

Mapping Address + Port

Martin Gysi, 26.10.2015

viele Slides © Cisco

Die 2 (*) Herausforderungen der Service Provider

- IPv6 ausrollen.
 - Gebt den Hennen ihre Eier
 - So viel Verkehr wie möglich soll so rasch wie möglich weg von IPv4
 - IPv4 Adressen sparen.
 - NAT, IPv4 Adressen multiplexen, d.h. mehrfach vergeben.
-
- (*) Es soll noch ein paar weitere geben...

Wie macht Swisscom das?

- IPv6: Mit RFC 5969, 6rd

Internet Engineering Task Force (IETF)
Request for Comments: 5969
Category: Standards Track
ISSN: 2070-1721

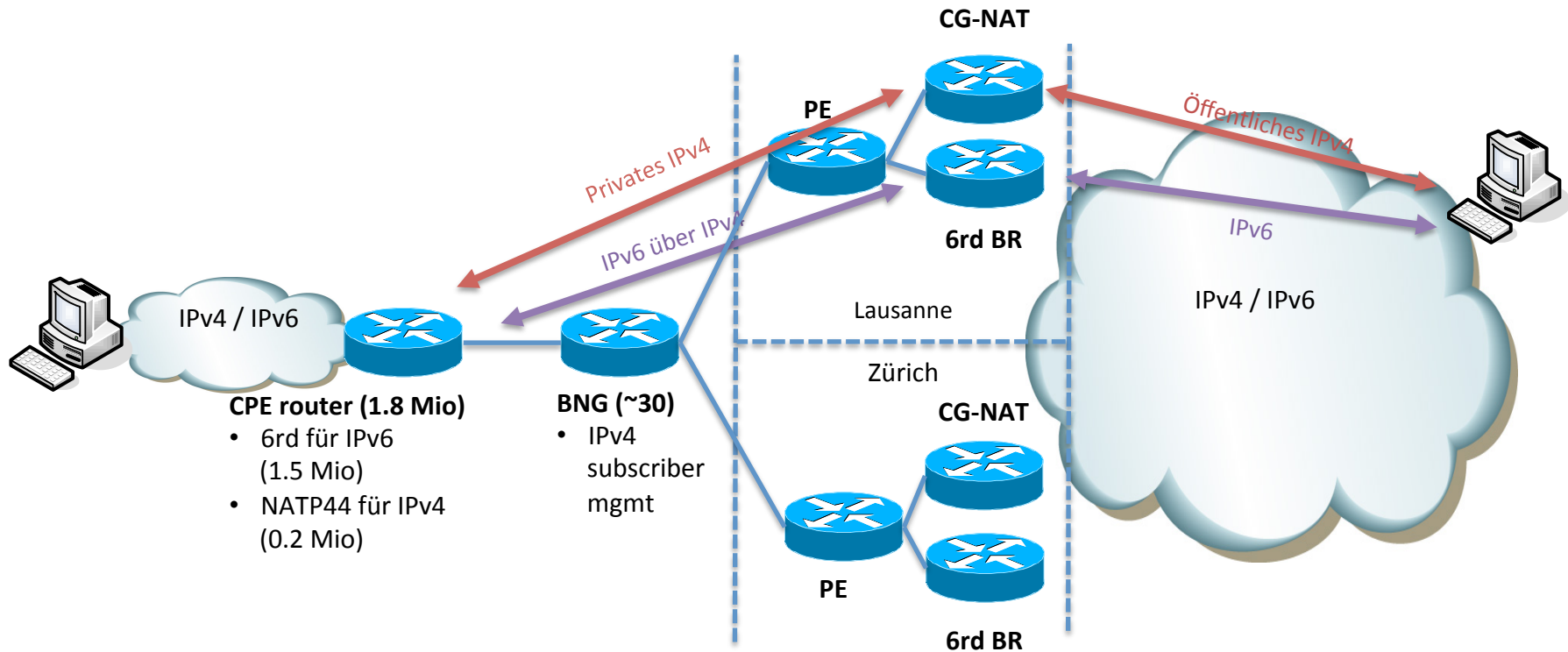
W. Townsley
O. Troan
Cisco
August 2010

IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) --
Protocol Specification

- IPv4: NAT444 (doppeltes NAT)
 - Beim Kunden von RFC 1918 auf RFC 6598 (100.64.0.0/10) Adressen
 - Beim Provider von RFC 6598 auf öffentliche Adressen

Netzdesign

- BR und CG-NAT sind an zentralen Orten redundant platziert



6rd funktioniert zustandslos

- Das IPv6 Präfix wird aus der IPv4 Adresse generiert, mit einer einfachen Regel

Swisscom 6rd Parameter

6rd prefix: 2a02:1200::/28

Border Relay Adresse: 193.5.29.1 (6rd.swisscom.com)

IPv4 Adresse: Alle 32 bits enthalten



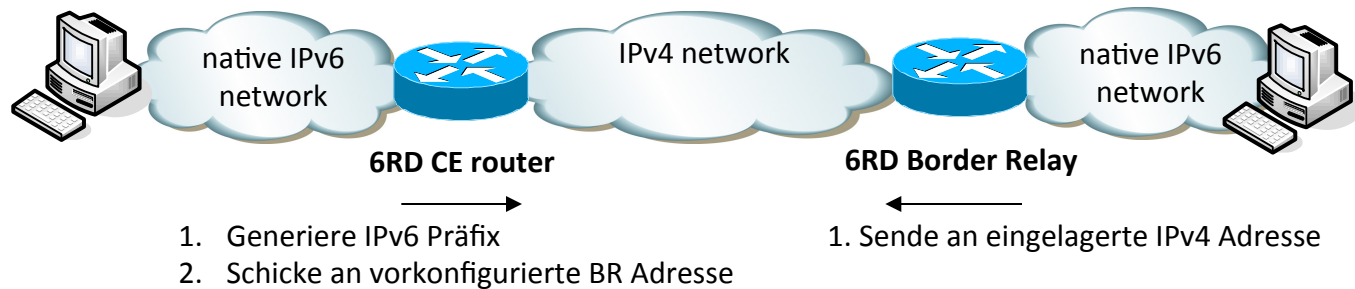
Wird vorgängig
provisioniert

- Mit dieser Regel
 - Generiert das CPE IPv6 Präfixe, welche seine IPv4 WAN Adresse enthalten
 - Schickt das CPE IPv6 Pakete an die IPv4 Adresse des BR
 - Schickt der BR IPv6 Pakete an die IPv4 Adresse des CPEs
- Der BR führt dies, wie normales Routing, in „Hardware“ aus
 - Nur der tatsächliche Verkehr limitiert den Durchsatz

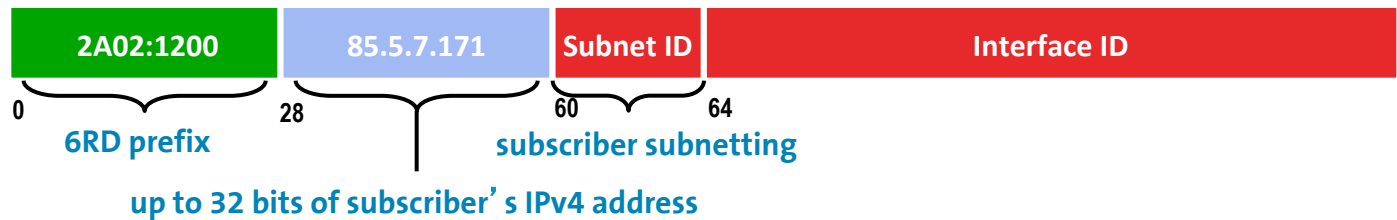
6RD is a stateless tunnel technology, embedding the CE's IPv4 address into the IPv6 prefix.

