Redpill Linpro

# An IPv6 primer

Tore Anderson
CG Security and Networking
Redpill Linpro
Altibox, Stavanger, June 2011

Redpill Linpro

# Do feel free to play around with the demo WLAN

Redpill Linpro

Do feel free to stop me with questions or comments

Redpill Linpro

# Today's agenda

1) Introduction
2) IPv4 depletion and the consequences thereof
3) Dual-stack brokenness and «*World IPv6 Day*»
4) IPv6 tech update for the IPv4 savvy
5) IPv6 deployment scenarios for ISPs
6) IPv6 deployment scenarios for content providers
7) Finishing discussion

Redpill Linpro

# Introducing myself

- Working for Redpill Linpro in Oslo for the last 10 years
  - Before: UNIX sysadmin + jack of all trades
  - Now: Mainly IP/storage networking and data centres
- IPv6 became a professional hobby for me back in 2008
- These slides are available from: **http://fud.no/talks**
- Contact information:
  - tore.anderson@redpill-linpro.com
  - @toreanderson
  - +47 95 93 12 12

Redpill Linpro

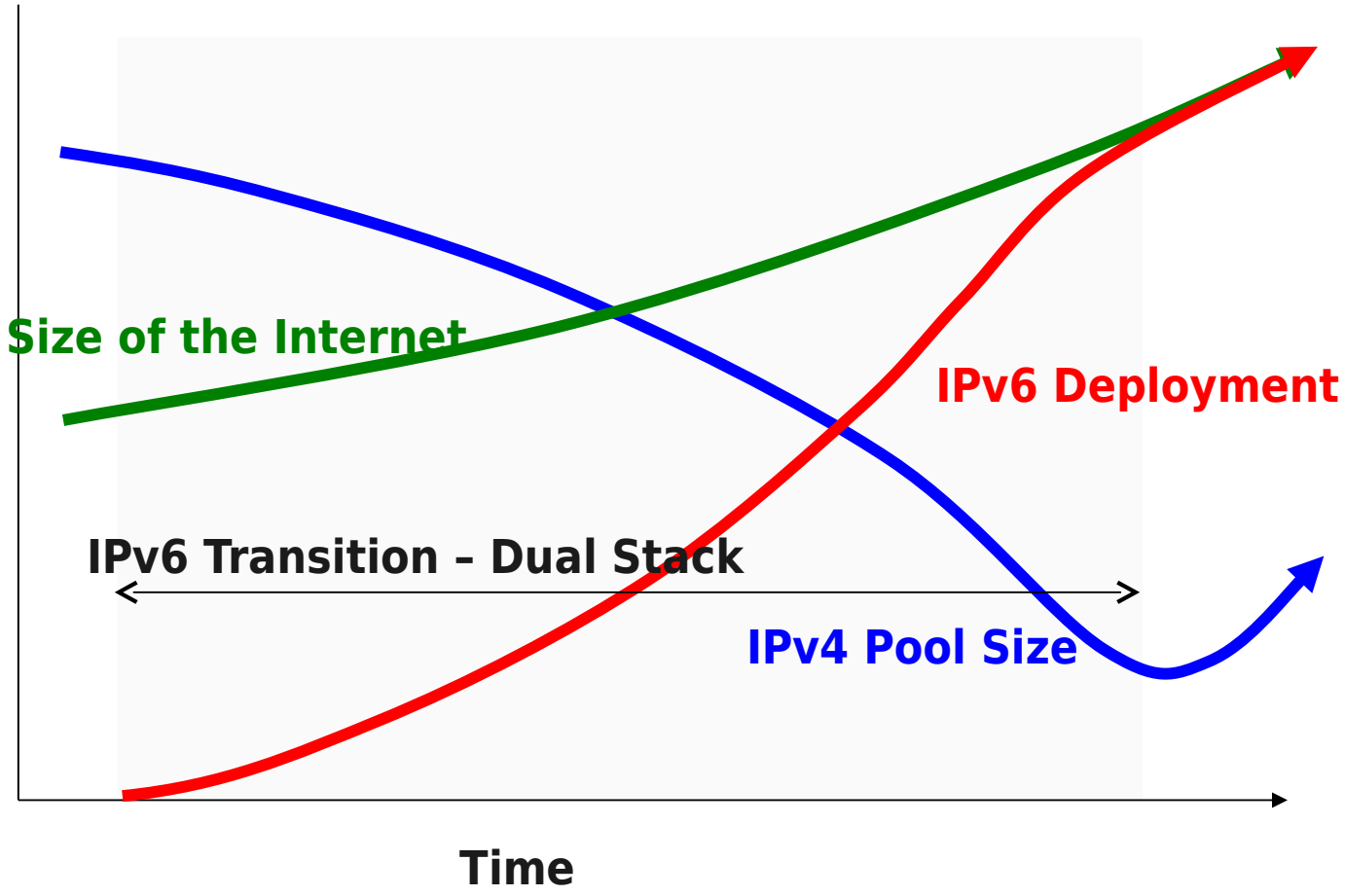
# Introducing my employer

- Does pretty much anything that involves open-source software
- Offices all over the Nordic countries, customers world-wide
- **Managed Services** hosts and maintains customers' IT systems
  - Design and set up the customers' application stacks
  - Data centre hosting and internet connectivity
  - 24/7/365 server/OS/application maintenance and monitoring
  - Same SLA for IPv6 as for IPv4, of course
- **http://www.redpill-linpro.com**

Redpill Linpro

# IPv4 depletion

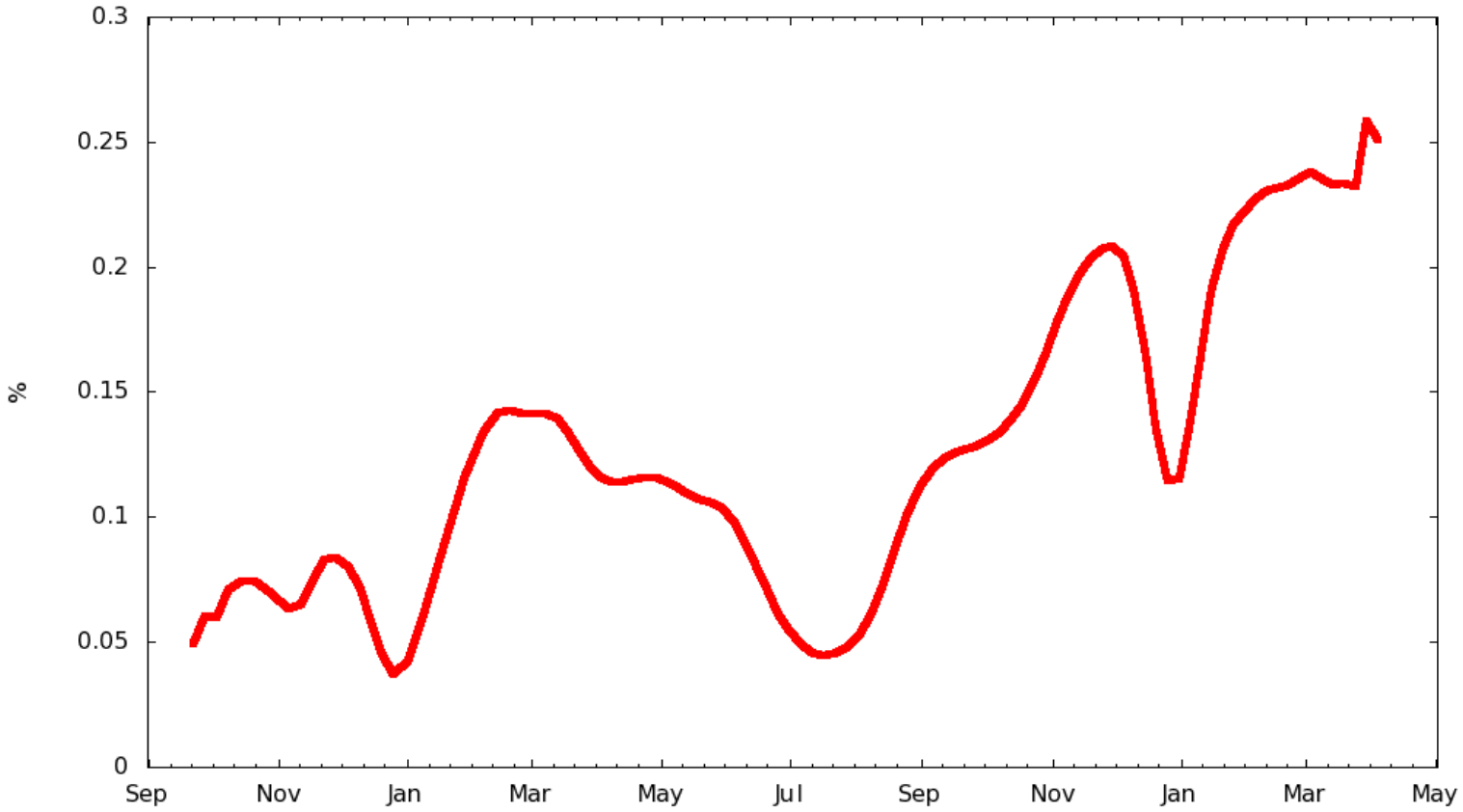
# The IPv6 Transition Plan




*With thanks to Geoff Huston/APNIC*

# Reality check 1

Approx. **1 out of 400** Norwegians have IPv6 connectivity




*(Connections using native IPv6 to A-pressen Digitale Medier and VG Nett, own data)*

# Reality check 2

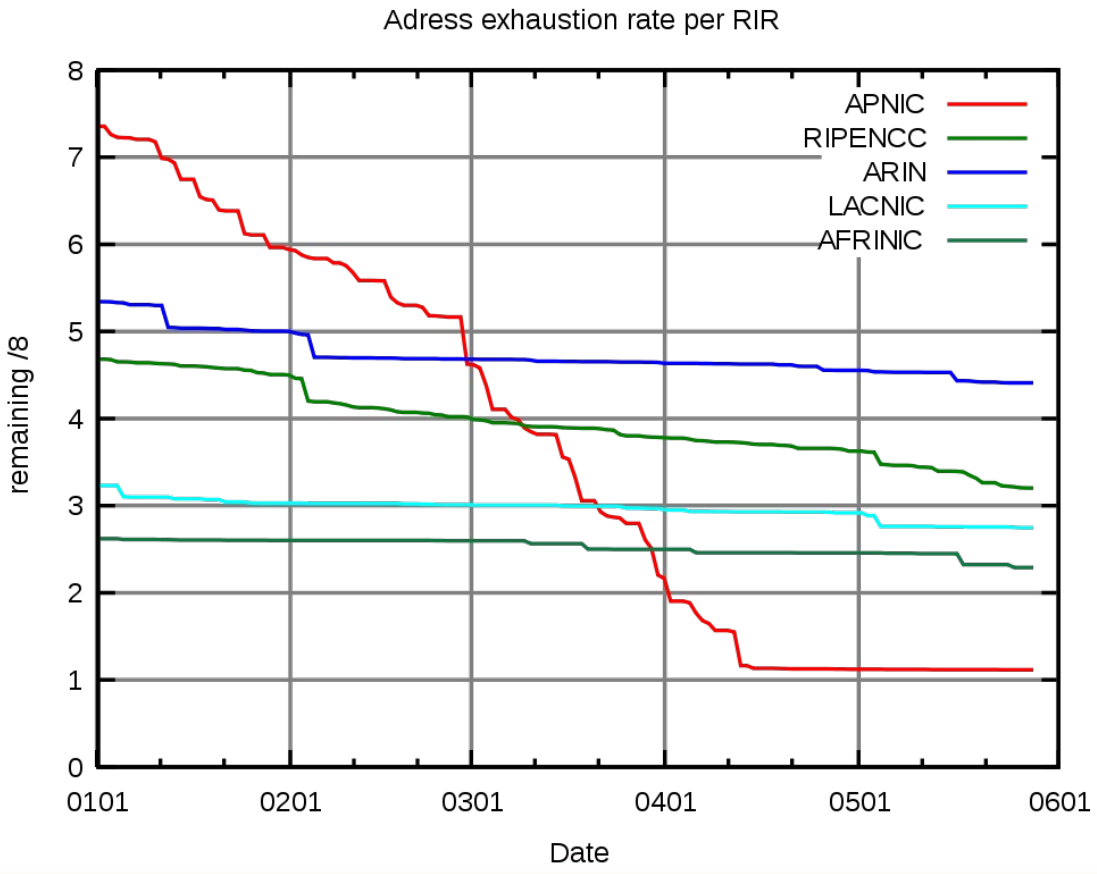
- The IANA global reserve depleted on February 3rd
- APNIC (Asia/Pacific region) followed suit on April 15th


