BGP, IPv6, IETF

Nivo
Weesp

17 mei 2016
Iljitsch van Beijnum

1996

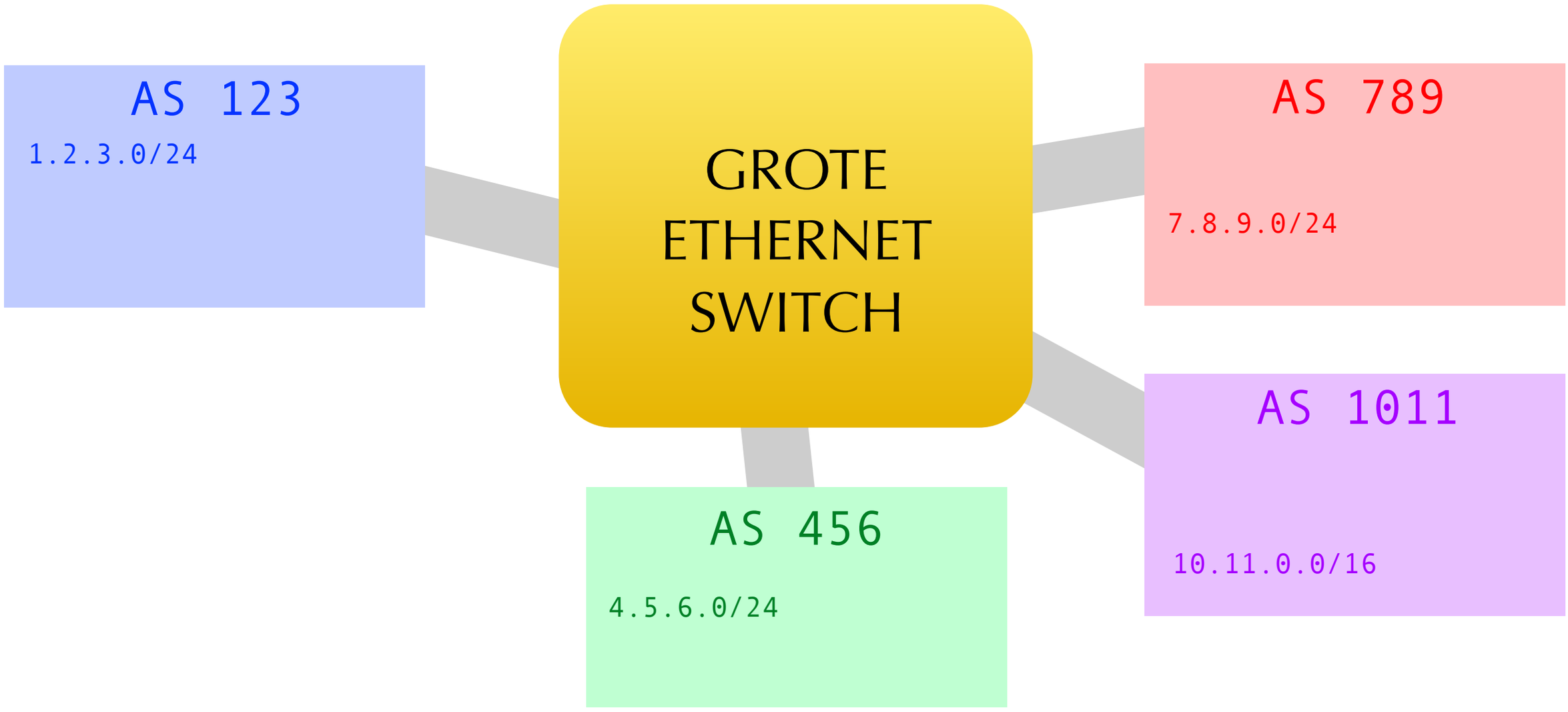
2002

2002

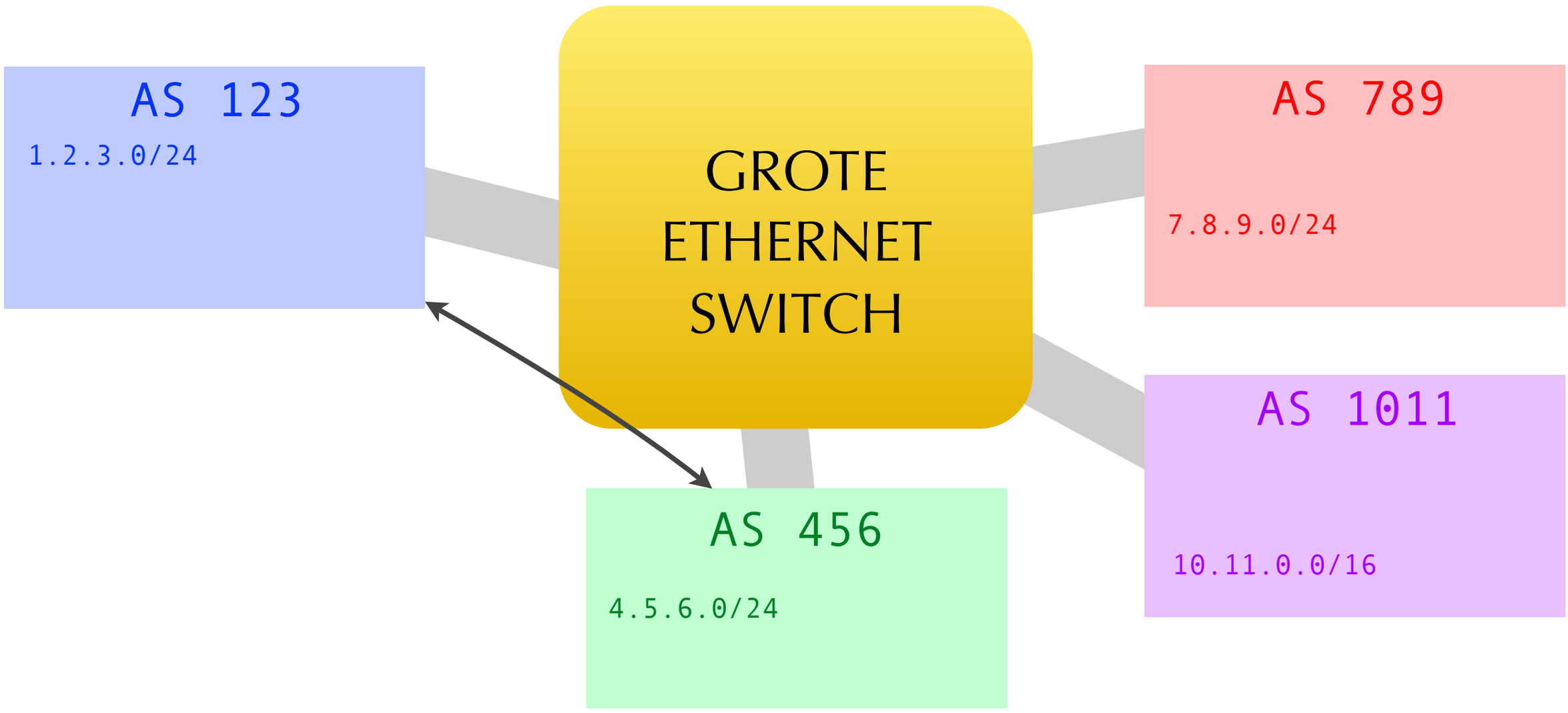

Wireless-G 2.4GHz
54g
Broadband Router
LINKSYS
2.4GHz
54g


Cisco Systems
CISCO CRS-1

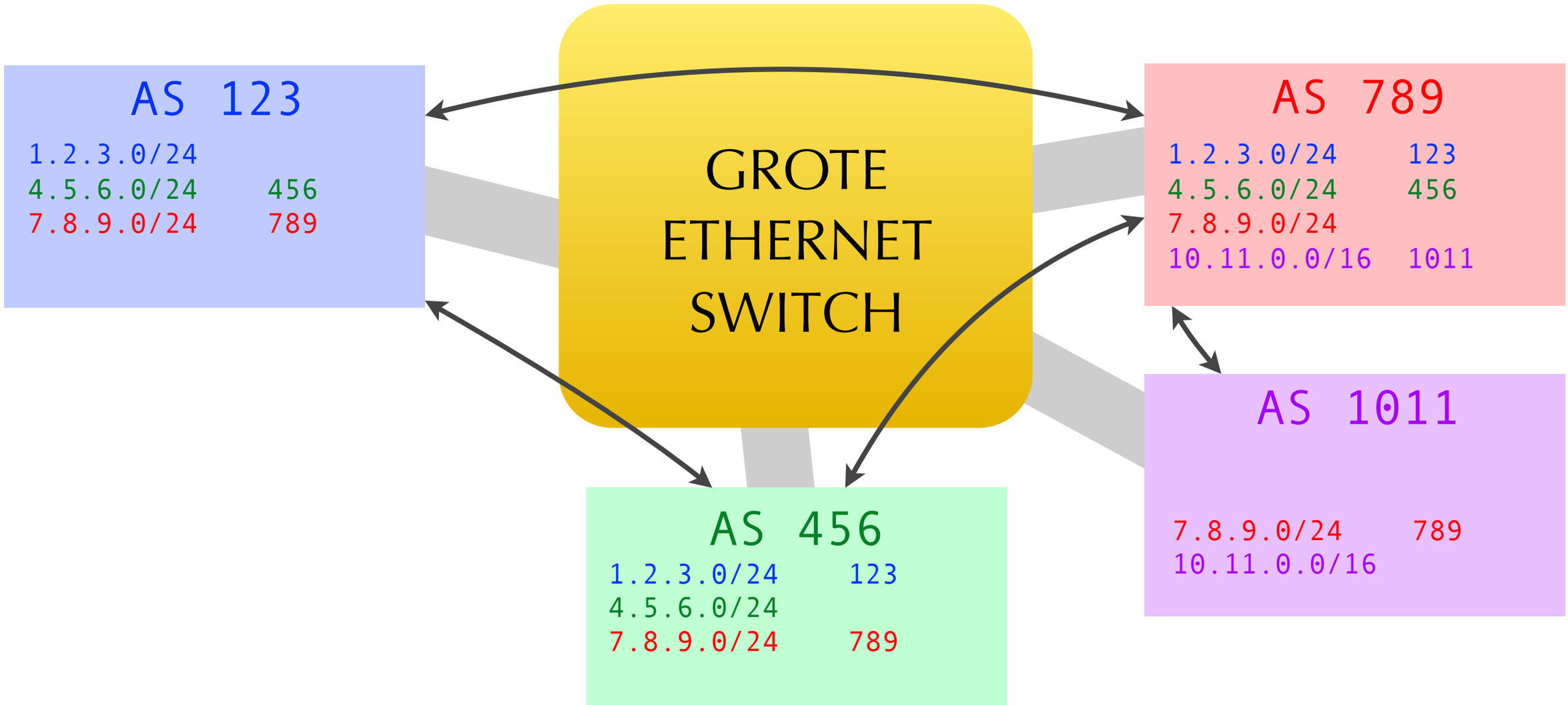
Internet exchange



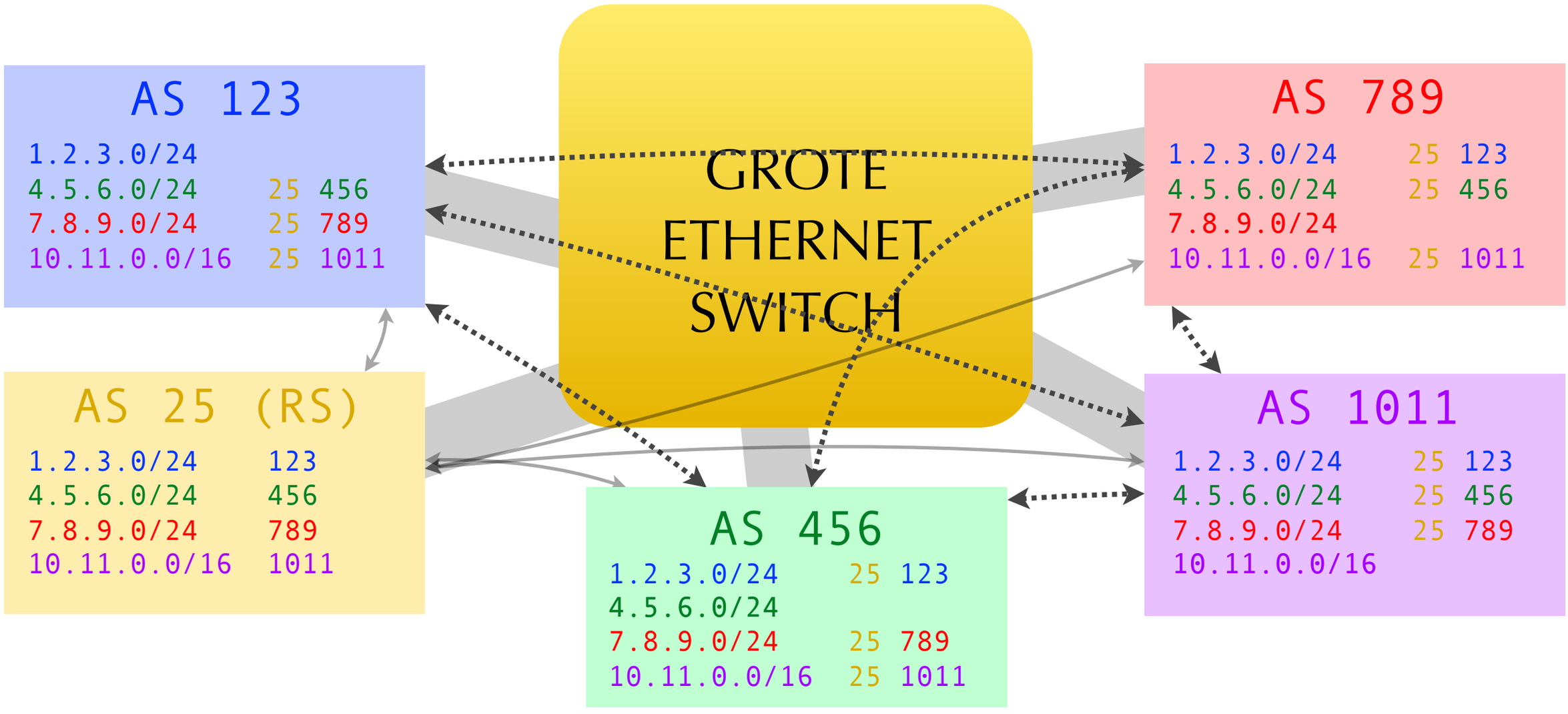
IX peering


AS 123
