Redpill Linpro

# Inciting Norwegian IPv6 deployment

Tore Anderson
CG Security and Networking
Redpill Linpro
IPv6-Kongress, Frankfurt am Main, May 2011

Redpill Linpro

# Agenda

1) Introduction
2) Deployment experiences and dual-stack brokenness
3) Status of IPv6 deployment in Norway
4) Our IPv6-only service deployment plans
5) Questions/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
- **http://www.redpill-linpro.com**

Redpill Linpro

# First steps

- Engineer-driven project, no mandate from management
- 1H 2008: Acquire IPv6 PA prefix from RIPE NCC and route it
- 2008-1H 2009: Dual-stack our backbone and data centre cores
- Key challenges:
  - Get quality native IPv6 transit that's properly supported
  - Lack of Netflow v9 support on Juniper MX-DPC line cards
  - Expensive licences for OSPFv3 on Juniper EX switches
  - Home-made IPv4-only Linux/iptables firewalls
- Demand IPv6 support in all new aquisitions
- Started at the outer border, worked our way inwards

Redpill Linpro

# Getting customers aboard

Redpill Linpro

# NAT444 – The Number Of The Beast

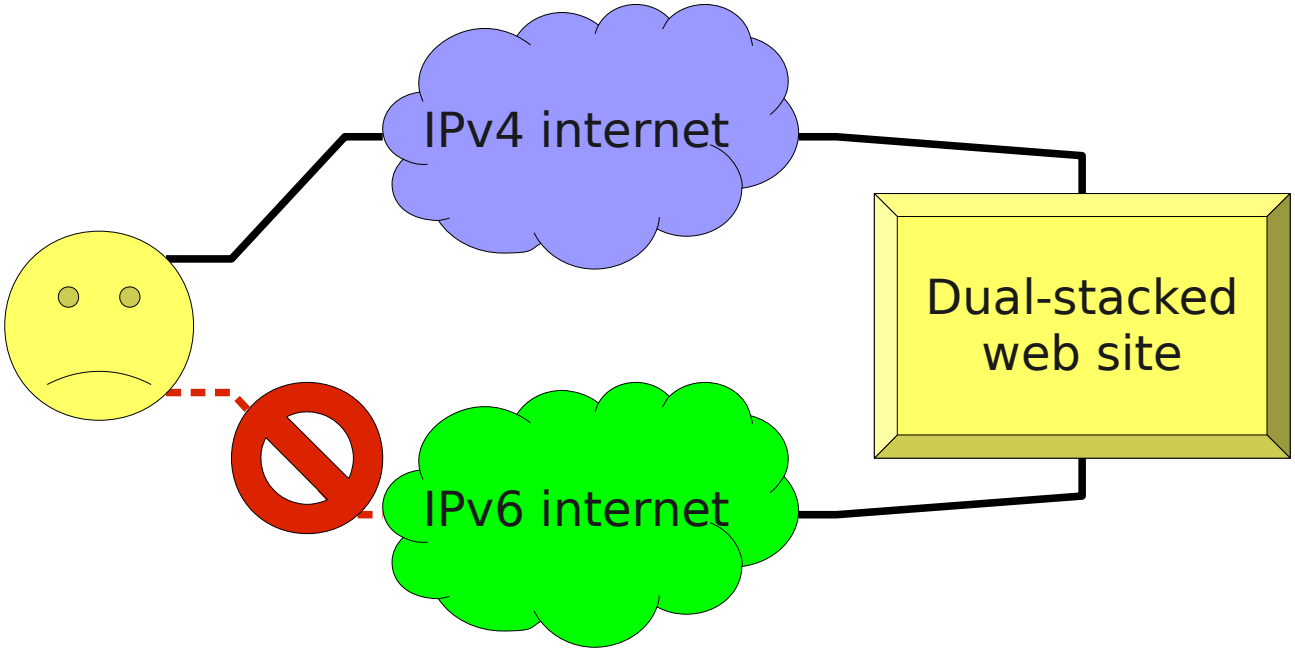
Redpill Linpro

# IPv4 will deteriorate

- We know for a fact that Norwegian ISPs are preparing for **NAT444** or similar **Carrer Grade NAT** systems out of necessity
- The performance and functionality of our customers' services will be adversely affected when these are deployed
- We need IPv6 as an alternative and faster path between our content and the end users
- Plan: Deploy early and hope the ISPs will follow suit

Redpill Linpro

# IPv6 isn't perfect, however




- **Dual-stack brokenness** causes a bad user experience
- End-user's OS/web browser incorrectly thinks there's IPv6 connectivity
- Long timeouts before fail-over to working IPv4 happen
  - Live demo: **http://broken.redpill-linpro.com**
- Prevents dual-stack deployment – it's **less reliable** than IPv4-only

Redpill Linpro

# Researching dual-stack brokenness

- We enrolled two of our customers in an experiment to identify common causes of brokenness and quantify the number of affected users
- **VG** – http://www.vg.no
  - Tabloid newspaper; Norway's largest web site
- **A-pressen Digitale Medier** – http://www.apdm.no
  - ~70 regional newspapers; Norway's 4th largest site
- Both are hosted in one of our Oslo data centres

# Measurement setup


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

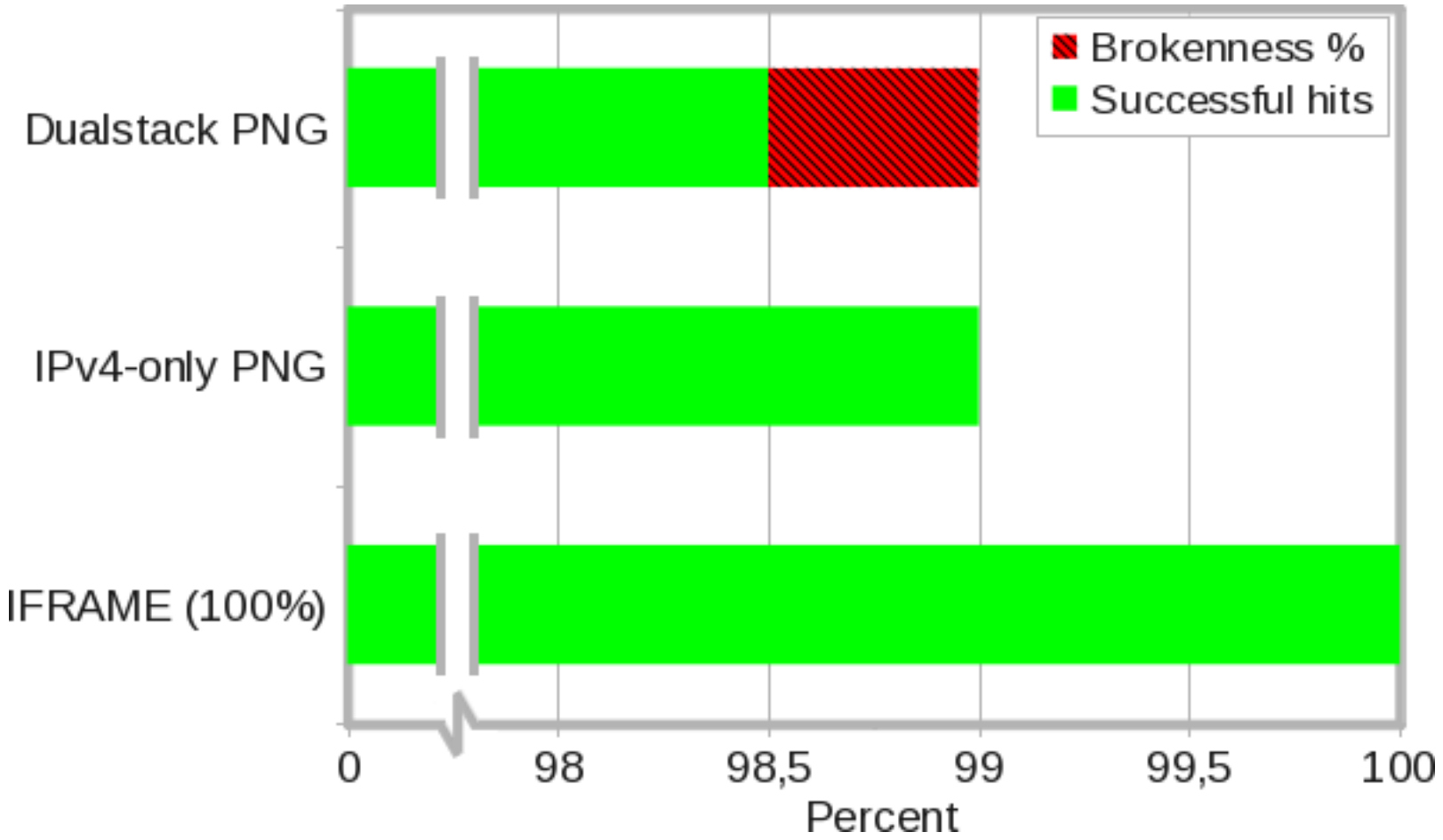
- 1x1.png
- IPv4-only

- 1x1.png
- Dual-stack

*Tip: Want to try this on your own site? Check out Éric Vyncke's http://www.vyncke.org/testv6/ !*