*(Available/unused IPv4 addresses in the top two distribution tiers.)*

*With thanks to Geoff Huston/APNIC*

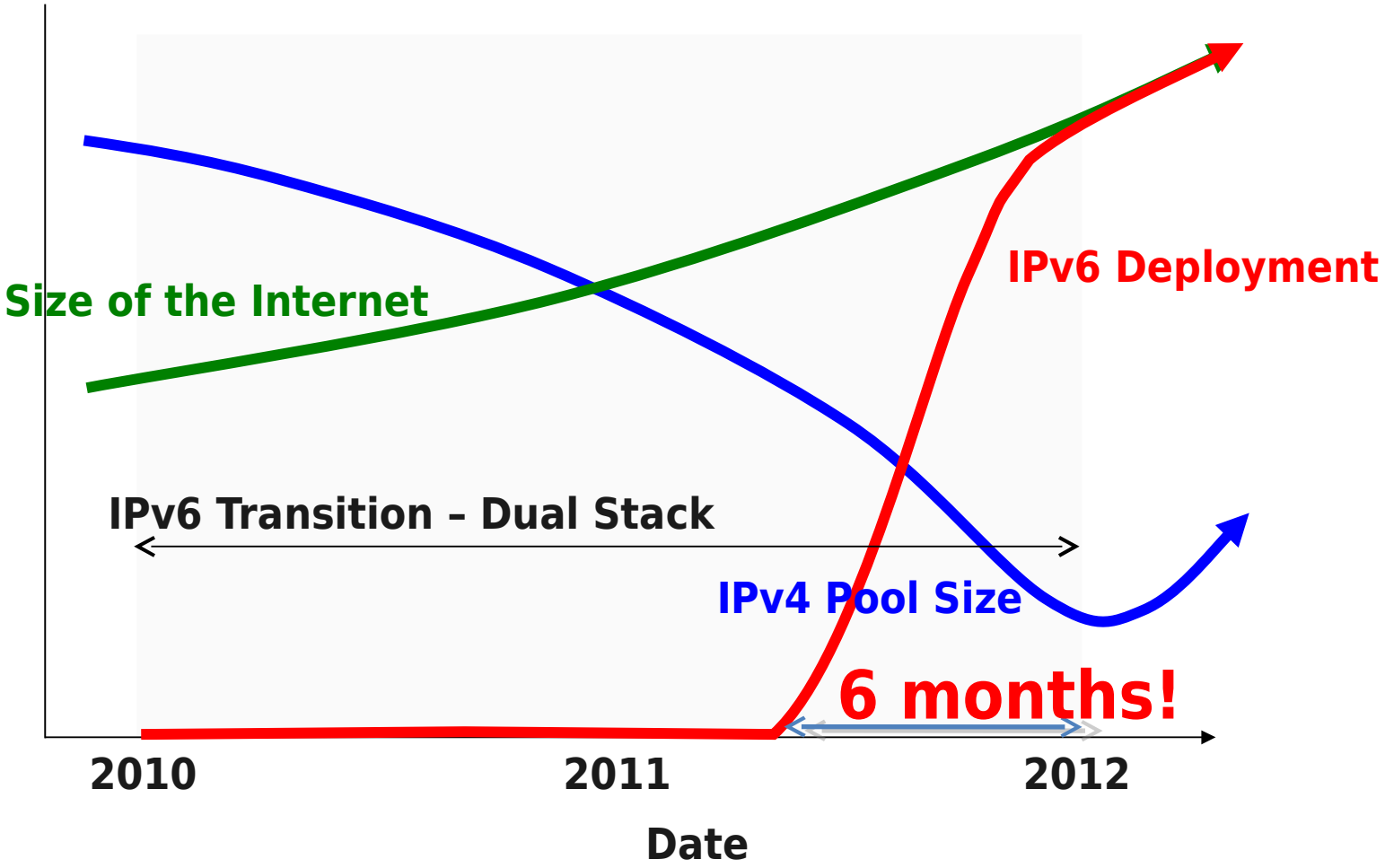
# Zooming in

- APNIC, the Asia-Pacific Network Information Centre, should really have changed its name to Pacific-Asia Network Information Centre in the beginning of March...
- Increased demand in the RIPE region also, but no hoarding – yet
- Anticipated RIPE exhaustion before the end of the year, or shortly after




*With thanks to Wikipedia*

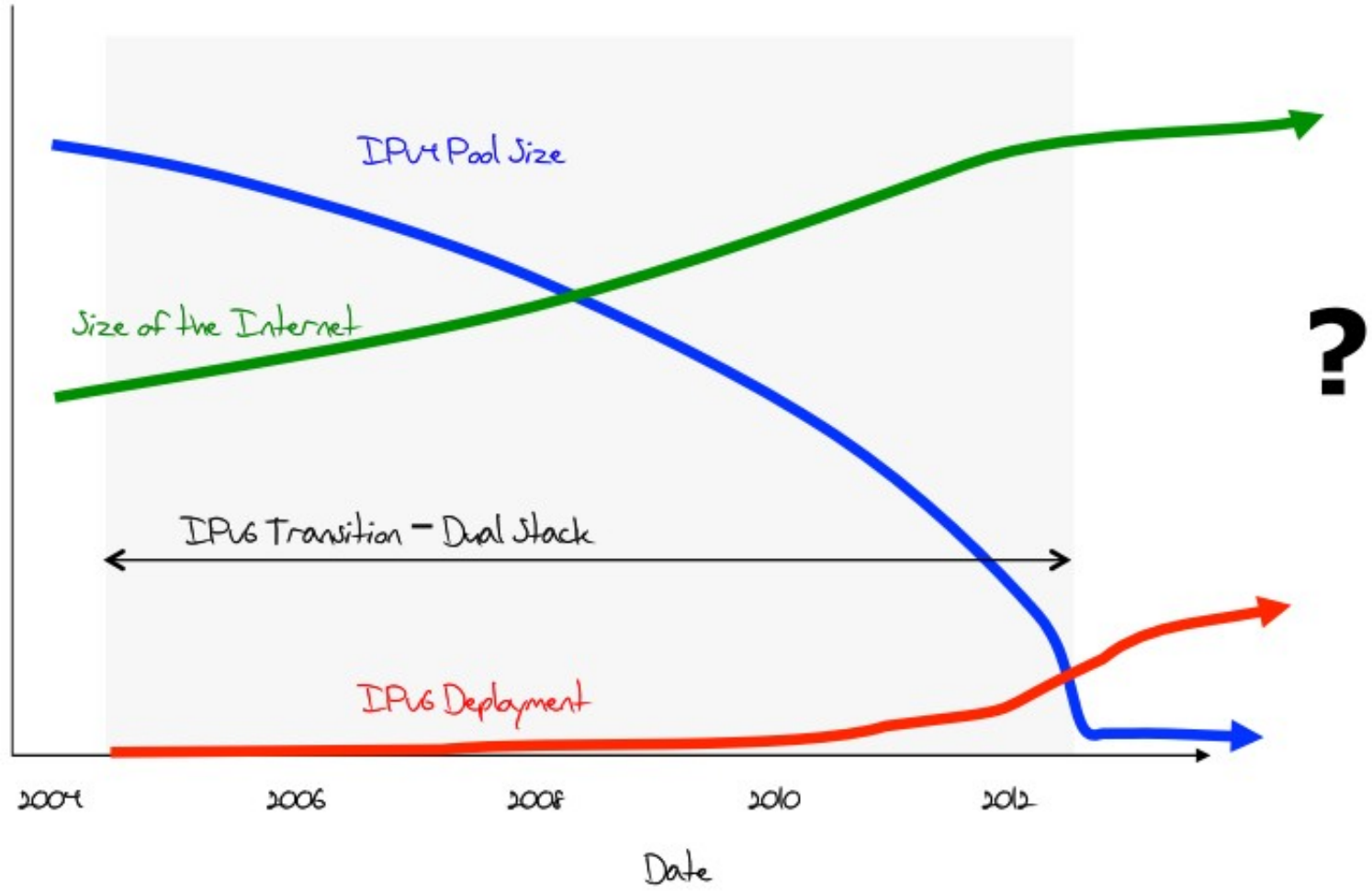
# The IPv6 Transition Plan - V2.0




*With thanks to Geoff Huston/APNIC*

Redpill Linpro

# What's **actually** about to happen




*With thanks to Geoff Huston/APNIC*

# Consequences of IPv4 depletion

- ISPs won't be able to acquire enough addresses to sustain growth
- Available solutions:
  - Deploy an additional layer of NAT; share IPv4 addresses between multiple subscribers
  - Buy IPv4 addresses – starting at US$11.25 in bulk quantities
- Content providers are very worried about the effects «Carrier Grade NATs» will have on network performance and functionality


*Extensive sharing? No problem, all payload will get there eventually!*
*(Picture from sales material of Honest Hank's Hardly Used MCs & CGNs.)*

- They therefore desperately want IPv6 as an alternative and faster path between the content and the end users – and want ISPs to do their part

# Who goes first?

- For long, IPv6 suffered from a chicken & egg dilemma
  - Why should an ISP invest in IPv6, when there's no content?
  - Why should a content provider, when there are no users?
- This is **no longer the case**, there's plenty of content today!















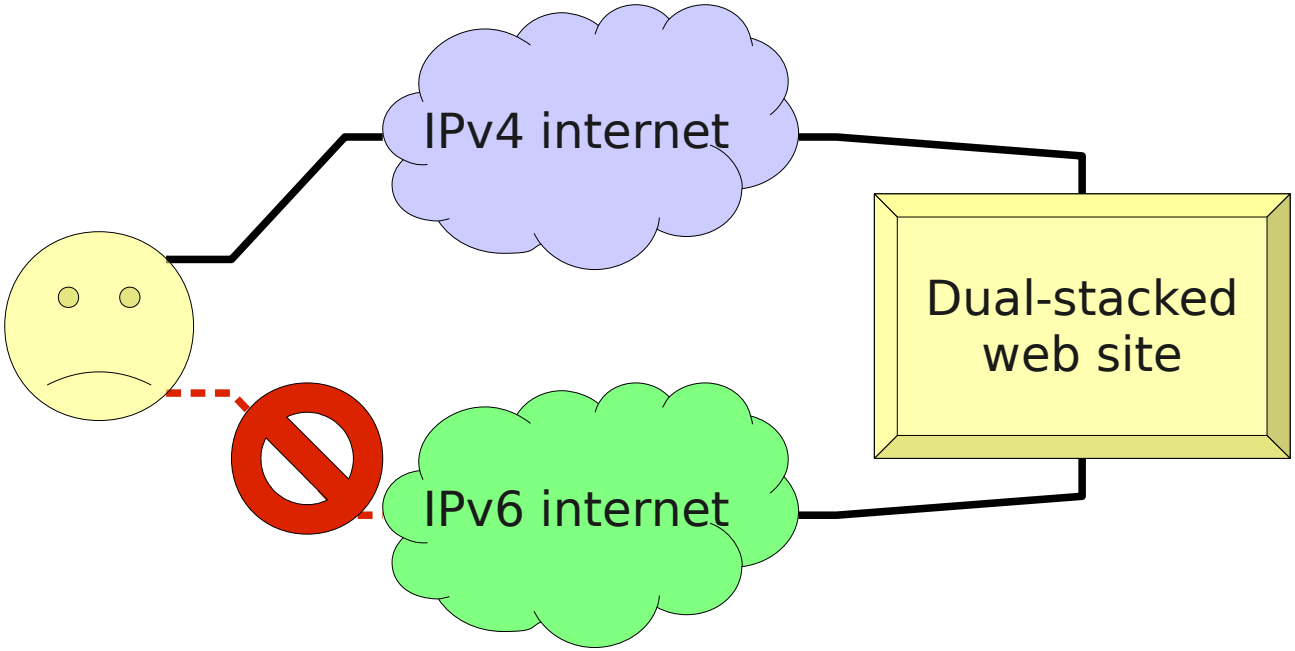


*With thanks to all of the above*

Redpill Linpro

# Dual-stack brokenness and «World IPv6 Day»

# IPv6 isn't entirely problem-free




- **Dual-stack brokenness** causes a bad user experience
- End-user's OS or web browser incorrectly thinks there's IPv6 connectivity
- IPv6 is tried first and fails; long timeouts before fail-over to IPv4
  - Live demo: **http://broken.redpill-linpro.com**
- Adding IPv6 results in a overall **less accessible** service than IPv4-only
  - → Economic disincentive for content providers to deploy IPv6

Redpill Linpro

# Researching dual-stack brokenness

- In 2009 we enrolled two of our customers in an experiment
  - **VG Nett**
  - **A-pressen Digitale Medier**
- Purpose of the experiment was to:
  - Determine whether or not deploying IPv6 was (sufficiently) safe
  - Find out if there were any systemic failures
    - And if so, try to get them fixed
- Results openly shared with the ISP and content communities:
  - http://fud.no/ipv6

# Measurement setup

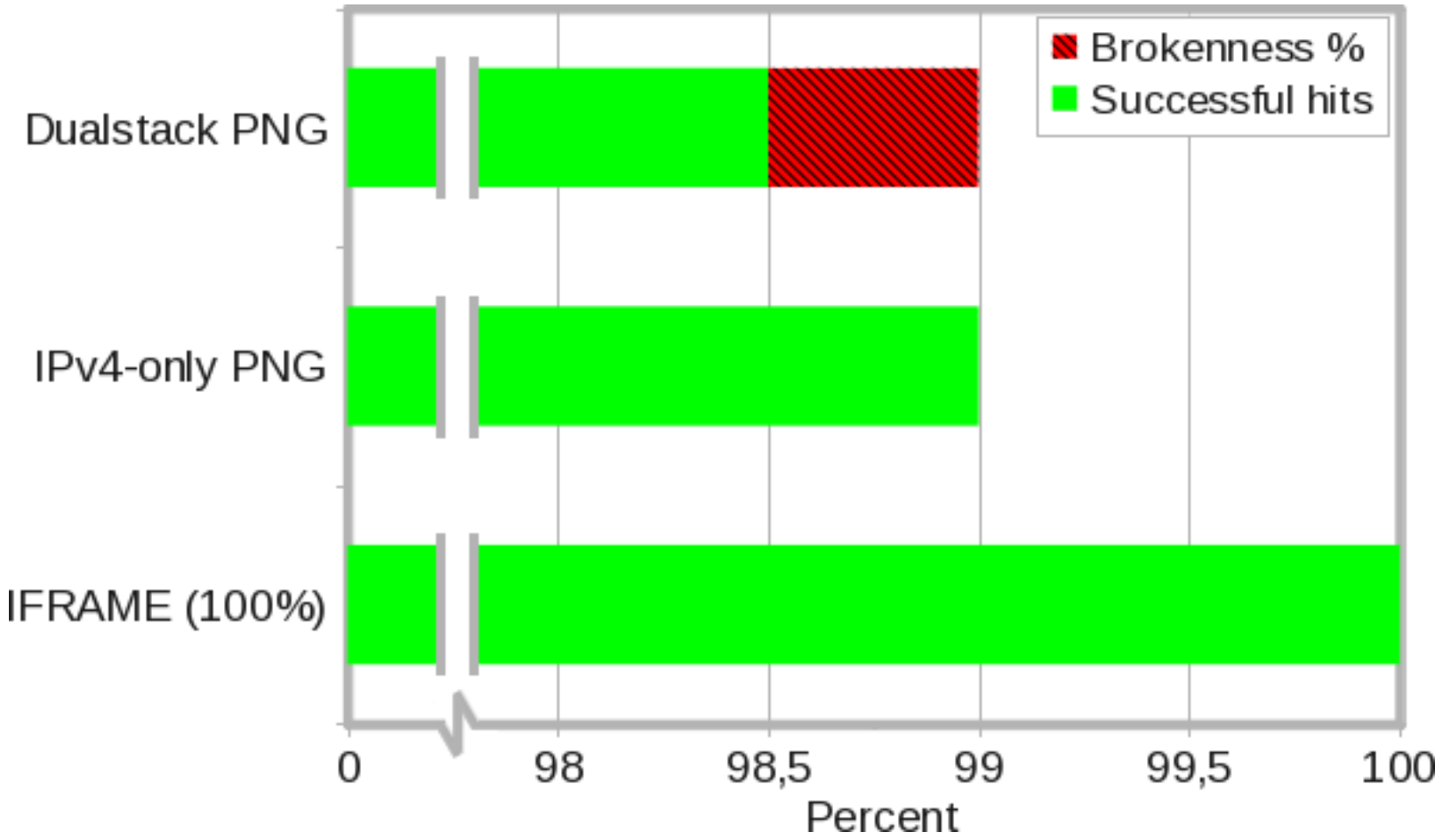

- Invisible IFRAME embedded in customer's HTML templates
- Single stack IPv4 only
- IMG links in random order

- 1x1.png
- IPv4-only

- 1x1.png
- Dual-stack

**Assumption:** We should see the same amount of hits to the two 1x1 PNGs. If not, we're seeing brokenness.

# Definition of «brokenness»




- The brokenness percentage is the spread, in percentage points, between the amount of successful hits to the IPv4-only PNG and to the dual-stacked PNG. In this example: **0.5%**.

