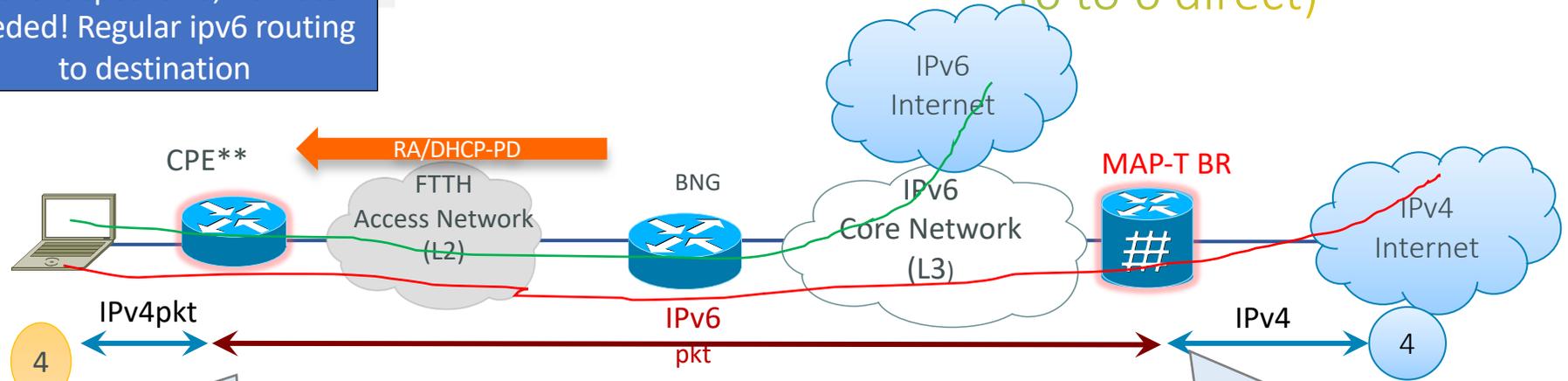


# MAP-T solution quick shot view (4 to 4 over 6)

If client speaks v6, no xlate needed! Regular ipv6 routing to destination

(6 to 6 direct)

Client ipv6 packet  
Client ipv4 packet

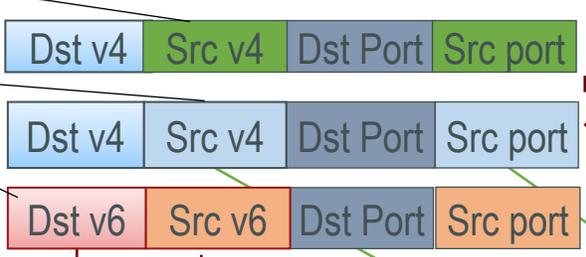
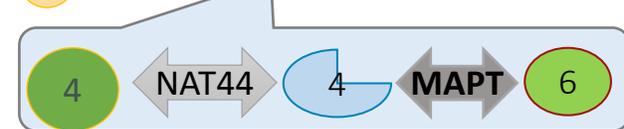


L3-Src is private ip from client

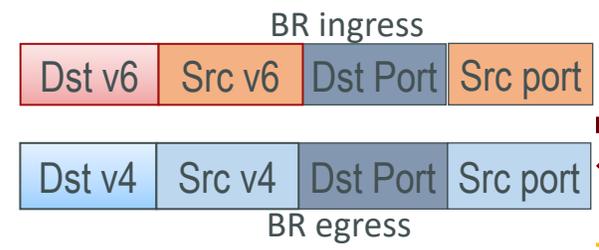
L3-Src is now from public pool

L3-Dst is BR v6 dest prefix and v4 appended

L3-Src is the subscriber v6 prefix with the pub v4 ip source embedded



NAPT44  
Public Source port from assigned range  
Stateless Translation



Stateless Translation



\*\*CPE receives v4 pub and port range through DHCPv6

Subscriber is identified by:  
1) Pub v6 prefix  
2) Pub v4 L3 PLUS L4 port range, provided in DHCPv6 offer

# Port range allocation defines subscriber to v4 ratio



- Average users require 100 ports on average at any given time



- Using v6 eliminates services that need translation drastically!