IPv6 Rapid Deployment on IPv4 Infrastructures (RFC 5969)

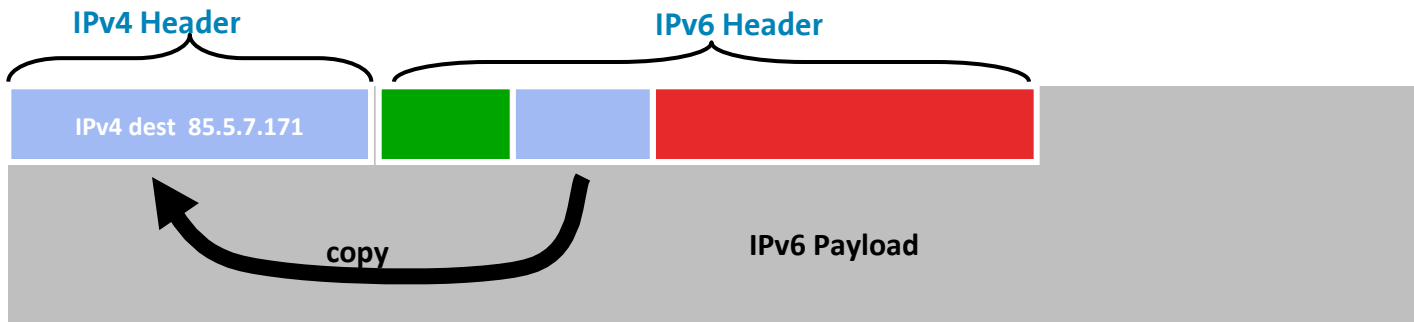
Network topology



IPv6 address format for 6RD

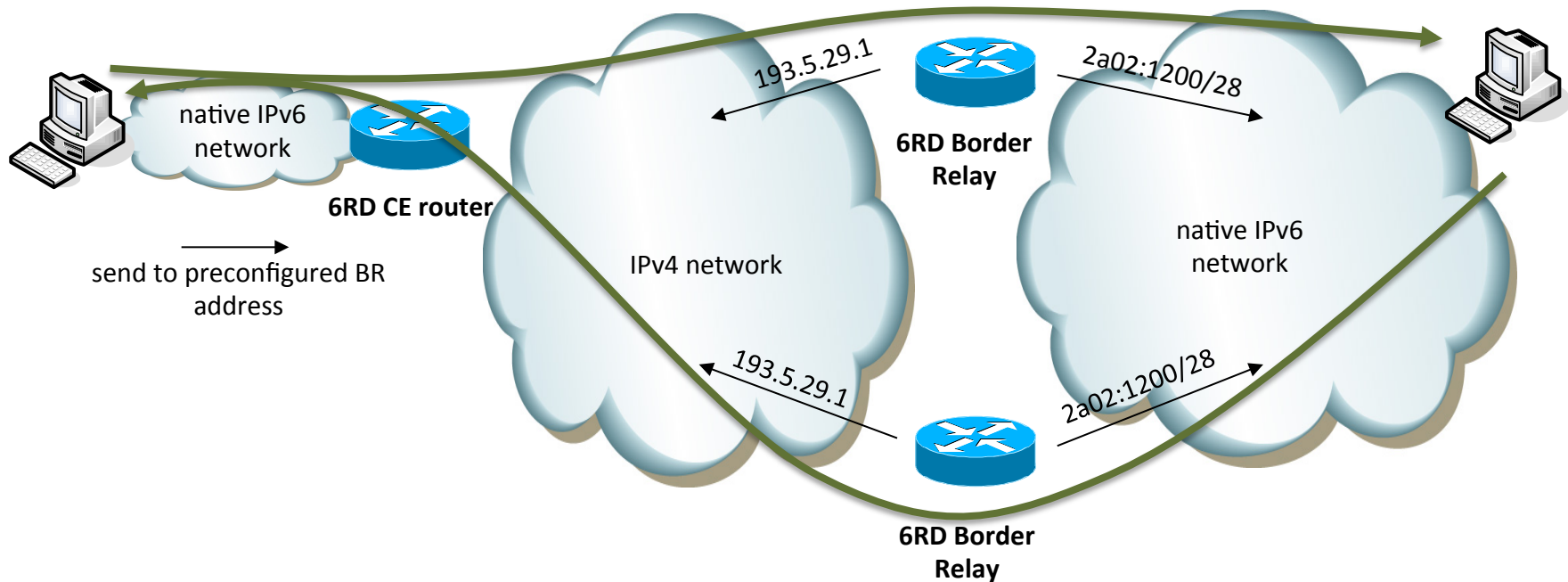


IPv4 header & encapsulated IPv6 packet (downstream)



Redundanz mit 6rd

- Kein Problem: Es gibt ja keinen Zustand auf dem BR und dem CPE.

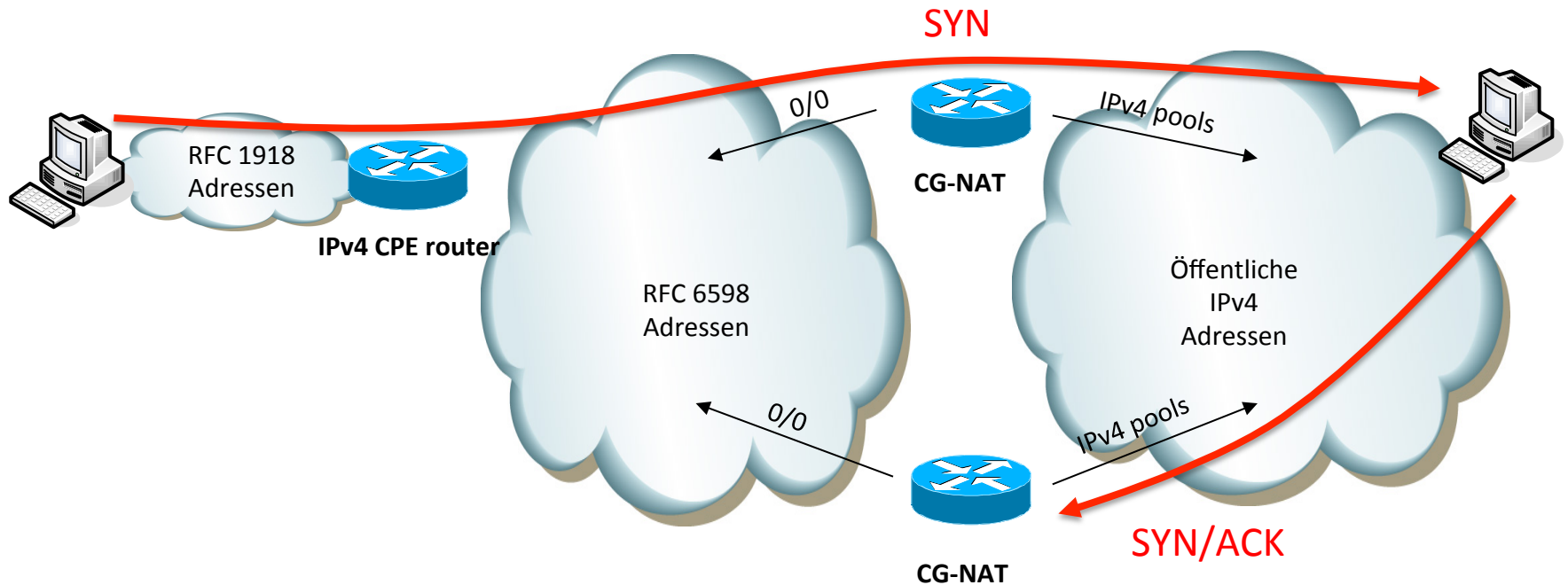


NAT444 ist zustandsbehaftet

- Wie können viele private Adressen auf wenige öffentliche Adressen abgebildet werden?
- Traditionelles NAT: Führt eine dynamische Tabelle von Quell- und Zieladresse, Quell- und Zielport
 - Limitierter Sessionsaufbau (new sessions/s)
 - Limitierte Anzahl Sessions
 - Jede Session muss geloggt werden

Redundanz mit NAT444

- Leider nein, das untere NAT kennt den Zustand des oberen nicht.