**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%**.

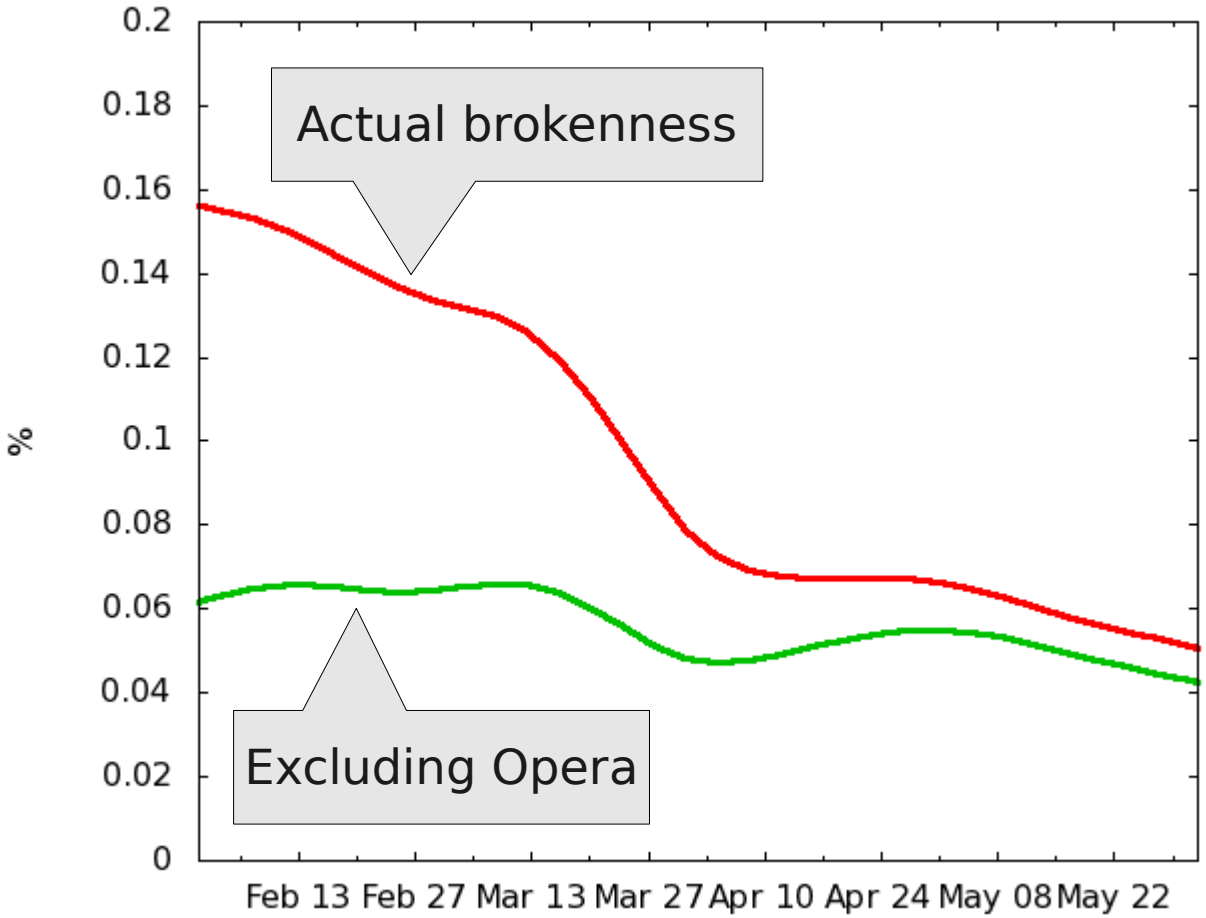
Redpill Linpro

# Initial findings – Q4 2009

- **0.2-0.3%** brokenness
  - ...a complete non-starter
- Certain sources of brokenness were standing out
  - Opera web browser on Windows
  - Mac OS X
  - Certain networks (enterprises, universities), ISPs
- 70-80% of IPv6 traffic was 6to4 and Teredo
  - ..which runs on top on IPv4, so can't possibly be more reliable
  - There's no real reason to use either in preference to IPv4

# Opera web browser on Windows

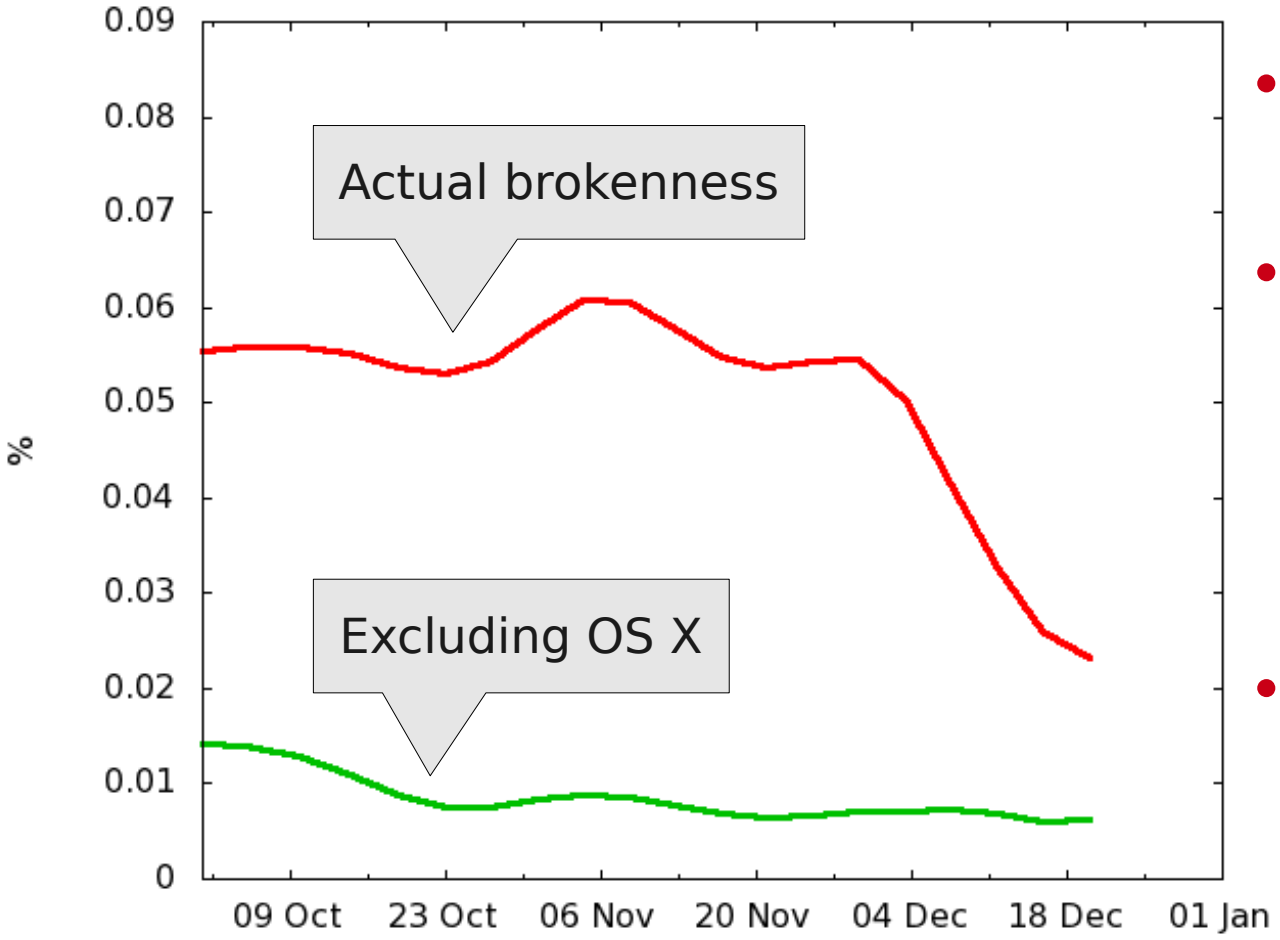
- Recent Windows will **automatically** enable 6to4 and/or Teredo
  - ..but de-prefers their use in the system resolver (RFC 3484)
- Opera, however, used its own built-in resolver