1.2.3.0/24
GROTE ETHERNET SWITCH
AS 789
7.8.9.0/24
AS 1011
10.11.0.0/16
AS 456
4.5.6.0/24

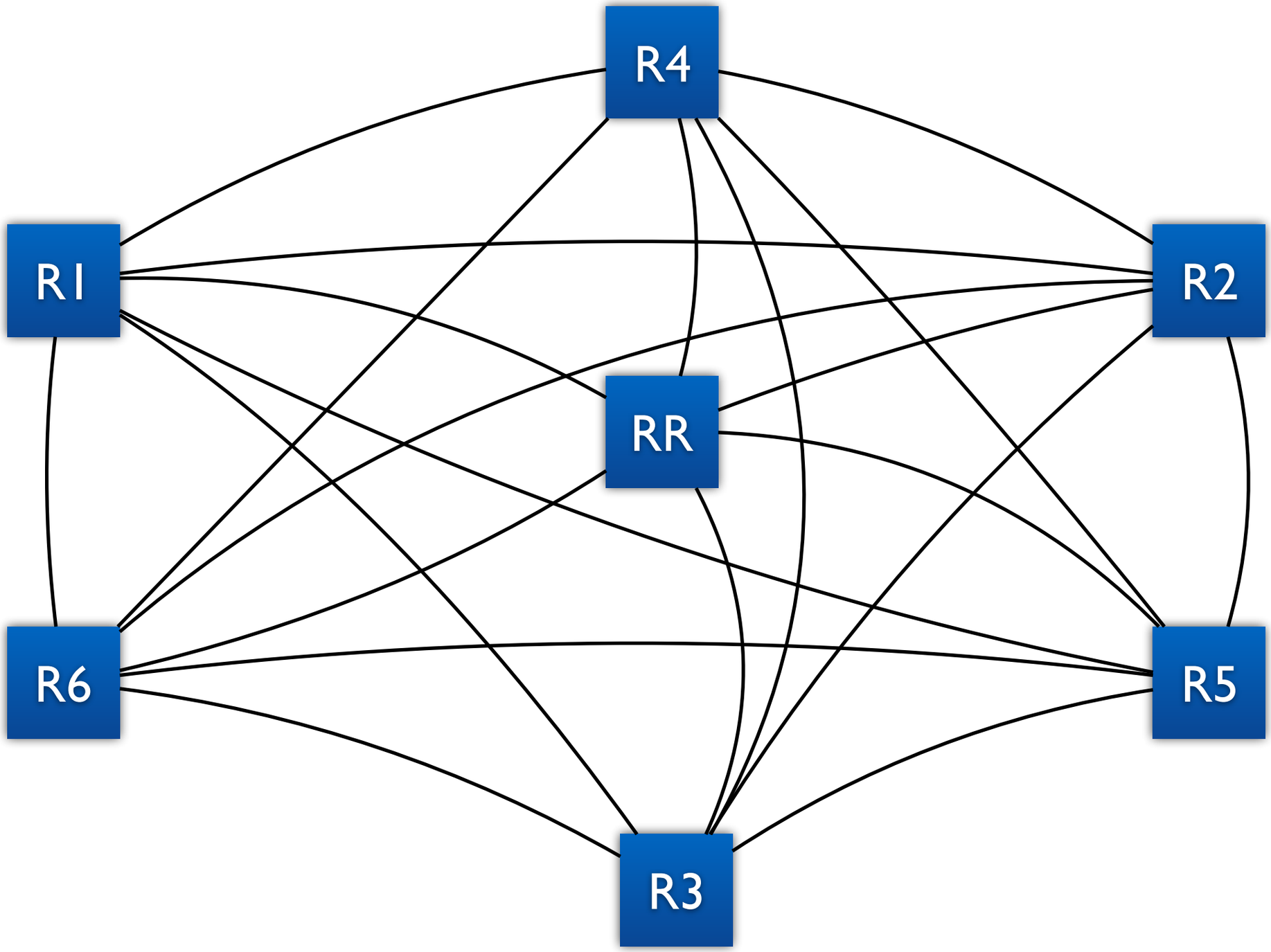
IX peering (2)


AS 123
1.2.3.0/24
4.5.6.0/24 456
7.8.9.0/24 789
GROTE ETHERNET SWITCH
AS 789
1.2.3.0/24 123
4.5.6.0/24 456
7.8.9.0/24
10.11.0.0/16 1011
AS 1011
7.8.9.0/24 789
10.11.0.0/16
AS 456
1.2.3.0/24 123
4.5.6.0/24
7.8.9.0/24 789

