Putting **IPv6** to work

North American IPv6 Summit

Plaza Tower One Conference Facilities

Greenwood Village, CO

April 22-23, 2015

Rocky Mountain IPv6 Task Force

RMv6TF

Mapping of Address and Port (MAP) an ISPs Perspective

E. Jordan Gottlieb
Principal Engineer | Charter Communications
jordan.gottlieb@charter.com

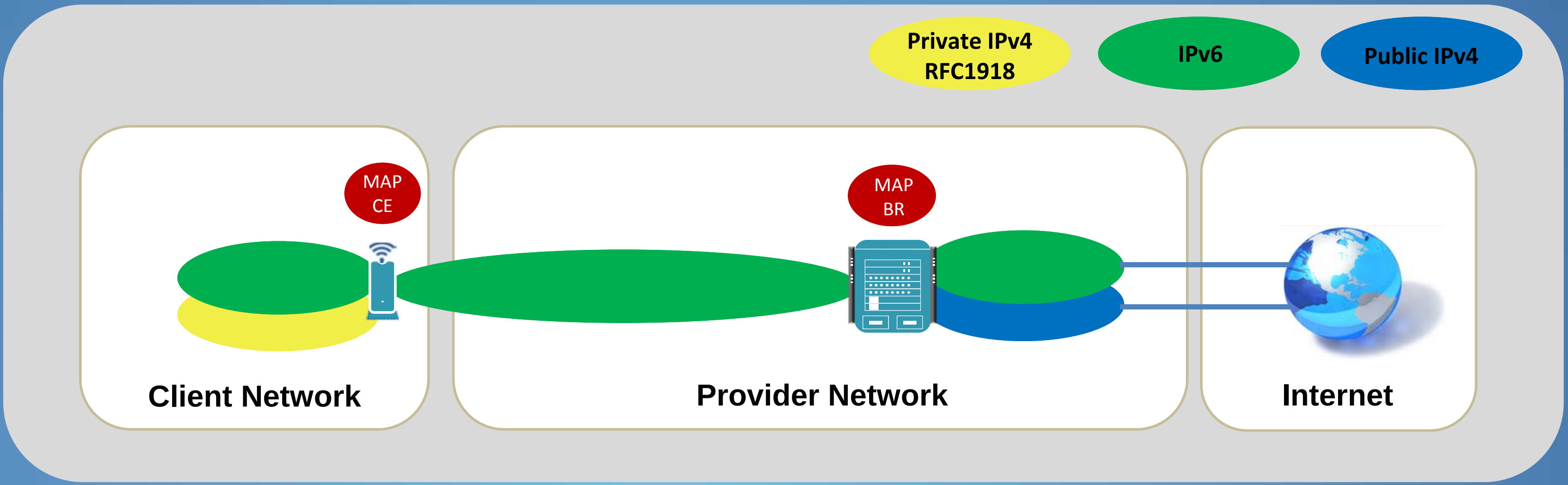
Agenda

- What is MAP?
- Benefits of MAP
- MAP in Action
- MAP Algorithms in Action
- MAP Deployment Considerations
- Resources and Attributions

What is MAP?

- Enables a service provider to support IPv4 based Internet access to customers over an IPv6 network domain
- Primary application is N:1 IPv4 address sharing
  - Competing technology in the CGN/LSN solution space
  - Can be deployed without IPv4 address overloading
- MAP defines two transport modes
  - Mapping of Address and Port with Encapsulation (MAP-E)
    - Transports MAP traffic across IPv6 domain by encapsulating IPv4 traffic within a IPv6 header as specified in RFC2473
  - Mapping of Address and Port using Translation(MAP-T)
    - Transports MAP traffic by translating IPv4 headers to IPv6 headers using RFC6145 stateless translation
- Extremely complementary to native IPv6 deployments

MAP Components in Context



- MAP Customer Edge (CE)
  - A home gateway or router with MAP support
  - Responsible for forwarding native IPv6 traffic and translating between IPv4 and IPv6
- MAP Border Relay (BR)
  - Provider side network gear that supports MAP
  - Can be existing (upgraded) equipment or dedicated appliance
  - Responsible for forwarding IPv6 traffic and translating MAP traffic between IPv4 and IPv6