- 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



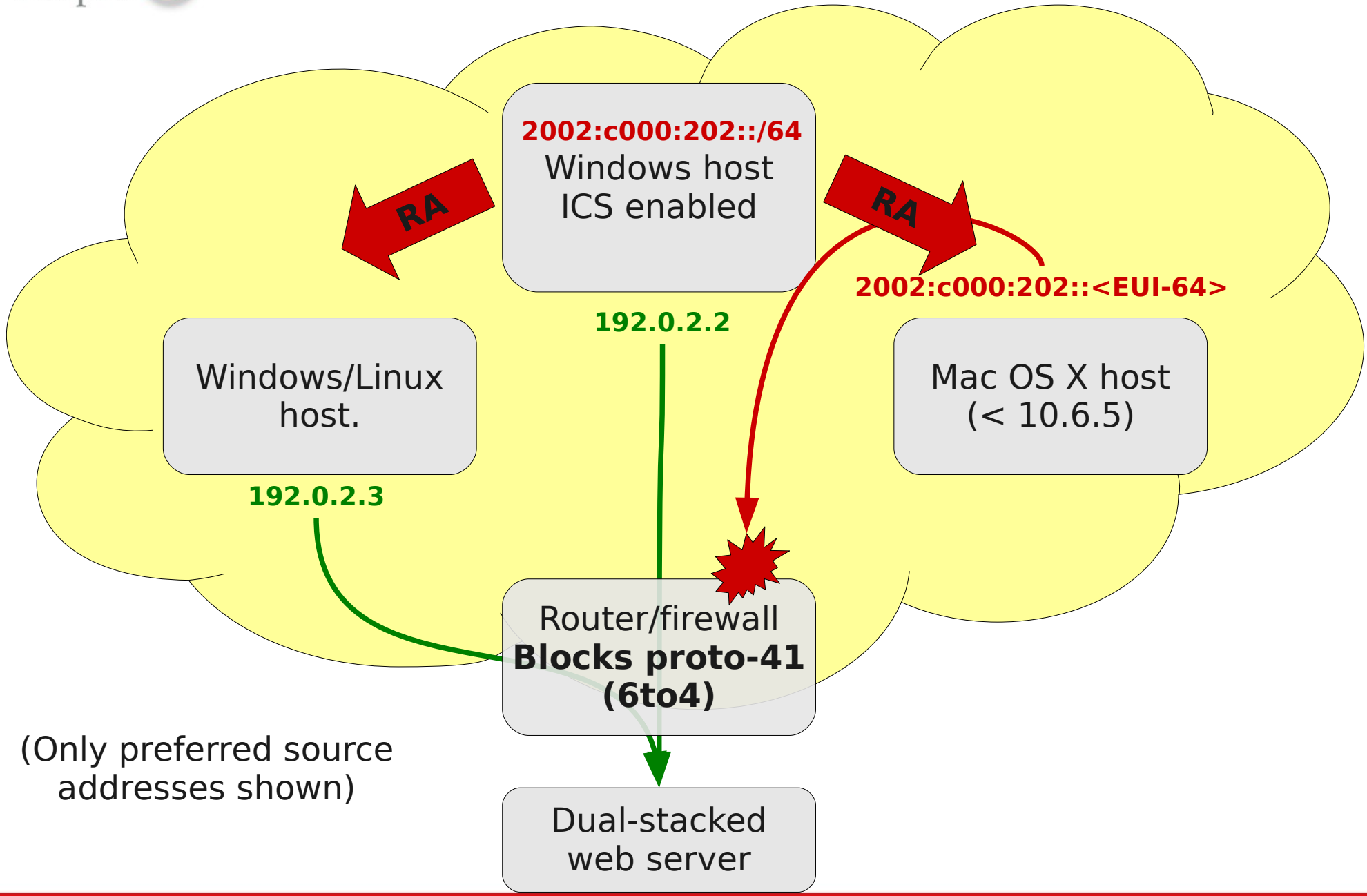

- Started nagging them about it
- Version 10.6.5, released 10th of November, de-prefers IPv6 completely if local 6to4 addresses are present
- No upgrade path for one-fourth of their users (running 10.4 and 10.5)