V6 is cheap

Single stack subs  
*(higher BNG scale)*

Highest reuse of v4

Allocating 128 or 256 ports allows for 250 or 500 subs per pub v4!  
Any range possible

*Both CGN and MAP-T provide port \*blocks\*  
Unused ports are not shared/redistributed.  
MAP-T bonus is native v6, which saves for translations*

***Lesser ports means  
higher savings!***

# MAP-T Advantages

- SP network can be only one domain – IPv6
- Decouples operator's transition to IPv6 from that of the IPv4 users
  - If needed the solution is v4 transparent to the end user!
- Current solutions in the space are stateful - DS-Lite and NAT64
- Stateless is better
  - Scales very well. reliable, robust.
  - Network architecture simplification
    - (dual stack subs become single stack v6 subs)
  - Network dimensioning
  - No new or extensive logging requirements

# Exception packet handling MAP-T

Exception packets for MAPT are:

- **IPv4 with options:** (no application uses source route or record etc) NO IMPACT
- **Fragmentation required:** (only when BR needs to frag, increase MTU) Can be prevented, no impact.
- **ICMP on BR other than echo:** (impact only for pMTUd \*if\* BR is lowest mtu hop) use MSS adjust or increase MTU egress on BR

# MAP-T benefits

- Stateless border relay means it can easily be hw accelerated
- Fragment forwarding at v6, (but not at 4)
- Doesn't require a special core only IPv6 enabled
- No designated logging required besides rules and dhcp information
- Nearly transparent to end user
- Relative Simple integration with existing deployment , only CPE sw needed