Rocky Mountain IPv6 Task Force
RMv6TF

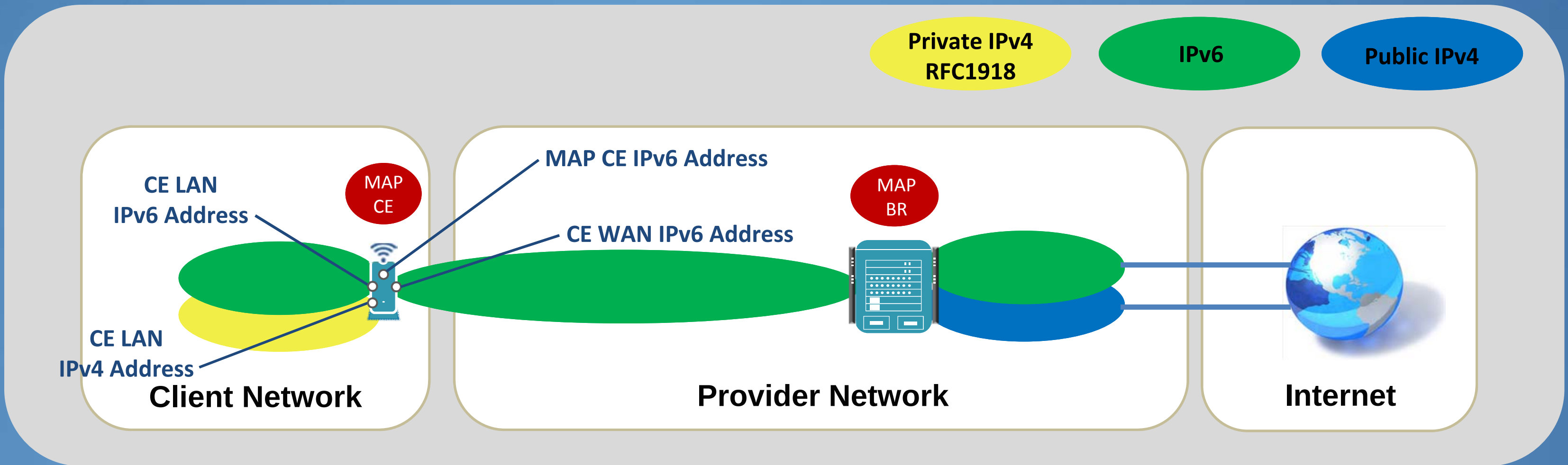
MAP Domains

- MAP domains: A collection of interconnected MAP CEs and BRs that share a common set of MAP parameters
  - Rule IPv6 Prefix: Customers are assigned end-user prefixes out of this block for the purpose of defining mapping rules
  - Rule IPv4 Prefix: The pool of available IPv4 addresses used to define mapping rules
  - Embedded Address (EA): A value representing the number bits used to represent the mapping of address and ports
  - BR Address or Prefix: Depending on transport mode a BR is reachable via a BR address (MAP-E) or BR prefix (MAP-T)

MAP Rules

- Base Mapping Rule (BMR)
  - Uses rule IPv6 prefix, rule IPv4 prefix, and EA length to define how the CEs derive their MAP related addressing and assigned port allocations from their IPv6 end-user prefix
- Forward Mapping Rule (FMR)
  - Optional rules to allow direct MAP CE to CE communications
- Default Mapping Rule (DMR)
  - Used by MAP-T to reach destinations outside the MAP domain through the BR.
    - MAP-E defines a BR address to achieve the same capability

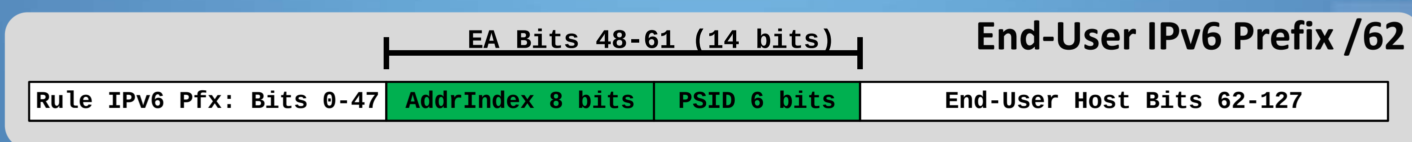
MAP CE Addressing



- IPv4 CE Addressing
  - Customer controls assignment on CE LAN segment (typically RFC1918)
- IPv6 CE Addressing
  - WAN address assigned by provider and is independent of MAP
  - MAP address is derived from assigned end-user prefix
  - LAN prefix is allocated from assigned end-user prefix