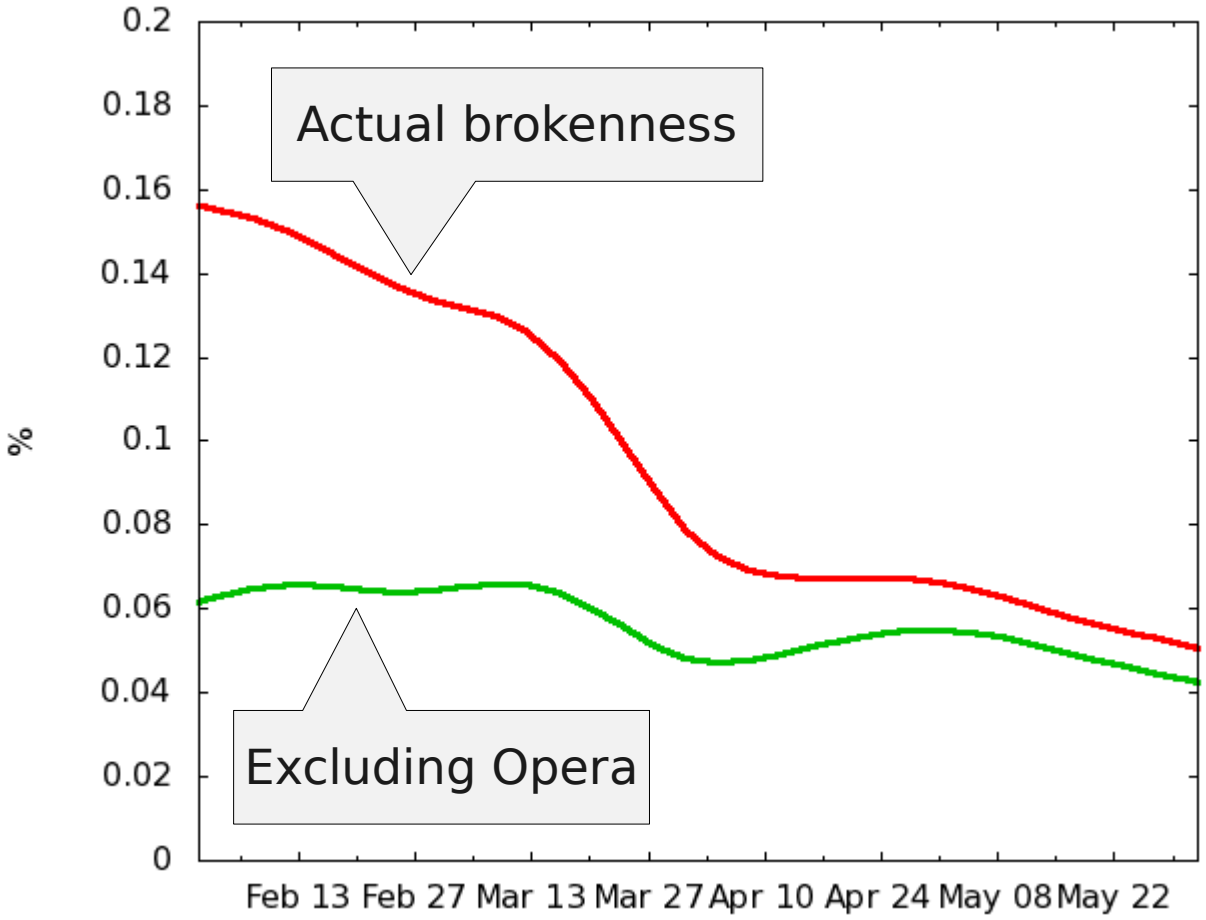
Redpill Linpro

# Initial findings

- **0.2-0.3%** brokenness
- Certain sources of brokenness were standing out
  - Opera web browser, running on Windows
  - Mac OS X
  - Certain networks (enterprises, universities), some ISPs
- 70-80% of IPv6 traffic was «transitional IPv6»; 6to4 and Teredo
  - IPv6 tunnelled inside IPv4, so can't possibly be more reliable
  - There's no real reason to use either in preference to IPv4
- The results were considered by VG and APDM to be too broken, deployment therefore postponed until situation had improved
  - The complete lack of IPv6-enabled users didn't help either

# Opera web browser on Windows

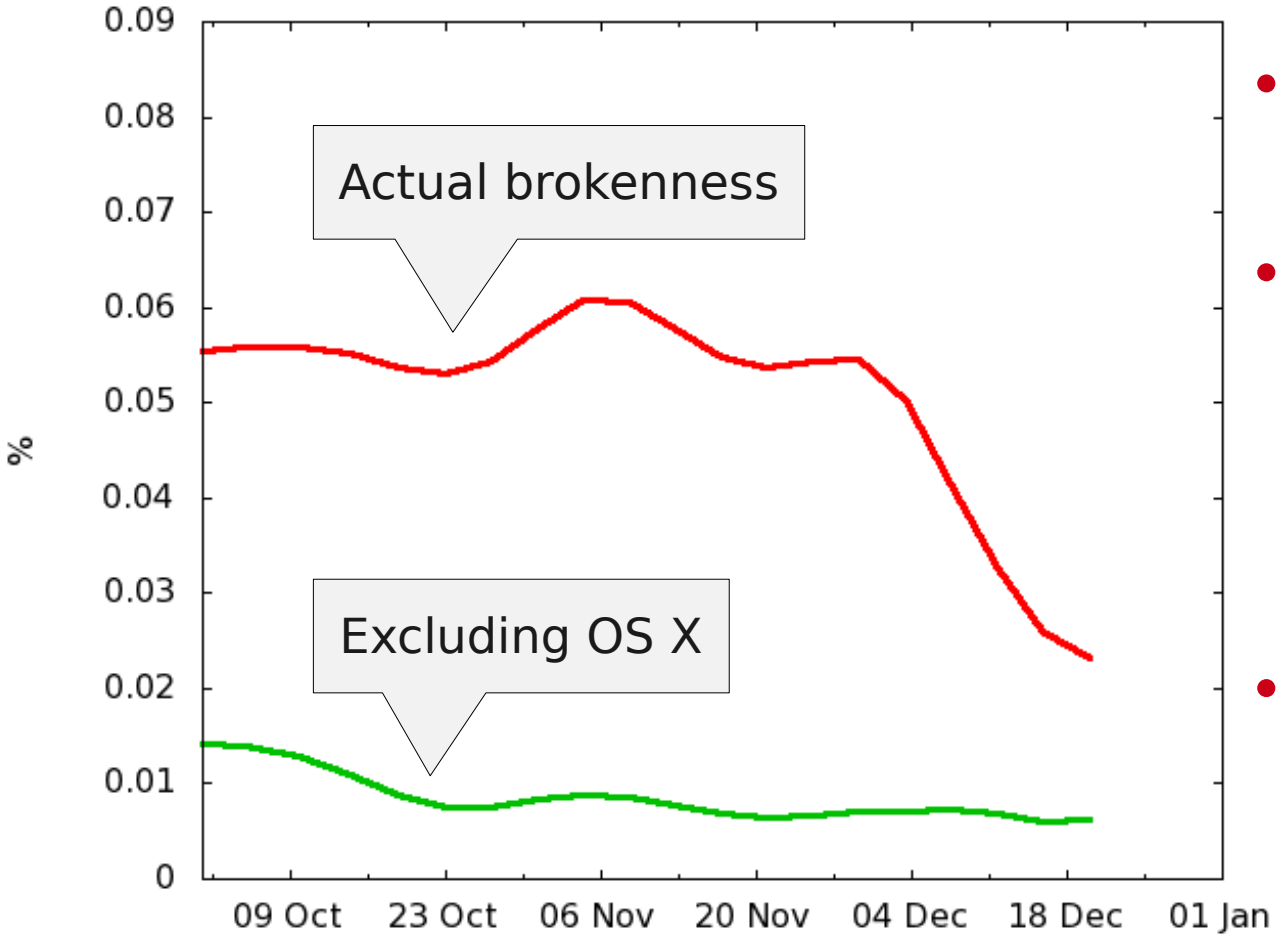
- Recent Windows will **automatically** enable 6to4 and/or Teredo
  - ..but de-prefers their use in the system resolver (RFC 3484)
- Opera, however, used its own built-in resolver




- Started nagging them about it
- Version 10.50, released the 22nd of March, fixed the problem
- Brokenness halved within a few weeks
- Also less 6to4/Teredo traffic

# Mac OS X

- Mac OS X does not implement RFC 3484 and unconditionally preferred IPv6, including 6to4 and Teredo, above IPv4
- Does not automatically enable 6to4 but is duped by Rogue RAs




- Started nagging them about it
- Version 10.6.5, released 10th of November, de-prefers IPv6 completely if local 6to4 addresses are present
- No upgrade path for one-fourth of their users (running 10.4 and 10.5)

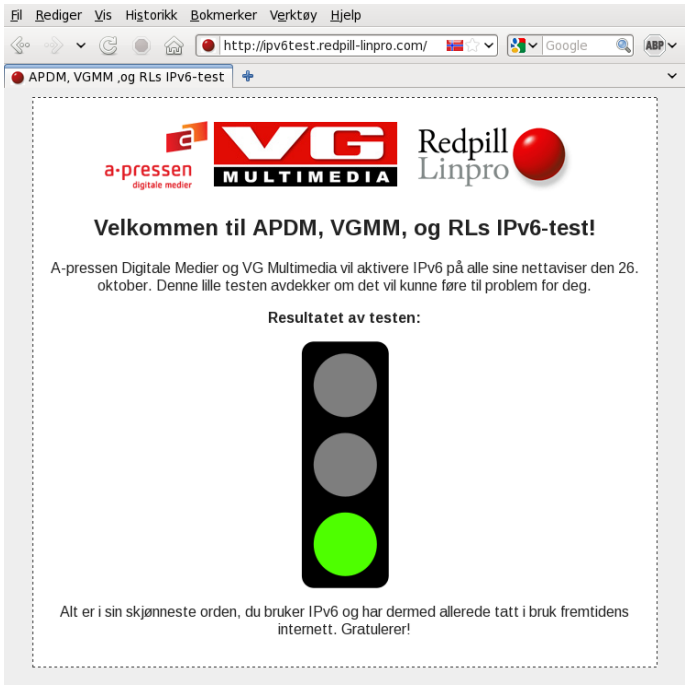
Redpill Linpro

# Rogue RAs

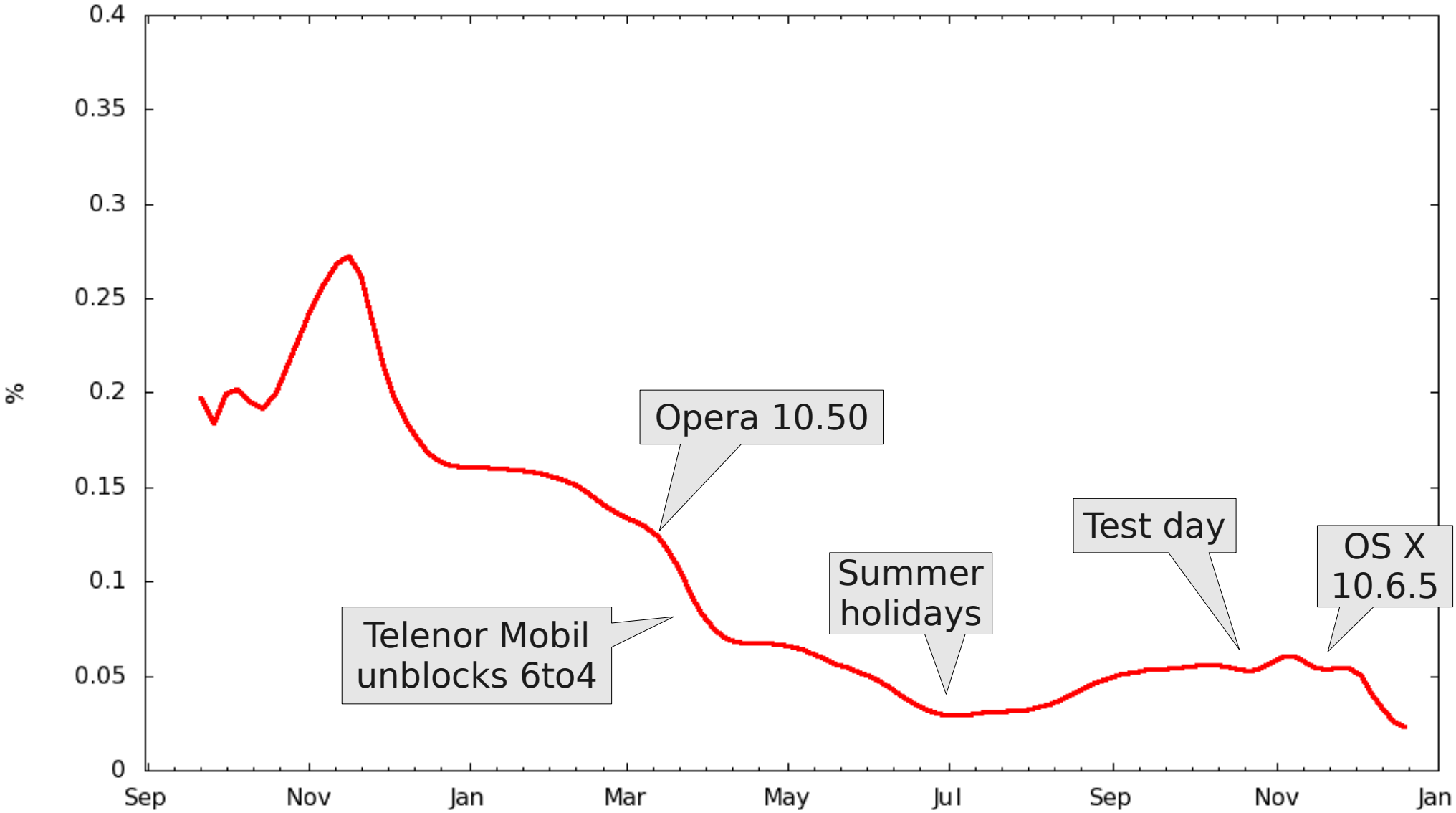
- Hosts that falsely announce themselves to the local network as IPv6 routers, often using the 6to4 prefix
  - Breaks dual-stack for all the old Mac OS X hosts on the LAN
  - Observed **10%** brokenness from certain campus networks
- Most common cause is Windows **Internet Connection Sharing**
  - Microsoft has not yet published a patch for this bug
- Routers that do 6to4 by default – championed by Microsoft
  - http://msdn.microsoft.com/en-us/windows/hardware/gg463251.aspx#EZC
- The IETF is about to deprecate 6to4 entirely – best to avoid it
  - http://tools.ietf.org/html/draft-ietf-v6ops-6to4-to-historic
  - http://tools.ietf.org/html/draft-ietf-v6ops-6to4-advisory