- Entweder aktiv / standby und Neuaufbau aller Sessions bei Failover,
- Oder aktive Statussynchronisation zwischen den Geräten

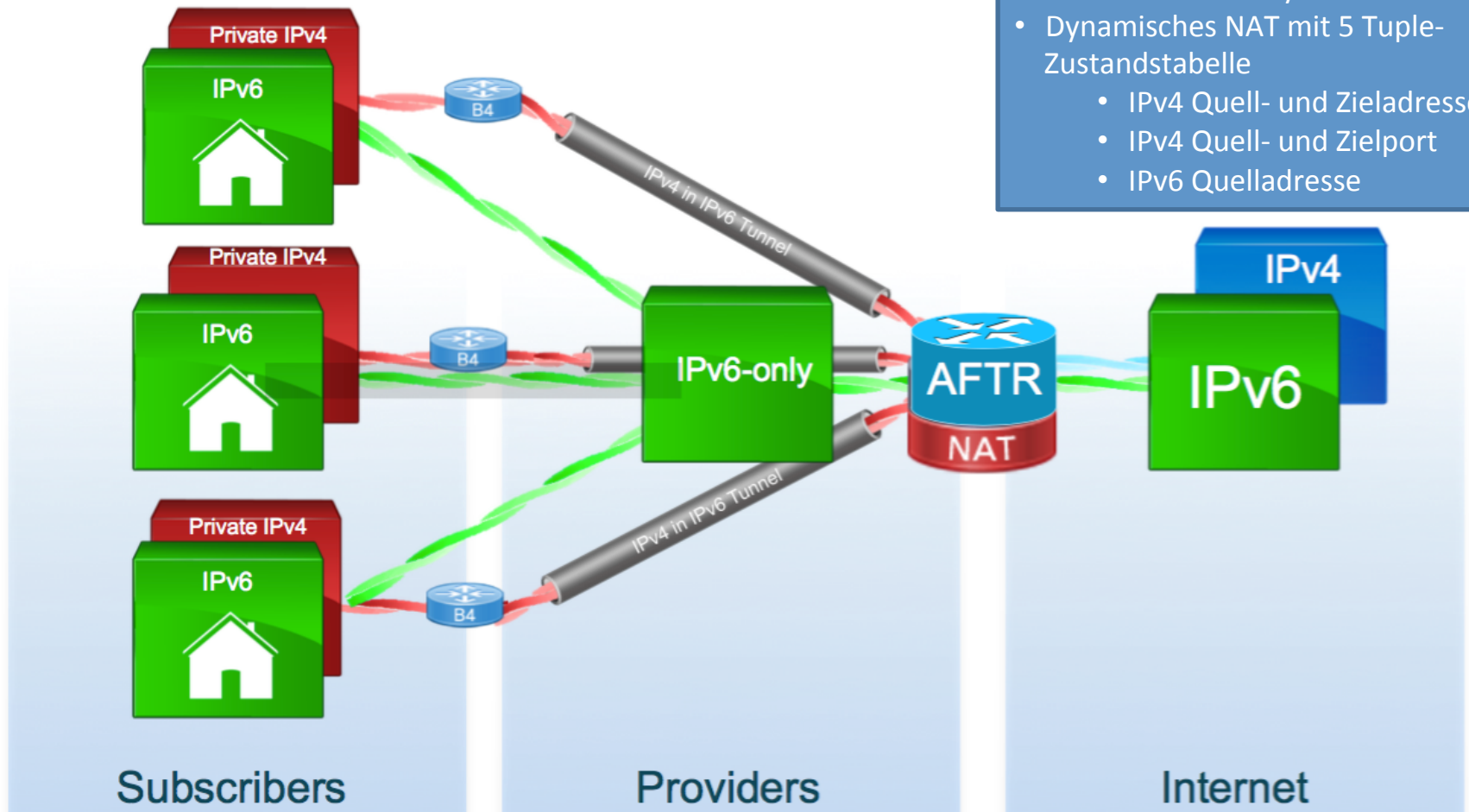
Zustandslos versus zustandsabhängig

- Zustandslos / stateless: Kein dynamischen, durch den Verkehr verursachte Zustände. Deswegen aber nicht konfigurationslos!
 - IP Routing, 6rd- und MAP-Border Relay
- Zustandsabhängig / stateful: Verkehr erzeugt dynamisch Status, welcher unterhalten und wieder abgebaut werden muss.
 - Firewall, NAT44, DS-lite, Session Border Controller
- Eigenschaften zustandsloser Protokolle : Bessere Skalierbarkeit, grössere Flexibilität
- Kosten für 1 Gb/s Durchsatz:
 - 6rd: 1650 Fr.
 - CG-NAT: 8000 Fr. (ohne Kosten für Logging)

Wie kann eine Lösung basierend auf IPv6 aussehen?

- Ziele:
 - Nur noch IPv6 im Netz
 - IPv4 als Service über IPv6
 - Integrierte Lösung fürs IPv4-Adresssparen (NAT)
 - Nur zustandslose Protokolle (soll)
- Aktuell gibt es drei Lösungsansätze welche die ersten drei Punkte erfüllen:
 - DS-Lite (RFC 6333)
 - Lightweight 4 over 6 (RFC 7596)
 - MAP (RFC 7599, RFC 7597)

DS-Lite



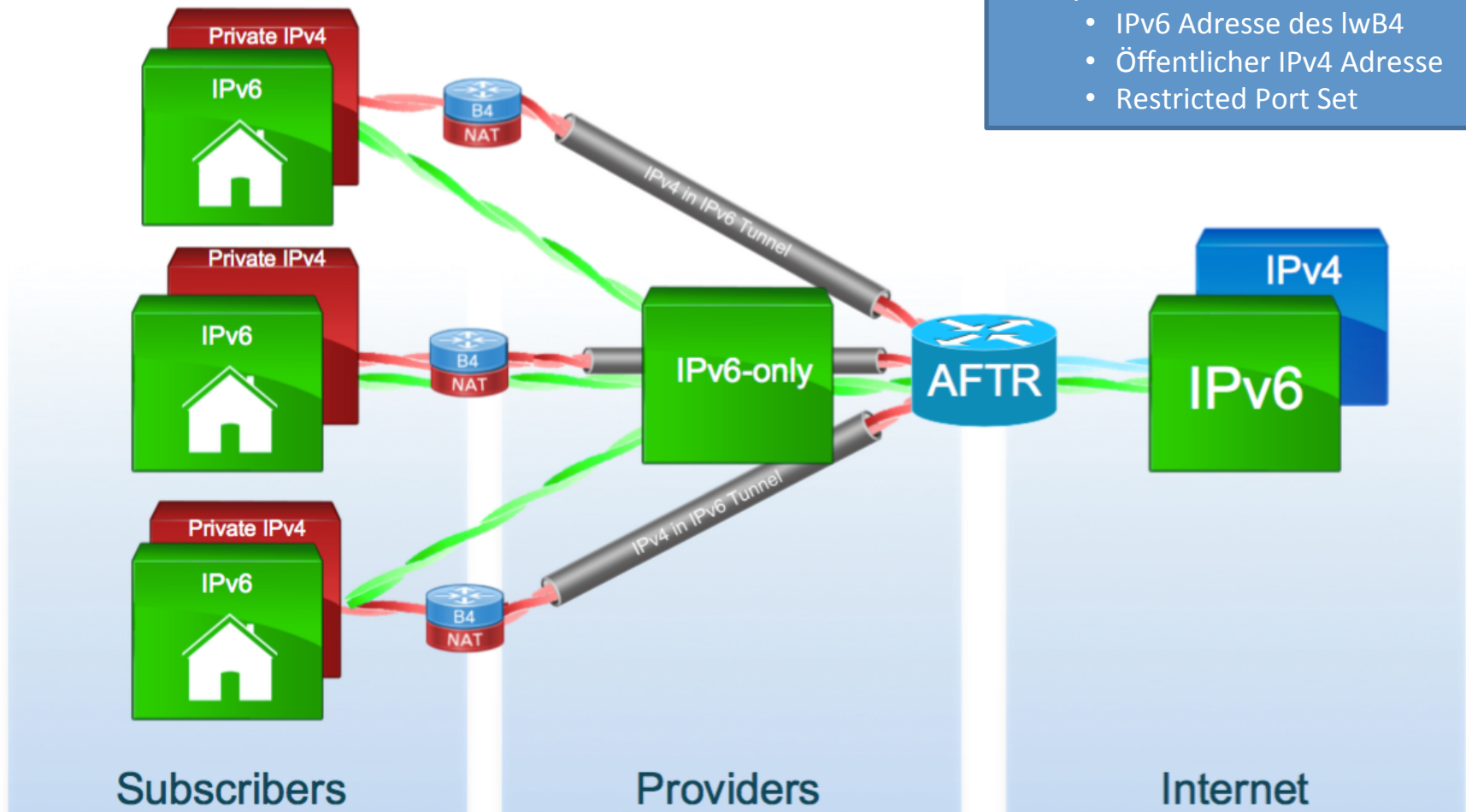
B4: Basic Bridging BroadBand Element

- Schickt privat adressierte IPv4 Pakete direkt an AFTR (Netzwerkbrücke)
- Kein NAT