Mapping of Address and Port

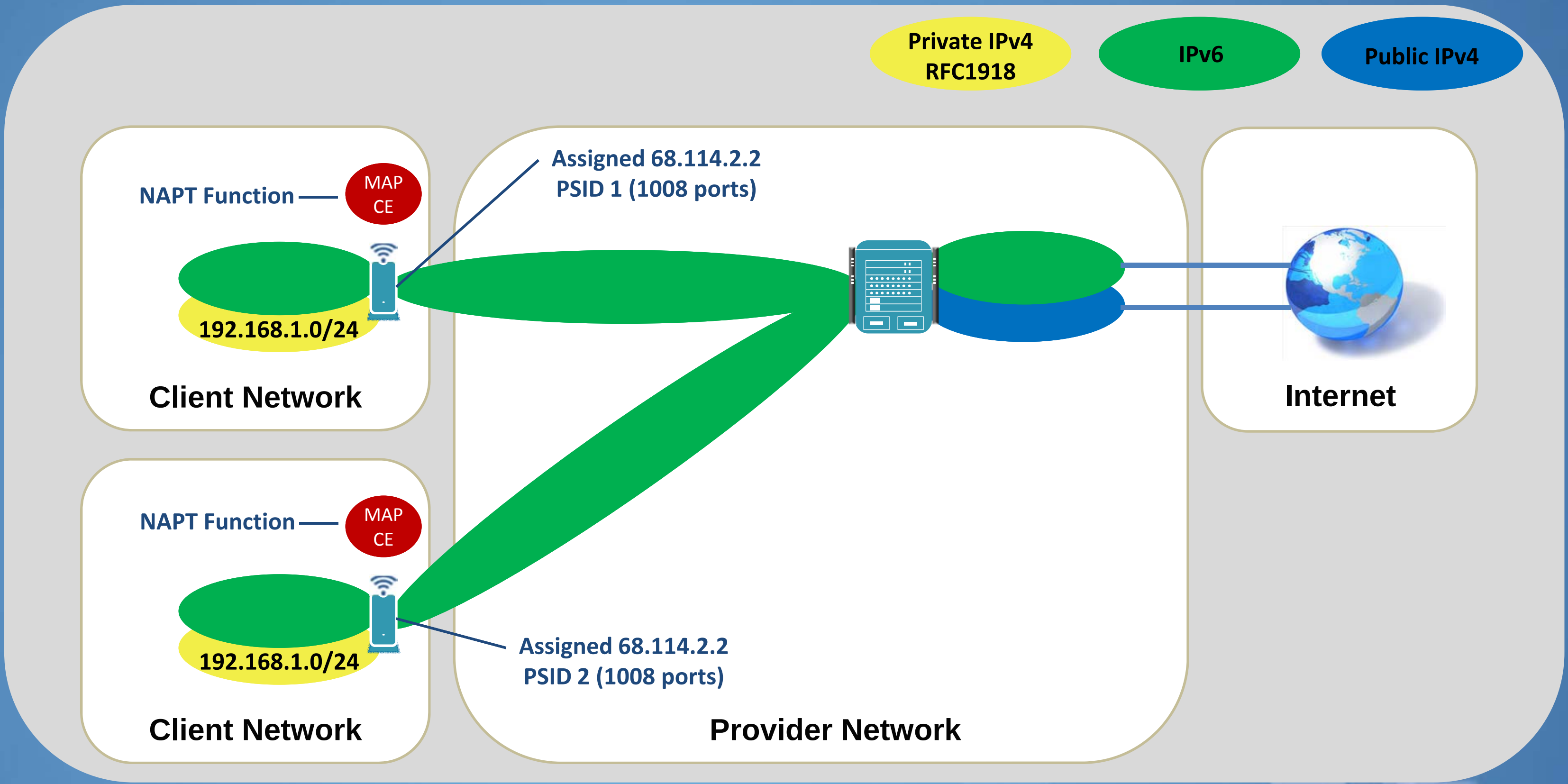
- A MAP CE is assigned a logical “Public” IPv4 Address
  – The BMR is used to obtain this address by extracting an index value from the EA bits within the end-user IPv6 prefix
  – When a “Public” IPv4 address is shared the EA bits also contain an index value into a set of ports a CE can use for the NAPT function
    - Port-set ID (PSID): The index value for a particular unique set of algorithmically assigned layer-4 ports and ICMP messages with the “Identifier” field
- Given a rule IPv6 prefix length /48, a rule IPv4 prefix length /24, and an EA length of 14 (sharing ratio of 64 to 1):



Benefits of MAP

- Traditional Broadband CPE Paradigm
- Stateless Border Relays
- Minimal Logging Requirements
- Aligns with Native IPv6 Deployments

