A Practical (and Personal) Perspective on IPv6 for Servers

Geoff Huston

June 2011

Lets look at ...

- Why we need IPv6
- The differences between IPv4 and IPv6
- Some practical hints for Dual Stack Services

Why?

Because we've run out of addresses!

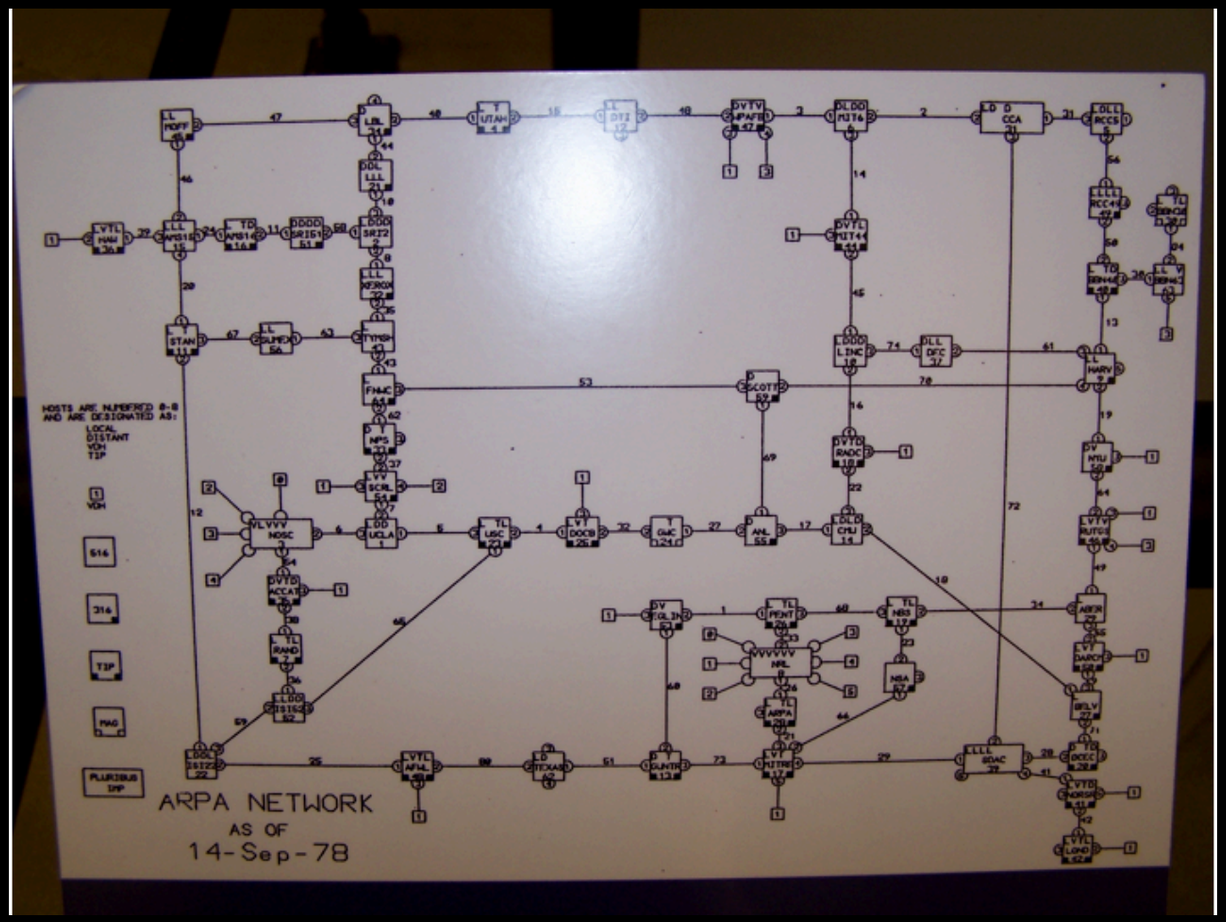
again!

A bit of history...

The original ARPAnet design of 1969 used the NCP protocol, which used 8 bit addresses

- Maximum network of 256 nodes
- Enough, yes?

ARPAnet IMP



ARPAnet - September 1978

Transition V1.0

- Turns out that 8 bits of addresses was not enough for the next generation of mini computers
- ARPAnet undertook a transition from NCP to a new protocol: TCP/IP
  - Expansion from 8 to 32 bit addresses
  - Flag Day: 1 January 1983
  - Shutdown and reboot every node into the new protocol