Route server



Route *reflector*



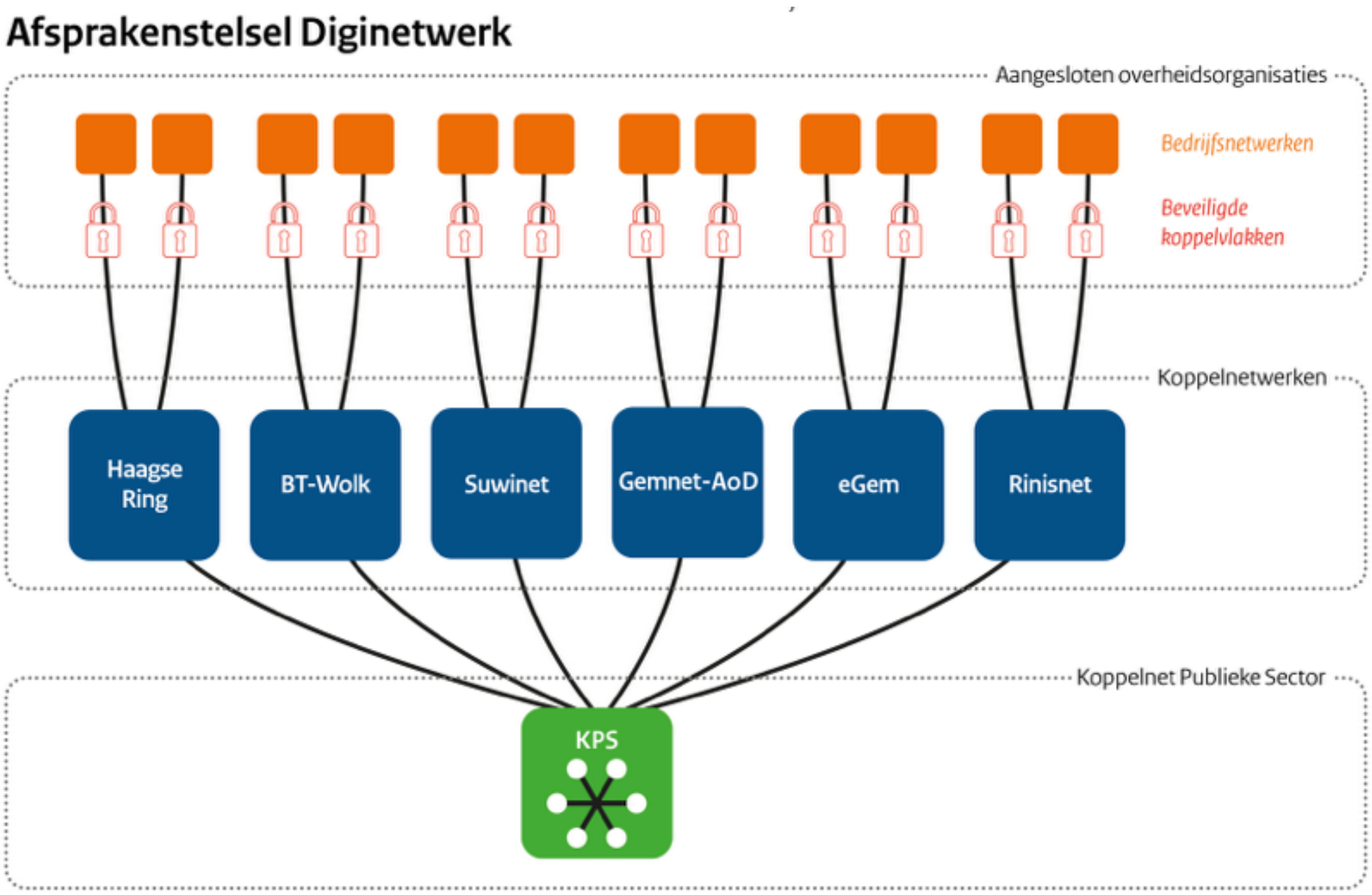
Publieke peering Netflix

Public Peering Exchange Points

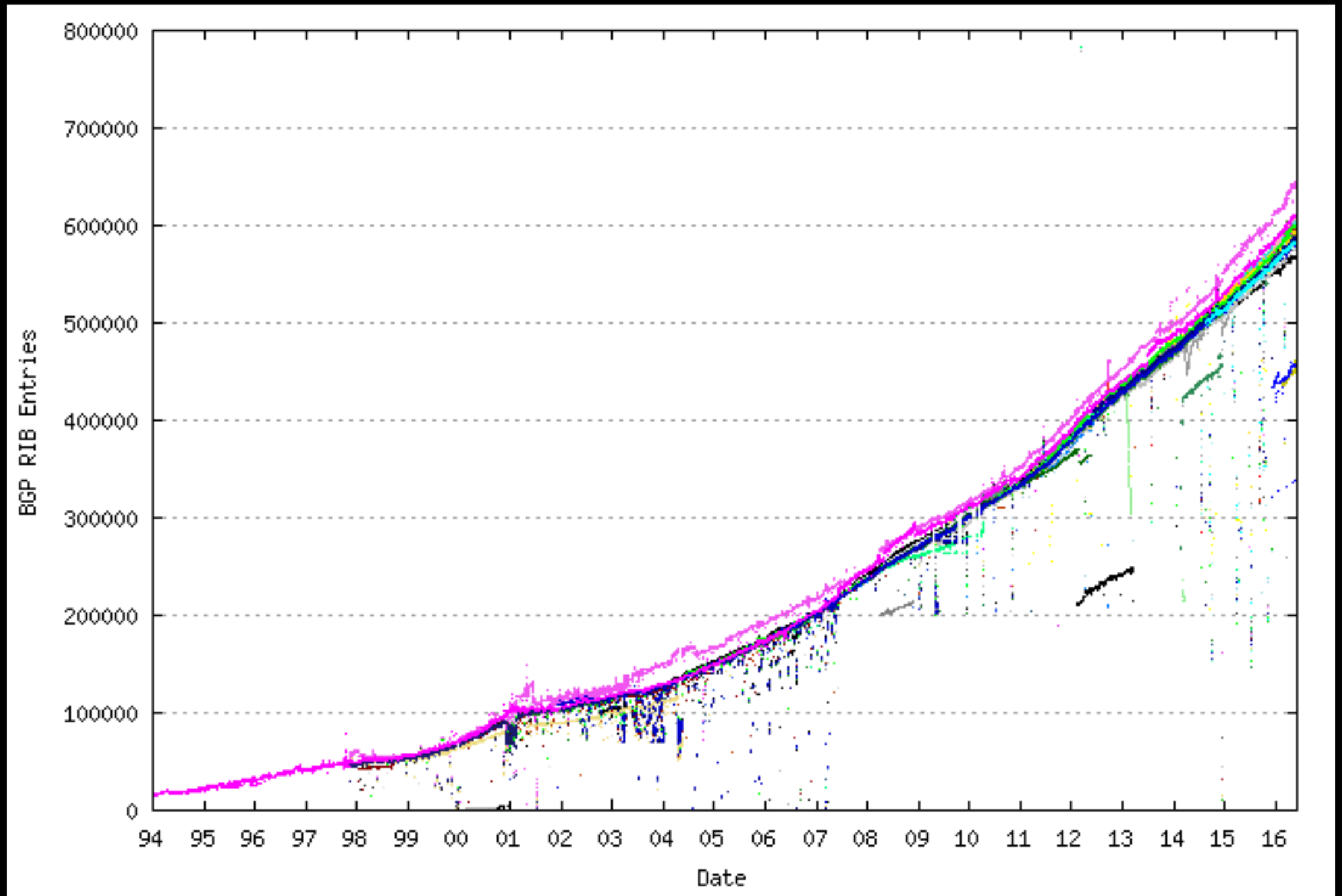
Filter

Exchange ASN	IPv4 IPv6	Speed RS P…
AMS-IX 2906	80.249.210.250 2001:7f8:1::a500:2906:1	100G
AMS-IX 2906	80.249.211.250 2001:7f8:1::a500:2906:2	100G
AMS-IX BA 2906	206.41.106.21 2001:504:3d:1:0:a500:2906:2	10G
BBIX Hong Kong / Singapore 2906	103.231.152.76 2001:df5:b800:bb00::2906:3	10G
BBIX Hong Kong / Singapore 2906	103.231.152.77 2001:df5:b800:bb00::2906:4	10G
BBIX Hong Kong / Singapore 2906	103.231.152.74 2001:df5:b800:bb00::2906:1	10G
BBIX Hong Kong / Singapore 2906	103.231.152.75 2001:df5:b800:bb00::2906:2	10G
BBIX Tokyo 2906	218.100.6.117 2001:de8:c::2906:1	10G
BBIX Tokyo 2906	218.100.6.119 2001:de8:c::2906:2	10G
BCIX 2906	193.178.185.80 2001:7f8:19:1::b5a:1	20G
BCIX 2906	193.178.185.81 2001:7f8:19:1::b5a:2	20G
Boston Internet Exchange 2906	206.108.236.32 2001:504:24:1::b5a:1	10G
ChIX 2906	206.41.110.25 2001:504:41:110::25	10G
CoreSite - Any2 California 2906	206.72.210.215 2001:504:13::210:215	40G
CoreSite - Any2 California	206.72.210.226	40G

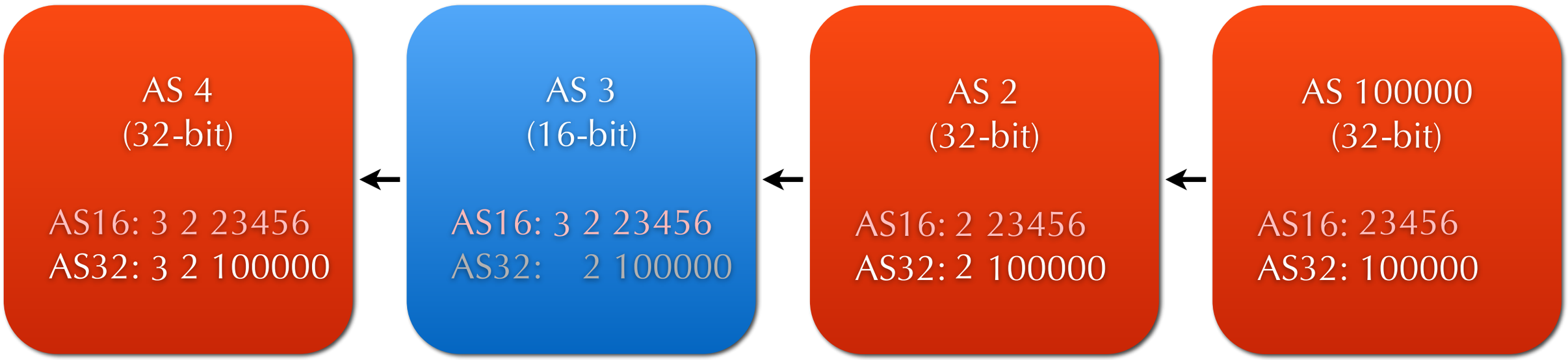
Koppelnet Publieke Sector



Groei BGP-tabel



32-bit AS-nummers



BGP-ontwikkelingen

- We draaien al > 20 jaar BGP-4
  - zelfs IPv6 over BGP-4 ook al is BGP-4 ouder dan IPv6!
- Actueel: beveiliging
  - RPKI
  - BGPSEC