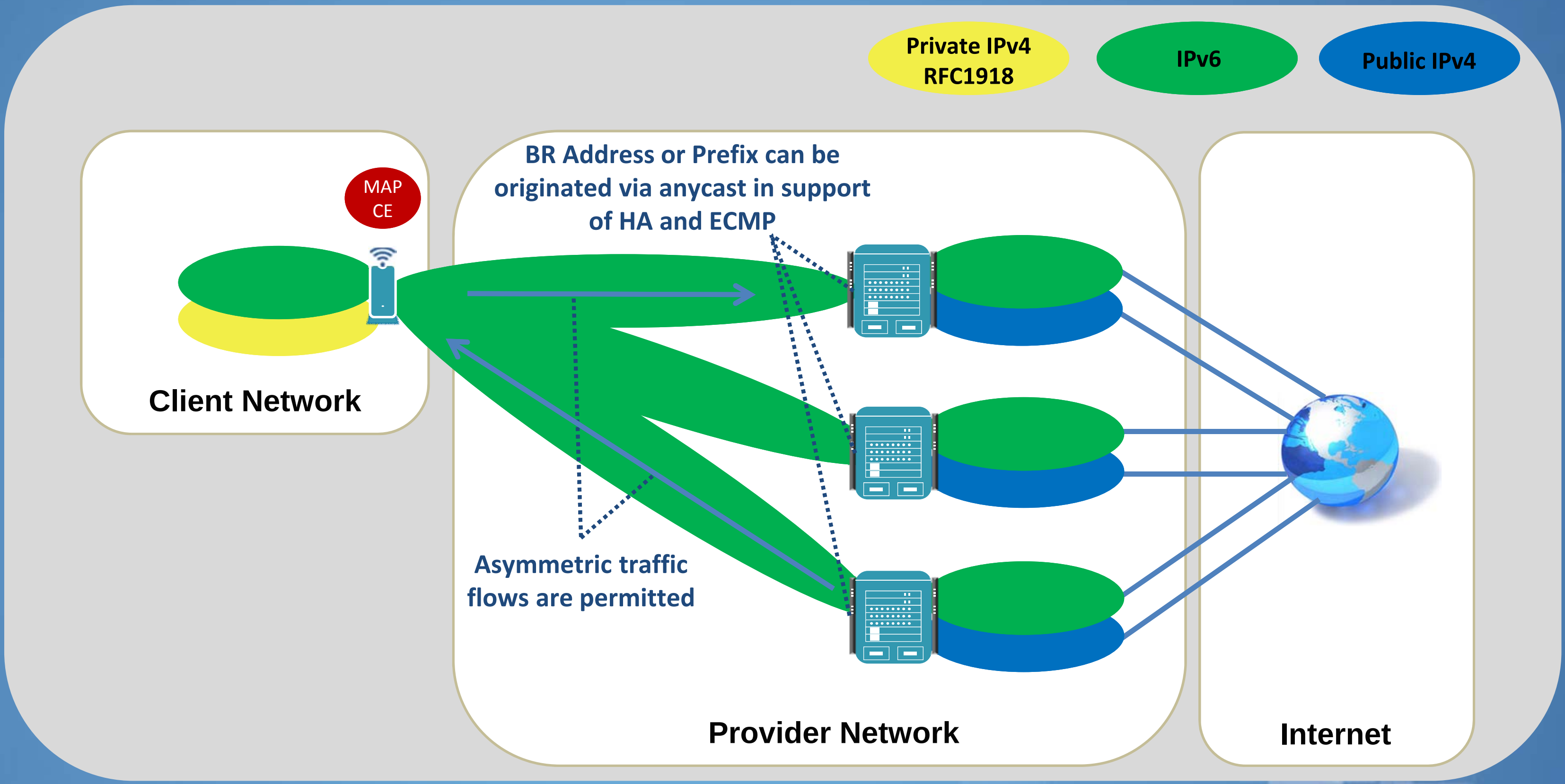
Traditional Broadband CPE Paradigm



- Looks and operates in similar fashion as traditional broadband services only with a reduced number of provisioned layer-4 ports
- Scales better than NAT444 with RFC6598 (100.64.0.0/10) addressing as forwarding to and from MAP CE is with a IPv6 address

Rocky Mountain IPv6 Task Force
RMv6TF

Stateless Border Relays



- Border relays behave more like routers and less like firewalls
- Easier to test than competing stateful LSN approaches (no ALGs, no state)
    - Testing consists of valid unidirectional flows