# Production for VG and APDM

- In October we did a 24 hour production test, inspired by Heise.de
- Broken users were warned and sent to a test site which shows instructions on how to fix and/or get in touch with us for help
- The users didn't complain, but didn't really fix the problems either
- **AAAA records permanently deployed the 21th of December**

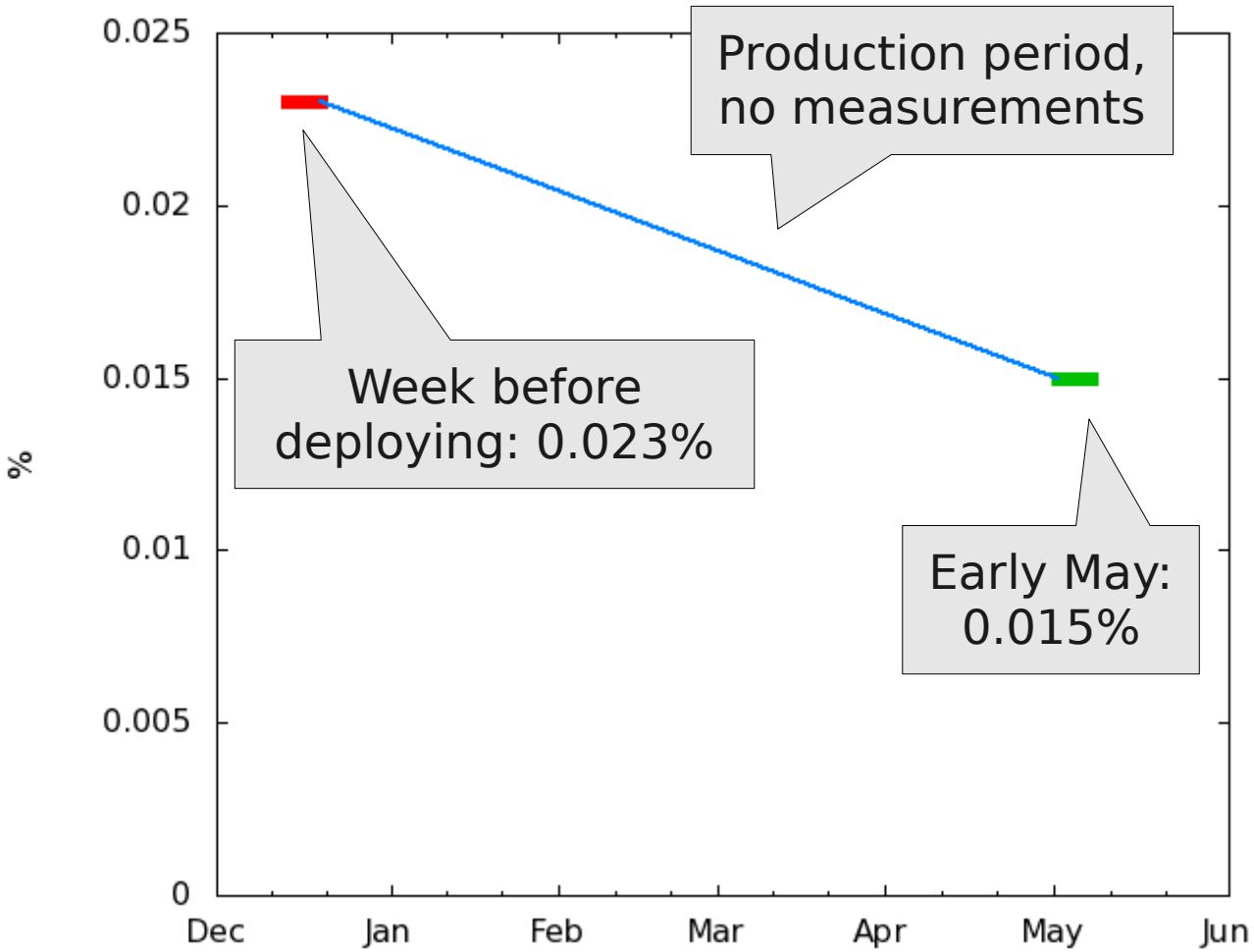



# From the start until production




Brokenness over the last seven days before production: **0.024%**

# Brokenness status right now




- 35% decrease in brokenness levels during the four months of production
- Known bugs in Opera and Firefox has been fixed
- We're expecting several more fixes (Windows, Mac OS X) in the coming months
- ***World IPv6 Day*** will hopefully also help out

*(Dual-stack was turned off temporarily for a week in order to perform this new measurement)*

Redpill Linpro

# World IPv6 Day

# What is it?

- Essentially a copy of the IPv6 Day we had last October with VG and APDM, and an earlier «German IPv6 Day» with Heise Online
- Organised by non-profit ISOC
- Google, Yahoo, Bing, Facebook, Akamai, Limelight, as well as 400+ others participants, will deploy IPv6 access to their content simultaneously
- Tomorrow (8th of June 02:00 – 9th of June 02:00 local time)
- Fate-sharing between competitors makes it easier to do
  - A broken user that experiences Google as being down will do so for Google's competitors Bing and Yahoo aswell
  - «The Internet doesn't work» instead of «Facebook doesn't work»
- http://www.worldipv6day.org
- Redpill Linpro participates as a facilitator for other participants

# What to expect tomorrow

- My best guess: Not much difference from any other day
  - Remember VG and APDM are already dual-stacked, so every day so far in 2011 has been a «*Norwegian IPv6 Day*»
- But perhaps there's a few customers calling in with complaints
  - Please don't say «wait until tomorrow» - in all likelihood, several participating sites will stay permanently IPv6-enabled
  - There's several web sites that can detect brokenness
    - http://test-ipv6.com
    - http://ipv6test.google.com
    - Or simply ask if http://ipv4.google.com works fine and http://www.google.com doesn't

# Blanket recommendations

- Ensure browser and operating system is patched, especially:
  - Opera 11.10
  - Firefox 4.0
  - Mac OS X 10.6.5
  - Firmware of Apple AirPort/TimeCapsule to 7.5.2
  - Firmware of Cisco Linksys E-series to 1.0.04
- Windows users should disable tunneled IPv6
  - http://go.microsoft.com/?linkid=9732984 (FixIt #50433)
- Windows users should also disable Internet Connection Sharing
  - Note: The ICS bug affects **other** hosts on the LAN, esp. Macs

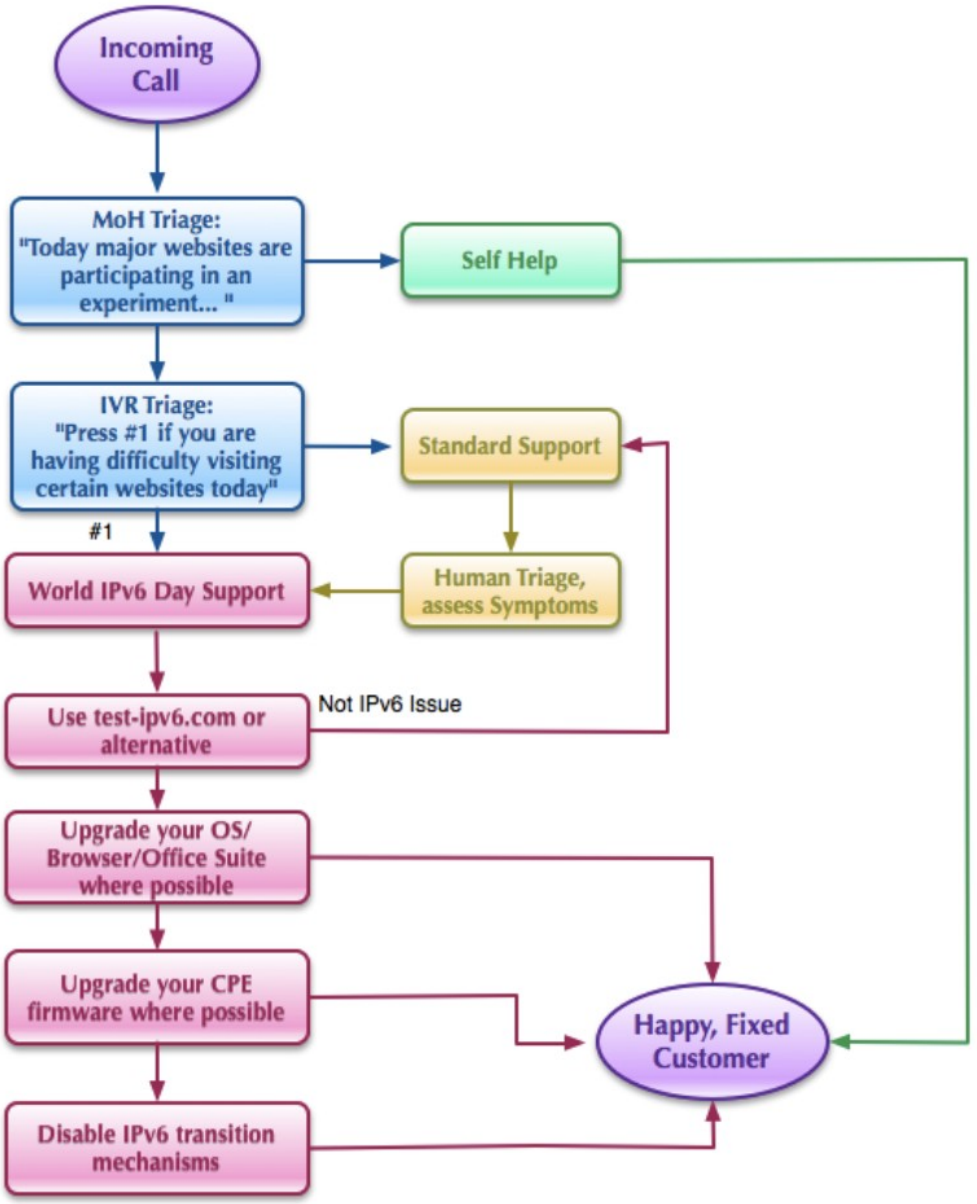
Redpill Linpro

# Still problems?

- The latest version of the Google Chrome browser glosses over brokenness by initiating IPv4 and IPv6 connections almost simultaneously, and picking the one that succeeds first
- There's a few other but less frequent problems too. I've documented all that I'm aware of in ARIN's IPv6 Wiki:
  - http://getipv6.info/index.php/Customer_problems_that_could_occur
- If all fails, simply disable IPv6 outright... :-(
  - Firefox: about:config -> network.dns.disableIPv6 -> True
  - Windows: http://go.microsoft.com/?linkid=9732985 (FixIt #50444)
  - Mac OS X: System Preferences -> Network -> Advanced -> TCP/IP -> Configure IPv6 -> Off
- Do let me know if you happen to come across a new and unknown bug!

Redpill Linpro

# An example process flow….




RIPE

*With thanks to Jan Žorž*

Redpill Linpro

# But enough with the scaremongering


- So there's really no reason to call in sick tomorrow

Redpill Linpro

# But enough with the scaremongering


- So there's really no reason to call in sick tomorrow, unless:
  - It's nice weather

Redpill Linpro

# But enough with the scaremongering


- So there's really no reason to call in sick tomorrow, unless:
  - It's nice weather, or
  - It's hard to get out of bed due to crappy weather

Redpill Linpro

# But enough with the scaremongering


- So there's really no reason to call in sick tomorrow, unless:
  - It's nice weather, or
  - It's hard to get out of bed due to crappy weather, or
  - You just got a new PS3 or Xbox game

Redpill Linpro

# But enough with the scaremongering


- So there's really no reason to call in sick tomorrow, unless:
  - It's nice weather, or
  - It's hard to get out of bed due to crappy weather, or
  - You just got a new PS3 or Xbox game, or
  - You're sick.

Redpill Linpro

# IPv6 tech update

Redpill Linpro

# IPv6 is

# HUGE

- 96 new bits of address space, 128 bits in total
- Allows the internet to continue growing well into the future
  - *«The killer application of IPv6 is the survival of the open Internet as we know it.»* - Lorenzo Colitti, Google
- Removes the need for NAT
  - End users will receive **minimum** 64 bits of address space
  - 72 or 80 bits appears to become the industry norm

# 96 extra bits. That's it?

- Yep, that's the **only** «must-have» feature. Really.