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

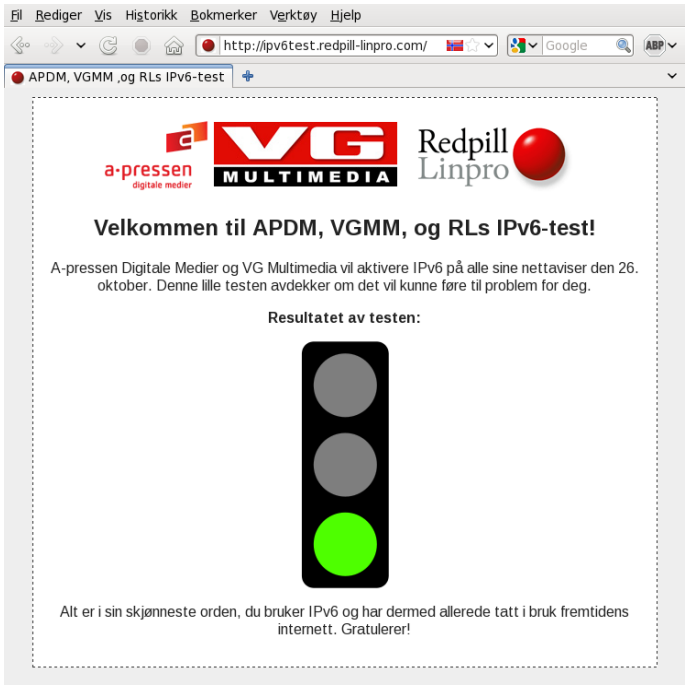
# Rogue RA-infested network



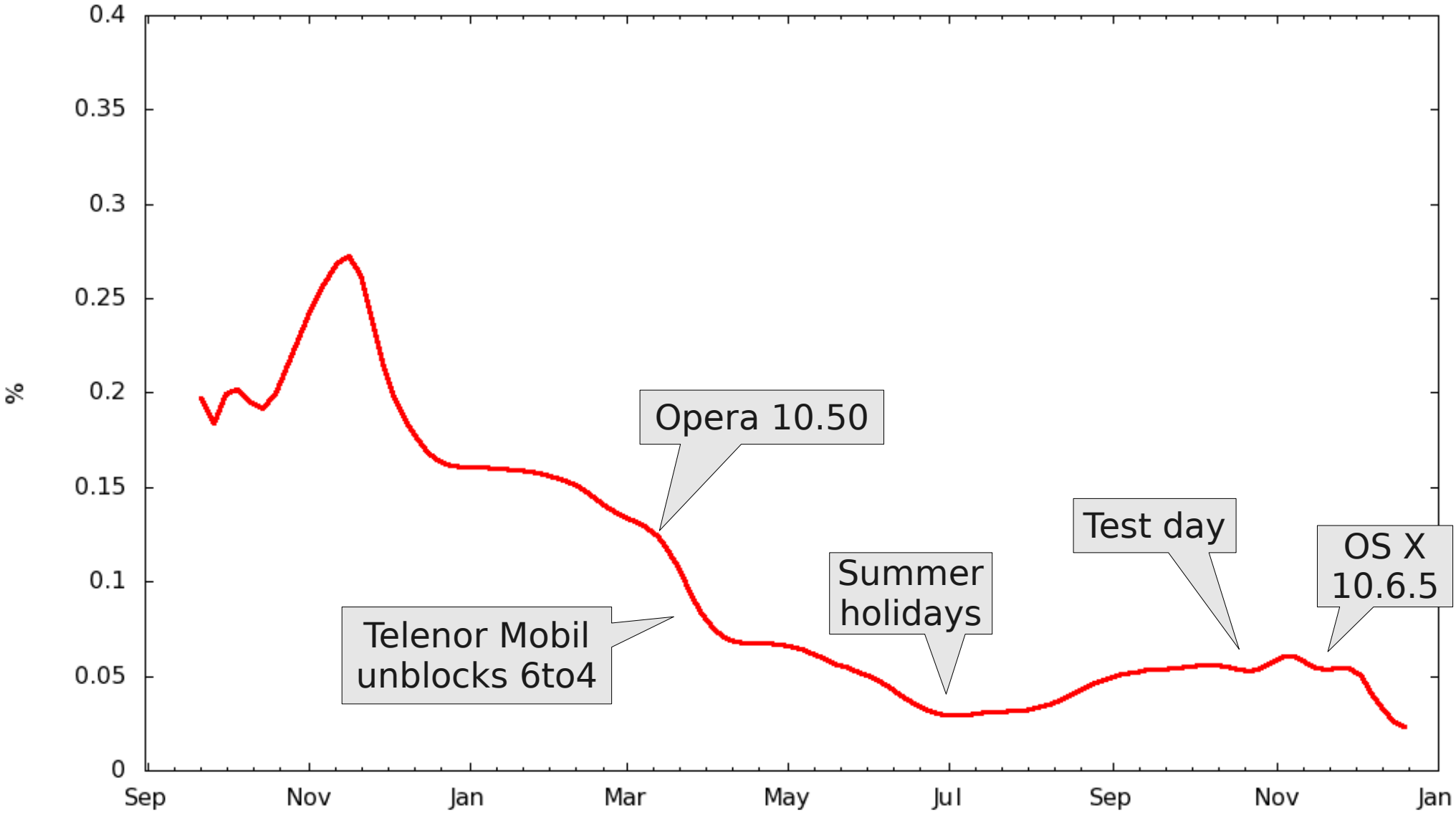

(Only preferred source addresses shown)

# Production for VG and APDM

- In October we did a 24 hour production test, inspired by Heise.de
- Broken users are warned and redirected 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**

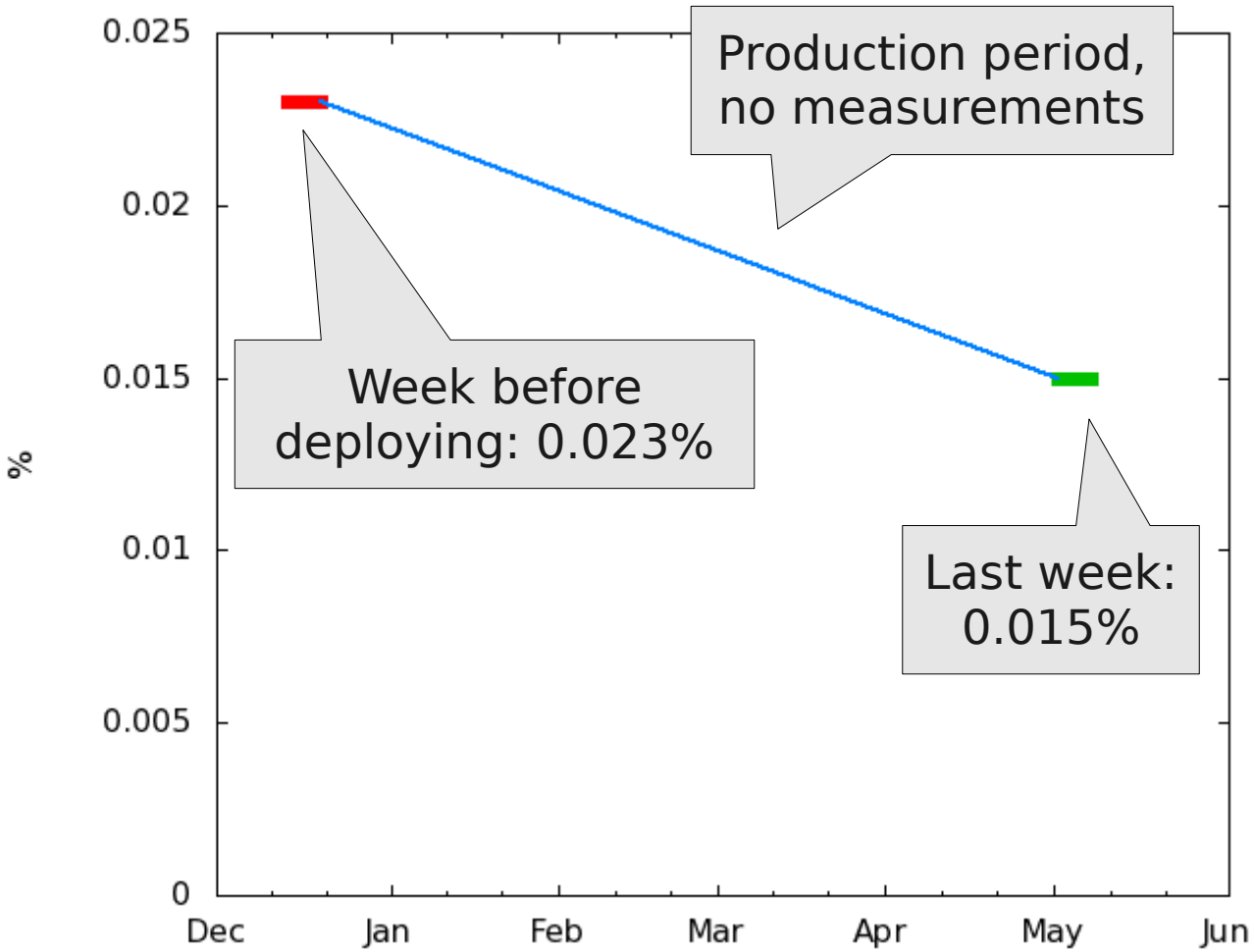



# From the start until production




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 last week in order to perform this new measurement)*

Redpill Linpro

# Fear of brokenness is a thing of the past)

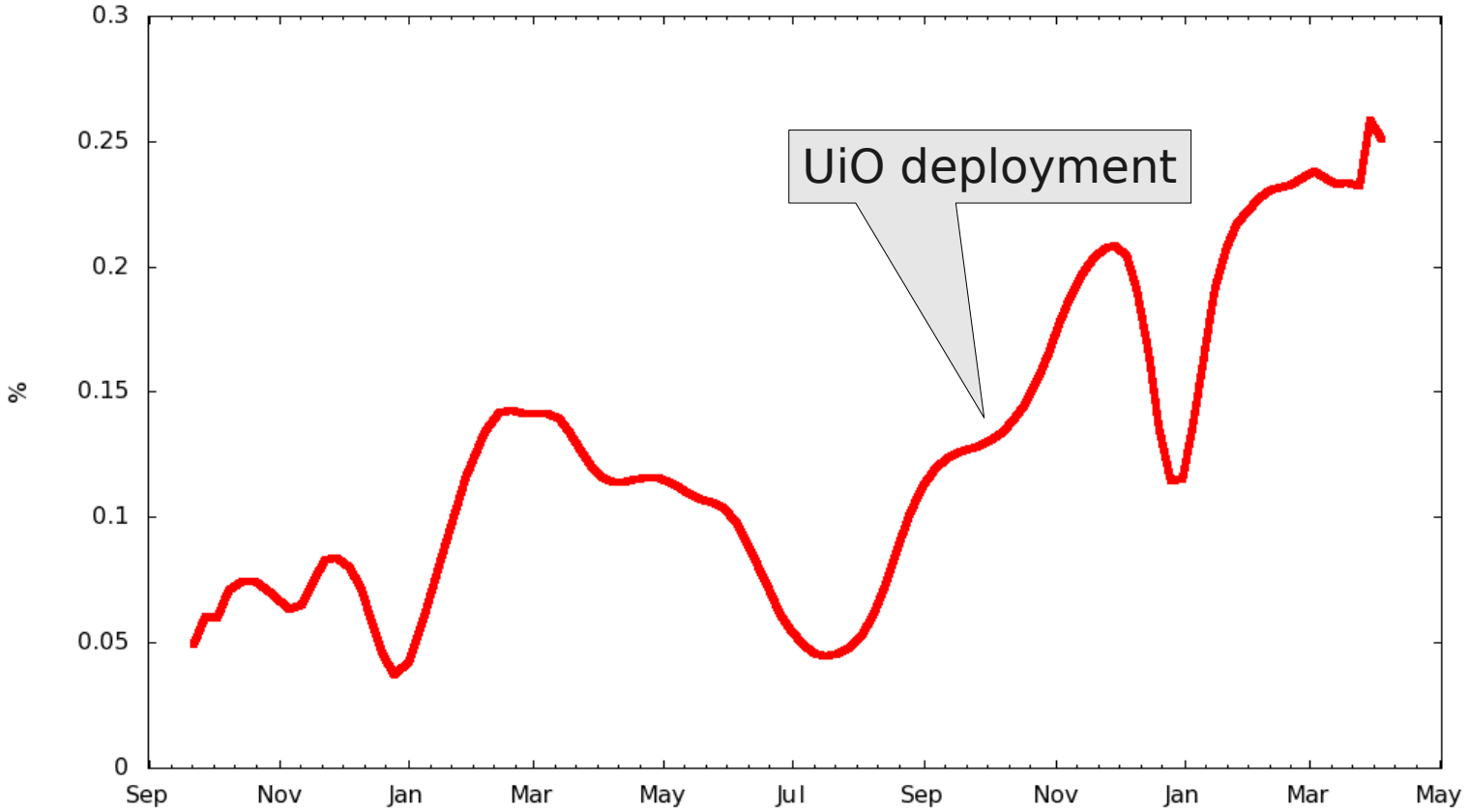
- Approximate visitor numbers for www.vg.no, per month:
  - ~12 million unique web browsers/devices
  - ~3 million unique individuals
- Population of Norway:
  - 4,9 million
- IPv6-related complaints since deploying:
  - Less than 50
- Broken users that could not be easily helped:
  - **0**
- Every day is «*Norwegian IPv6 Day*»!