Minimal Logging Requirements

- Low logging overhead in support of law enforcement and abuse requests
  - Non-shared public IPv4 addresses can be algorithmically mapped back to the assigned IPv6 end-user prefix
  - Shared public IPv4 address and port combination can be algorithmically mapped back to the assigned IPv6 end-user prefix
- Even a network engineer can do it:
  - https://github.com/ejordangottlieb/pyswmap: A Python 3 module I wrote

jgottlieb@demo: ~/examples/pyswmap

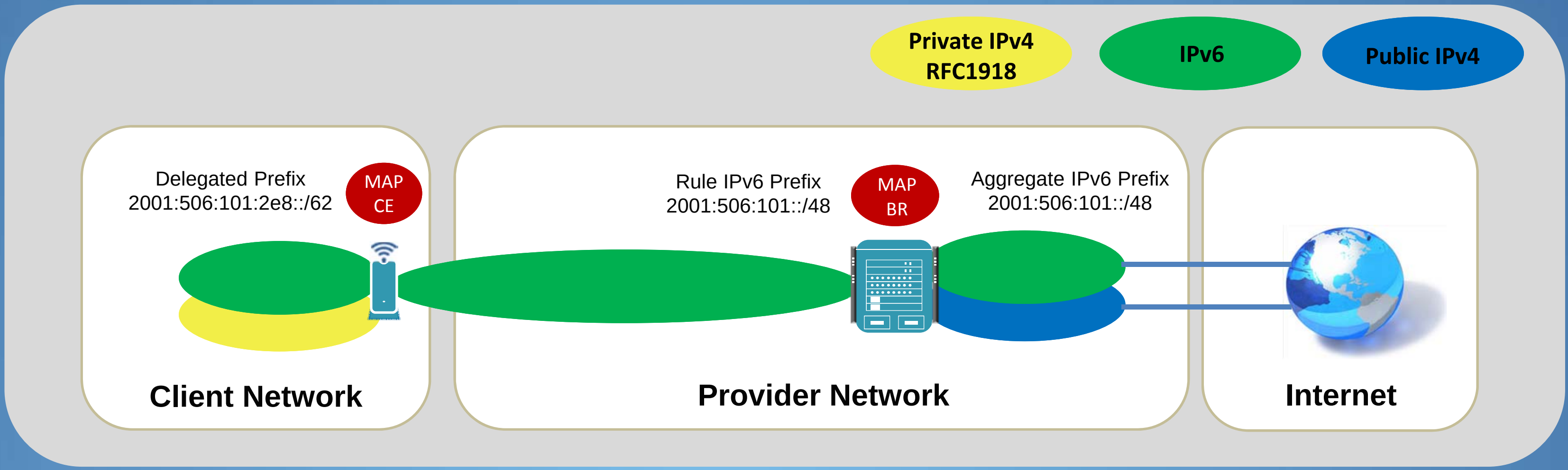
```
jgottlieb@demo:~/examples/pyswmap$ ./bmcheck.py 68.114.2.252 54183



------------------------------------------------
PD for this client is: 2001:506:101:fce8::/62
MAP CE Address is: 2001:506:101:fce8:0:4472:2fc:3a
------------------------------------------------


jgottlieb@demo:~/examples/pyswmap$
```

Aligns with Native IPv6 Deployments



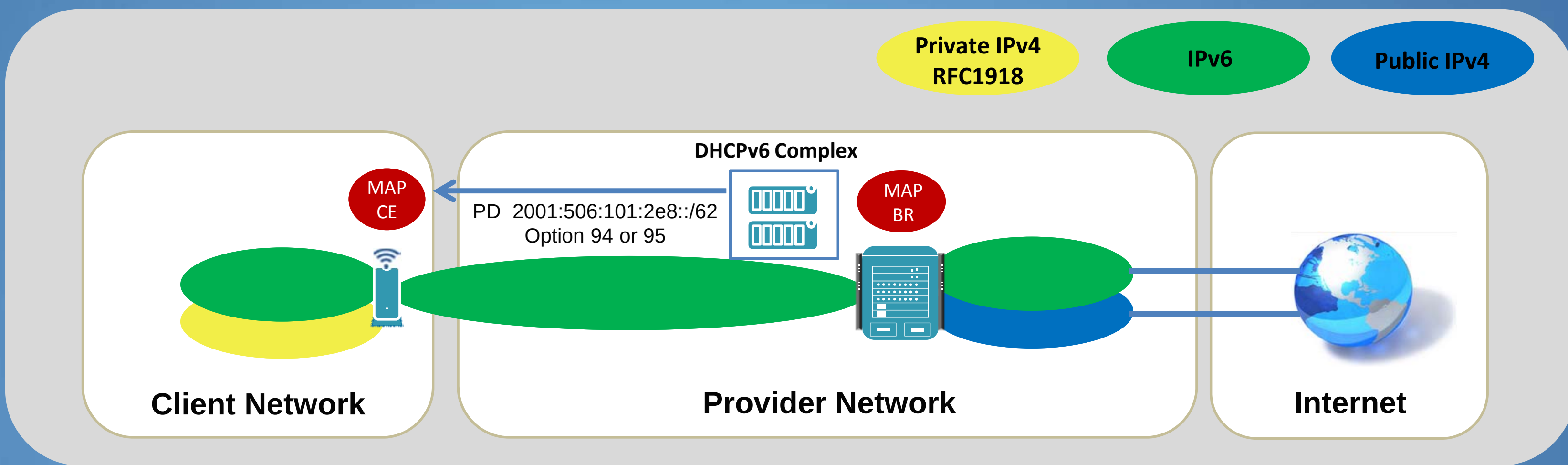
- Basically a native IPv6 deployment where an IPv6 prefix is assigned to the customer LAN
  - Delta: Requires CPE and BR with MAP support and provisioning of MAP domain parameters
- Derives IPv4 address and port allocations without requiring new IPv6 aggregation and customer prefix allocations