- Does not require same entry point into network as stateful solutions
  - (eg doesn't need symmetric routing)
- Doesn't require any routing tricks
- IPv4-Mapped IPv6 Address (is easy to analyze and troubleshoot)
  - Natural easy addressing scheme, easy for me old mind 😊
- Native v6 services will not require nat saving ports!

# NAT/CGN comparison

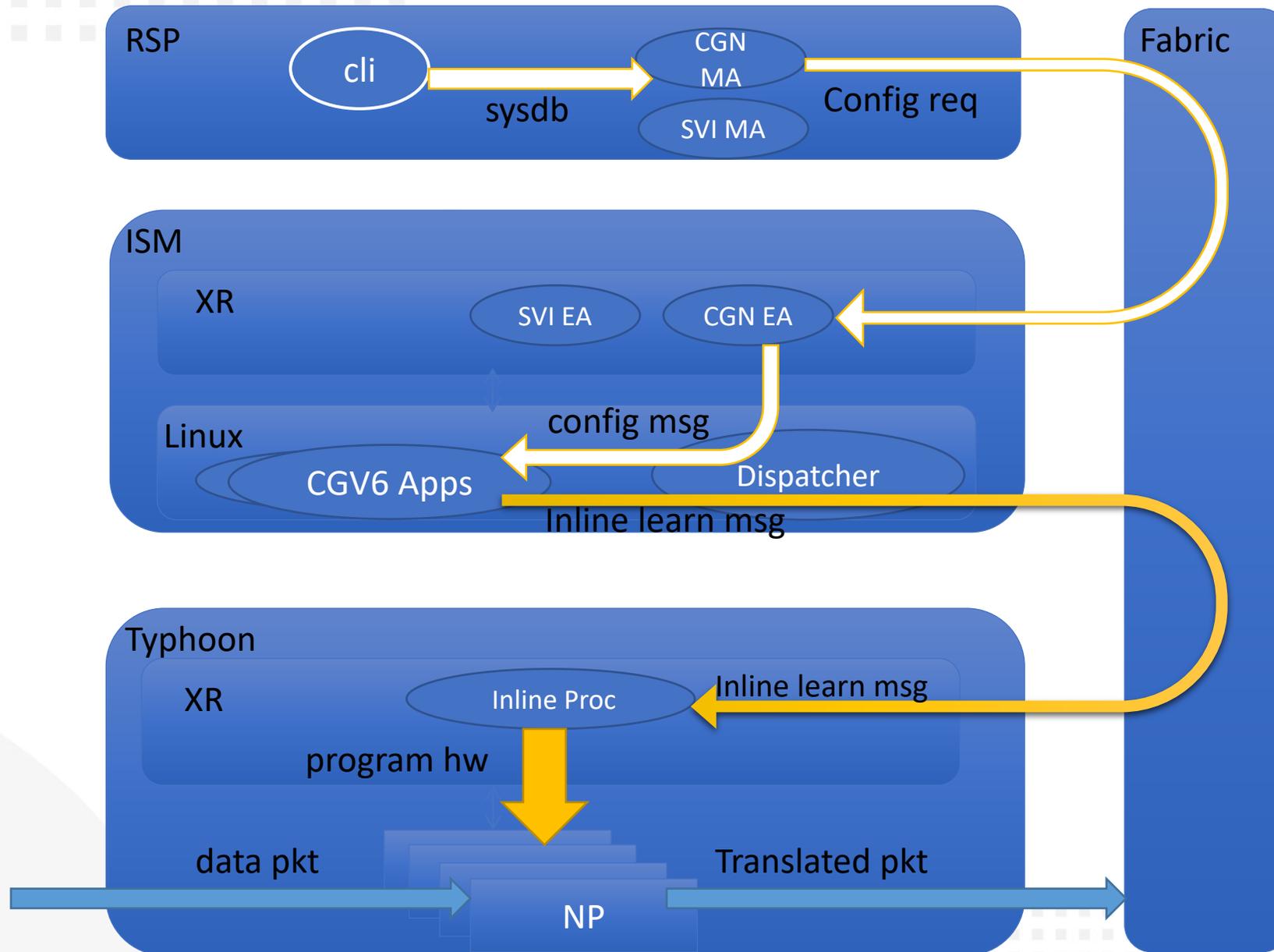
Item / Technology	MAP-T	NAT44(4)	MAP-E	DS-Lite
Concept	V4 addr embedded in v6 with v4 src addr/port xlate at cpe	Port overload pub v4 for requesting private v4 source	Encapsulate (natted) v4 addr into new v6 header (tunnel)	V6 tunnel between cpe and gateway. GW performs v4 NAT
Requires symmetric routing	No (any gw with ruleset can do it)	Yes	No	No
Layer 4 aware	Yes	Yes	No	No
Specific CPE requirement	Yes (MAP capable CPE + nat)	No (standard nat44)	Yes (MAP capable CPE + NAT)	Yes v6 tunnel only
Port/Address conservation	High (v6 native services)	Low due to high port count per sub	medium	Medium/low
Availability	OpenWRT	common	OpenWRT	
Port allocation flexibility	Fixed, at initial deployment not easily changeable	Dynamic allocation (grab what you want)	Fixed, but allows for easier allocation scheme changes	Dynamic, no hard setting
Provides native 46 xlate	Yes	No	No	No
Packet overhead	28 Bytes	None	40 Bytes	40 Bytes
HW accelerated	Yes	No	Yes	No
integration with NAT or gateway/relay	No (v4 may need ALG)	Yes - PCP /ALG	No Yes ALG	Yes - PCP ALG
Cost of ownership	Low (reuses gateway)	High (dedicated hw)	Low reuses gateway	Medium