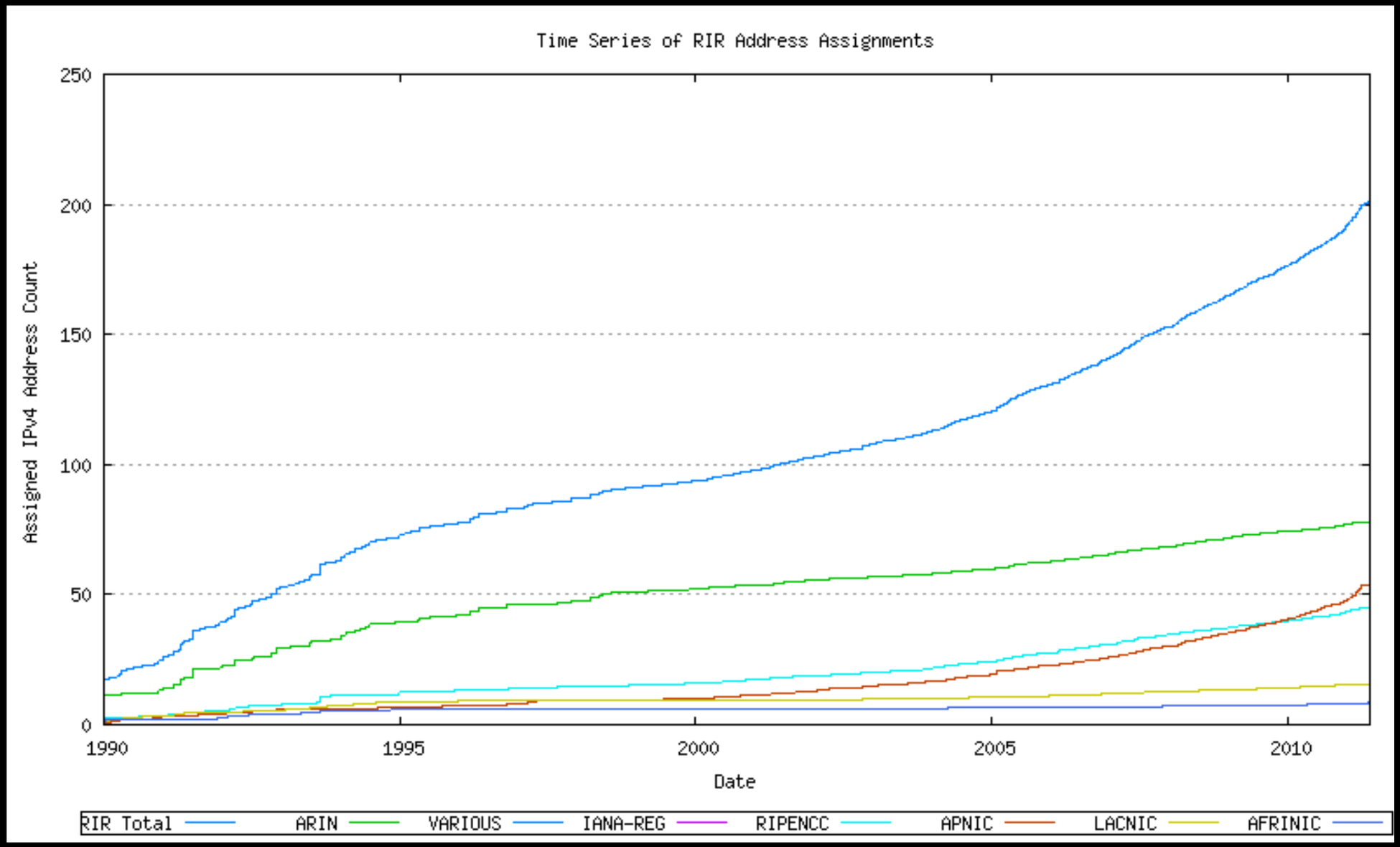
Vint Cerf, APRICOT, Feb 2011

“This time, for sure!” *

* Actually Vint Cerf did not say that!

IP Version 4

- 32 bit address field
  - That's 4,294,967,296 addresses
- We've used this to build today's Internet:
  - some 360,000 networks
  - Around 900 million connected devices
- Now, 28 years later, we've run out of addresses - again!



IPv4 Address Allocations

IPocalypse?

Maybe not

- Many ISPs have been stockpiling IPv4 addresses
- Address “recovery” programs are underway
- So it’s not a sudden halt

- But the network grew by more than 250 million services last year
  - Which was the largest year so far for the Internet

It’s more like this!

Transition V2.0

- A “Flag Day” switchover is impossible
- Piecemeal replacement won’t work either as IPv6 is not backward compatible with IPv4
- So we need to run both protocols in tandem “for a while”
- But bear in mind that one protocol has already run out of addresses
- And network growth continues at record levels

Transition V2.0

We need to :

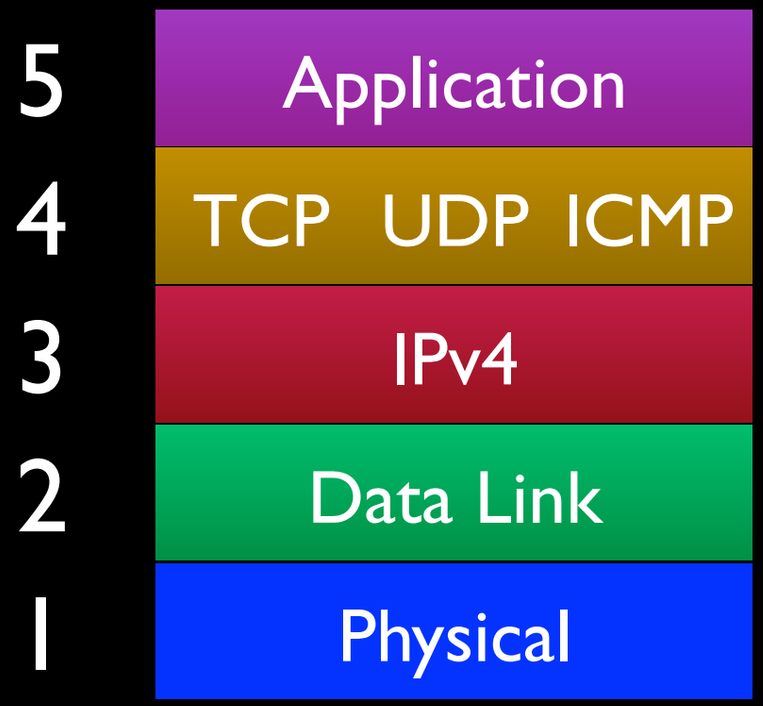
- deploy IPv6 in parallel with IPv4
- deploy ever more stringent IPv4 address conservation measures within the network
- allow the network to expand at an ever increasing rate

All at the same time!