Rocky Mountain IPv6 Task Force
RMv6TF

MAP in Action

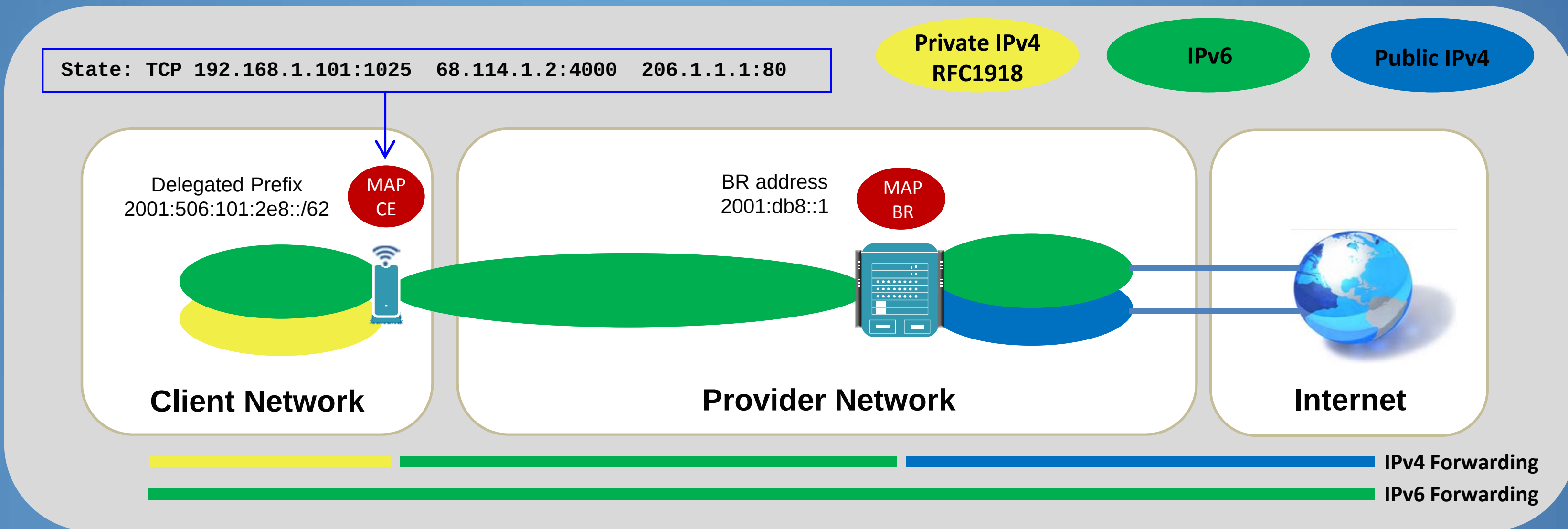
- Using our example PD of 2001:506:101:2e8:/62
  - Rule IPv6 Prefix 2001:506:101::/48
  - Rule IPv4 Prefix 68.114.2.0/24
  - EA length 14 (sharing ratio 64 to 1)
  - PSID 58 (derived from our assigned PD)
- DHCPv6 Provisioning Example
- MAP-E Traffic Flow Example
- MAP-T Traffic Flow Example

DHCPv6 Provisioning Example



- Nearly identical to DHCPv6 with Prefix Delegation
  - Option 94 (MAP-E) and Option 95 (MAP-T) provide domain values
    - Rule IPv6 and IPv4 Prefix
    - EA Length
    - PSID Offset
    - BR address for MAP-E or BR prefix for MAP-T