- There's several other technical adjustments:
  - New address syntax and DNS records
  - Link-local addresses are mandatory
  - Multicast is mandatory and extensively used by ICMPv6
  - Trimmed and extensible protocol header format
  - Fragmentation only done by end hosts, not the network
  - New methods of address auto-configuration
  - Address/router lifetimes and intrinsic multinetting support
  - De-facto standard subnet size of 64 bits

# IPv6 syntax

- 32 hexadecimal digits, divided into eight four-digit groups:

**2a02:00c0:1002:0011:0000:0000:0000:0002**

- Like IPv4, may be shortened by omitting leading zeroes in the groups:

**2a02:00c0:1002:0011:0000:0000:0000:0002**

↓

**2a02:c0:1002:11:0:0:0:2**

- Can be further shortened by compacting consecutive fields with the value zero to a set of double double colons:

**2a02:c0:1002:11:0:0:0:2**

↓

**2a02:c0:1002:11::2**

(The IPv4 equivalent of the above address would be **42.2.0.192.16.2.0.17.0.0.0.0.0.0.0.2)**

# IPv6 syntax, continued

- New **AAAA** record for forward DNS lookups:

***www.redpill-linpro.com. IN AAAA 2a02:c0:1002:11::2***

- New **ip6.arpa.** domain for reverse DNS lookups:

***2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1.1.0.0.2.0.0.1.0.c.0.0.2.0.a.2.ip6.arpa.***
***IN PTR www.redpill-linpro.com.***

- Must be enclosed by square brackets in URIs to avoid ambiguity:

***http://[2a02:c0:1002:11::2]:443/***

- A subnet or prefix is specified using CIDR prefix length notation:

***2a02:c0:1002::/48***

- The interface may be explicitly specified after a percent sign:

***fe80::1234:abcd%eth0***

# Link-local addressing and EUI-64

- Every IPv6-enabled interface require a link-local address
- Initialized automatically by the operating system's IPv6 stack
- Host part usually derived from the MAC address using **EUI-64:**

*1. Start with the 48-bit MAC address of the interface*

***00:30:1b:bc:7f:23***

*2. Pad it up to 64 bits by inserting **ff:fe** in the middle*

***0030:1bff:febc:7f23***

*3. Flip bit 7 (the universal/local bit)*

***0230:1bff:febc:7f23***

*4. Append to the well-known prefix **fe80::/64***

***fe80::230:1bff:febc:7f23***

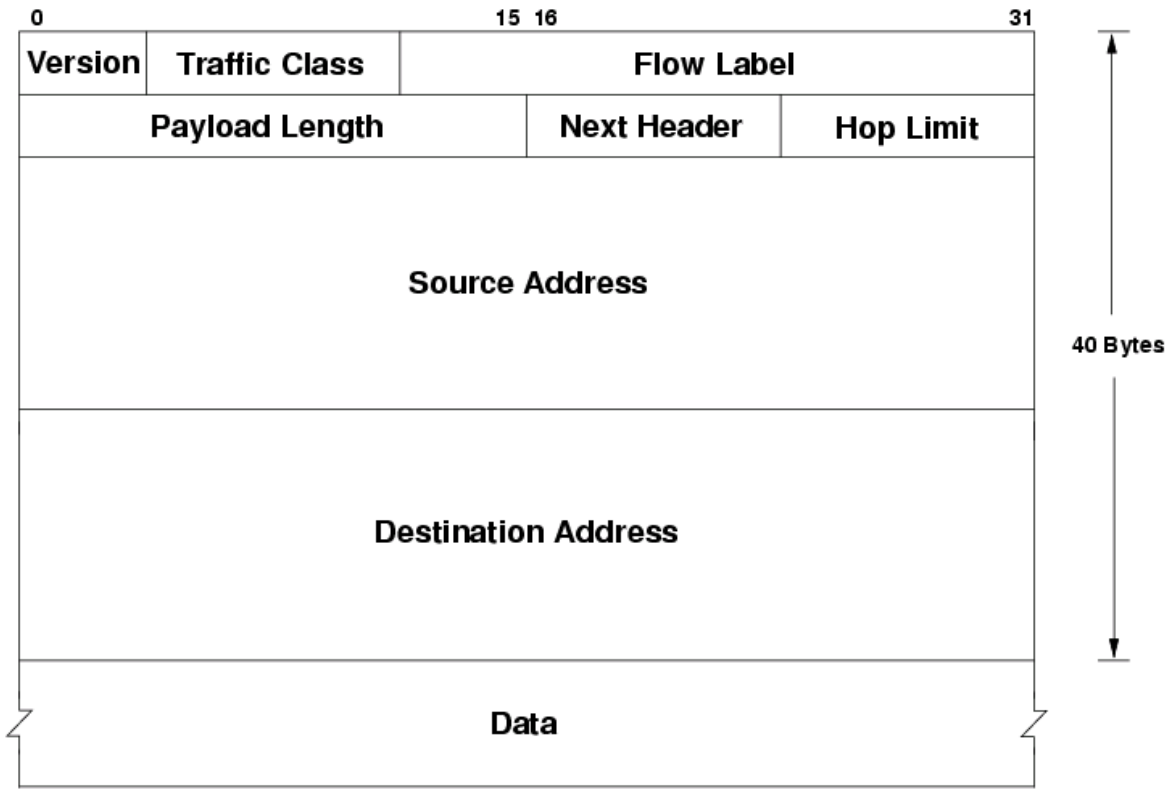
Redpill Linpro

# Multicast and ICMPv6

- Multicast is mandatory in IPv6
  - All IPv6 hosts automatically join the group **ff02::1** (*All Nodes*)
  - This replaces broadcast in IPv4
  - Try for example ***ping6 ff02::1%wlan0***
- **ICMPv6 Neighbor Discovery Protocol** replaces IPv4 ARP
  - Since all IPv6 hosts have addresses assigned and are listening to multicast, this no longer has to be «layer 2.5» function
- **ICMPv6 Duplicate Address Detection** prevents address collisions by keeping the newly-added address in a tentative state until its uniqueness on the link has been verified

# The IPv6 header format

- Removed since IPv4:
  - ***Header Length***, ***Options***
  - ***Identification***, ***Flags***, ***Fragment Offset***
  - ***Header Checksum***
- IPv4 ***Total Length*** replaced by ***Payload Length***
- IPv4 ***Time To Live*** renamed to ***Hop Limit***




- ***Next Header*** (replaces IPv4 ***Protocol***) specifies the type of the next header that follows in the payload: a L4 header, or an IPv6 Ext. Header
- The use of ***Flow Label*** is still undecided
- An IPv6 header is 20 bytes larger than a minimum-sized IPv4 header

Redpill Linpro

# IPv6 extension headers

- The ***Next Header*** field may contain a protocol number assigned to an **IPv6 Extension Header** instead of a layer 4 protocol (like TCP)
- An Extension Header always begins with another ***Next Header*** field, then a ***Length*** field, then the Extension Header Payload
- Multiple Extension Headers will form a **linked list** that eventually leads to the layer 4 protocol header
- Routers ignore them, except for the special **Hop-by-Hop** extension header
- Used by e.g. IPSEC, Mobile IPv6, and fragmentation control
- **The good**: Makes IPv6 very easily extensible
- **The bad**: Makes it hard for intermediate devices to do filtering at layer 4 and above, as the upper-level protocol header offset in the IPv6 packet is undefined – the entire header chain must be walked

Redpill Linpro

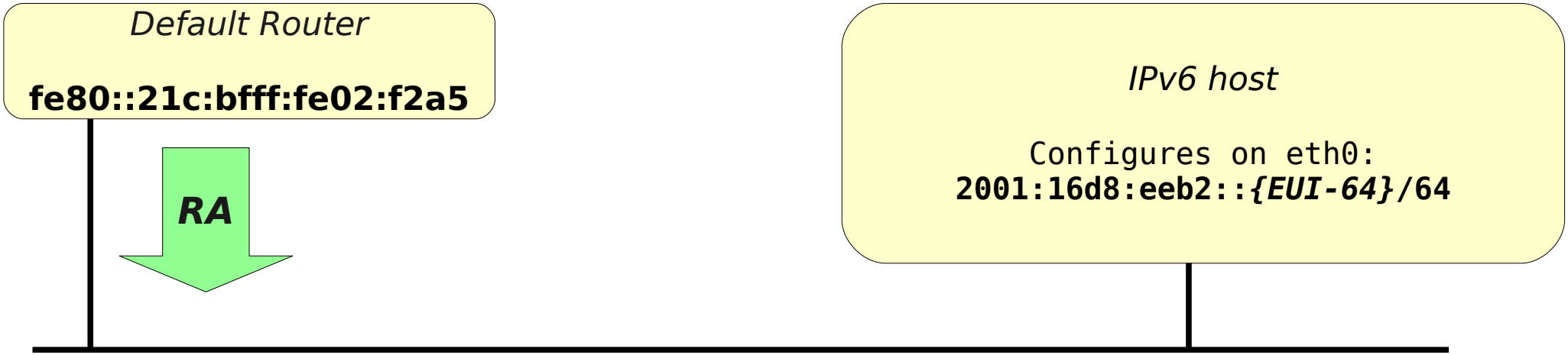
# Configuration of hosts

- Routers announce their presence by periodically transmitting ICMPv6 **Router Advertisements** to the *All Hosts* multicast group on a link
  - At init, hosts transmit an ICMPv6 **Router Solicitation** to the *All Routers* multicast group in order to receive an RA immediately
- An **RA** will usually tell a host how configure its IPv6 connectivity
  1. Contains the link-local address of the default router
  2. Contains information about addressing
     A) A prefix the host can freely choose an address from (**SLAAC**)
     B) An instruction to acquire a leased address from DHCPv6
  3. Contains information about DNS servers
     A) A list of recursive DNS servers to be used
     B) An instruction to ask DHCPv6 for DNS servers (and more)

# 1) Default router configuration

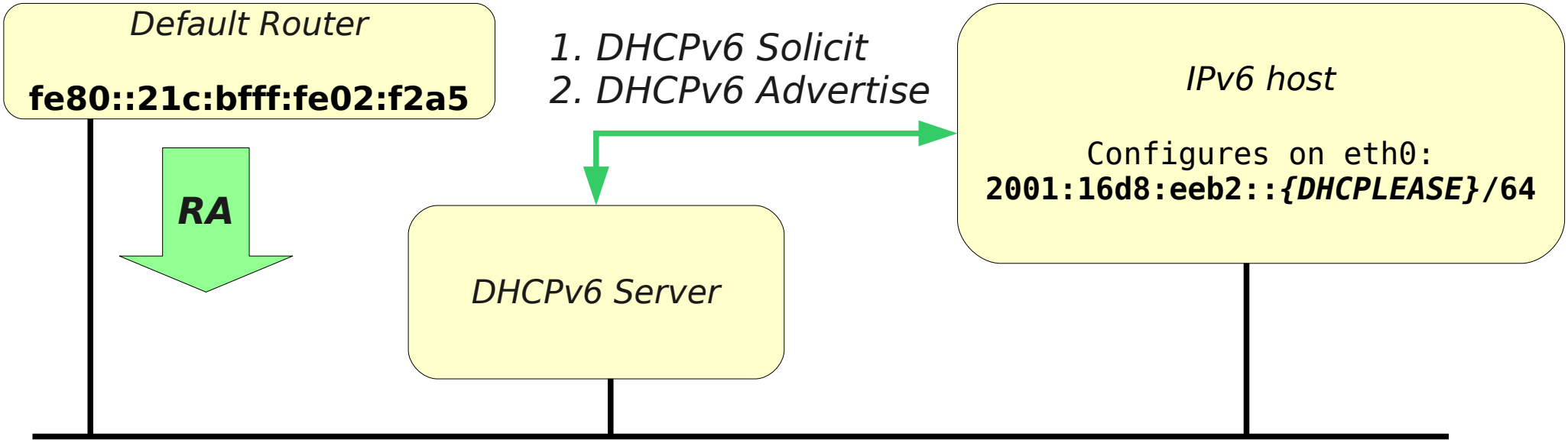

```
IP6 (hlim 255, next-header ICMPv6 (58) payload length: 80) fe80::21c:bfff:fe02:f2a5 >
ff02::1: [icmp6 sum ok] ICMP6, router advertisement, length 80

    hop limit 64, Flags [managed, other stateful], pref high, router lifetime 3600s,
reachable time 0ms, retrans time 0ms

      prefix info option (3), length 32 (4): 2001:16d8:eeb2::/64, Flags [onlink,
auto], valid time 86400s, pref. time 14400s

      rdnss option (25), length 24 (3):  lifetime 3600s, addr:
2001:16d8:eeb2:0:ffff::6106

      source link-address option (1), length 8 (1): 00:1c:bf:02:f2:a5
```

# 2A) Stateless Adress Autoconf.



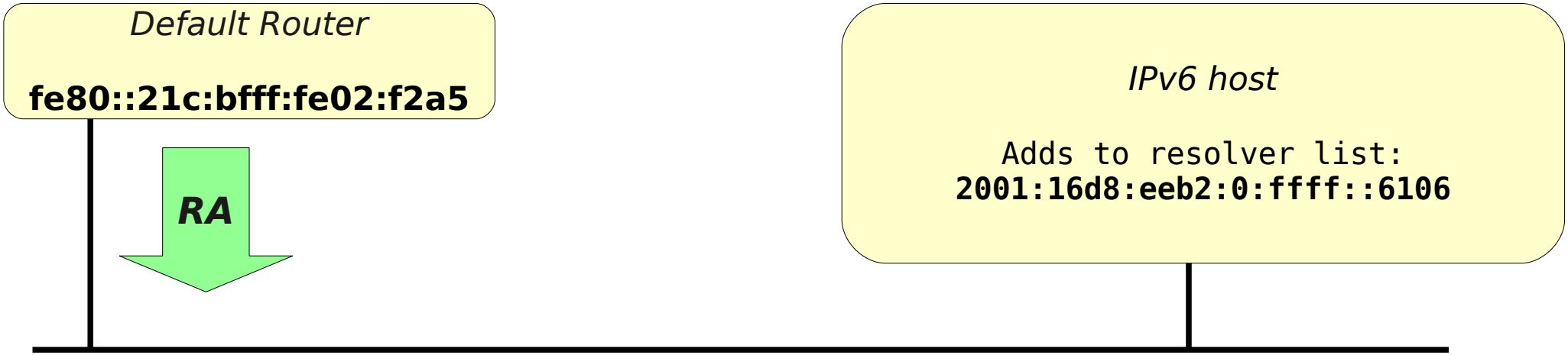

```
IP6 (hlim 255, next-header ICMPv6 (58) payload length: 80) fe80::21c:bfff:fe02:f2a5 >
ff02::1: [icmp6 sum ok] ICMP6, router advertisement, length 80

    hop limit 64, Flags [managed, other stateful], pref high, router lifetime 3600s,
reachable time 0ms, retrans time 0ms

      prefix info option (3), length 32 (4): 2001:16d8:eeb2::/64, Flags [onlink,
auto], valid time 86400s, pref. time 14400s

      rdnss option (25), length 24 (3):  lifetime 3600s, addr:
2001:16d8:eeb2:0:ffff::6106

      source link-address option (1), length 8 (1): 00:1c:bf:02:f2:a5
```