AFTR: Address Family Transition Router

- Dynamisches NAT mit 5 Tuple-Zustandstabelle
 - IPv4 Quell- und Zieladresse
 - IPv4 Quell- und Zielport
 - IPv6 Quelladresse

Lightweight 4 over 6



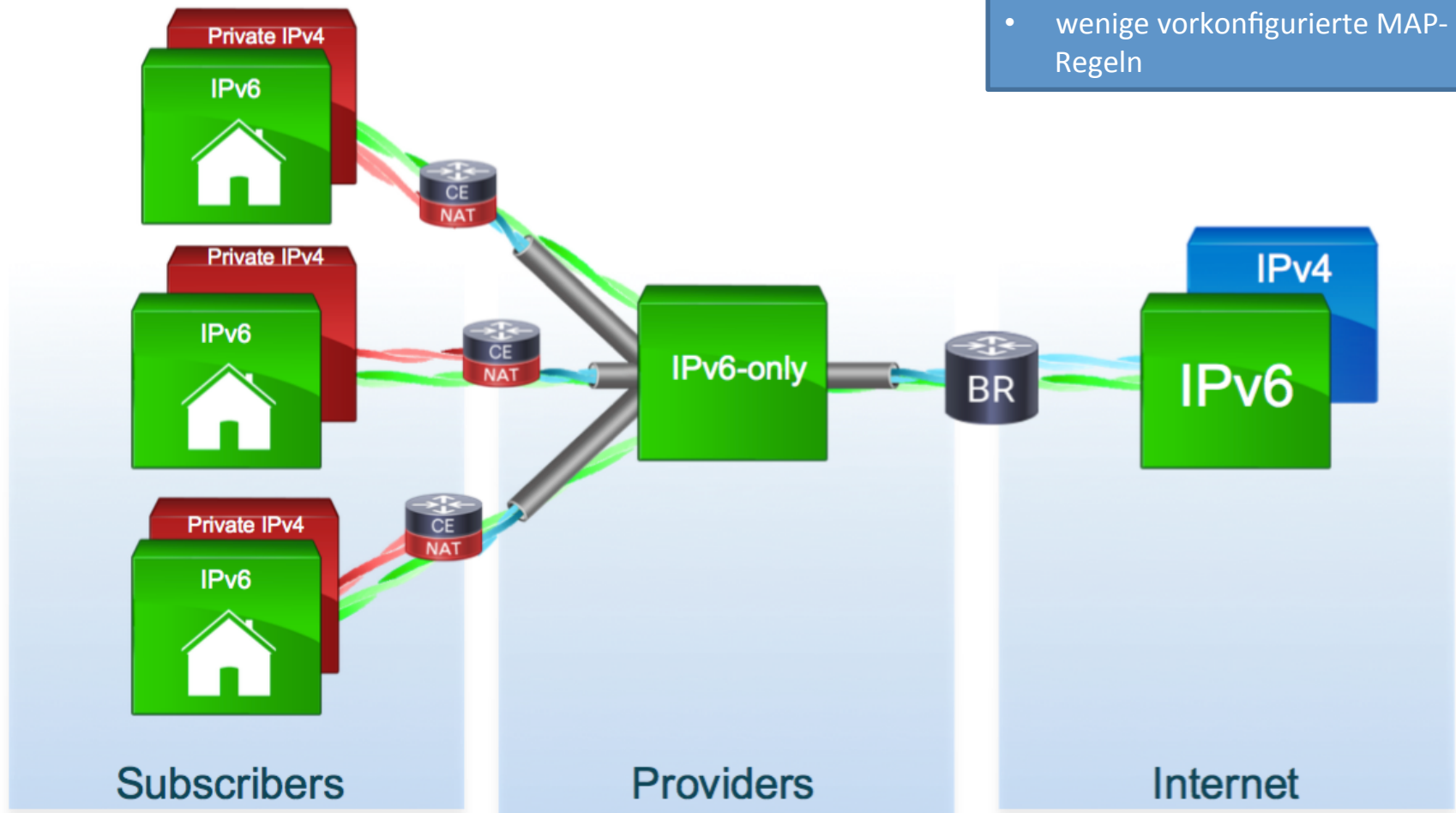
lwB4:

- dynamisches Port-restricted NAT

lwAFTR:

- dynamische Zustandstabelle mit
 - IPv6 Adresse des lwB4
 - Öffentlicher IPv4 Adresse
 - Restricted Port Set

MAP (-T, -E)



MAP-CE:

- dynamisches Port-restricted NAT
- zustandsloses NAT46

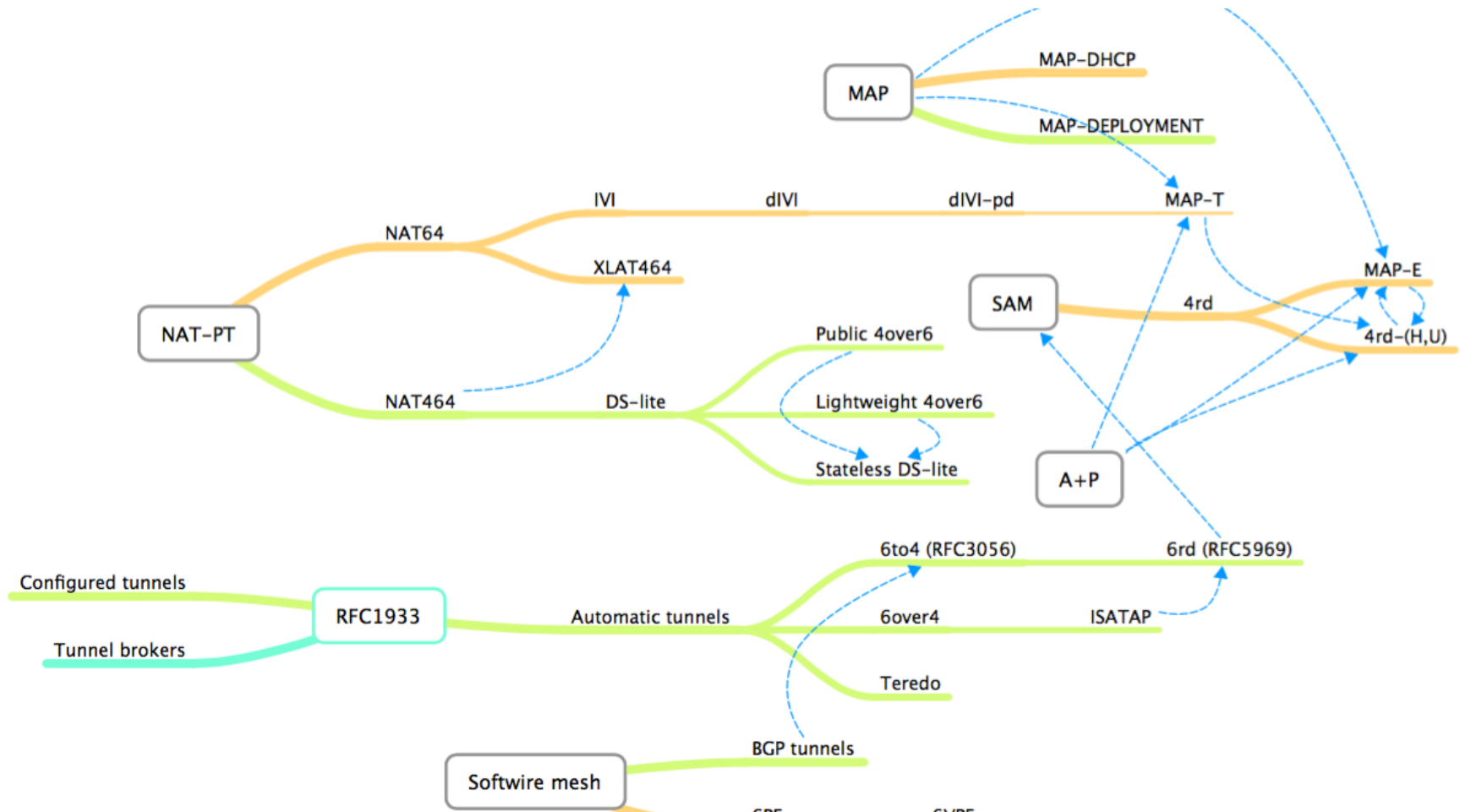
MAP-BR:

- wenige vorkonfigurierte MAP-Regeln

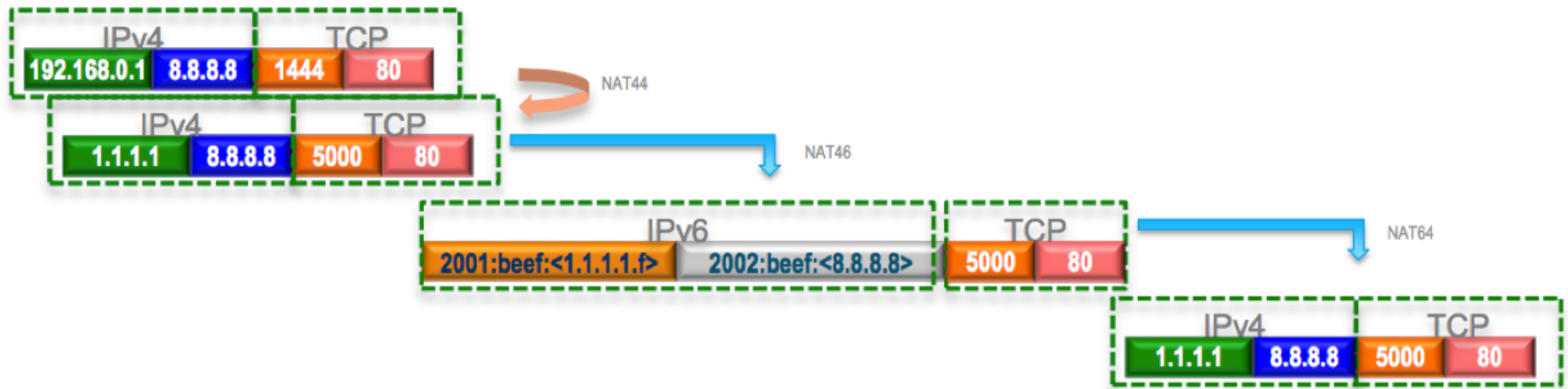
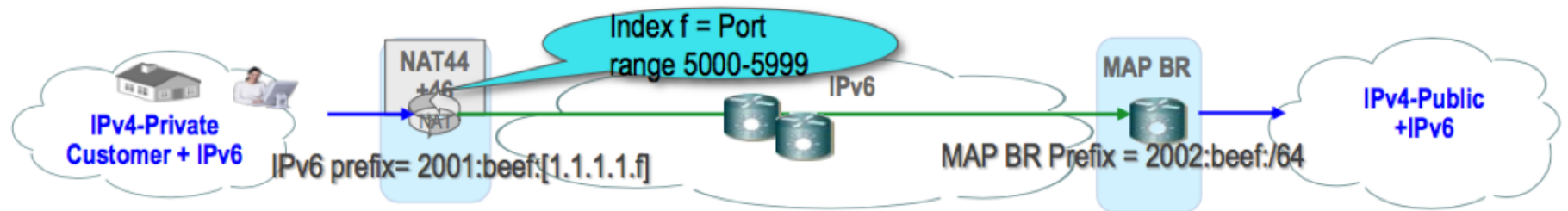
Lösungen im Überblick

	Lösungen	IPv6 Forwarding	IPv4 Forwarding	ISP Gateway (Stateful)	Home NAT (Stateful)	Home Networks
1	Dual-Stack E2E	Native	Native	Yes*	Yes	Dual-stack
2	6rd (v6 over v4)	Tunneled	Native	Yes*	Yes	Dual-stack
3	NAT64 (IVI)	Native	Translated	Yes	-NA-	Single-stack
4	DS-Lite	Native	Tunneled	Yes	No	Dual-stack
5	MAP-T	Native	Translated	No	Yes	Dual-stack
6	MAP-E	Native	Tunneled	No	Yes	Dual-stack
7	LW 4o6	Native	Tunneled	Yes	Yes	Dual-stack
8	CGN44/4	-NA-	Translated	Yes	Yes	Single-stack