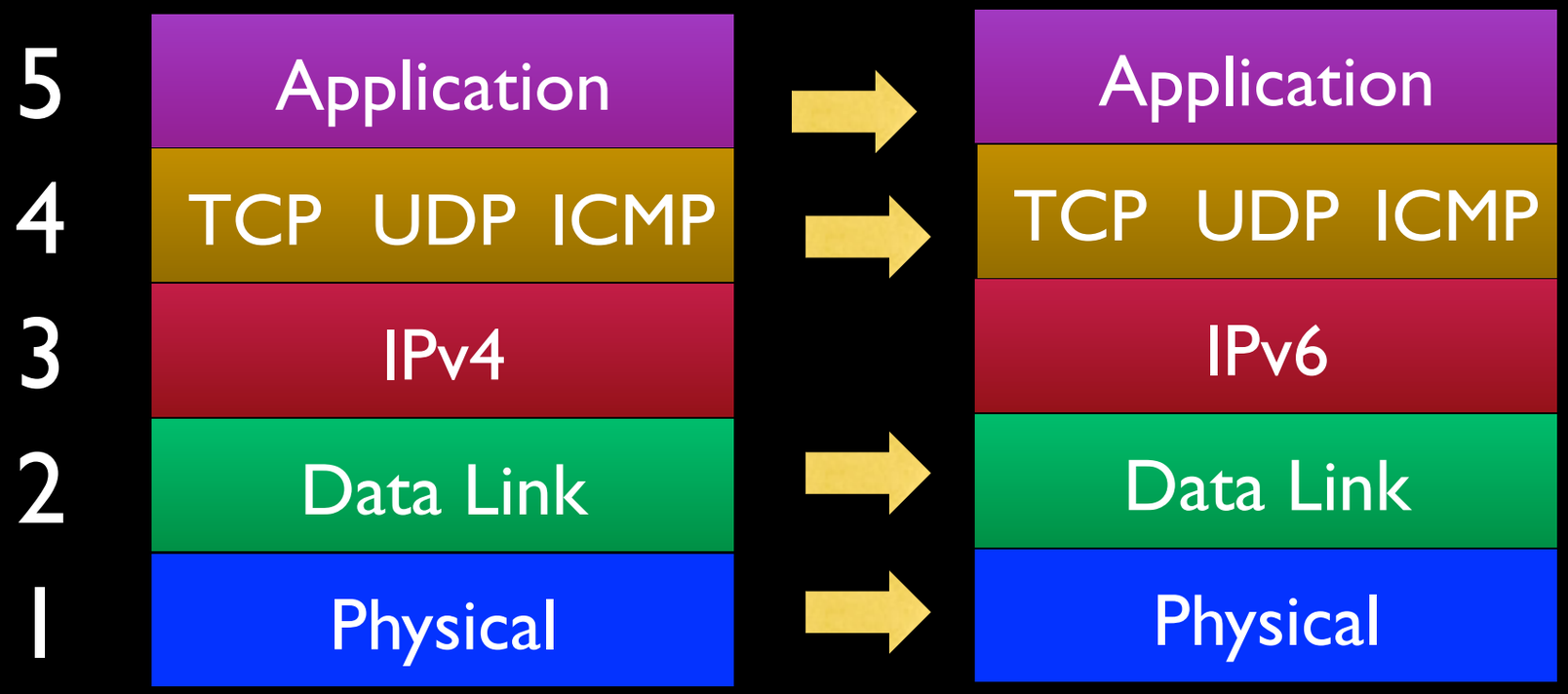
Or maybe it's like this!

What changes with IPv6?

Layer-3 Protocol



Layer-3 Protocol



Only the IP layer changes – nothing else!

Addresses Notation

- IPv4 addresses: 32 bits, looks like
  `192.168.123.234`

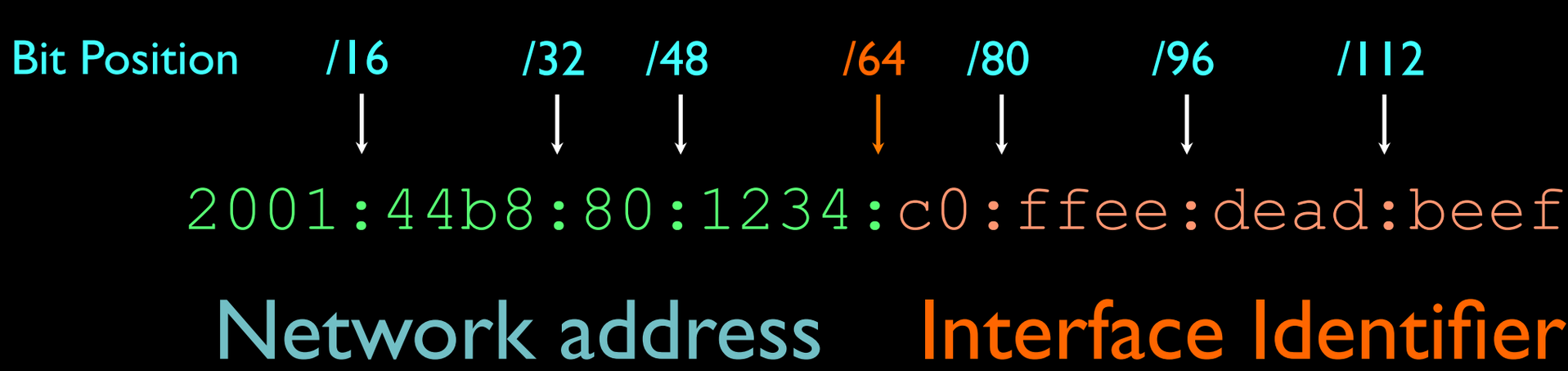
  Four decimal octets separated by “dots”

- IPv6 addresses: 128 bits, looks like
  `2001:44b8:80:12a6:c0:ffee:dead:beef`

  Eight hexadecimal hextets separated by colons

Address Notation changes! Use the DNS, Luke!

IPv6 address structure



Back to a fixed network / host part boundary!

Layer-4 protocols

- In the IPv4 world: TCP and UDP are the most common layer-4 protocols
- In the IPv6 world: TCP and UDP are the same as they are in IPv4

Plus ça change, plus c'est la même chose

ICMP

- In the IPv4 world:
  - ICMP is used for some "control plane" functions
    - Noteably "destination unreachable"
- In the IPv6 world:
  - ICMP6 is used for all "control plane" functions -- *more important than v4 ICMP*.

No Broadcast!

IPv4:

- Sending a packet to the “all-ones” address in a subnet produces a broadcast.

IPv6:

- No such thing as a broadcast.
- Heavy reliance on multicast for everything that requires the interaction of more than two nodes.
- Multicast is *required*

Look! No ARP!