# 2B) Stateful DHCPv6 address conf.



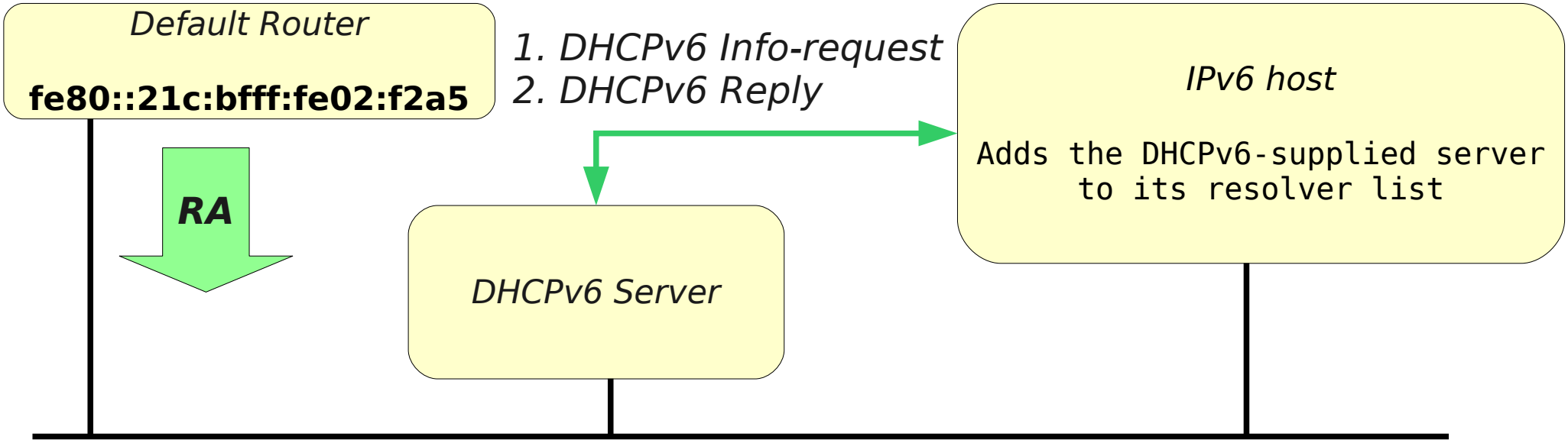

```
IP6 (hlim 255, next-header ICMPv6 (58) payload length: 80) fe80::21c:bfff:fe02:f2a5 >
ff02::1: [icmp6 sum ok] ICMP6, router advertisement, length 80

    hop limit 64, Flags [managed, other stateful], pref high, router lifetime 3600s,
reachable time 0ms, retrans time 0ms

      prefix info option (3), length 32 (4): 2001:16d8:eeb2::/64, Flags [onlink,
auto], valid time 86400s, pref. time 14400s

      rdnss option (25), length 24 (3):  lifetime 3600s, addr:
2001:16d8:eeb2:0:ffff::6106

      source link-address option (1), length 8 (1): 00:1c:bf:02:f2:a5
```

# 3A) Recursive DNS server in RA




```
IP6 (hlim 255, next-header ICMPv6 (58) payload length: 80) fe80::21c:bfff:fe02:f2a5 >
ff02::1: [icmp6 sum ok] ICMP6, router advertisement, length 80

    hop limit 64, Flags [managed, other stateful], pref high, router lifetime 3600s,
reachable time 0ms, retrans time 0ms

      prefix info option (3), length 32 (4): 2001:16d8:eeb2::/64, Flags [onlink,
auto], valid time 86400s, pref. time 14400s

      rdnss option (25), length 24 (3):  lifetime 3600s, addr:
2001:16d8:eeb2:0:ffff::6106

      source link-address option (1), length 8 (1): 00:1c:bf:02:f2:a5
```

# 3B) Information-Only DHCPv6




```
IP6 (hlim 255, next-header ICMPv6 (58) payload length: 80) fe80::21c:bfff:fe02:f2a5 >
ff02::1: [icmp6 sum ok] ICMP6, router advertisement, length 80

    hop limit 64, Flags [managed, other stateful], pref high, router lifetime 3600s,
reachable time 0ms, retrans time 0ms

      prefix info option (3), length 32 (4): 2001:16d8:eeb2::/64, Flags [onlink,
auto], valid time 86400s, pref. time 14400s

      rdnss option (25), length 24 (3):  lifetime 3600s, addr:
2001:16d8:eeb2:0:ffff::6106

      source link-address option (1), length 8 (1): 00:1c:bf:02:f2:a5
```

# OS support for IPv6-only networks

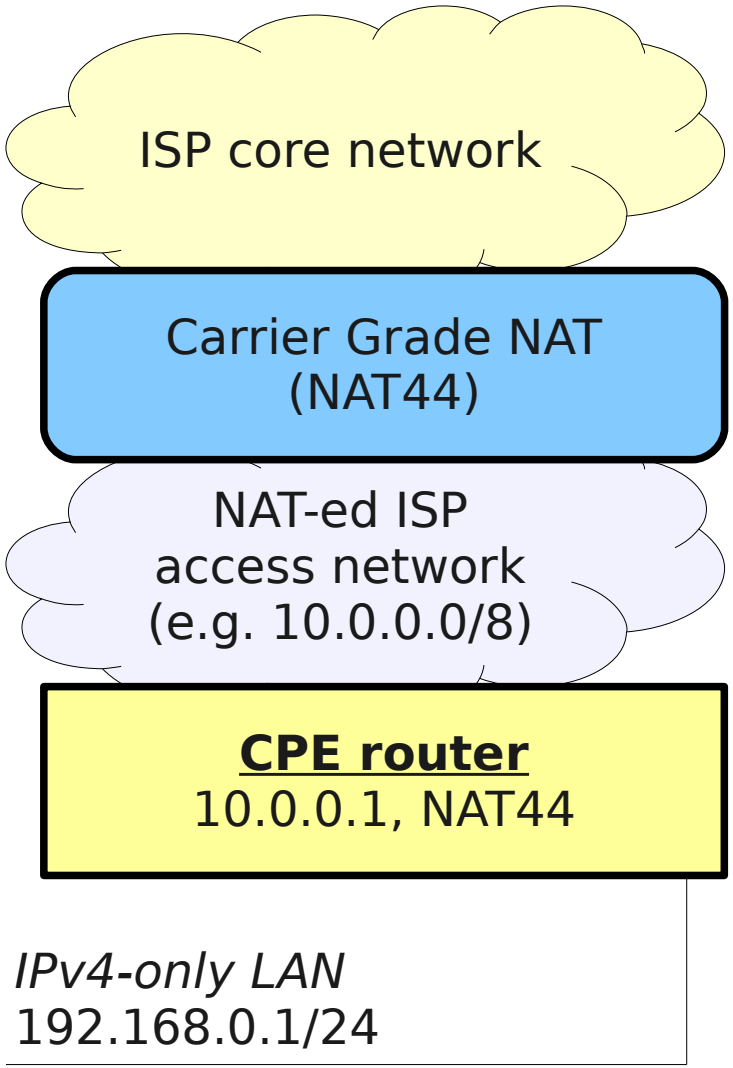
- Linux
  - By default kernel-level SLAAC only (no DNS, no DHCPv6)
  - GNOME NetworkManager can be configured to support everything
    - But fails to connect to IPv6-only networks by default
  - When NM is set up with IPv6 mode auto and IPv4 not required:
    - DHCPv6 responses dropped in default firewall on Fedora 14
    - Ubuntu 10.10 has no DHCPv6 client *(it stupidly tries DHCPv4 instead..)*
  - Android can't connect to IPv6-only networks **at all** – IPv4 required
- Apple
  - Mac OS X supports SLAAC only (no DNS, no DHCPv6)
  - Mac OS X is unable to learn any IPv6 DNS servers!
    - Third-party WIDE-DHCPv6 client can be used
  - iPhone and iPad support at least SLAAC and DNS in RA
- Microsoft Windows
  - Supports SLAAC and DHCPv6 (no DNS in RA)
  - Only major desktop OS with usable out-of-the-box IPv6-only support!

Redpill Linpro

# ISP deployment scenarios

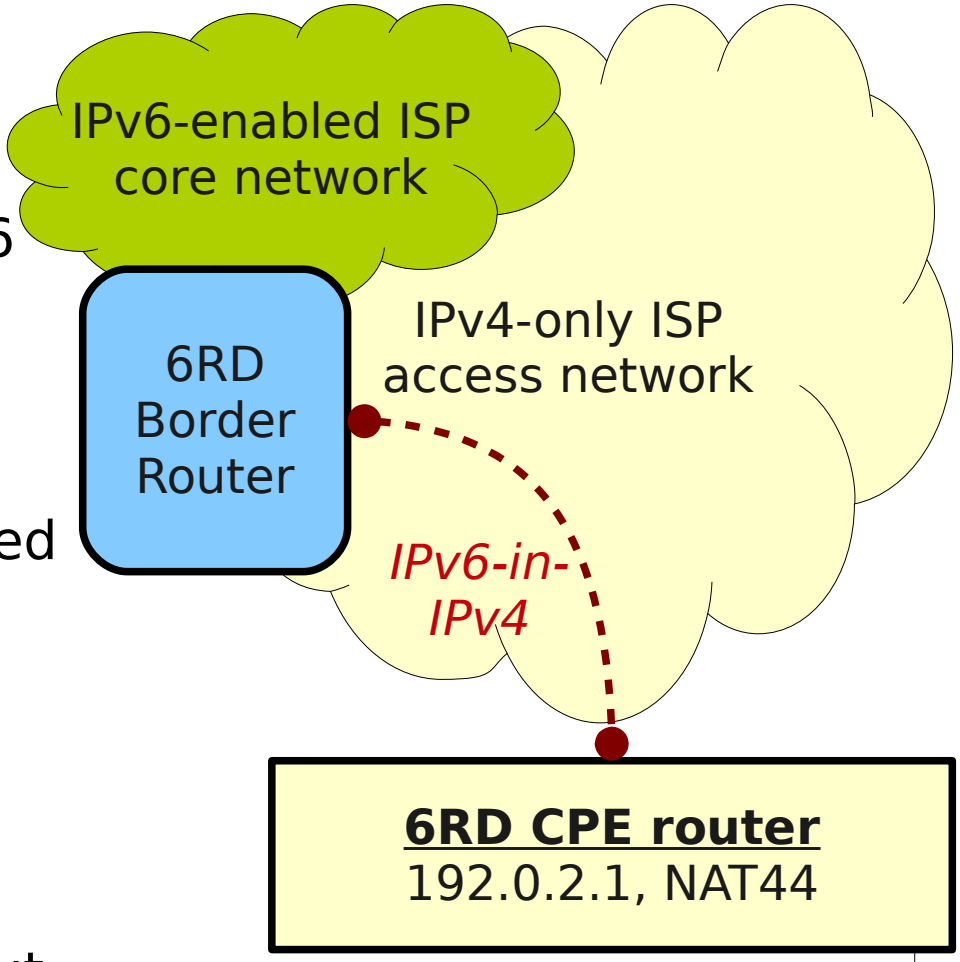
# NAT444 Carrier Grade NAT

- Two stacked layers of NAT44
- Advantages
  - Helps with IPv4 depletion
  - Easy to retro-fit to existing networks
- Disadvantages
  - **No IPv6**, no IPv4 exit strategy
  - The CGN is a performance bottleneck
  - Breaks applications that rely on NAT traversal (e.g. UPnP and NAT-PMP)
  - Breaks home servers / port forwards
  - Breaks geolocation
  - Breaks abuse handling



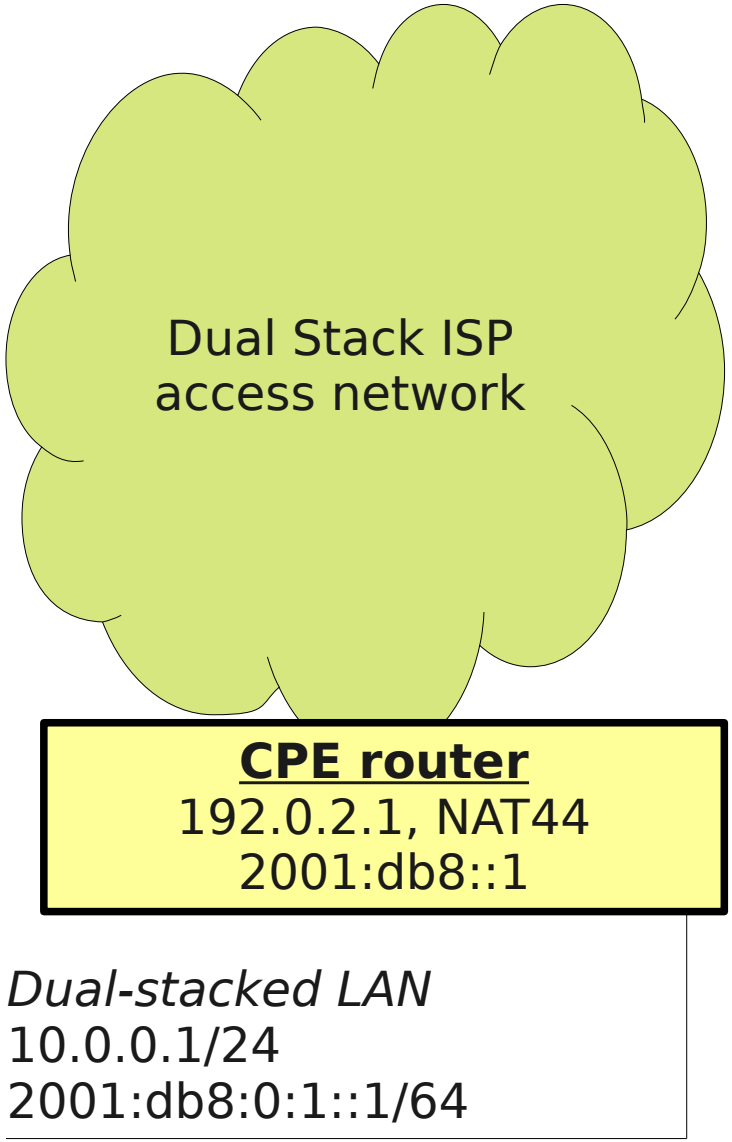
# IPv6 Rapid Deployment (6RD)

- IPv6 is tunneled across an IPv4-only access network to the end user
- IPv4 tunnel endpoint embedded in IPv6 prefix delegated to end user
- Advantages
  - No change to access network needed
- Disadvantages
  - CPE must support 6RD tunnelling
  - MTU and performance overhead
  - Doesn't help with IPv4 depletion (but may be combined with IPv4 CGN)
  - Provides no IPv4 exit strategy



# Dual Stack

- IPv6 and IPv4 are equal citizens
- The CPE learns its assigned IPv6 prefix using **DHCPv6 Prefix Delegation**
- Advantages
  - No compromise on performance for either protocol
- Disadvantages
  - Might require massive infrastructure upgrades in the ISP's access network
  - Doesn't help with IPv4 depletion (but may be combined with IPv4 NAT444)
  - Double work for many things needed (monitoring, troubleshooting, etc.)