Redpill Linpro

«Inciting»..?

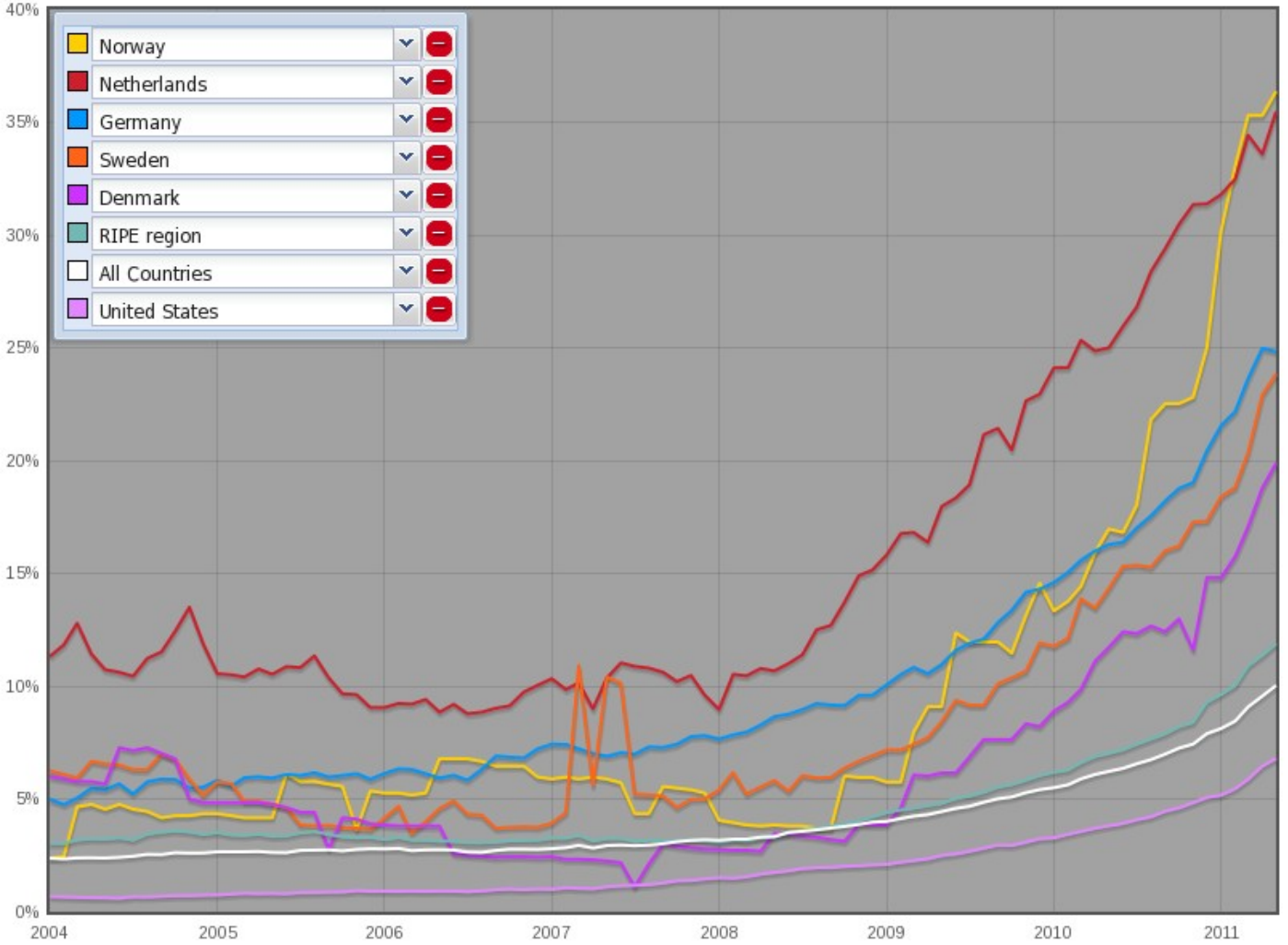
# Not as much as we'd hoped

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




- University of Oslo deployed IPv6 on their student dorm networks as a direct response to our AAAA test day
- Drops during Christmas/Easter/Summer holidays

# But things are happening




- Over the last 18 months Norway has rapidly become the world leader in IPv6-enabled autonomous systems
- Hopefully that means IPv6-enabled end users will follow soon

Redpill Linpro

# Other initiatives are popping up too

- The Norwegian Post and Telecoms Authority is taking an active role in encouraging IPv6 adoption
- The trade organisation for the ICT industry as well
  - Arranges bi-annual IPv6 conferences
- At least one major broadband ISP (Altibox) has publically announced plans to provide IPv6 as a standard service by the end of the year.
- Other content providers are following our lead
- And I'll be dualstacking our remaining customer base as fast as my schedule will permit me to :-)