MAP-E Example



IPv4 Forwarding

1025	80	TCP	192.168.1.101	206.1.1.1

NAPT and IPv6 Encapsulation at MAP CE

4000	80	TCP	68.114.2.2	206.1.1.1

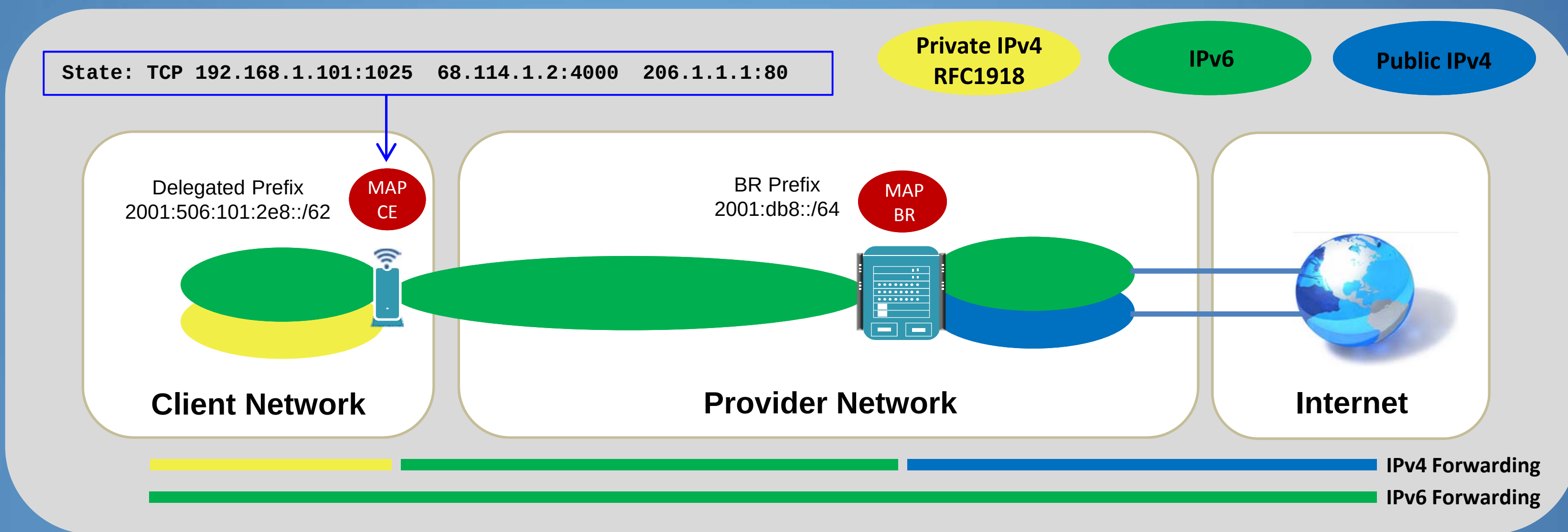
4000	80	TCP	68.114.2.2	206.1.1.1	2001:506:101:2e8:0:4472:202:3a	2001:db8::1

4000	80	TCP	68.114.2.2	206.1.1.1

IPv6 Forwarding

4000	80	TCP	2001:506:101:2e8:0:4472:202:3a	2605:ab:cd::1

MAP-T Example



IPv4 Forwarding

1025	80	TCP	192.168.1.101	206.1.1.1

NAPT and Stateless IPv6 Translation at MAP CE

4000	80	TCP	68.114.2.2	206.1.1.1
4000	80	TCP	2001:506:101:2e8:0:4472:202:3a	2001:db8::ce:101:0100:0

4000	80	TCP	68.114.2.2	206.1.1.1

IPv6 Forwarding

4000	80	TCP	2001:506:101:2e8:0:4472:202:3a	2605:ab:cd::1

MAP Algorithms in Action

- MAP Basic Mapping Rule (BMR)
- PSID Algorithm

MAP Basic Mapping Rule (BMR)

- Using our example PD of 2001:506:101:2e8::/62
  - Rule IPv6 Prefix 2001:506:101::/48
  - Rule IPv4 Prefix 68.114.2.0/24
  - EA length 14 (sharing ratio 64 to 1)

2001:506:101:[0000001011101000]::/62 = 2001:506:101:2e8::/62

68.114.2.0/24 index 2 PSID = 58
= 68.114.2.2

PSID Algorithm

- Using our example PD of 2001:506:101:2e8::/62
  - Rule IPv6 Prefix 2001:506:101::/48
  - Rule IPv4 Prefix 68.114.2.0/24
  - Sharing Ratio 64 to 1
  - PSID Offset: 6