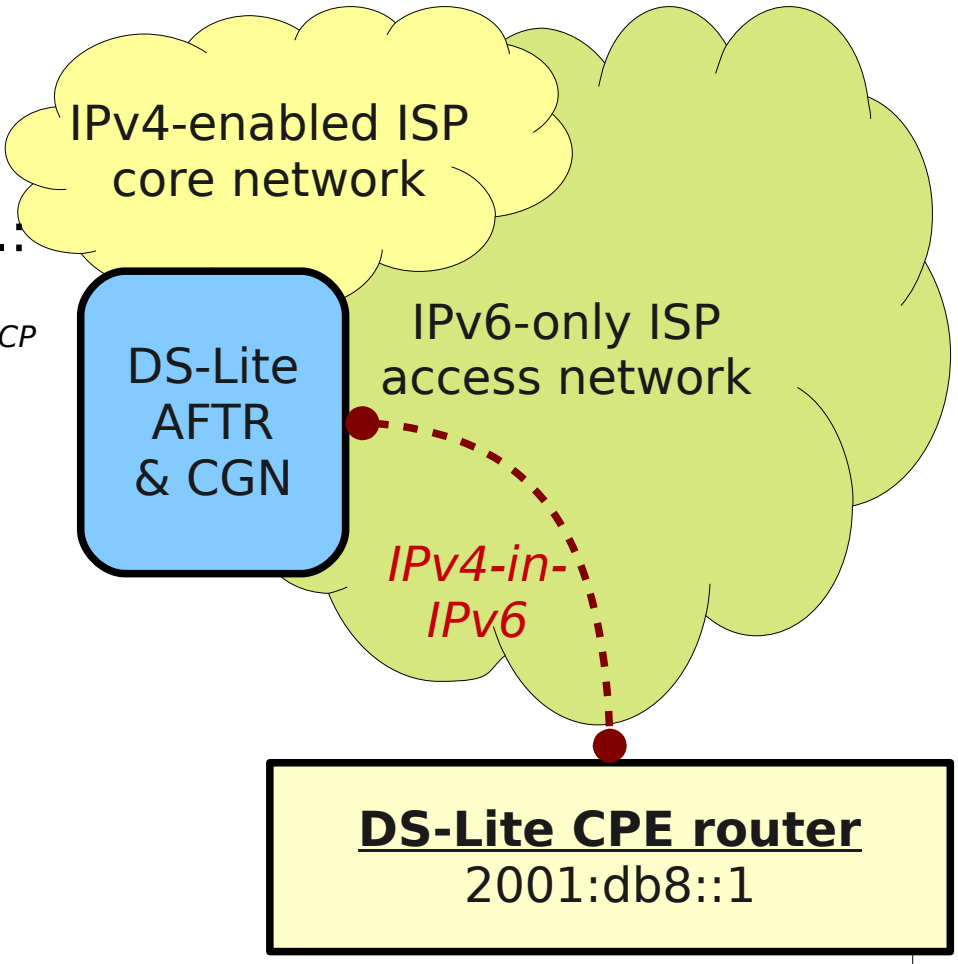


# Dual-Stack Lite (DS-Lite)

- IPv4 is tunneled across an IPv6-only access network to the end user
- Has a CGN that includes the IPv6 tunnel end point in the state table, e.g.:

*SRC={**2001:db8::1**,192.168.0.10:61234} DST=192.168.0.123:443 TCP*

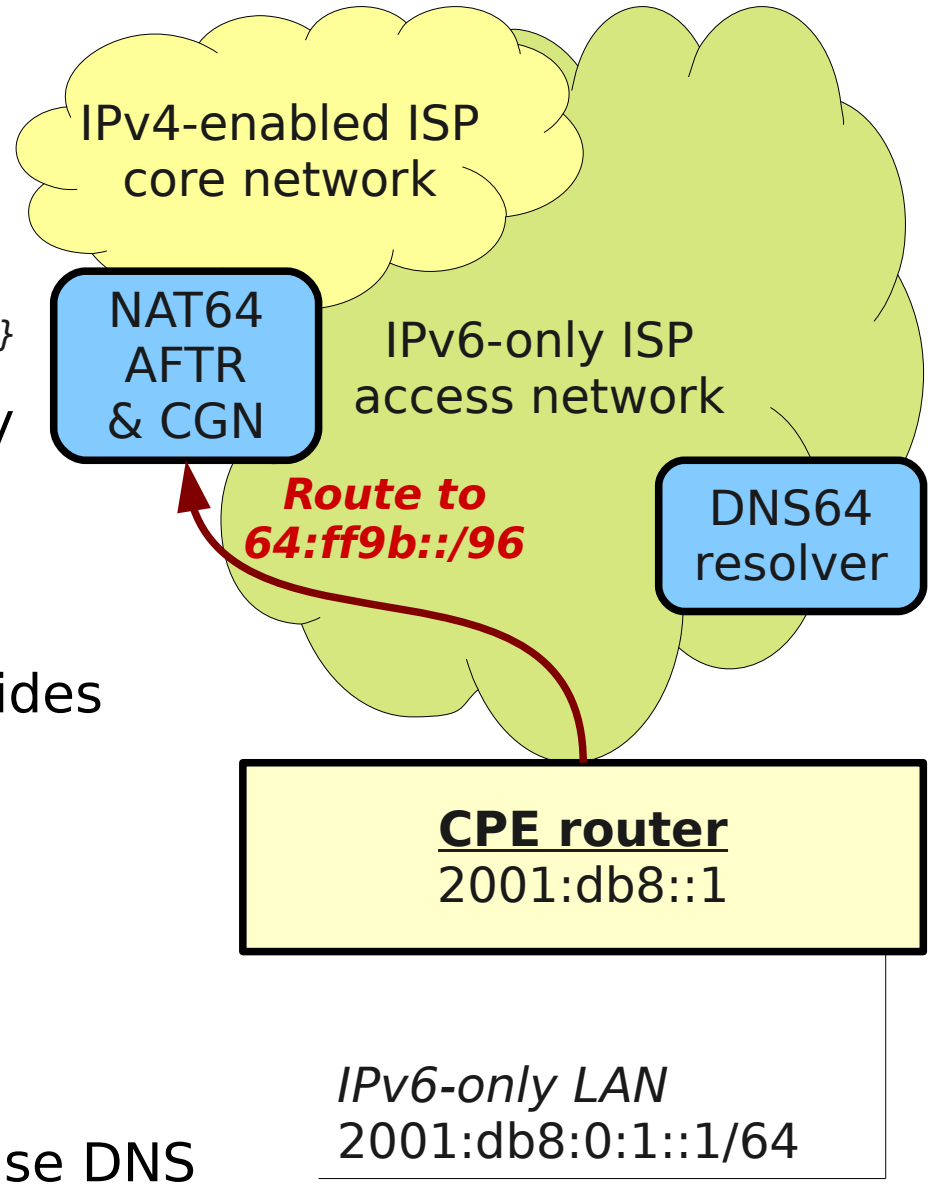
- Advantages
  - Helps with IPv4 depletion and provides clear exit strategy from it
- Disadvantages
  - CPE must support DS-Lite
  - CGN (but only single level, so NAT traversal may continue to work)
  - Access network upgrades necessary
  - MTU and performance overhead



# IPv6-only + NAT64/DNS64

- Native IPv6 connectivity only
- DNS64 resolver synthesizes AAAA records for IPv4-only host names by prepending an IPv6 /96 prefix, e.g.:

  *www.ipv4-only-domain.tld. IN AAAA* ***64:ff9b::****{192.0.2.123}*
- NAT64 device performs Address Family Translation and stateful NAT
- Advantages
  - Helps with IPv4 depletion and provides clear exit strategy from it
- Disadvantages
  - CGN, no NAT traversal
  - OS and apps must all support IPv6
  - Only works with applications that use DNS

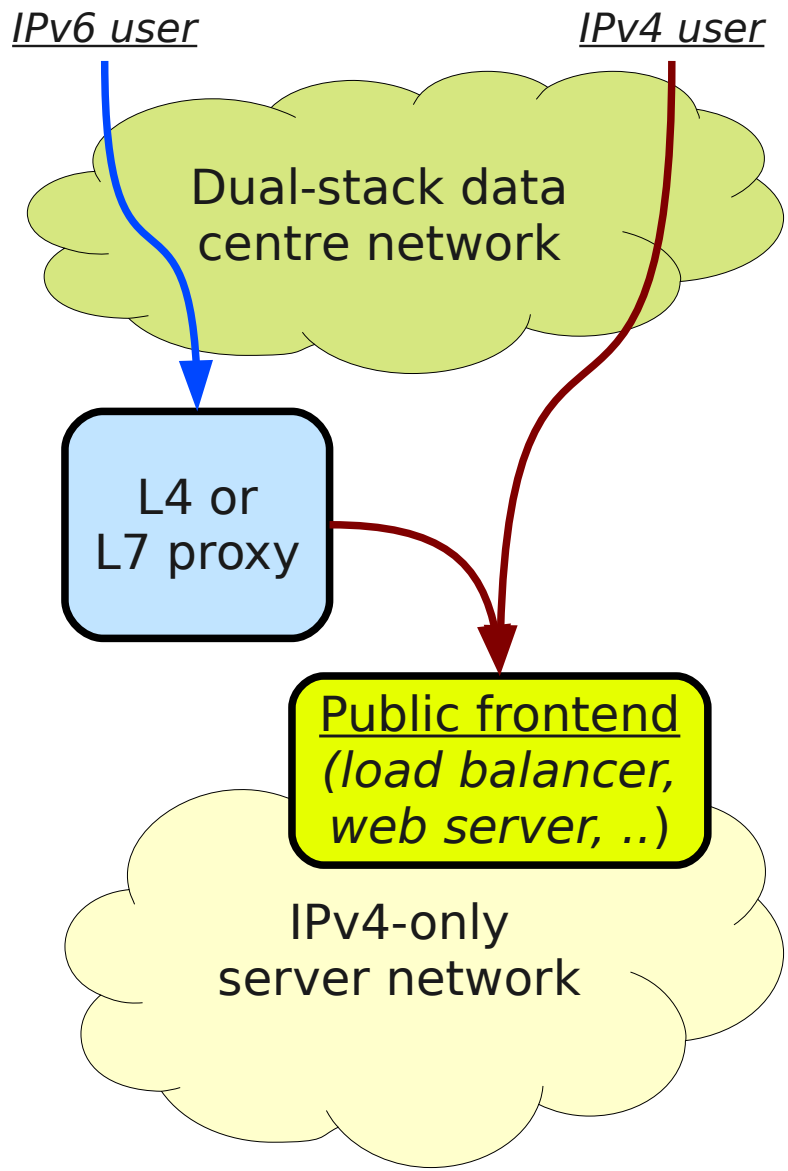


Redpill Linpro

# Content deployment scenarios

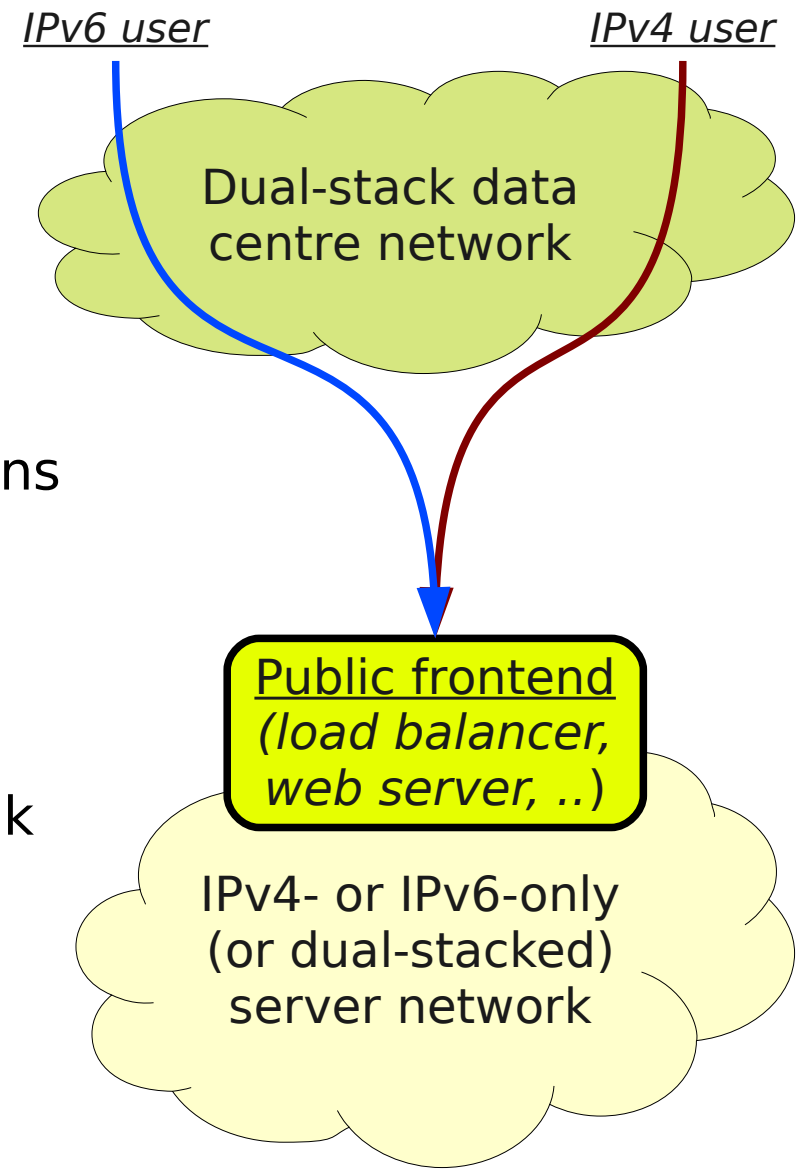
# IPv4 + application proxy

- IPv6 is routed to a system that translates incoming IPv6 traffic to IPv4
  - E.g. Varnish, HAProxy, Nginx, ...
- Advantages
  - Simple to retrofit to existing installations
  - Translator function may be outsourced
- Disadvantages
  - Possible performance bottleneck
  - Loss of insight into client source address
    - L7 headers (e.g. **X-Forwarded-For**) might take care of that
  - No exit path from IPv4
- www.hoyre.no does this (their real web servers are hosted by another provider)