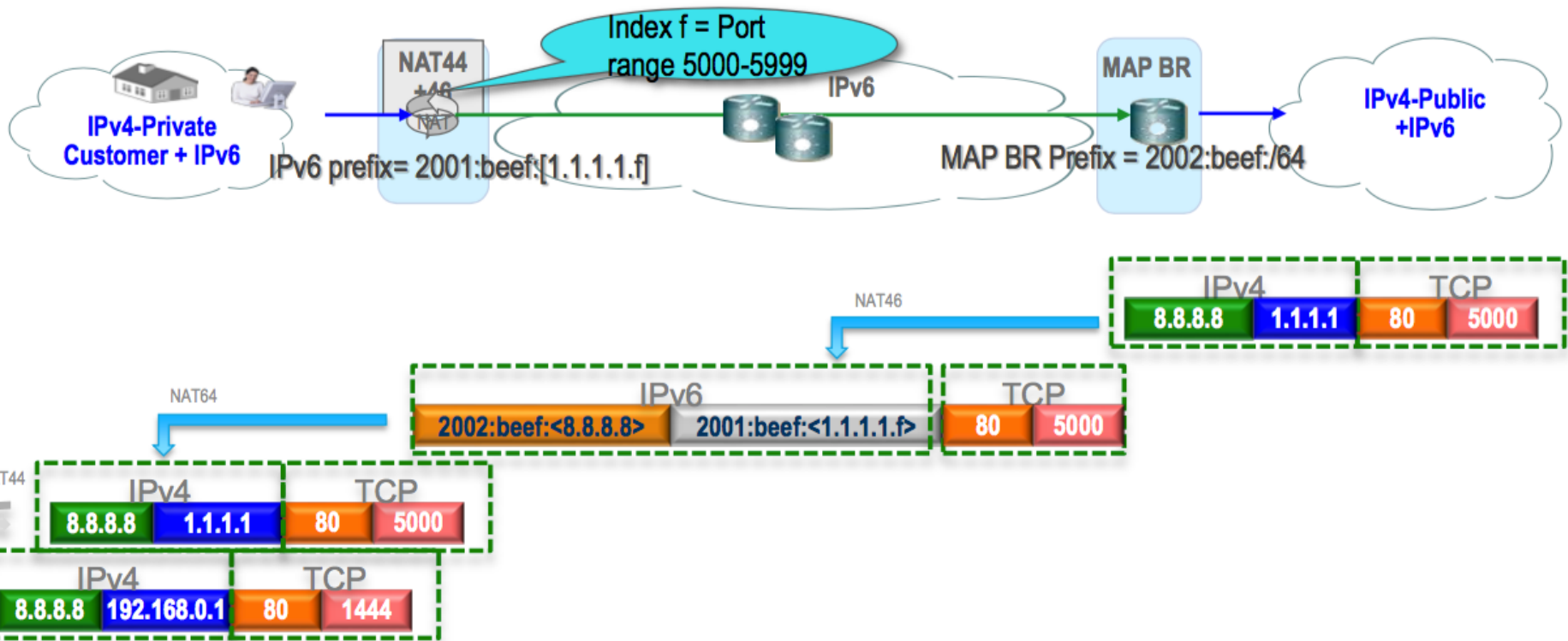
Die Evolution der Lösungen



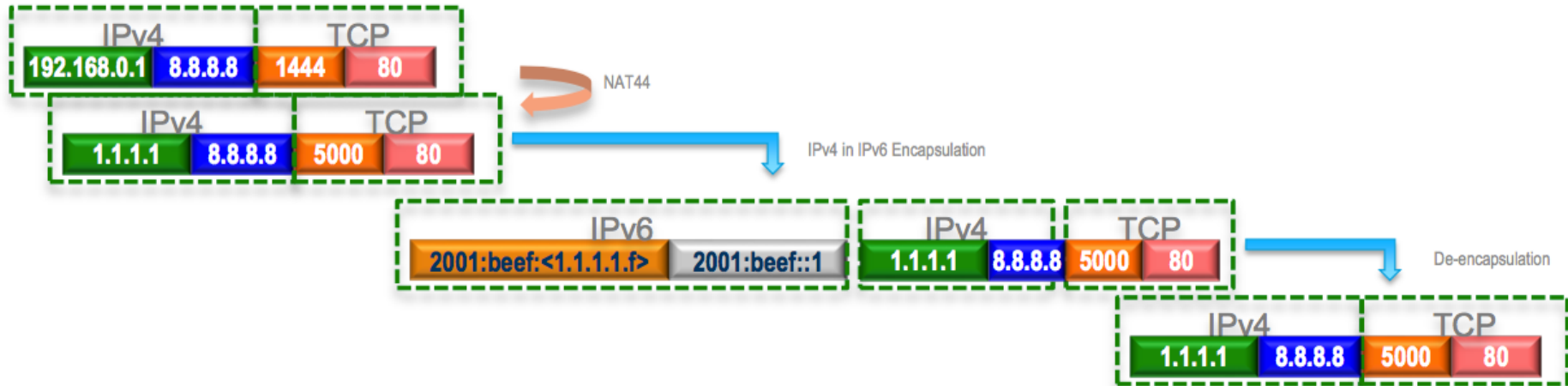
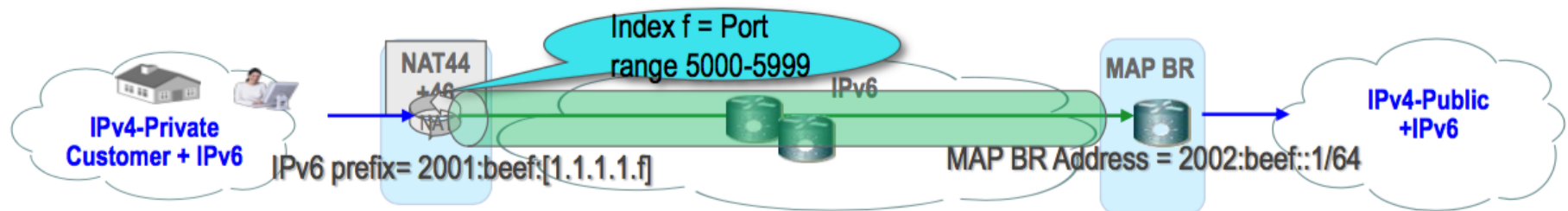
Upstream-Paketweiterleitung bei MAP-T



Downstream-Paketweiterleitung bei MAP-T

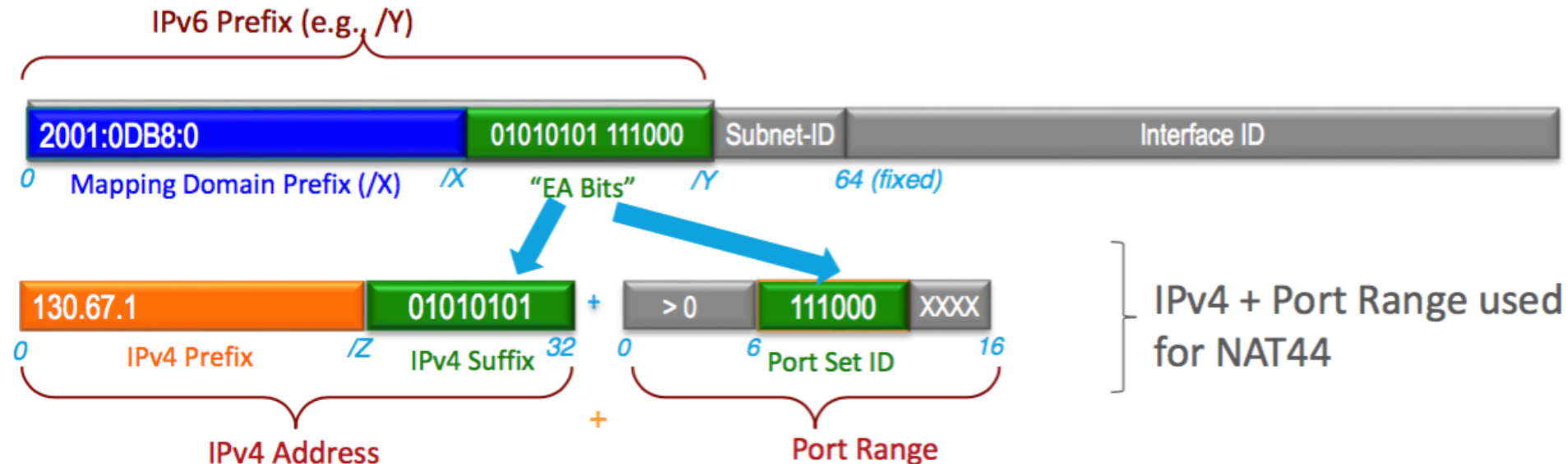


Paketweiterleitung bei MAP-E



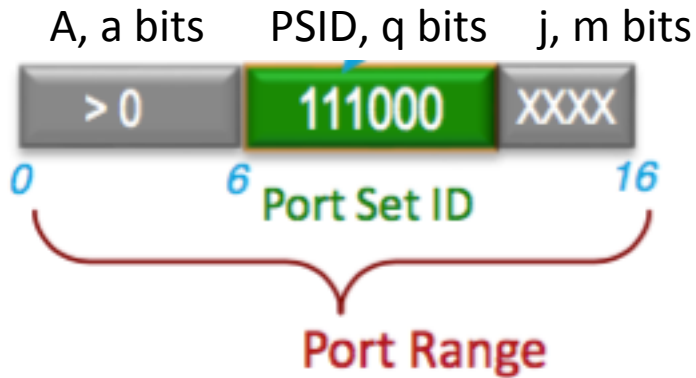
Wie funktioniert das Mapping zwischen IPv6 und IPv4?

- Ein CPE bekommt mittels DHCP-PD ein IPv6 Präfix der Länge /Y.
- Woher kommen die IPv4 Adresse und der Portbereich? → Basic Mapping Rule

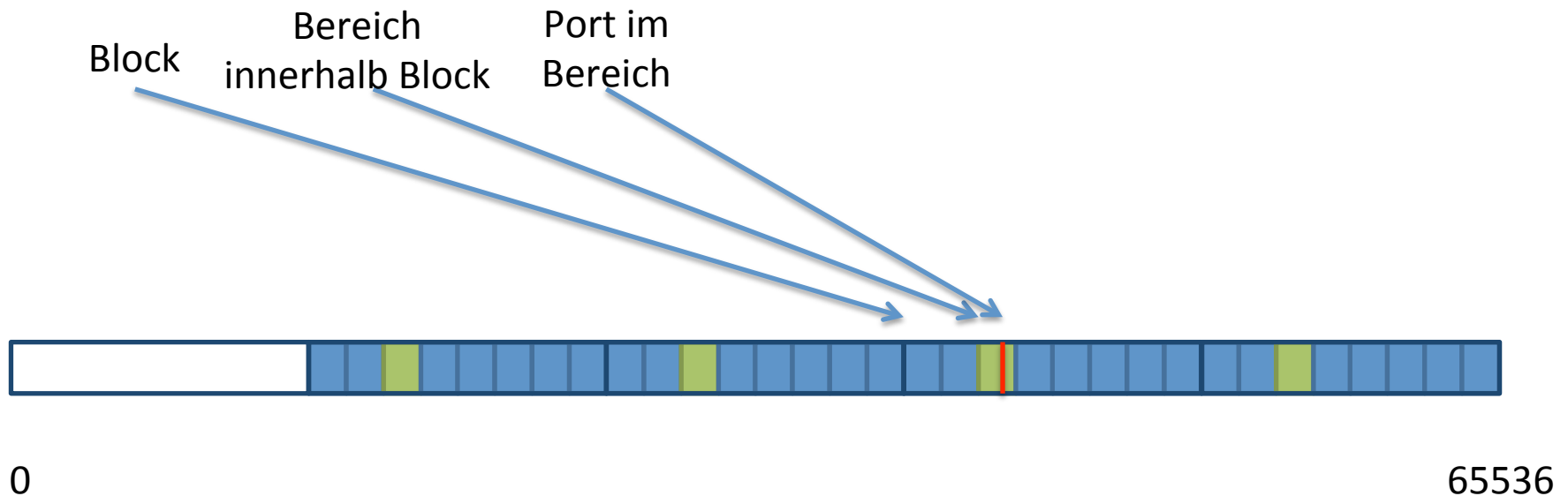


- Mittels DHCP Option (OPTION_S46_RULE) [oder TR-69 oder Netconf/Yang]:
 - Länge der „EA Bits“, /Y - /X: ea-len [0 .. 48]
 - IPv4 Prefix: ipv4-prefix, /Z: prefix4-len [0 .. 32]
 - Mapping Domain Prefix: ipv6-prefix, /X: prefix6-len