- PSID = **58** from our previous example

  0000001110100000 ← **16 bits representing all possible layer-4 port values**

  A Y M

- Allowed ports = (**A** >= 000001), **Y**=PSID, and any value **M**
- A >= 000001 (because PSID Offset = 6) is => 1024
- A = 000001 and M = 0000 thru A = 111111 and M = 1111
- **1024** + **928** + **0** = **1952** is first available port
- **64512** + **928** + **15** = **65455** is last available port

MAP Deployment Considerations

- IPv4 BR Fragment Handling
- Payload Considerations
- Layer-3 Transparency
- Rule IPv6 Prefix Sizing
- MAP-E or MAP-T?

BR Fragment Handling

- Border relay forwarding is not 100% stateless
  - Initial fragment contains layer-4 information and can be forwarded to the correct CE MAP address using the mapping algorithm
  - Subsequent packet fragments do not include layer-4 information and must be forwarded by IPv4 destination and fragment ID which requires state maintenance

Payload Considerations

- Encapsulation and translation as transport mechanism come with varying levels of overhead
  - MAP-E overhead is >= 40 bytes over original transmitted IPv4 packet
  - MAP-T overhead is >= 20 bytes over original transmitted IPv4 packet
- Jumbo frame support in the IPv6 domain is desirable
  - A MAP CE with a 1500 byte LAN MTU and a jumbo WAN MTU can mimic a traditional IPv4 1500 MTU broadband environment

Rocky Mountain IPv6 Task Force
RMv6TF

Layer-3 Transparency

- Layer-3 Transparency refers to the ability to retain all the values stored in the IPv4 header
  - MAP-E achieves full transparency by encapsulating the IPv4 header
  - MAP-T achieves a high level of transparency by mapping to equivalent IPv6 header values through address translation techniques
- MAP-T Propagation of DF and MF Flags
  - RFC6145 describes stateless translation algorithms used by MAP-T
    - DF Bit = 0 include a IPv6 fragment header in translation process
    - DF Bit = 1 do not include a IPv6 fragment header in translation process
    - MF Bit = 1 implied when IPv6 fragment header is present
    - No way to do DF=1 and MF=1 which is a recommendation made in RFC4821 section 8

Rule IPv6 Prefix Sizing

- MAP domain rule IPv6 scope and sharing ratio drives the IPv4 prefix requirement
  - MAP Domains defined on top level IPv6 aggregation boundaries will require a large contiguous IPv4 block
    - A domain defined on a regional boundary with a /32 IPv6 prefix allocation, a /62 user prefix, and a 64:1 ratio requires a /7 IPv4 prefix allocation
  - Smaller IPv4 prefix allocations drive more MAP domain instances per region
    - A domain defined with a /46 IPv6 prefix allocation, a /62 user prefix allocation, and 64:1 ratio requires a /22 prefix allocation

Rocky Mountain IPv6 Task Force

RMv6TF

MAP-E or MAP-T

- I am an advocate of MAP-T because of the full 5-tuple visibility which allows
  - Traditional deployment of edge filters and classifiers
  - Traffic engineering for IPv4 destinations outside the MAP domain
  - The ability to deploy IPv4-mapped IPv6 addressed servers and services that support IPv4 customer endpoints. BR is not necessary.
  - Superior levels of granularity for link aggregation and ECMP
  - Utilization of DoS countermeasures based on BGP Flowspec

Resources and Attributions

- Resources
  - http://map46.cisco.com/MAP.php Cisco MAP Simulation Tool and Other Links
  - https://github.com/ejordangottlieb/pyswmap A MAP Module for Python 3
  - http://enog.jp/~masakazu/vyatta/map/ Vyatta based MAP BR or CE
- Attributions
  - https://tools.ietf.org/html/draft-ietf-softwire-map-t MAP-T
  - https://tools.ietf.org/html/draft-ietf-softwire-map MAP-E
  - https://tools.ietf.org/html/draft-ietf-softwire-map-dhcp MAP DHCP Options
  - https://tools.ietf.org/html/rfc6052 IPv6 Addressing of IPv4/IPv6 Translators
  - https://tools.ietf.org/html/rfc6145 IP/ICMP Translation Algorithm
  - https://tools.ietf.org/html/rfc4821 Packetization Layer Path MTU Discovery
  - http://www.cisco.com/c/en/us/solutions/collateral/ios-nx-os-software/enterprise-ipv6-solution/whitepaper_C11-729800.html Cisco Technical Guide to Mapping of Address and Port