- IPv4:

  Broadcast: ARP: Who has IP address <X>
  <X> answers with an ARP response

- IPv6:

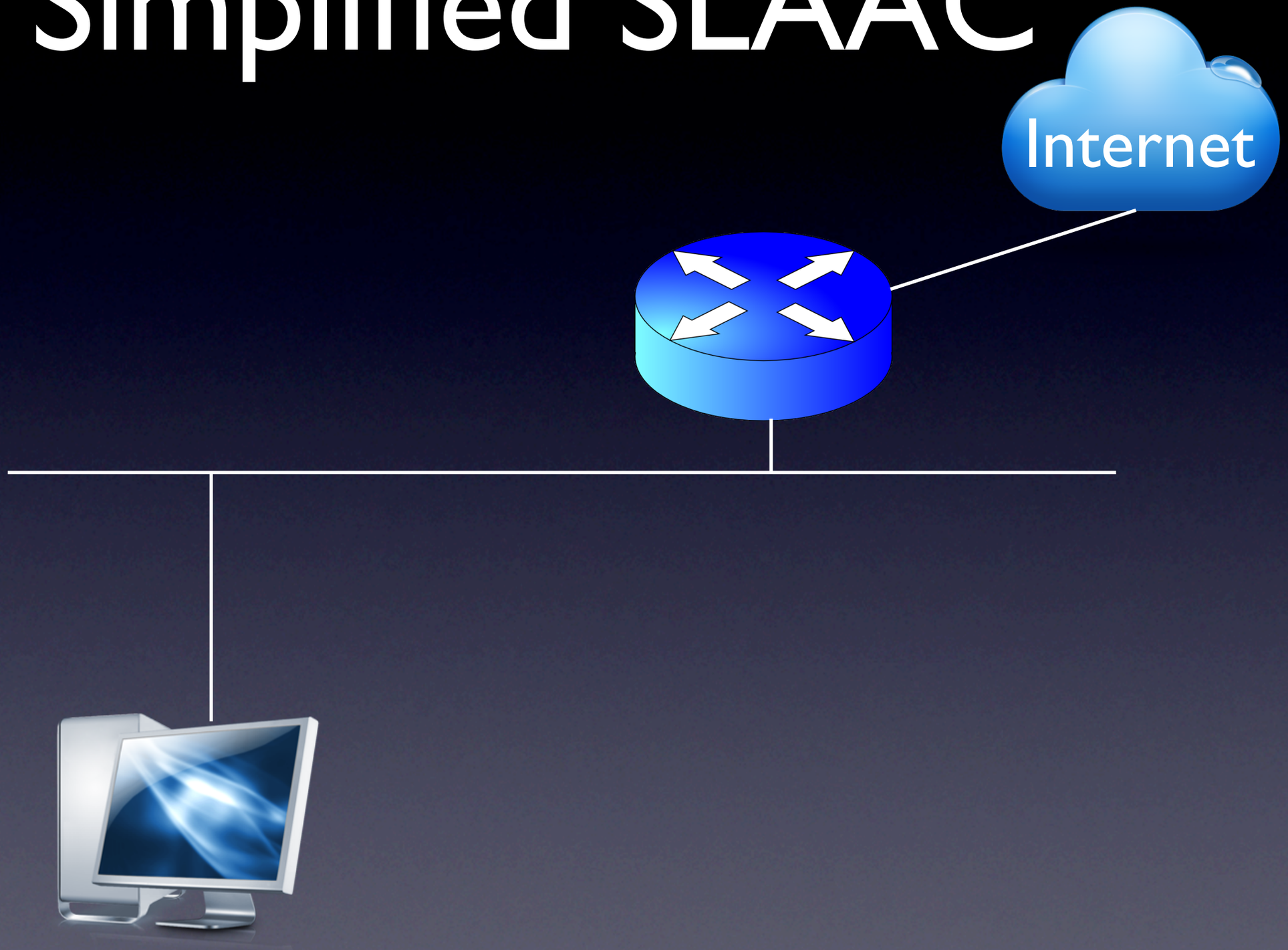
  Multicast: solicited node multicast: ICMPv6 Neighbour Solicitation request for <X>
  <X> answers with an ICMP6 Neighbour Advertisement message

Address assignment

- IPv4: Almost all addresses are assigned statically (configuration on the host) or via DHCP.
- IPv6: *Almost all addresses are assigned dynamically using a method intrinsic to the protocol* - SLAAC.
  - DHCPv6 is still available, but differs in several important areas from IPv4 DHCP.
  - Having multiple addresses on one host is totally normal and expected.