IPv6

Status IPv4

July 2015: waitlist

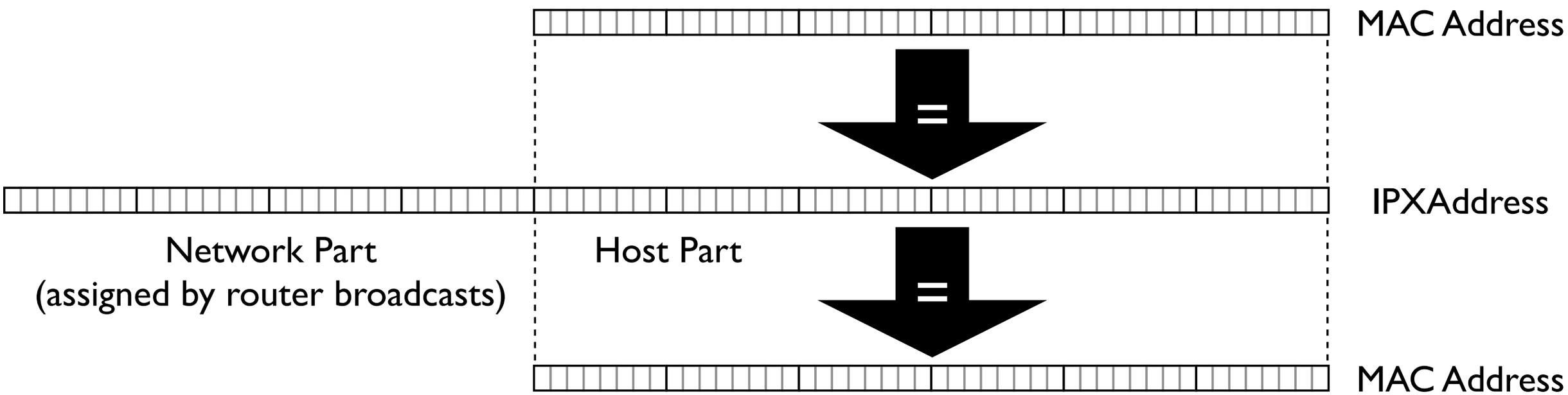
Sep 2012: final /8

Jun 2014: final /10

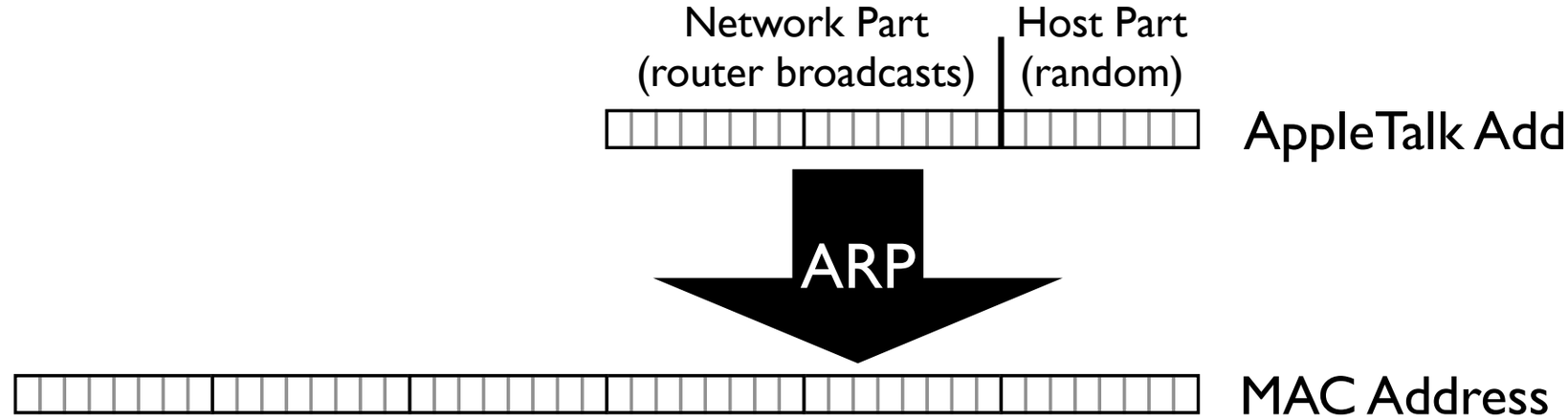
Apr 2011: final /8

AfriNIC
APNIC
ARIN
LACNIC
RIPE NCC

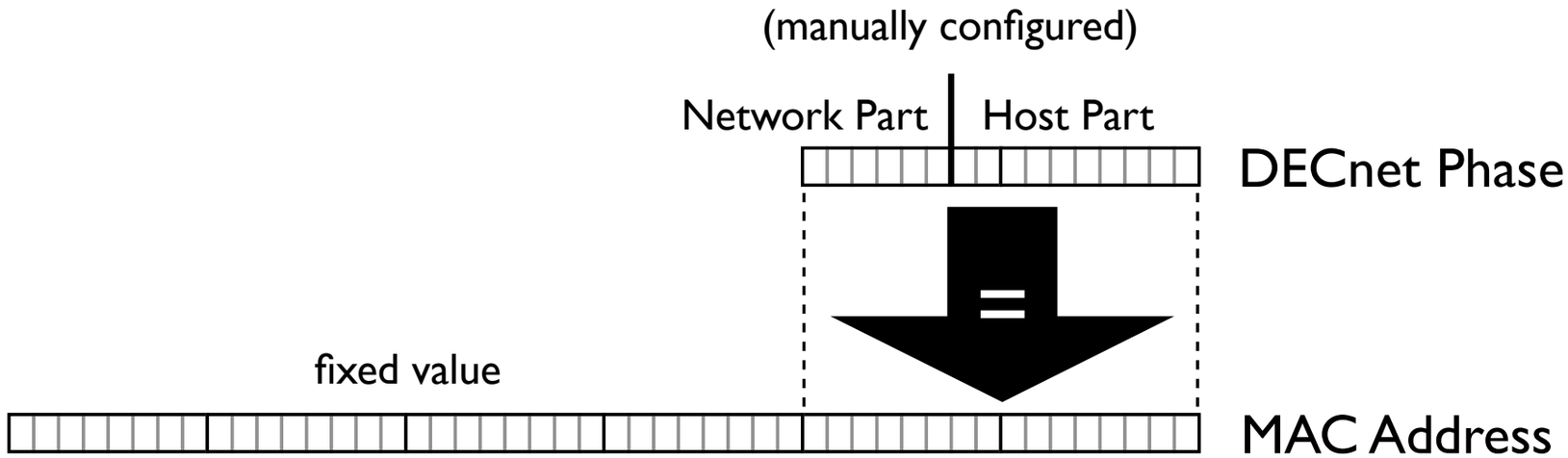
IPX


MAC Address
=
IPXAddress
Network Part
(assigned by router broadcasts)
Host Part
=
MAC Address

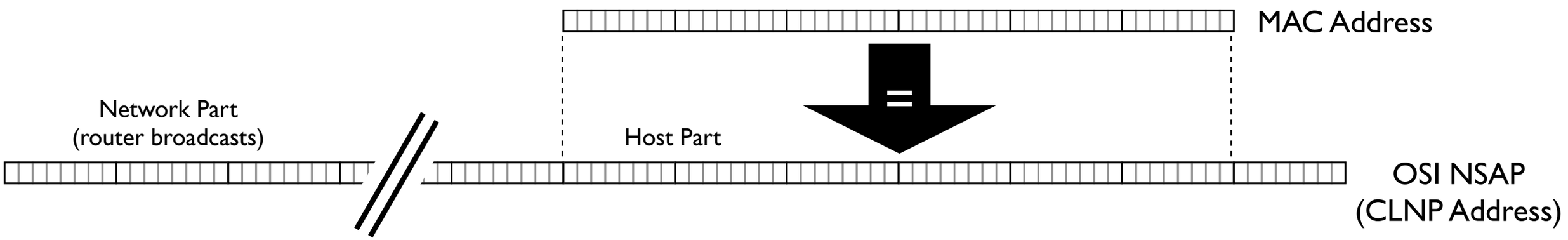
AppleTalk


Network Part
(router broadcasts)
Host Part
(random)
AppleTalk Address
ARP
MAC Address