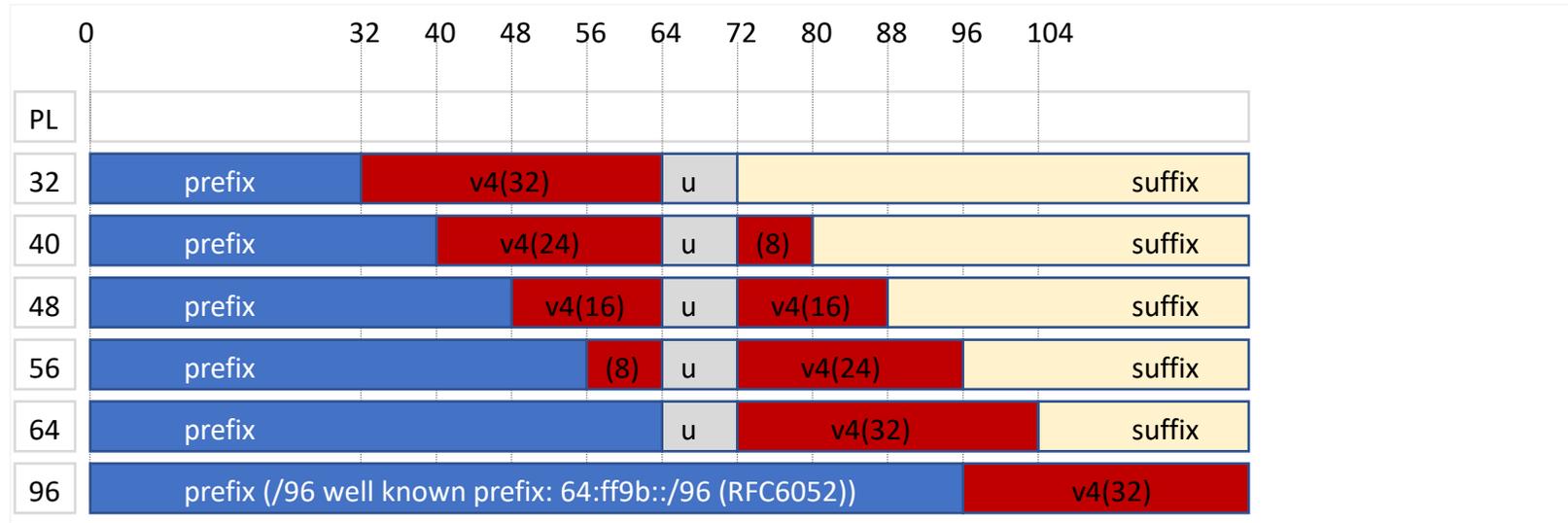
# Backup slides



# MAP-T Control path and data Path



# v4 in v6 address embedding



- RFC6052 specifies way to embed IPv4 address in IPv6 address.

# Sample CLI

```
RP/0/RSP0/CPU0:GANGES#sh running-config service cg6 cg6-1
service cg6 cg6-1
service-type map-t-cisco mapt1
cpe-domain ipv4 prefix length 24
cpe-domain ipv6 vrf SVRF-003
cpe-domain ipv6 prefix length 48
sharing-ratio 256
contiguous-ports 8
cpe-domain-name cpe1 ipv4-prefix 192.1.1.0 ipv6-prefix 2301:d01:1122::
ext-domain-name ext1 ipv6-prefix 3301:d01:1122::/48 ipv4-vrf VRF-1
```

# Config Explained

```
service cgn cgn1
service-inline interface Bundle-Ether2
service-inline interface tenGigE 0/2/0/0/2
```

- Traffic which come via the interfaces which are configured under service inline will be subjected to MAPT rules if it matches the class map that PBR will create internally based on CGN config
- Interfaces like physical ,vlan, and bundle-ether interfaces are supported.

```
service-type map-t-cisco gvk-100
cpe-domain
```

- All CPE domain parameters are grouped under here.

```
cpe-domain
ipv6-vrf SVRF-003 -> All EXT domains will be under only one VRF
ipv4-prefix-length 24 -> All CPE domain will have same ipv4 prefix length
ipv6-prefix-length 64 -> All CPE domain will have same ipv6 prefix length
contiguous-ports <>
sharing-ratio <>
```

And if (CPE domain v6 prefix length is 64) or CPE domain (v4 length is 32 and v6 prefix length is 128) sharing-ratio and contiguous-ports will not be considered in the translation and need not to be configured as well.

# Config Explained

```
cpe-domain
  ipv6-vrf SVRF-003
  ipv4-prefix-length 24
  ipv6-prefix-length 64
  domain-name cpe1 ipv4-prefix 100.0.2.0 ipv6-prefix 2001:db8:1122:1122::
  domain-name cpe2 ipv4-prefix 101.0.2.0 ipv6-prefix 2002:db8:1122:1122::
  .... (upto 1000 entries can be configured)
```

```
!
ext-domain
  domain-name ext1 ext-ipv6-prefix 2001:db8:1123:1123::/64 ipv4-vrf VRF-0003
  domain-name ext2 ext-ipv6-prefix 2002:db8:1123:1123::/64 ipv4-vrf VRF-0001
  ..... (upto 8000 entries can be configured)
!
```

There will be only one ipv6-vrf(CPE domain VRF) and all CPE domain networks will be under that. After translating it to as V6 packet ,packet will be forwarded to that V6 VRF.

For each ext-domain entry we give one ipv4-vrf (EXT domain vrf) which will indicate after V6->V4 translation ,to which vrf packet should be forwarded.

This way we will have V4 networks sitting in different VRFs and all V6 networks sitting in one VRF.

# Config explained

- This way under one mapt instance 8000 external domain and 1000 cpe domain networks are supported.
- Packet from any external domain can go to any cpe domain
- To achieve this in existing mapt model we would need  $8000 * 1000$  mapt instances.
- If each mapt instance has one cpe and ext domain entry ,then max scale of 1000 mapt instances can be achieved.

## MAP-T v4 to v6 translation

### In2out direction (done by CPE after NAT44)

- v4->v6 Destination address translation - RFC 6052  
external domain IPv6 prefix + IPv4 dest
- v4->v6 Source address translation – map-t algorithm  
cpe domain IPv6 prefix + IPv4 src + **port\_set\_id**  
port\_set\_id makes IPv6 source address unique to CPE

### Ou2in direction (done by provider BR router)

- map-t algorithm for v6 destination address
- RFC 6052 for v6 source address