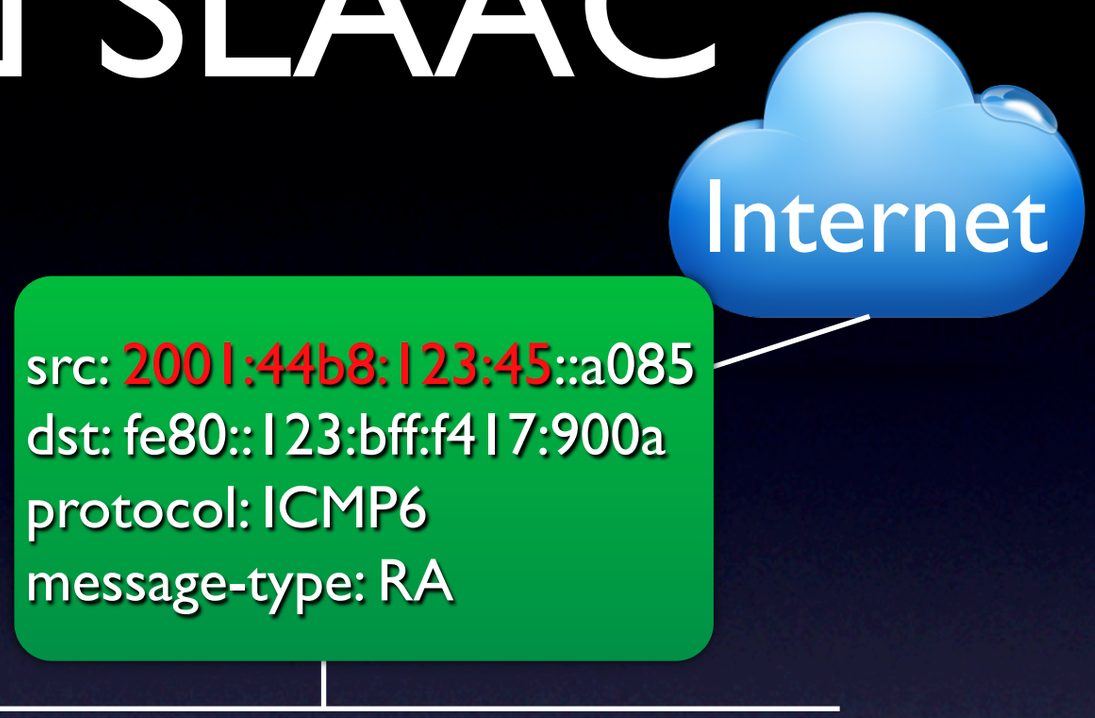
Simplified SLAAC

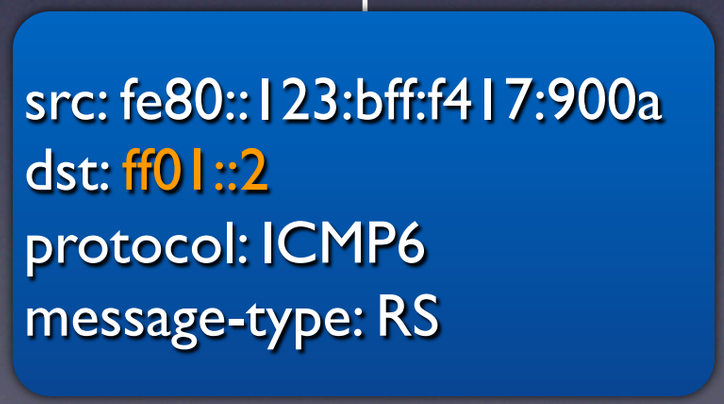
Simplified SLAAC



fe80::123:bff:f417:900a

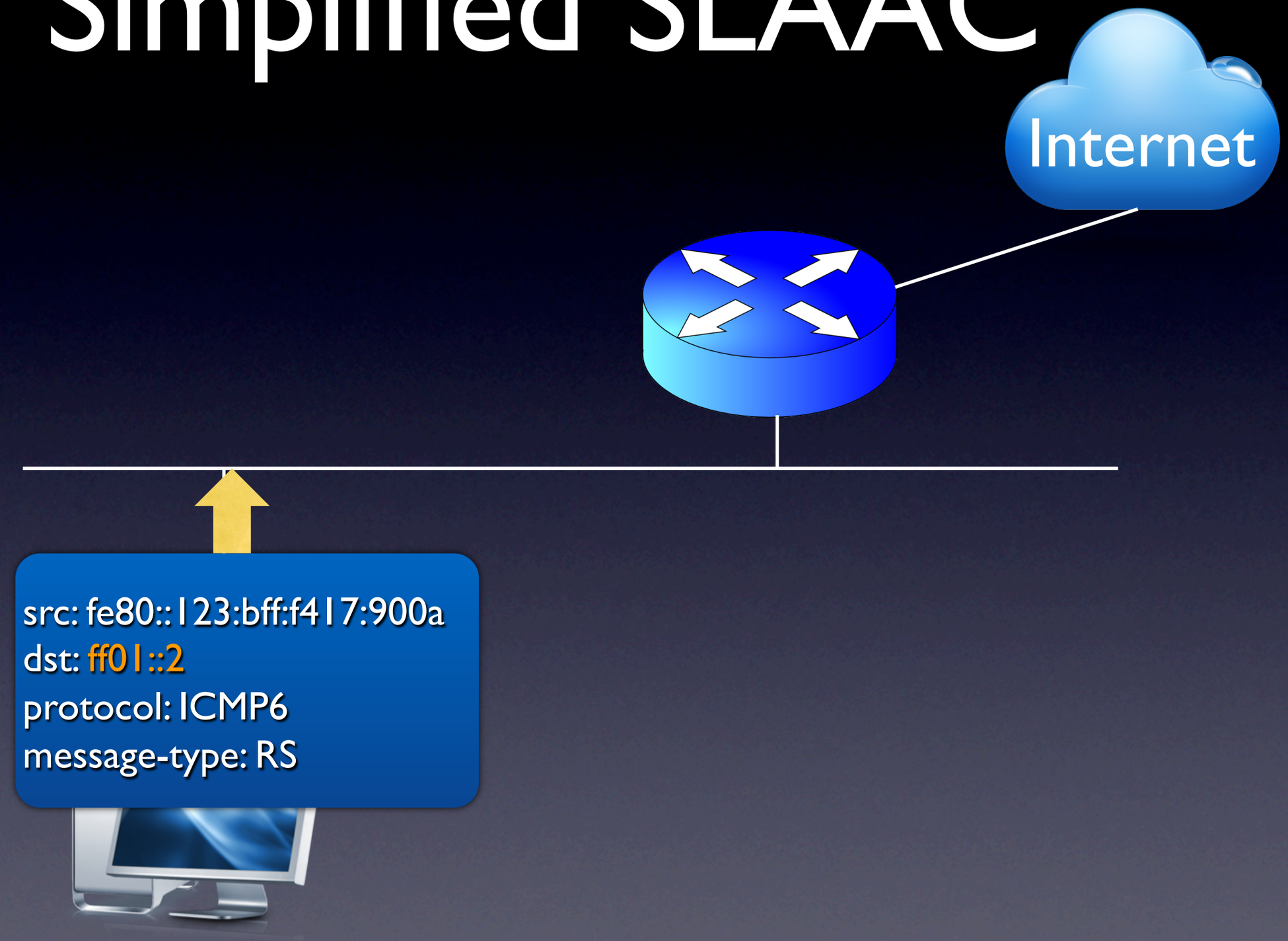
Simplified SLAAC



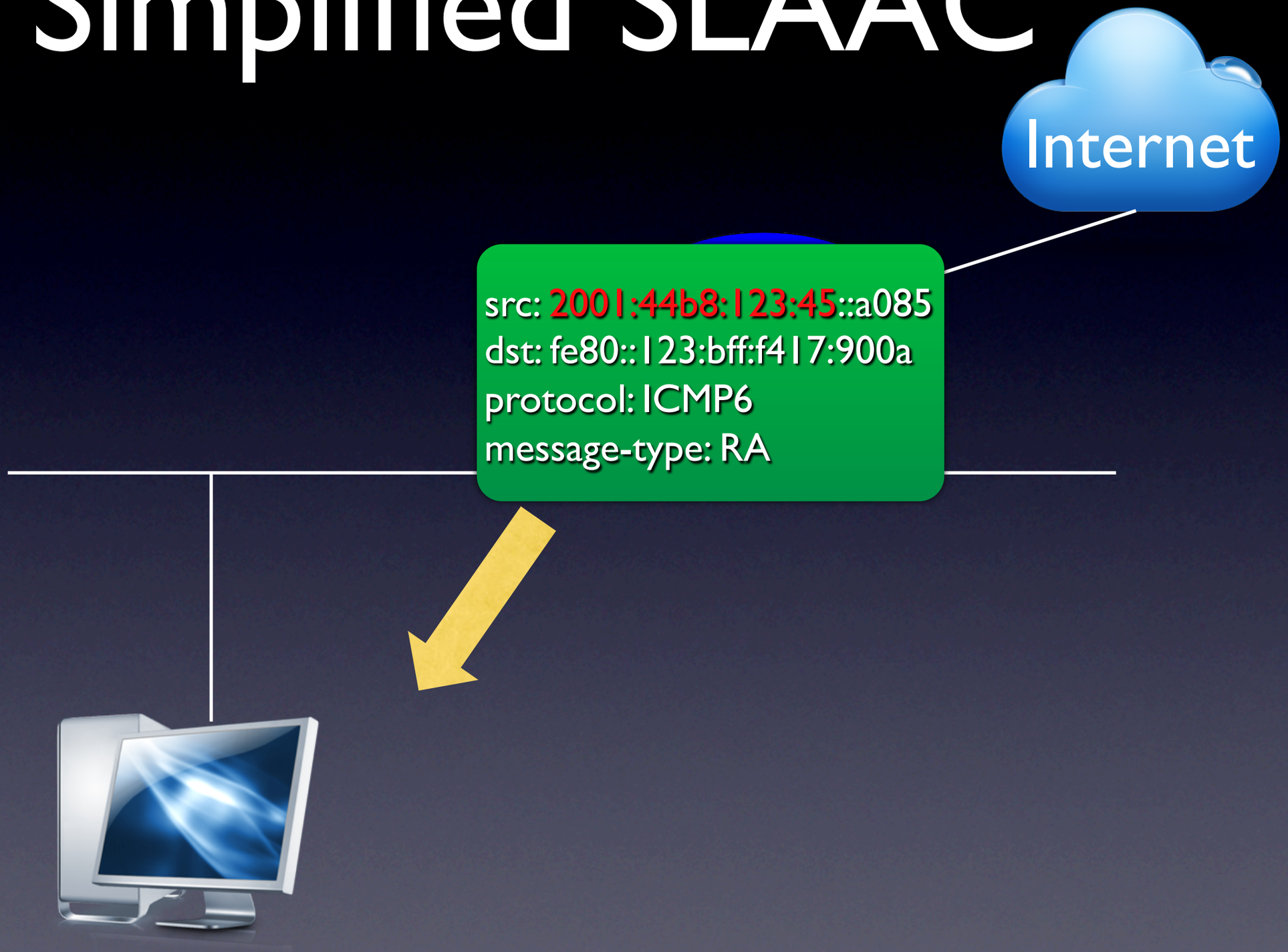


fe80::123:bff:f417:900a
2001:44b8:123:45:123:bff:f417:900a

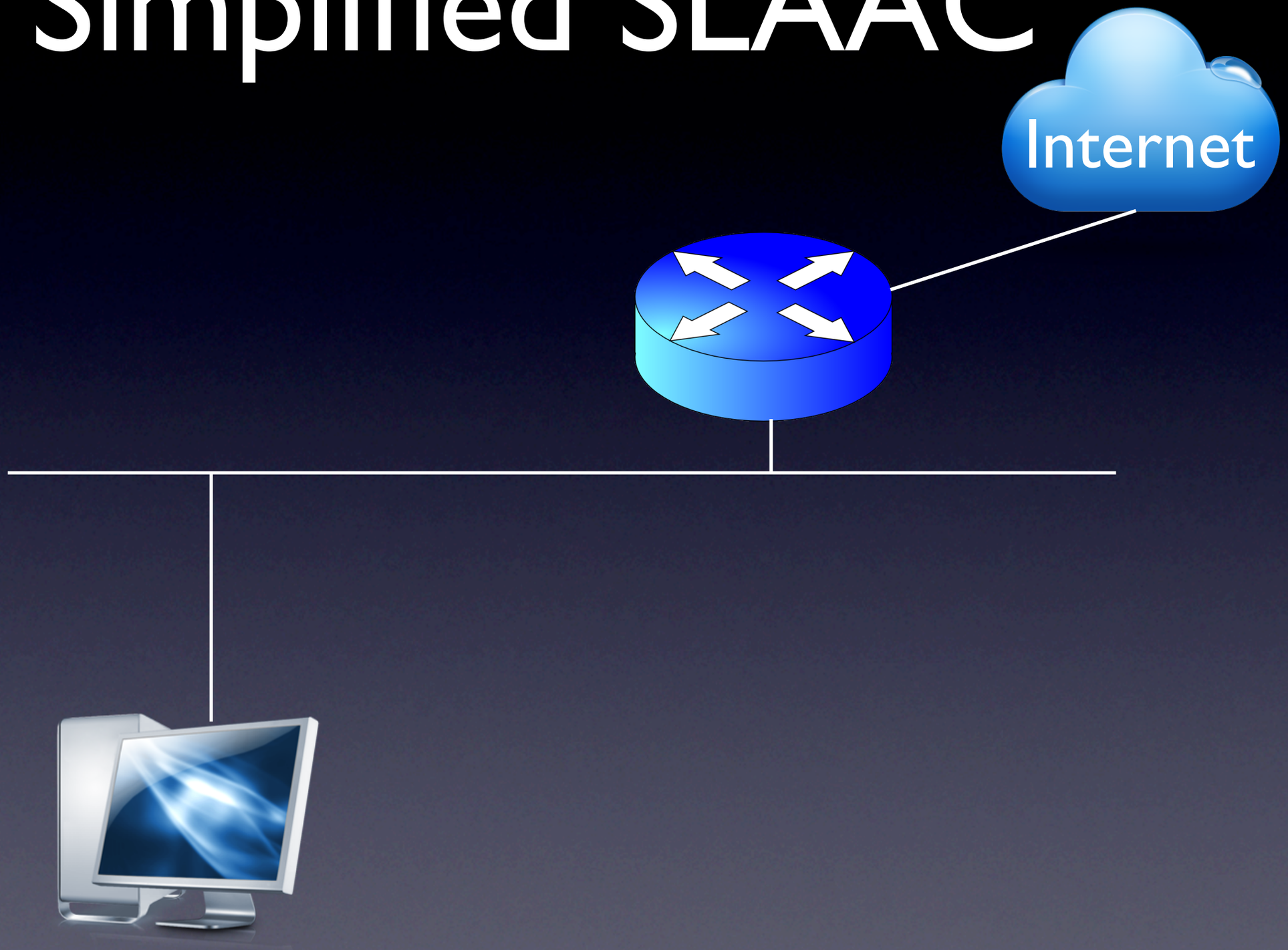
Simplified SLAAC



Simplified SLAAC



Simplified SLAAC



fe80::123:bff:f417:900a

2001:44b8:123:45:123:bff:f417:900a

Address Assignment

- Static configuration and DHCPv6 assignment are also available (not often used)
- A flag in the RA message indicates to the host whether DHCPv6 should be used
  - (so the router needs to know - unlike v4)
- Duplicate Address Detection (DAD) automatically rectifies cases where dynamic addresses collide.