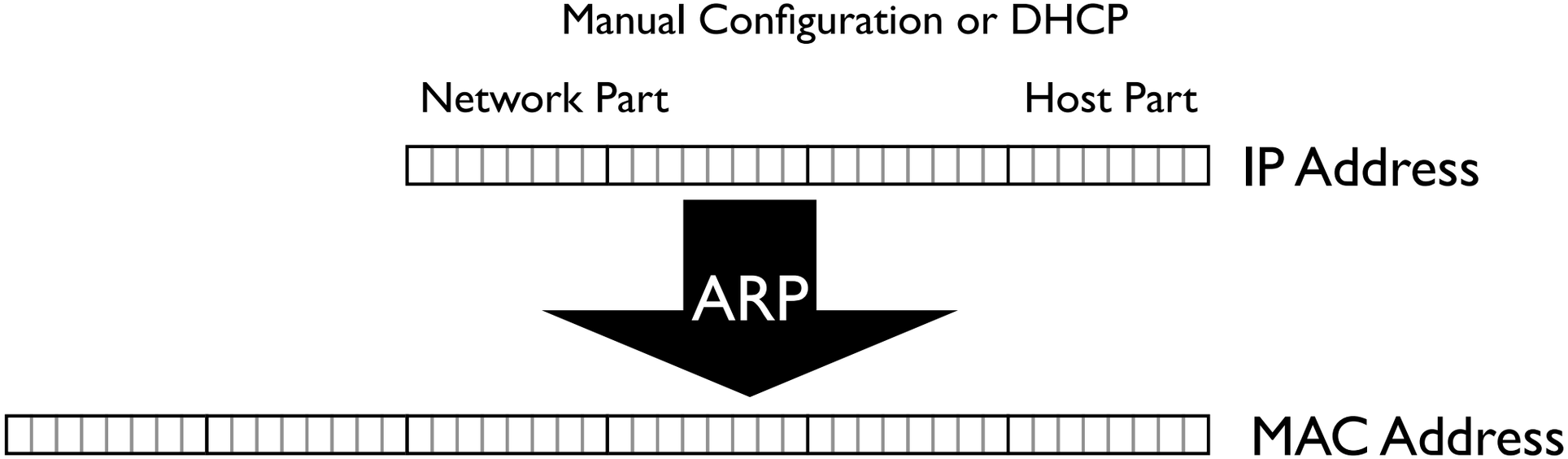
DECnet Phase IV



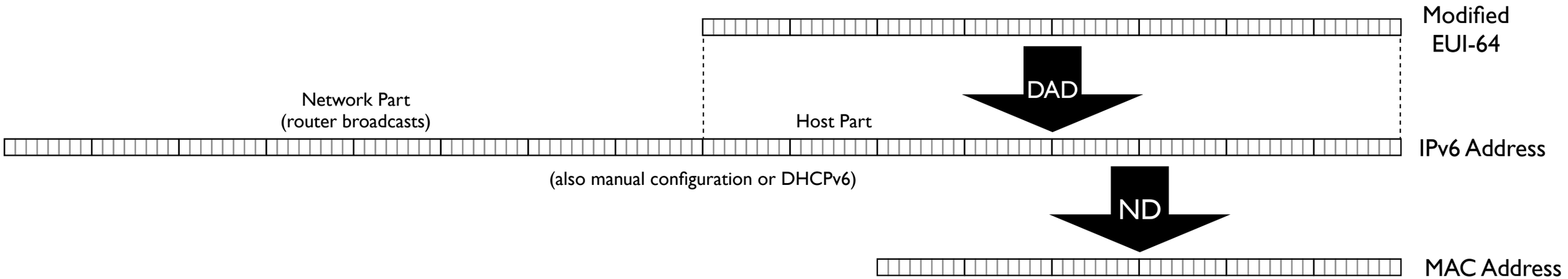
CLNP


MAC Address
=
Network Part
(router broadcasts)
Host Part
OSI NSAP
(CLNP Address)

IP(v4)


Manual Configuration or DHCP
Network Part
Host Part
IP Address
ARP
MAC Address

IPv6

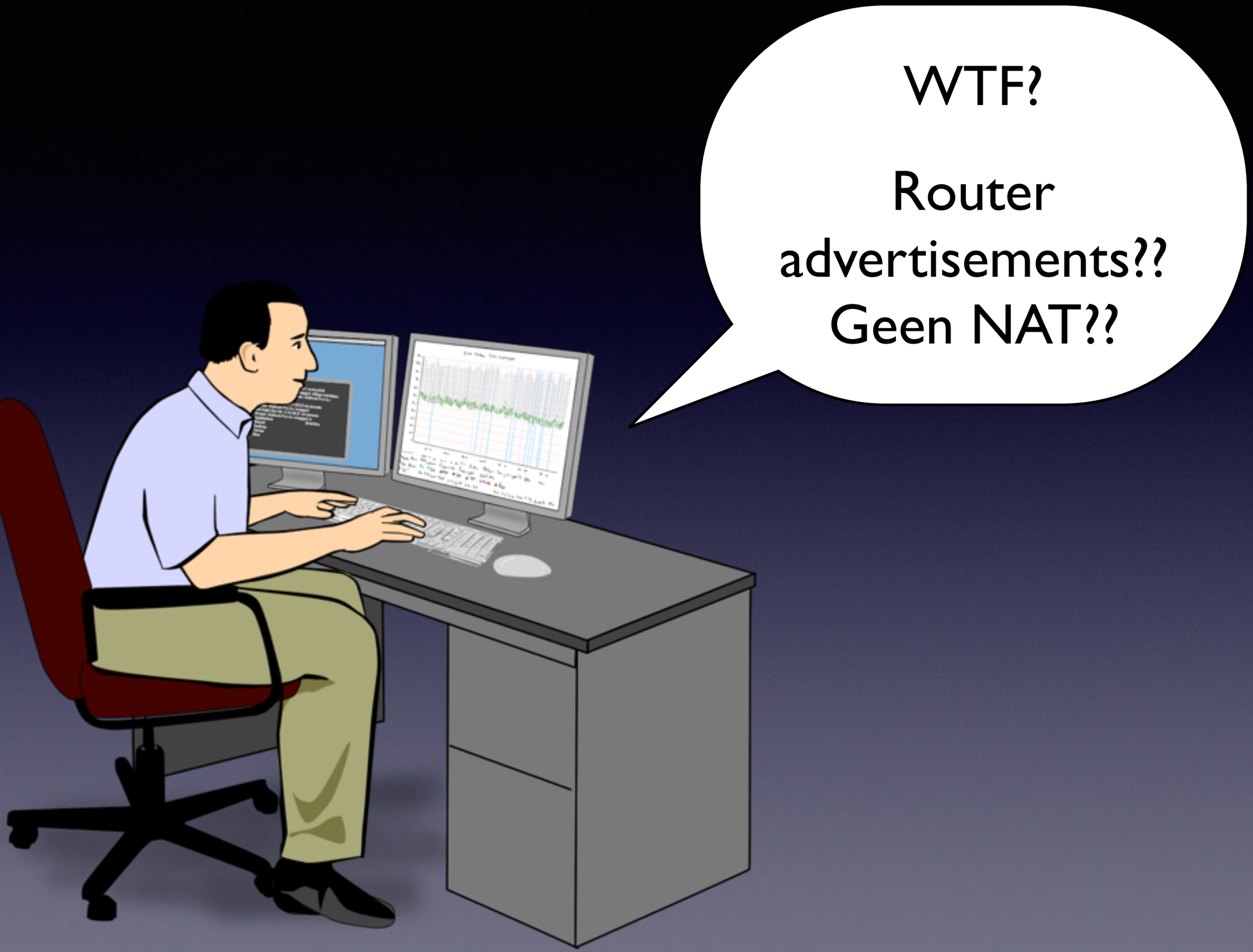

Network Part
(router broadcasts)
Host Part
(also manual configuration or DHCPv6)
Modified
EUI-64
DAD
IPv6 Address
ND
MAC Address

IPv6 adresconfiguratie

- Vanaf het begin: stateless autoconfig: router advertisement + MAC-adres
- Later: privacy-adressen: router advertisement + willekeurig nummer
- (Nog later: router advertisement + hash over public key)
- Veel later: DHCPv6

RA flags and DHCPv6

M	O	Prfx	A	Result
0	0	-		default gw but no address
0	0	yes	0	default gw but no address
0	0	yes	1	working IPv6 but no DNS
0	1	-		default gw + DNS but no address
0	1	yes	0	default gw + DNS but no address
0	1	yes	1	working IPv6
1	0	-		address+DNS, no subnet length (may not work)
1	0	yes	0	working IPv6
1	0	yes	1	working IPv6, 2 addresses
1	1	-		address+DNS, no subnet length (may not work)
1	1	yes	0	working IPv6
1	1	yes	1	working IPv6, 2 addresses


WTF?
Router advertisements??
Geen NAT??

Aaargghh!

Communicatie tussen router- en servermensen??


Zonder NAT kan iedereen m'n adres zien!

Als hosts hun eigen adres maken weet ik niet wie achter welk adres zit!