Redpill Linpro

# What's next?

Redpill Linpro

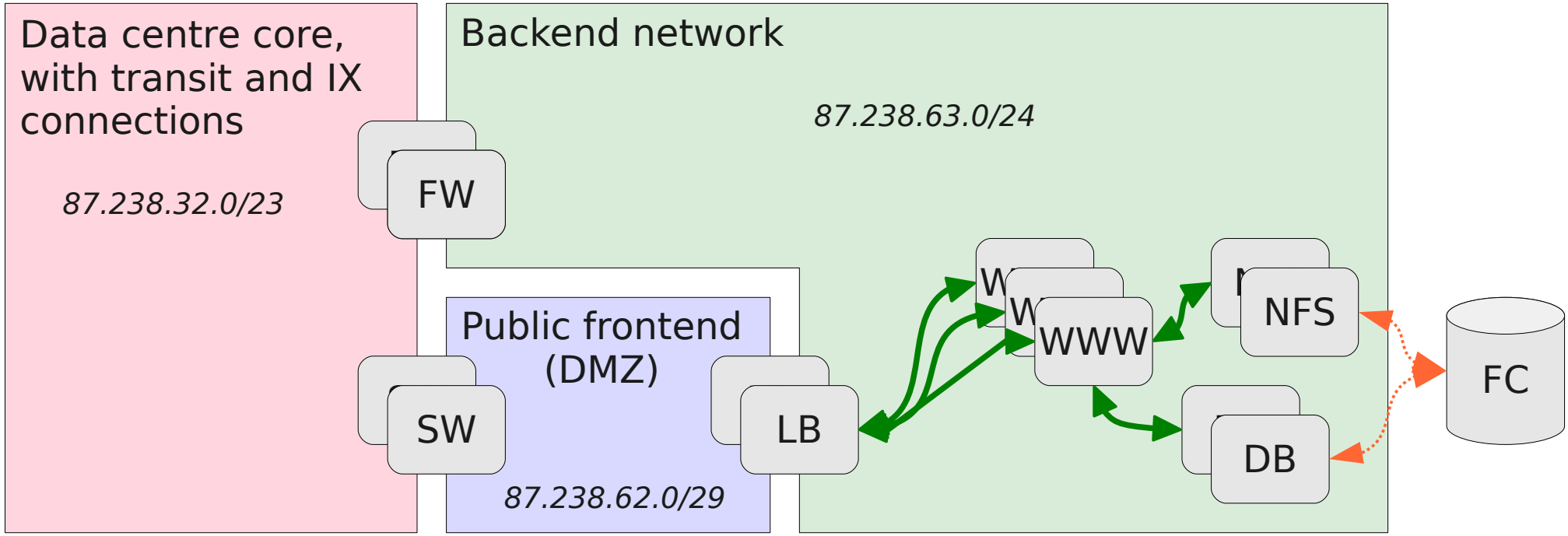
# Turning off IPv4

Redpill Linpro

# No, I'm not insane

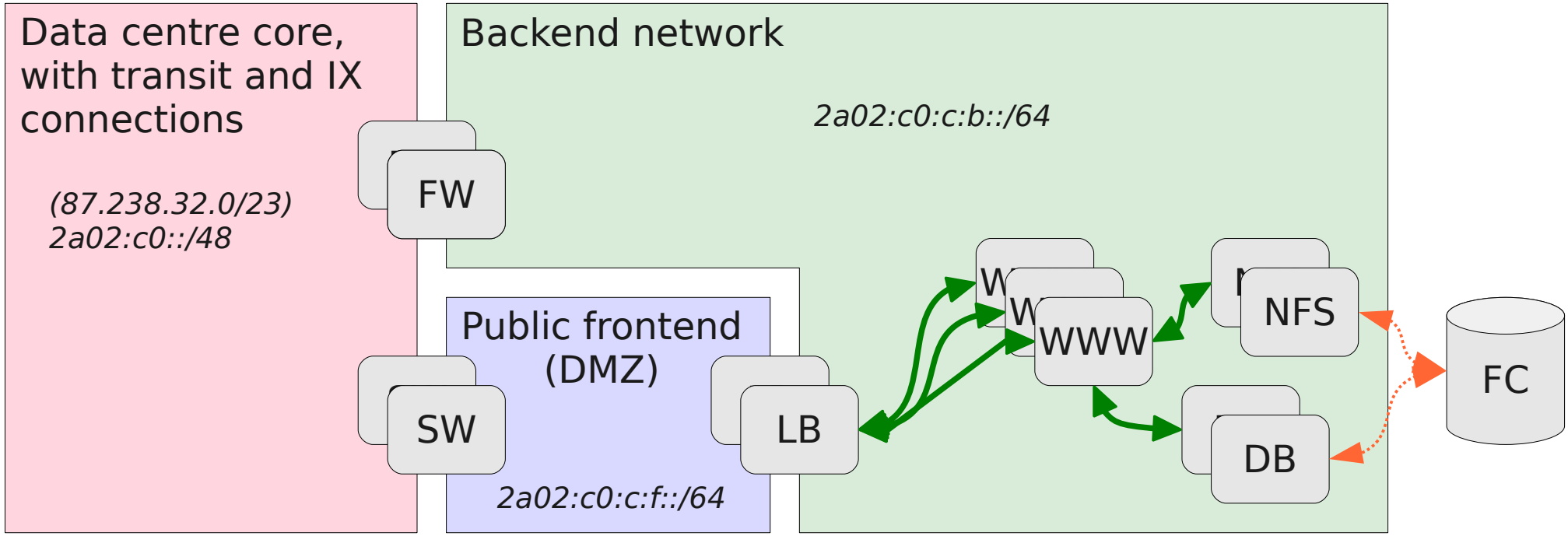
- Dual-stack equals operational overhead
  - Twice the amount of ACLs to configure
  - Twice the amount of services to monitor
  - Twice the amount of OSPF adjacencies to maintain
    - RFC 5838 will solve this eventually though
  - Twice the amount of routes to carry in your IGP
- More things that can 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

# A typical customer of ours



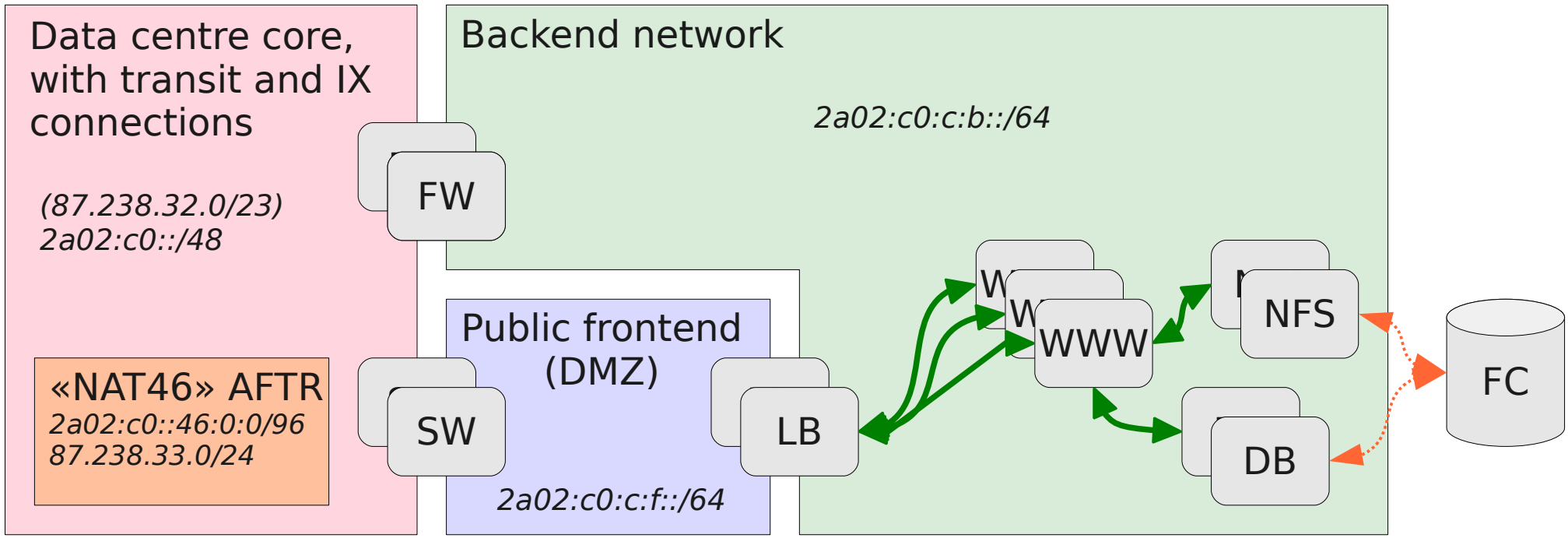

- Everything is IPv4
- Non-production traffic pass through a firewall (e.g. backup, ssh, ..)
- Production traffic (to load balancers, web caches, or similar) via a separate frontend network with no stateful devices in the path

# A green field IPv6-only customer




- No change in topology, except for the IP protocol version upgrade
- Can save us 100s of IPv4 addresses for a single large customer
- The core network needs to support IPv4 until all customers are IPv6-only – but IPv4 can be gradually turned off there too