Mehr zu den Port-Bereichen



- Ein CPE darf nur Ports in den grünen Bereichen verwenden. Diese haben alle dieselbe Port Set ID
- Um Ports 0 – 1023 nicht zu verwenden, ist a 6 bit lang und darf nicht Null sein.



$$\text{Port} = A * 2^a + \text{PSID} * 2^q + j$$

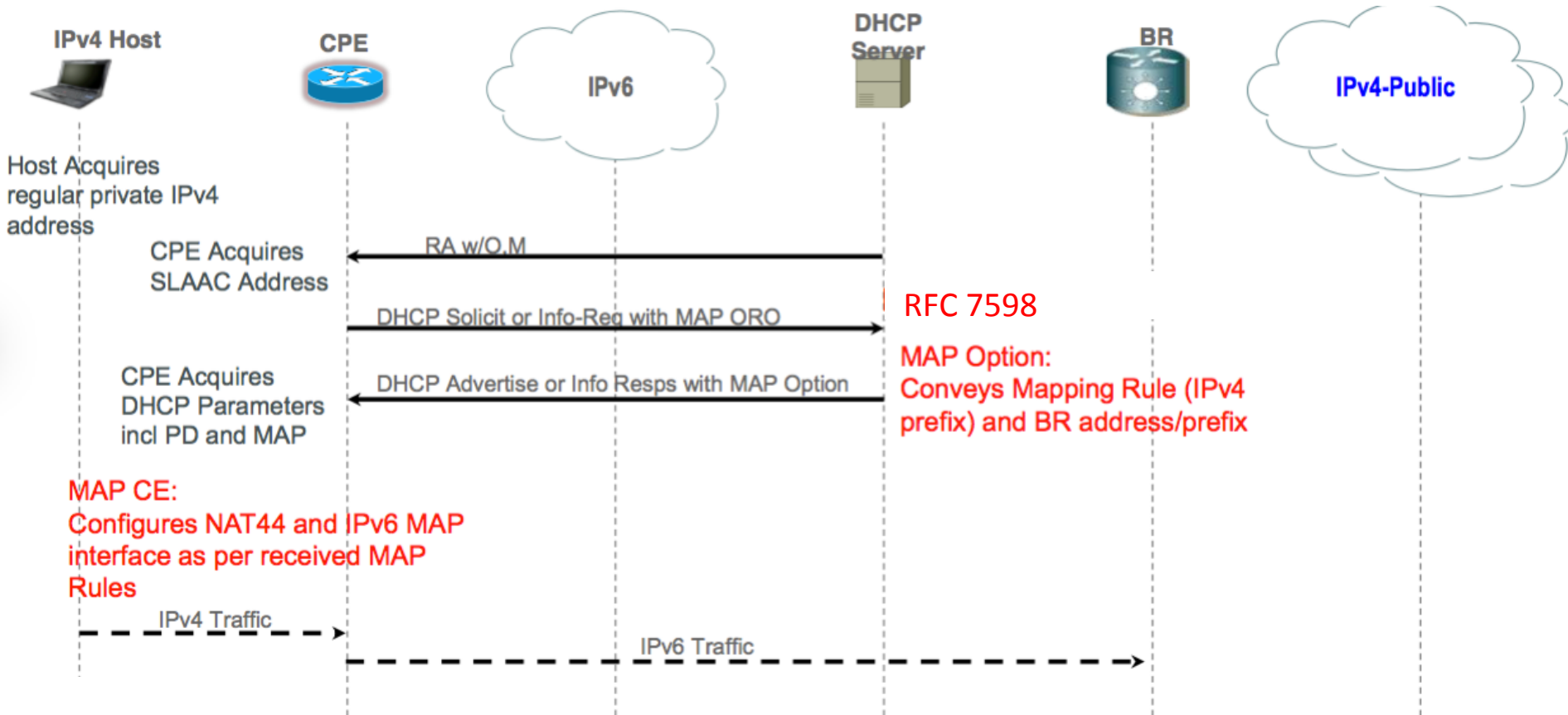
Wie funktioniert das Mapping zwischen IPv6 und IPv4?

Wohin soll ein CPE seinen Verkehr schicken? Was ist seine IPv4 Default Route?

MAP-T: Default Mapping Rule (DMR) definiert das IPv6 Präfix des Border Relays.

- Mittels DHCP Option (OPTION_S46_DMR)
 - BR Präfix: dmr-ipv6-prefix, BR Präfixlänge: dmr-prefix6-len

Provisionierung über DHCP



Provisionierung des BR

- Entweder über Netconf/Yang
- Oder CLI (hier ASR9k):

$$q = \log_2(16) = 4$$

$$m = \log_2(1024) = 10$$

$$\rightarrow a = 16 - q - m = 2$$

DMR Präfix

```
interface ServiceApp1
service cgn MAP-T-1 service-type map-t
ipv4 address 192.168.66.1 255.255.255.252
!
interface ServiceApp2
service cgn MAP-T-1 service-type map-t
ipv6 address 2001:918:ffff:ffe1::1/126
!
service cgn MAP-T-1
service-location preferred-active 0/0/CPU0
service-type map-t MAP-T-Instance-1
  cpe-domain ipv4 prefix 100.90.0.0/24
  cpe-domain ipv6 prefix 2001:918:ffff:5000::/52
  sharing-ratio 16
  contiguous-ports 1024
  external-domain ipv6 prefix 2001:918:ffff:100::/64
  address-family ipv4
    interface ServiceApp1
    !
  address-family ipv6
    interface ServiceApp2
    !
  !
!
!
router static
address-family ipv4 unicast
  100.90.0.0/24 ServiceApp1
!
address-family ipv6 unicast
  2001:918:ffff:100::/64 ServiceApp2 2001:918:ffff:ffe1::2
end
```