No Nats!

- There is no common private address space for IPv6
  - Unlike 10.0.0.0/8, etc in IPv4
  - Response 1 – use “site local” addresses
    Ooops!
  - Response 2 – use public addresses
    Sort of Ooops!
  - Response 3 – use ULAs

No Fragmentation

- IPv6 routers cannot fragment an IPv6 packet
- If a packet is too big to be forwarded then the router sends an ICMP6 “Packet Too Big” message to the source
- All IPv6 links MUST pass a 1280-octet IPv6 packet without fragmentation

Other Changes

- Flow-ID – more padding bits in the header!
- Option Handling – no significant changes

Summary

Significant IPv6 Changes:

- 128 bit address fields
- Replaced Broadcast and ARP with Multicast and SLAC
- Removed on-the-fly fragmentation with ICMP6 notification to source

Hints for Servers

It's Dual Stack

- IPv6 is not "backward compatible" with IPv4
- A network service MUST be capable of speaking IPv4 to IPv4 clients AND speaking IPv6 to IPv6 clients

Go Native

- Don’t rely on tunnelled IPv6-in-IPv4 to support your service
  - It’s unreliable
  - It’s inefficient
  - It’s difficult to debug
- Get a “native” IPv6 transit service

Site Security

There is no blind address scanning in IPv6 … but

- Don’t use low addresses for local hosts and routers
- Don’t be promiscuous with your DNS zone information
- Make sure your firewalls are IPv6-aware

IPv6 Address Plan

- Addressing each LAN with /64s
- Addressing point-to-point links with /127's

  (some folk use /124s, others use /64s, but my personal preference is for p-t-p links to be numbered from a /127 as ::0 and ::1)

Plug-n-Play Perils!

- IPv6 has auto-config capability via SLAAC and DHCPv6
  - But has no intrinsic defense against rogue Ras
  - And SEND has a really high overhead
    - And you may not want to play games with dynamic DNS for your service advertisements
- For server sites you may want to configure your servers manually
  - But be aware of the longer term implications of manually managing the address plan as the site grows over time

Firewalls

- ICMP6 is your friend – don't block it!
- Protocol 41 is used by IPv6-in-IPv4 tunnelling protocols – don't block it!
- There is NO NAT – if you want NAT-like asymmetric behaviour then you will need to explicitly configure it

Defensive MTU

- Don’t assume that everyone else knows what they are doing – they don’t!
- Set the MTU for all service delivery platforms to 1280 octets

Don’t forget the DNS

- Add AAAA records to the DNS for your service points
- Dual-Stack the servers that are authoritative for your DNS domains
- Add AAAA records to the nameservers’ records

Swinging infrastructure

- When you dual stack the DNS namespace, be aware that some applications may then switch to IPv6 transport with the DNS
  - Neither DNS (nor NTP) have no inbuilt IPv4 / IPv6 preference rules, So when you dual stack DNS (and NTP) services expect immediate IPv6 uptake from local clients
- SNMP IPv6 support should be examined carefully for your site

Be nice to the less fortunate

- Set up a local outbound 6to4 relay
  - Use the 192.88.99.1 anycast address as the IPv4 wrapper

    (ensure your transit will permit outbound packets with this anycast IPv4 source address)