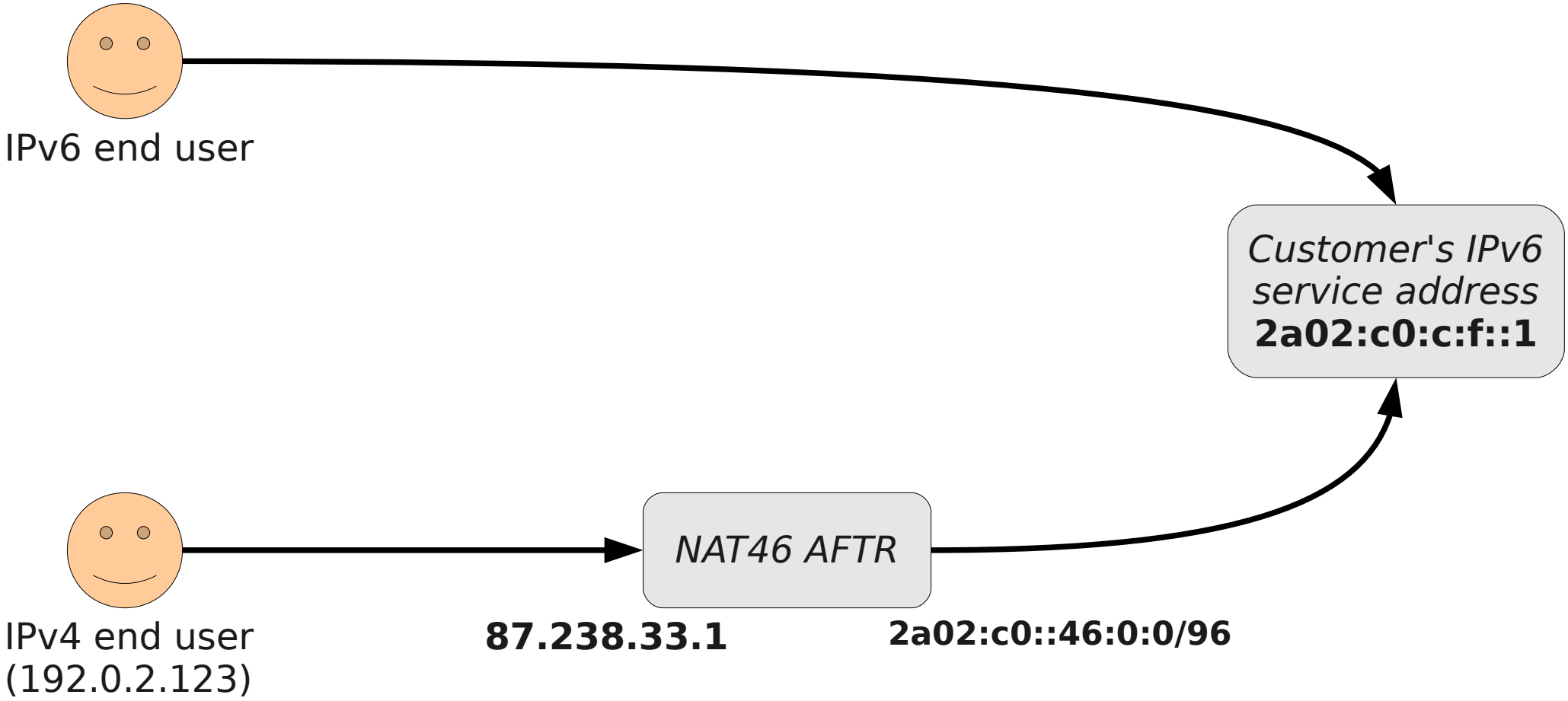
# Handling IPv4-only end users




- An **Address Family Translation Router** in the core network translates inbound connections from IPv4 clients to the customer's IPv6 service addresses
- Stateless, per-packet operation (according to RFC 6145 **SIIT**)
  - No performance impact for IPv4 clients!

# A closer look

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

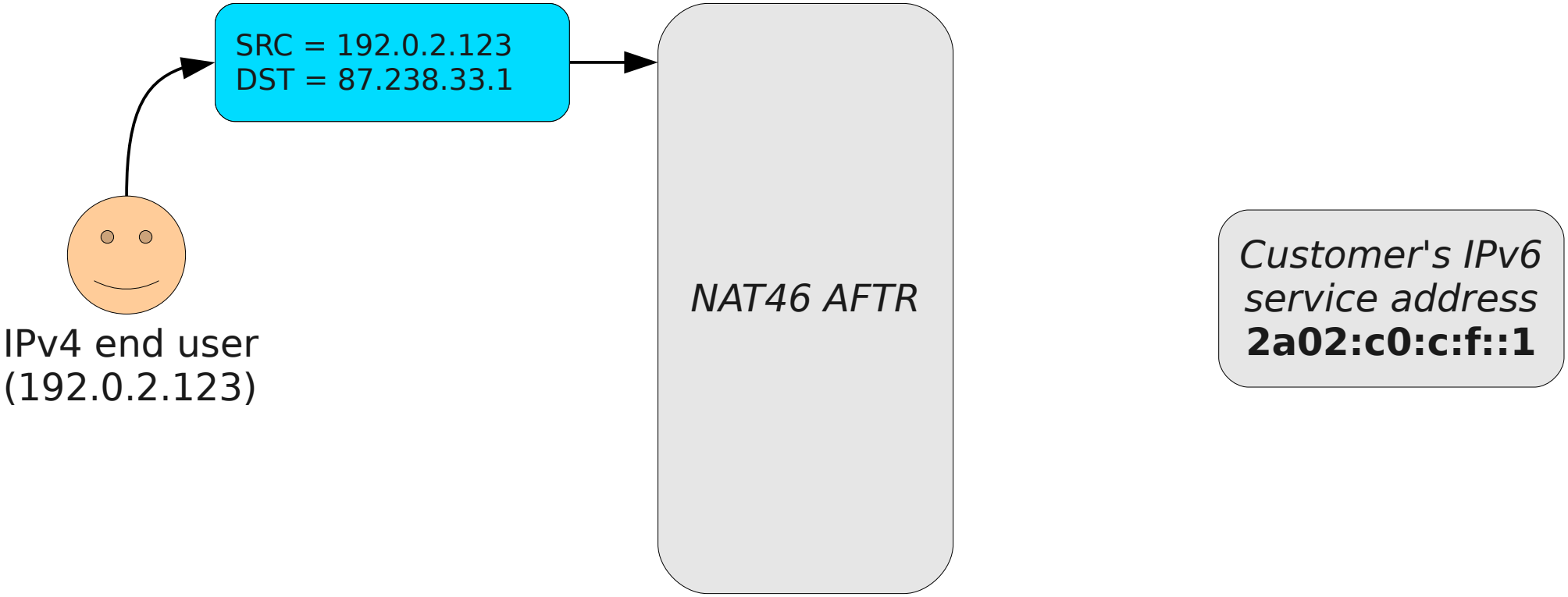



A mirrored pair of static mappings are configured on the AFTR

1) *When the destination is **87.238.33.1**, rewrite it to **2a02:c0:c:f::1***
2) *When the source is **2a02:c0:c:f::1**, rewrite it to **87.238.33.1***

# Packet flow 1

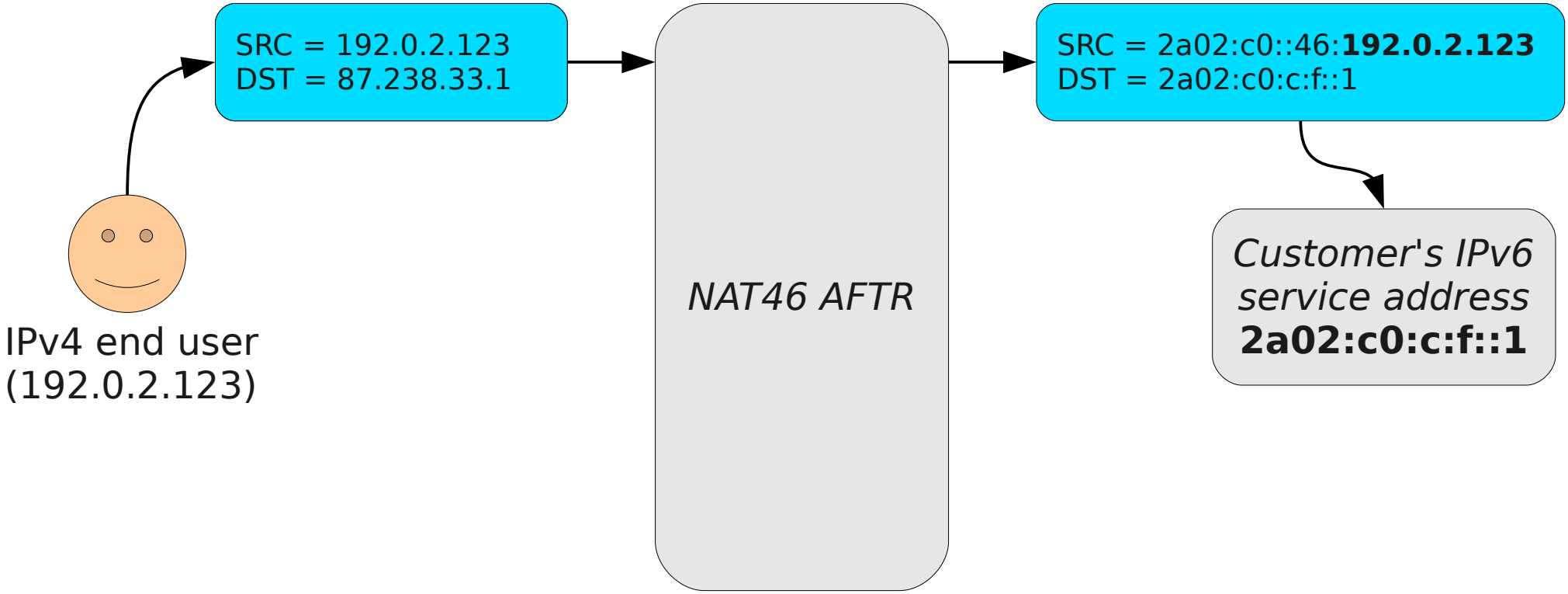
DNS setup:
www.cust.no. IN AAAA 2a02:c0:c:f::1
www.cust.no. IN A 87.238.33.1