Provisionierung des BR

- ASR1k

DMR Präfix

$$q = \log_2(128) = 7$$

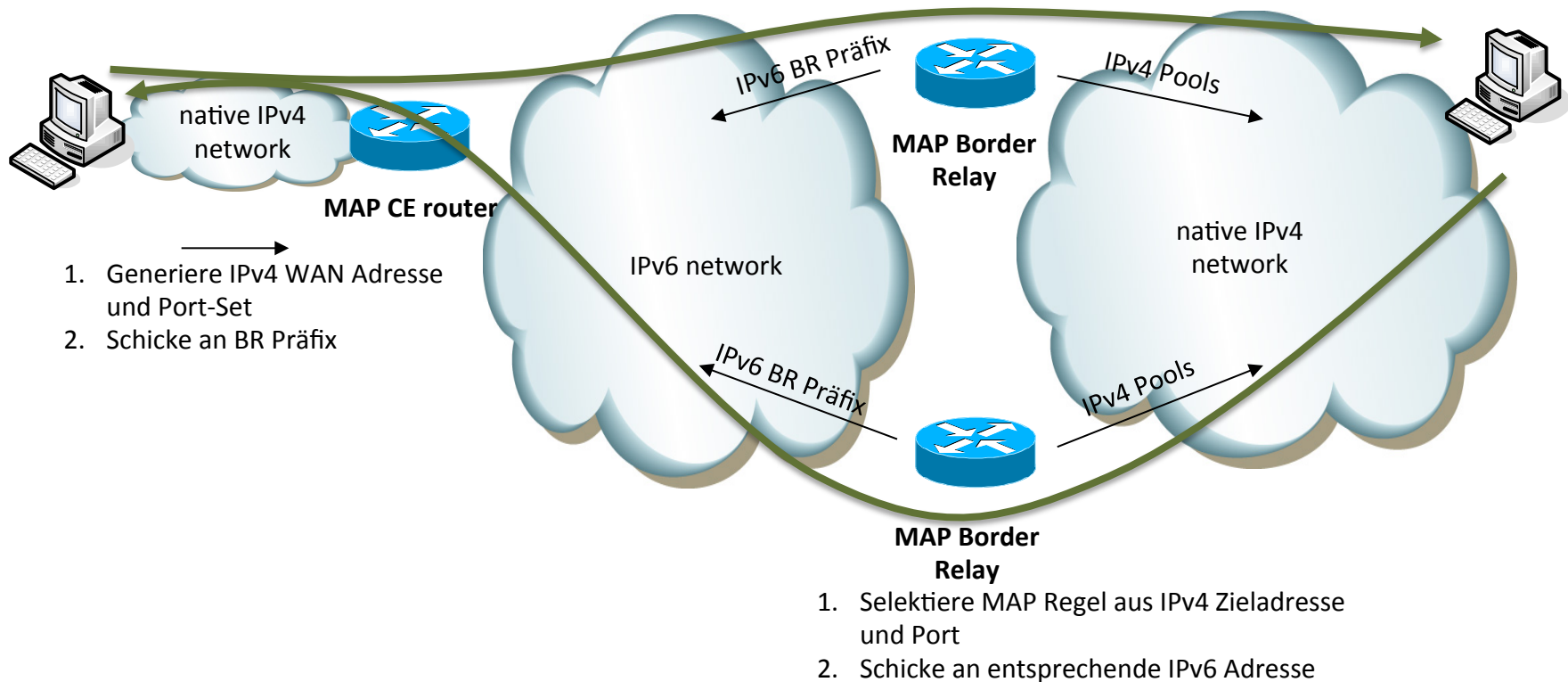
$$a = 16 - \log_2(1024) = 10$$

$$\rightarrow m = 16 - q - a = 3$$

```
nat64 map-t domain 1
default-mapping-rule 2001:918:FFFF:12::/64
basic-mapping-rule
  ipv6-prefix 2001:918:FFFF:3000::/52
  ipv4-prefix 194.209.152.60/31
  port-parameters share-ratio 128 start-port 1024
!
interface GigabitEthernet1
ip address 172.16.16.130 255.255.255.252
negotiation auto
nat64 enable
cdp enable
!
interface GigabitEthernet2
description **clients**
nat64 enable
ipv6 address 2001:918:FFFF:11::1000/64
ipv6 enable
ipv6 mtu 1520
ipv6 nd other-config-flag
ipv6 dhcp relay destination 2001:918:FFFF:99:20C:29FF:FEA7:72CA
!
```

Redundanz mit MAP

- Kein Problem: Es gibt ja keinen Zustand auf dem BR und dem CPE.



Wie könnte eine Adressierung bei Swisscom aussehen?

		shared ports 131'072 user		nicht-geteilte Ports 8'192 user		
Model:		IPv4 shared {BNG with 1:32 address sharing, 2a02:1210:0:0/39,188.62.32.0/20,17,6,1}		IPv4 dedicated {BNG with 1:1 addresses, 2a02:1210:8000:0/43,188.62.0.0/19,13,0,0}		IPv6 only
MAP rule:						
BNG #	IPv6 range	IPv6 range for shared IPv4	IPv4 range	IPv6 range for dedicated IPv4	IPv4 range	IPv6 range for IPv6 only
0	2A02:1210:0::/38	2A02:1210:0::/39	/20	2A02:1210:200::/43	/19	2A02:1210:400::/39
1	2A02:1210:800::/38	2A02:1210:800::/39	/20	2A02:1210:A00::/43	/19	2A02:1210:C00::/39
2	2A02:1210:1000::/38	2A02:1210:1000::/39	/20	2A02:1210:1200::/43	/19	2A02:1210:1400::/39
3	2A02:1210:1800::/38	2A02:1210:1800::/39	/20	2A02:1210:1A00::/43	/19	2A02:1210:1C00::/39
4	2A02:1210:2000::/38	2A02:1210:2000::/39	/20	2A02:1210:2200::/43	/19	2A02:1210:2400::/39
5	2A02:1210:2800::/38	2A02:1210:2800::/39	/20	2A02:1210:2A00::/43	/19	2A02:1210:2C00::/39
6	2A02:1210:3000::/38	2A02:1210:3000::/39	/20	2A02:1210:3200::/43	/19	2A02:1210:3400::/39
7	2A02:1210:3800::/38	2A02:1210:3800::/39	/20	2A02:1210:3A00::/43	/19	2A02:1210:3C00::/39
8	2A02:1210:4000::/38	2A02:1210:4000::/39	/20	2A02:1210:4200::/43	/19	2A02:1210:4400::/39
9	2A02:1210:4800::/38	2A02:1210:4800::/39	/20	2A02:1210:4A00::/43	/19	2A02:1210:4C00::/39
10	2A02:1210:5000::/38	2A02:1210:5000::/39	/20	2A02:1210:5200::/43	/19	2A02:1210:5400::/39

Bei 31 BNG bräuchte man

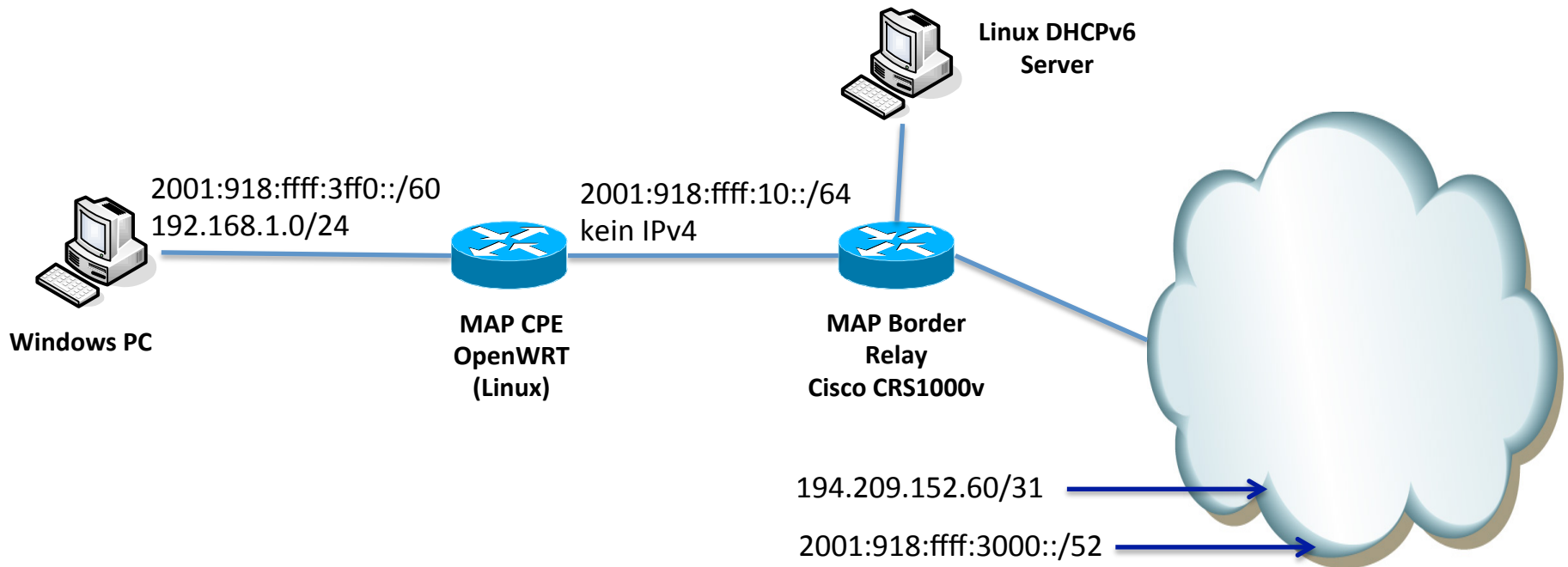
- 81'920 IPv4 Adressen für die Pools mit geteilten Port-Bereichen
- 163'840 IPv4 Adressen für die Pools mit eindeutigen IP Adressen

<http://map46.cisco.com/MAP.php>

Was plant Swisscom

- Wir planen mit Dual-Stack & CG-NAT
 - CG-NAT steht bereits
 - CG-NAT Verkehr ist klein und bleibt es hoffentlich (weil viel Verkehr über IPv6 geht)
 - Wir haben keine 245'760 IPv4 Adressen, um MAP-T einfach so einzuführen.
 - Deshalb müssten wir stückchenweise migrieren, was viel IT-Entwicklung mit sich bringen würde
 - Wir müssen Swisscom-TV und VoIP nicht sofort über IPv6 führen
- Falls beim BNG grosse Skalierbarkeitsprobleme entstehen, müssen wir dies überdenken.
 - Dann ist MAP-T Favorit
 - Müssen aber Swisscom-TV auf IPv6 bringen, um nicht den Hauptteil des Verkehrs tunneln zu müssen

Demo, alles virtuell und heute leider extrem langsam...



[Do it yourself](#)