- In the server environment direct 2002::/16 to the outbound 6to4 relay

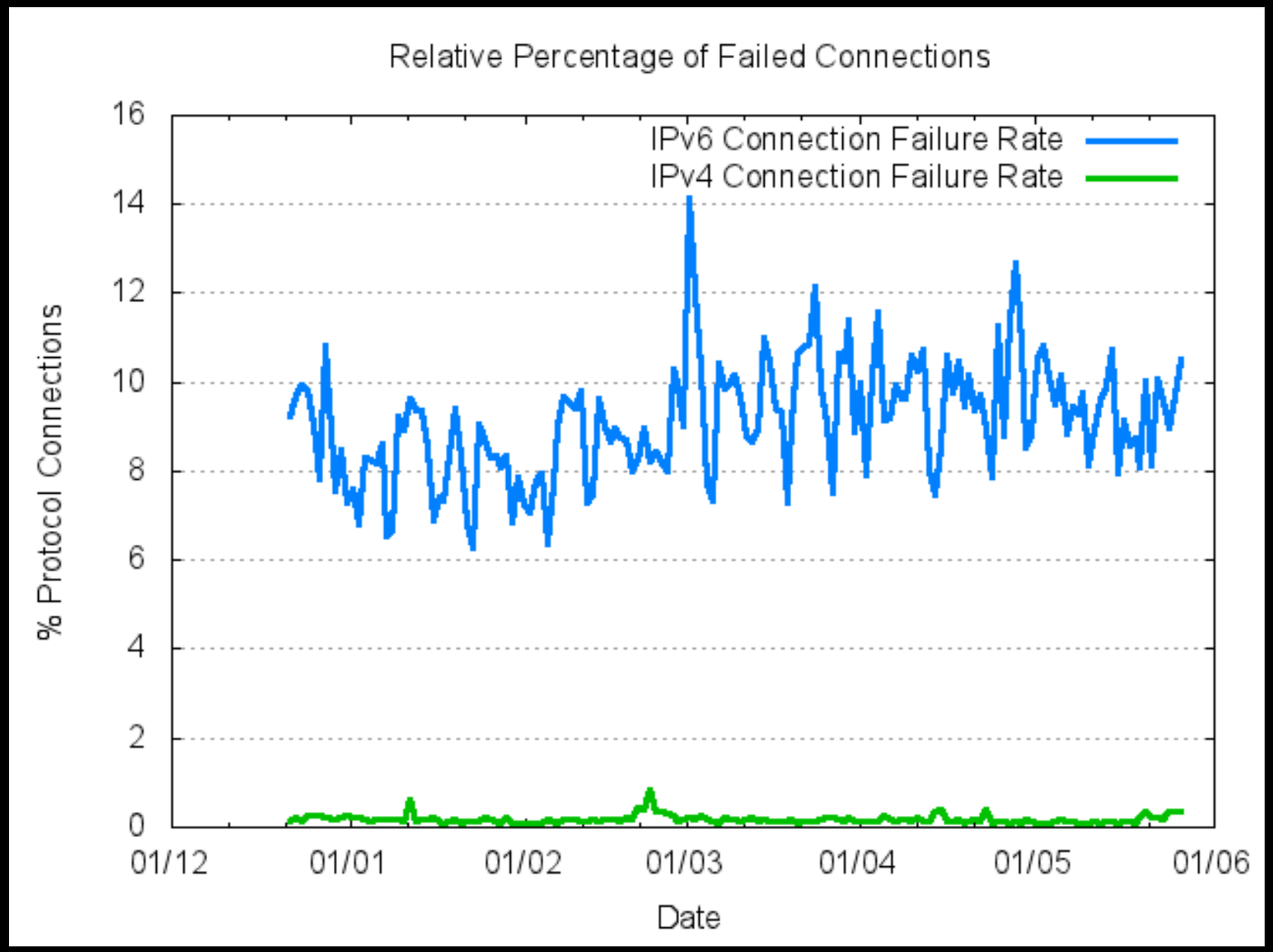
(Experimental) Option

- Set up a local Teredo Relay
- Use the Miredo packet to set up a Teredo local relay
- Set up a route to 2001::/32 via the outbound relay

Its not perfect – yet!

- IPv6 still has its wrinkles
- The metrics of connection failure are far higher than IPv4

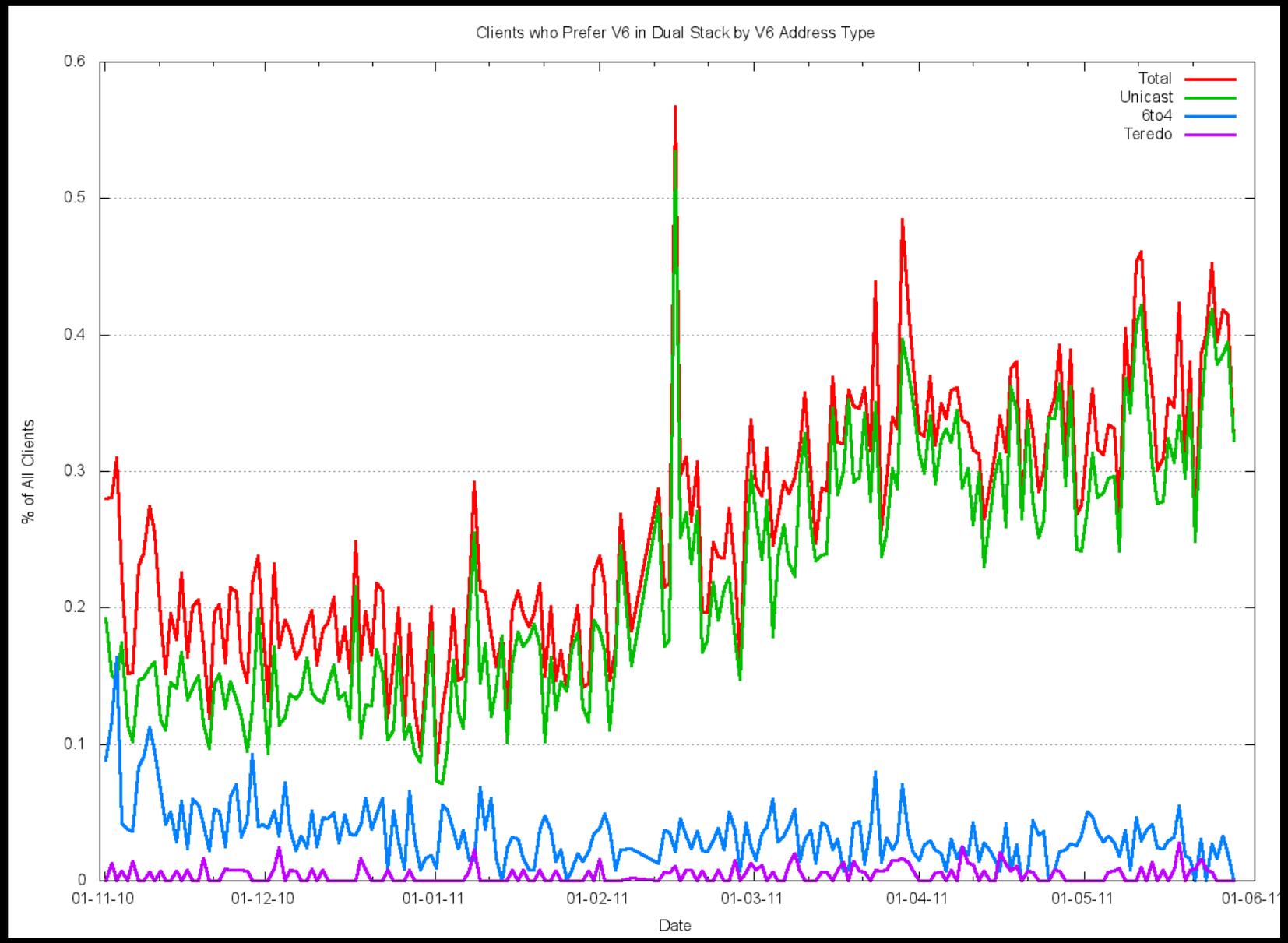
Connection Failure



Its not perfect – yet!

And the IPv6 use level is still low

IPv6 Connection Ratio



But we need to move quickly with IPv6

as there is no more IPv4!

Thank You

Acknowledgement

Many thanks to Mark Newton of Internode. I've used some of the material from his AUSCERT 2011 presentation on IPv6 Security in this pack.

(and Internode Rock! Their IPv6 service is truly awesome!)