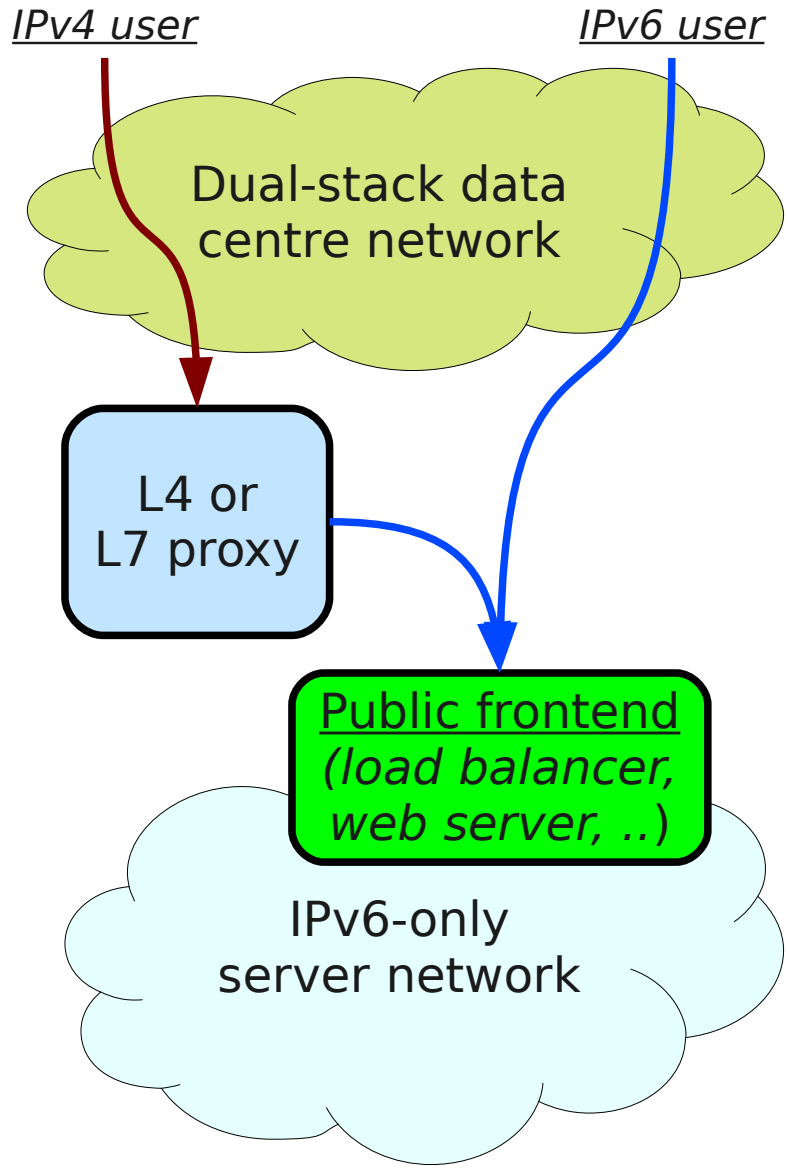
# Dual-stacked frontend

- Native dual-stack provisioned to the system's frontends
- Advantages
  - Simple to retrofit to existing installations
- Disadvantages
  - IPv6 support required on the public frontend devices/software
  - Operational overhead due to dual-stack (monitoring, troubleshooting, firewall filters, etc.)
- APDM, VG, and most other dual-stacked sites we host ourselves use this approach



# IPv6-only + application proxy

- IPv4 is routed to a system that translates incoming IPv4 traffic to IPv6
  - HAProxy, Varnish, nginx, ...
- Advantages
  - Clear exit path from IPv4
  - Frees up valuable IPv4 addresses
  - Translator function may be outsourced
- Disadvantages
  - Possible performance bottleneck
  - All apps and devices must support IPv6



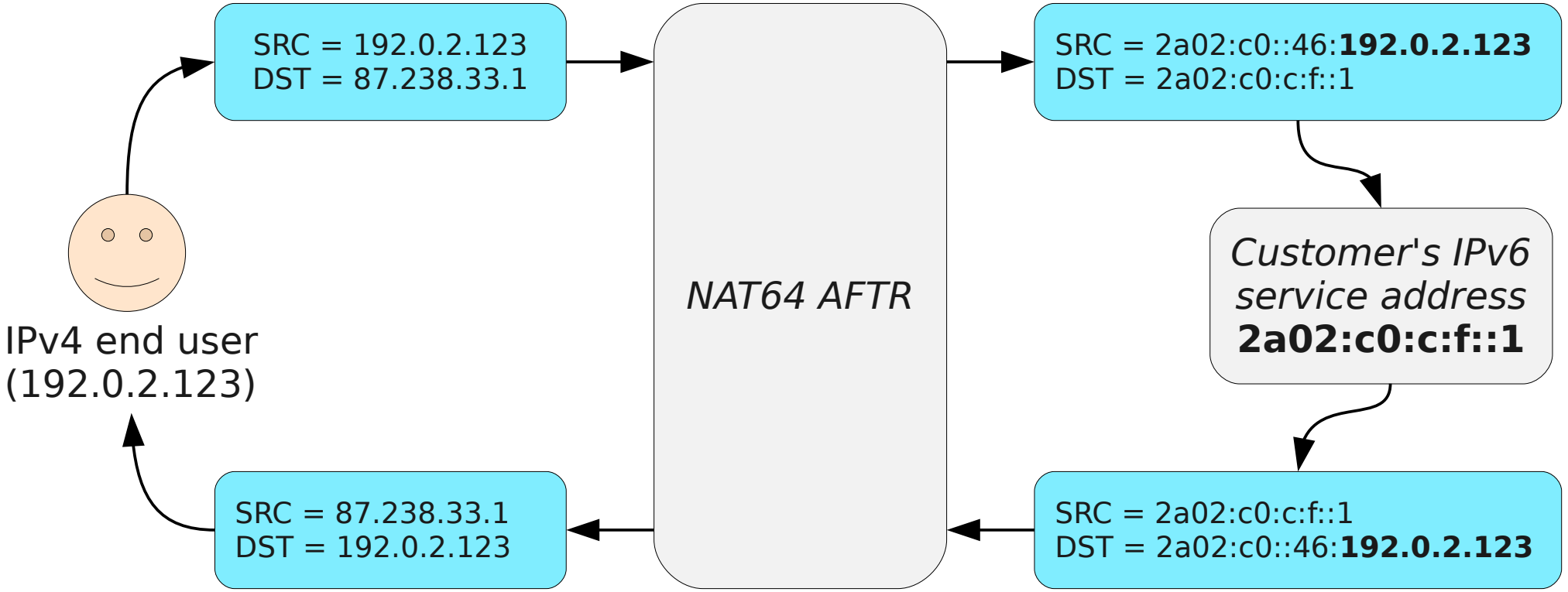
Redpill Linpro

# My favourite: Turning off IPv4

- Dual-stack equals operational overhead
  - Twice the amount of ACLs to configure
  - Twice the amount of services to monitor
  - Twice the amount of OSPF adjacencies to maintain
    - RFC 5838 will solve this eventually though
  - Twice the amount of routes to carry in your IGP
- More things that can go wrong and disrupt service
- And I simply don't believe the «servers must remain dual-stacked for the next 10 or 20 years» mantra

# Stateless NAT64

DNS setup:
www.cust.no. IN AAAA 2a02:c0:c:f::1
www.cust.no. IN A 87.238.33.1




- The NAT64 device(s) performs the following translations (+ in reverse):
  1) it rewrites the IP destination field by stripping the first 96 bits off the IPv6 address
  2) it rewrites the IP source field according to a static mapping
- The end user is not aware that the connection was translated
- The server can determine the IPv4 source address if it wants to

# Advantages of stateless NAT64

- Minimal operational overhead compared to dual-stack operation
  - Translator(s) can be centrally located in the core data centre network
- Stateless per-packet operation, no performance impact
  - Load balancing can be achieved with simple multipath routes
- The original IPv4 client address remains known to the application
  - Useful for Geo-location, ACLs, access logs, etc.
- Huge IPv4 address savings
  - One IPv4 address per **service** instead of one per server
  - Avoids unused addresses in a server LAN prefix – **100%** utilisation
- Forward-looking approach – why build services on a legacy foundation?
- I need a willing customer to partner with me in order to try...

Redpill Linpro

# Finishing words

Redpill Linpro

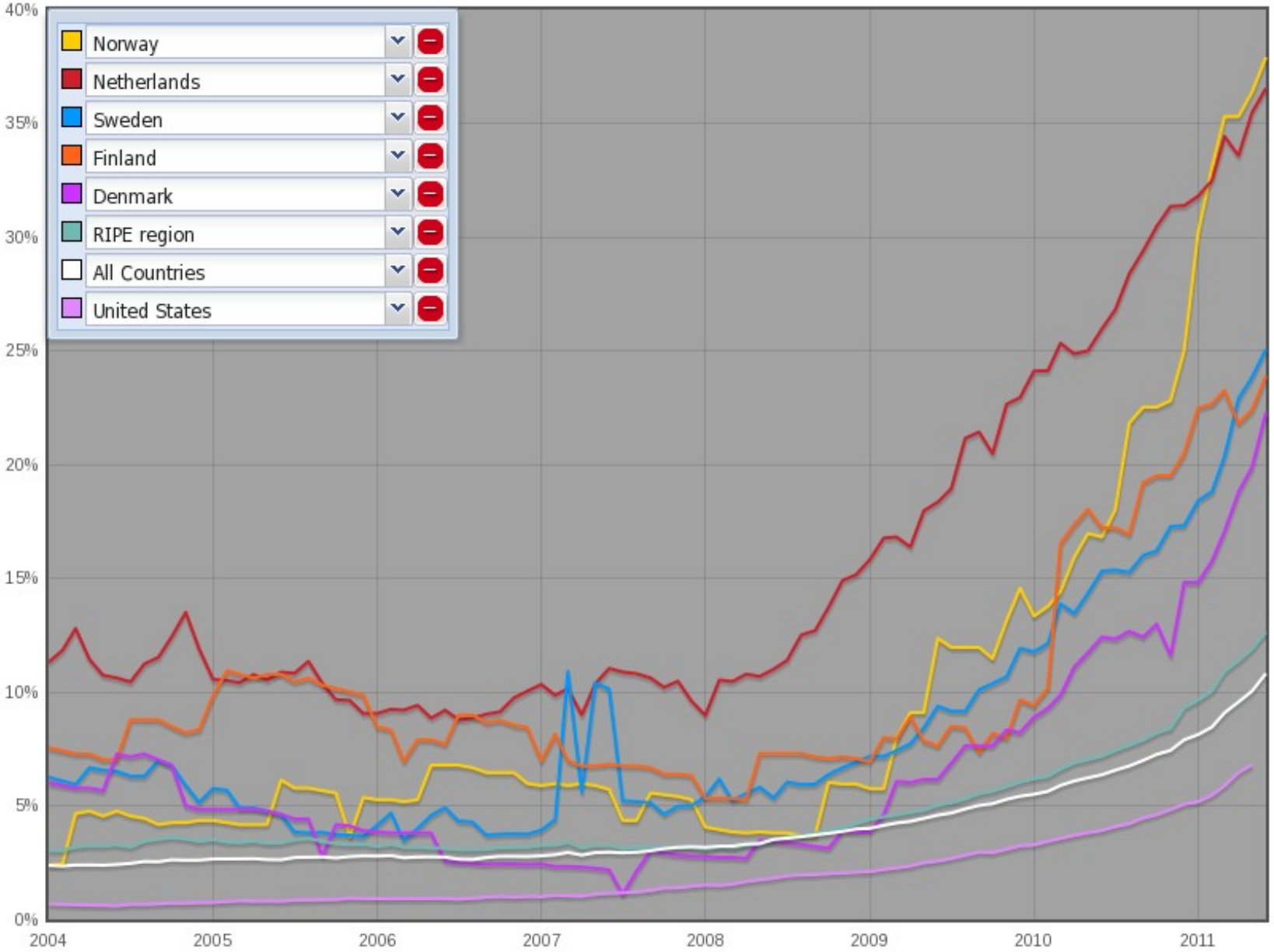
# Next steps for us

- Continue retro-fitting IPv6 capability to existing customers
- Hopefully do a pioneering IPv6-only deployment with a willing customer
  - And share the experiences with the community, of course!
- Identify remaining IPv4-only systems in our infrastructure and fix those
- Continue working to raise IPv6 awareness and accelerate deployment
- Wait for IPv6-enabled end users to appear...
  - There's still no direct reason for content folks to enable IPv6
  - I keep hearing Altibox engineers saying «in 2011» - I really hope that works out well for you; tvi, tvi!

Redpill Linpro

http://v6asns.ripe.net/v/6?s=NO;s=NL;s=SE;s=FI;s=DK;s=_RIR_RIPE_NCC;s=_ALL;s=US

This graph shows the percentage of networks (ASes) that announce an IPv6 prefix for a specified list of countries or groups of countries




- Many Norwegian ISPs have already perfomed the initial deployment steps!

*With thanks to Emile Aben/RIPE NCC*

Redpill Linpro

# Questions?