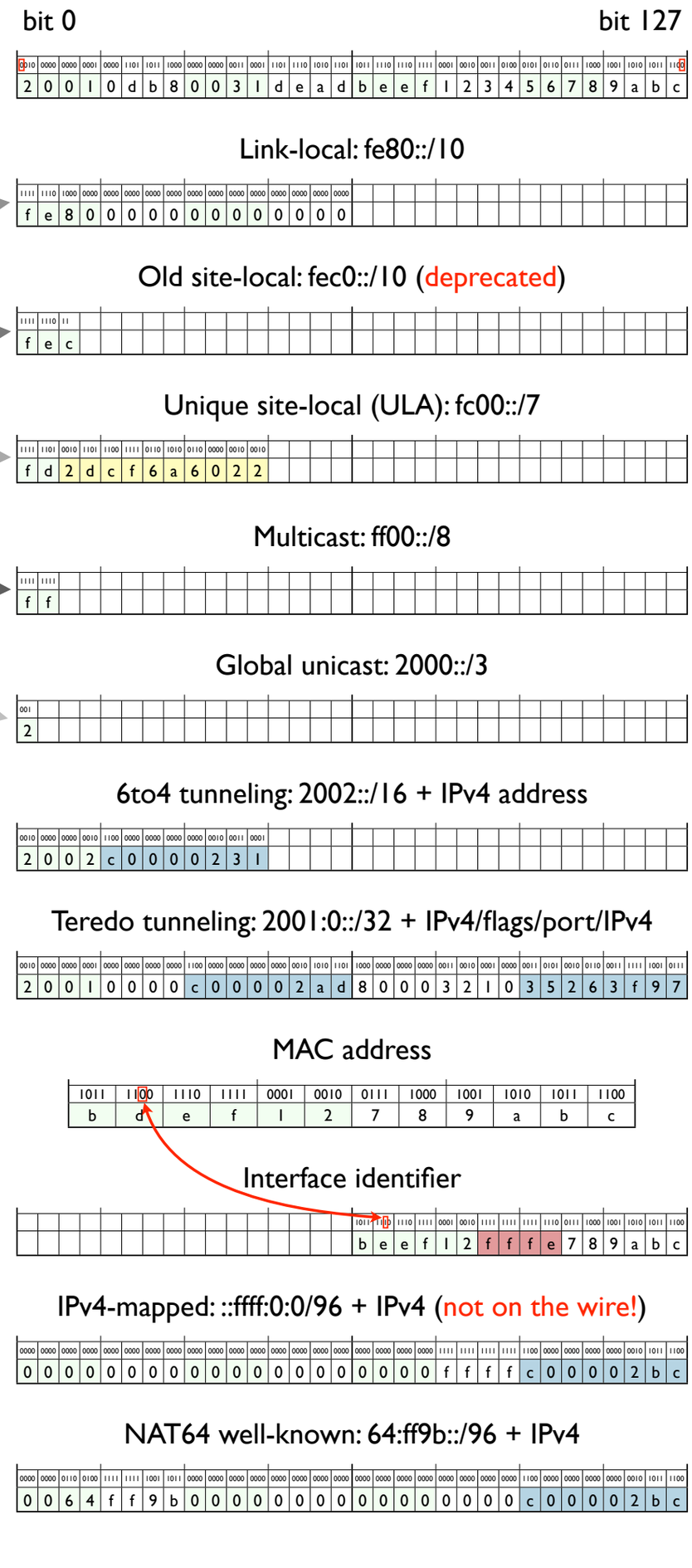
NAT breekt veel protocollen!
End-to-end is beter

Geen default gateway in DHCPv6: fate sharing (veel minder kans op fouten)

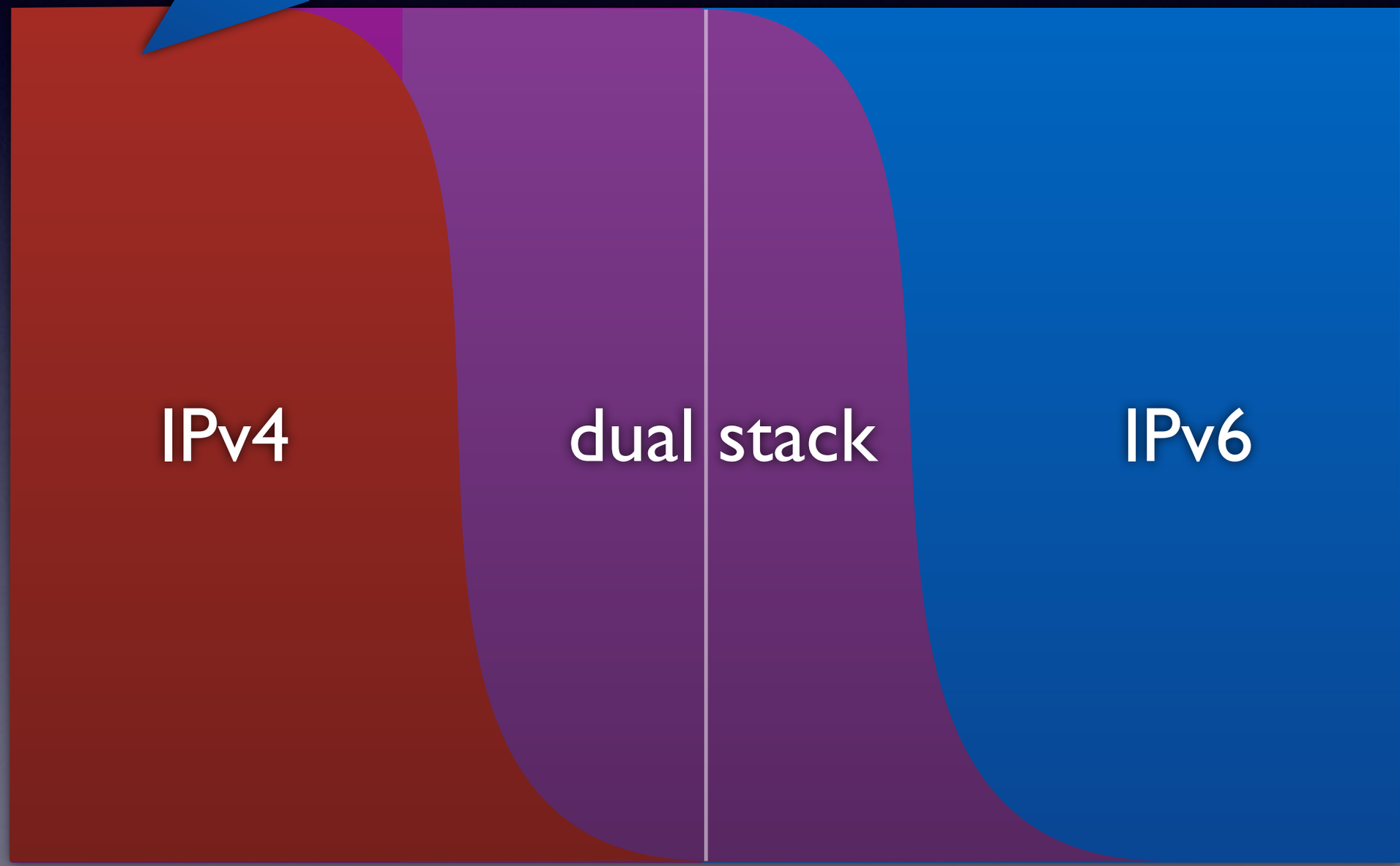
illustraties door Nitrozac

IPv6 Address Space

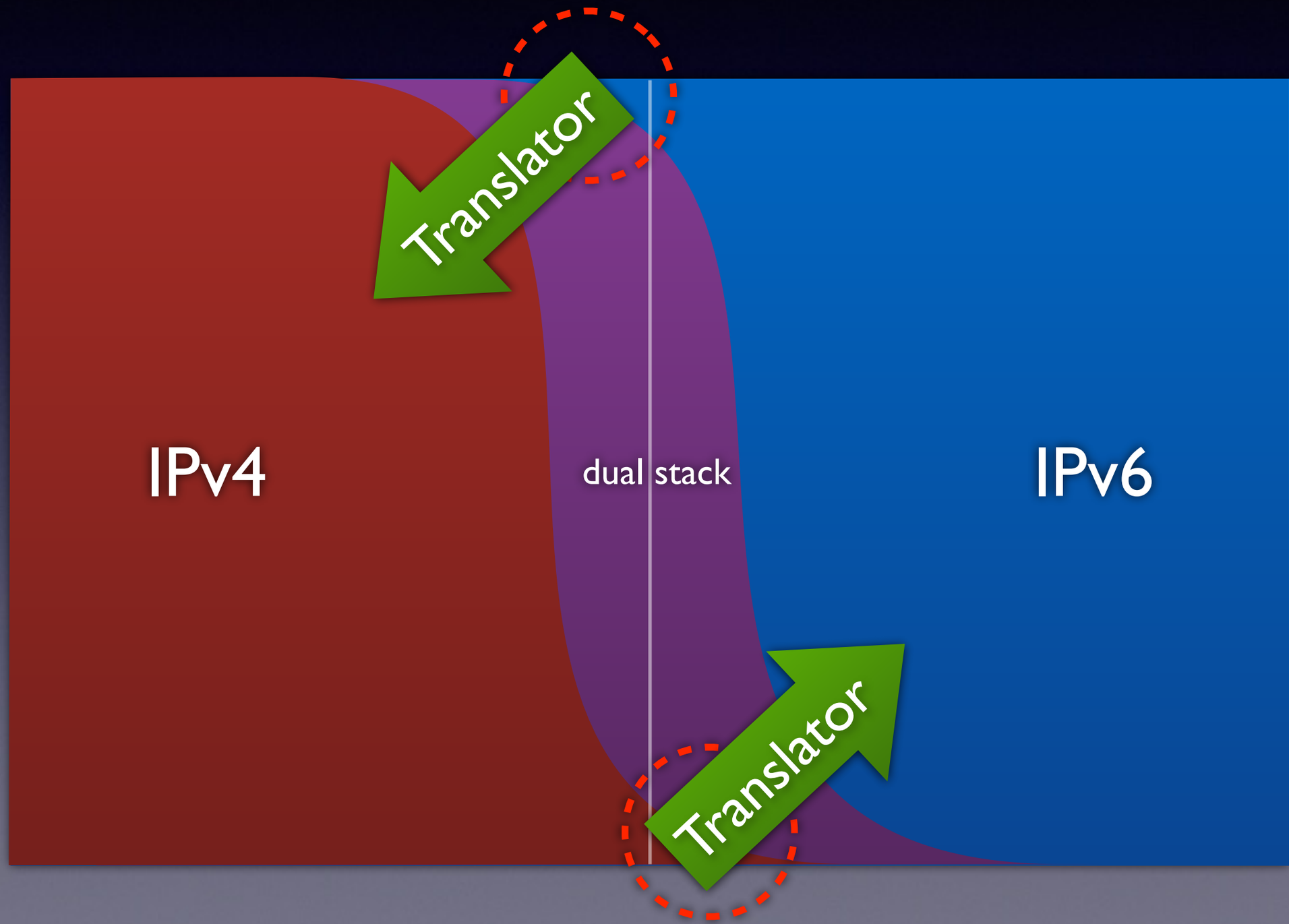
First Bits	Prefix	Purpose
000	::/3	Special uses
001	2000::/3	Global unicast
01 - 1111 1011	4000::/2 - fb00::/8	Reserved
1111 1100	fc00::/8	Unique site local
1111 1110 10	fe80::/10	Link local
1111 1110 11	fec0::/10	Old site local
1111 1111	ff00::/8	Multicast