- The end user sends a IPv4 packet to a service address that is routed to the AFTR and published in DNS

# Packet flow 2

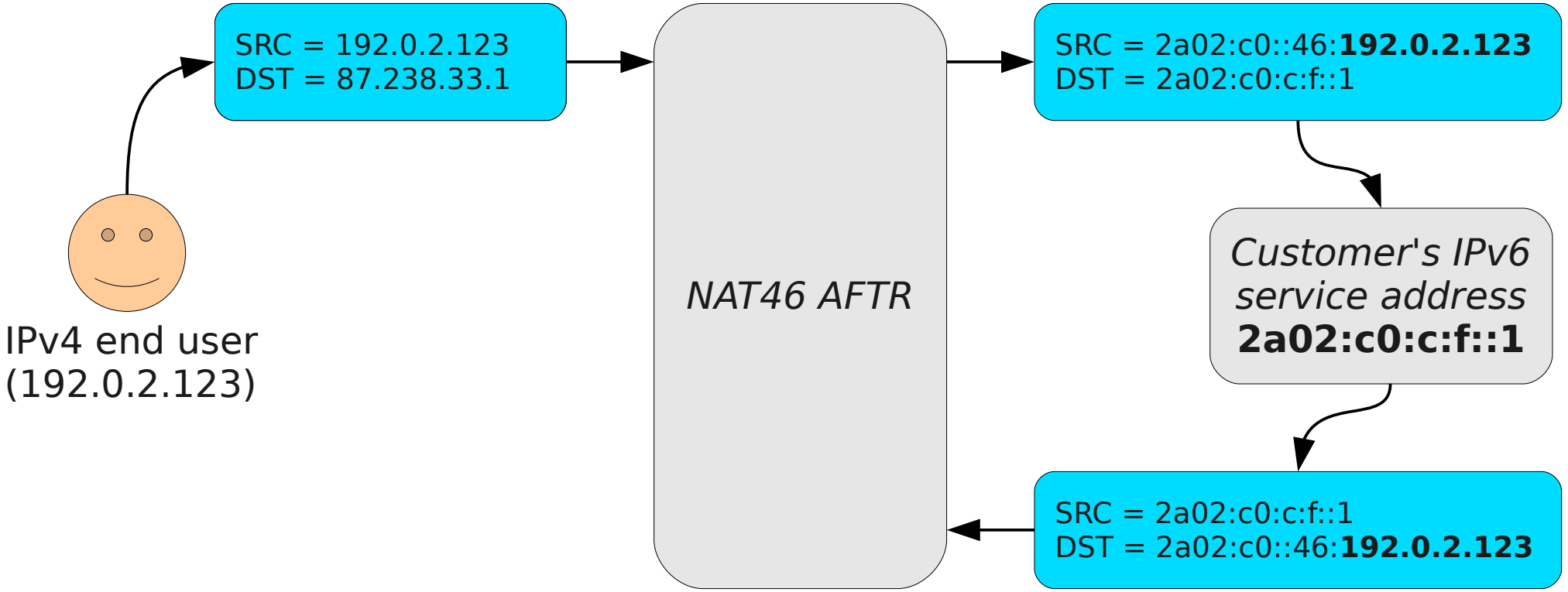
DNS setup:
www.cust.no. IN AAAA 2a02:c0:c:f::1
www.cust.no. IN A 87.238.33.1




- The AFTR device performs the following translations:
  1) it rewrites the IP destination field according to the static mapping
  2) it rewrites the IP source field by prepending its /96 prefix to the original IPv4 source address

# Packet flow 3

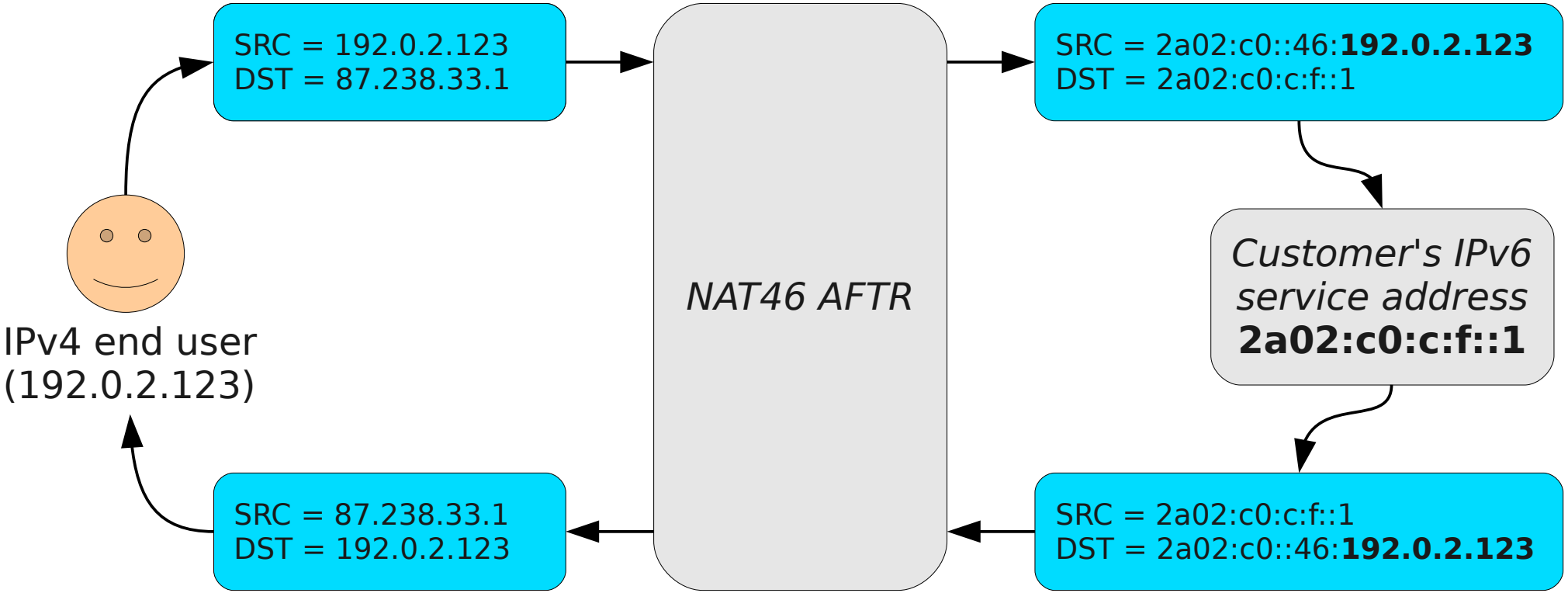
DNS setup:
www.cust.no. IN AAAA 2a02:c0:c:f::1
www.cust.no. IN A 87.238.33.1




- The customer's web server/load balancer/etc. responds to the packet exactly like it would do for any other native IPv6 client
- The destination address is routed to the AFTR device as part of a /96 route – the entire IPv4 internet is mapped into this prefix

# Packet flow 4

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




- The AFTR device performs the following translations:
  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 the static mapping
- The end user is not aware that the connection was translated

Redpill Linpro

# In summary

- Minimal operational overhead compared to dual-stack operation
- 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
- Clear forward-looking approach – why build new services on top of a legacy foundation?

Redpill Linpro

# Volunteers needed :-)

If you're a service owner that would like to participate in a pioneering IPv6-only deployment, and tell the industry about your experiences afterwards, do get in touch with me!

# Questions?

- Further reading:
  - http://fud.no/ipv6 (the report from our brokenness measurements)
  - http://getipv6.info/index.php/Customer_problems_that_could_occur (list of the bugs and other common causes for brokenness we found)
  - http://fud.no/talks (slides/video from this and my earlier talks)
  - http://v6asns.ripe.net (interactive graph over IPv6-enabled ASNs)
- My contact information:
  - tore.anderson@redpill-linpro.com
  - +47 95 93 12 12
  - @toreanderson
- Thank you for your attention!