Transition: dual stack

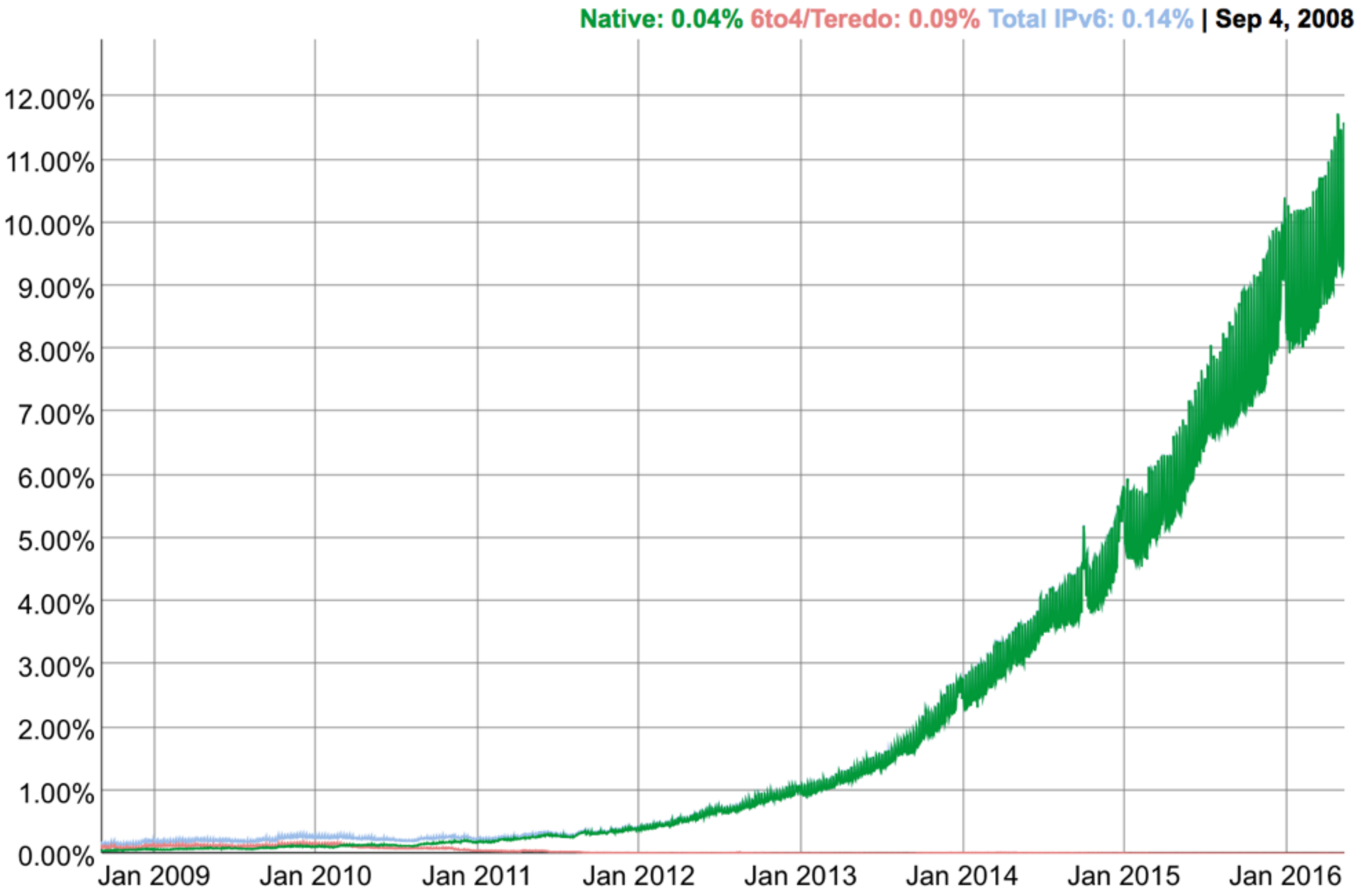


Transition: translation



IPv6 Adoption

We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.



NAT64

Internet Engineering Task Force (IETF)
Request for Comments: 6146
Category: Standards Track
ISSN: 2070-1721

M. Bagnulo
UC3M
P. Matthews
Alcatel-Lucent
I. van Beijnum
IMDEA Networks

April 2011

Stateful NAT64: Network Address and Protocol Translation
from IPv6 Clients to IPv4 Servers

Abstract

This document describes stateful NAT64 translation, which allows IPv6-only clients to contact IPv4 servers using unicast UDP, TCP, or ICMP. One or more public IPv4 addresses assigned to a NAT64 translator are shared among several IPv6-only clients. When stateful NAT64 is used in conjunction with DNS64, no changes are usually required in the IPv6 client or the IPv4 server.

Terminologie:

- NATXY: NAT van X naar Y:
  - NAT44: IPv4 → IPv4 (zoals thuis)
  - NAT64: IPv6 → IPv4
  - NAT46: IPv4 → IPv6
  - NAT444: IPv4 → IPv4 → IPv4 (door ISP) ook: CGN, Carrier Grade NAT
  - NAT66: IPv6 → IPv6...

Hoe werkt het?

DNS64

- Gewone DNS-server:
  - V: IPv6 adres van example.com?
  - A: bestaat niet
- DNS64:
  - V: IPv6-adres van example.com?
  - A: Pref64 + IPv4-adres

DNS64 (2)

- Pref64: een *prefix* die naar het NAT64-apparaat wijst
- gewoonlijk: 64:FF9B::/96
- dus 192.0.2.31 wordt 64:ff9b::c000:21f

De NAT64

- De Pref64 wordt naar de "translator" / converter gerouteerd
- Dus pakketten naar 64:ff9b::c000:21f gaan naar de NAT64
- NAT64 vertaalt IPv6 naar IPv4
  - 64:ff9b::c000:21f → 192.0.2.31
- NAT64 doet dan standaard IPv4 NAT

Beperkingen

- Werkt alleen:
  - *van* IPv6 clients
  - *naar* IPv4 servers
- Applicatie moet IPv6-compatible zijn
  - (dus geen Skype!)
- Werkt niet met IPv4-adressen in URL

DS-Lite en 464XLAT

- DS-Lite:
  - ISP (of bedrijfsnetwerk) runt IPv6, maar ook grote IPv4 NAT
    - eindgebruikers sturen IPv4-pakketten die *getunneld* worden over IPv6, dan geNAT
    - Dus compatible met IPv4-only apparaten/ applicaties
- 464XLAT:
  - zie DS-Lite maar pakketten worden in API onderschept

iljitsch@bgpexpert.com