

An introduction to IPv6

Tore Anderson
CG Security and Networking
Redpill Linpro
DSL-Partner, Oslo, October 2011

These slides are available from:
<http://fud.no/talks/>

Today's agenda

- What's new in IPv6
- IPv4 depletion
- ISP deployment approaches
- Security in IPv6
- Finishing discussion

What's new?

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
 - Assignments of /56 or /48 is becoming the industry norm

96 extra bits. That's it?

- Yep, that's the **only** new «must-have» feature. Really.
- There's several other changes though:
 - New address syntax and DNS records
 - Link-local addressing
 - Multicast mandatory (broadcast is gone)
 - Extended ICMP (ARP is gone)
 - Trimmed and more extensible protocol header format
 - Fragmentation is only performed by end hosts
 - New methods of address auto-configuration (ICMPv6/DHCPv6)
 - Address/router lifetimes, multinetting capabilities
 - De-facto standard subnet size of 64 bits (practically infinite)

IPv6 address 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 (may only be done once):

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



2a02:c0:1002:11::2

(The IPv4 equivalent of the above address would be **42.2.0.192.16.2.0.17.0.0.0.0.0.0.2**)

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:

*2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1.1.0.0.2.0.0.1.0.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

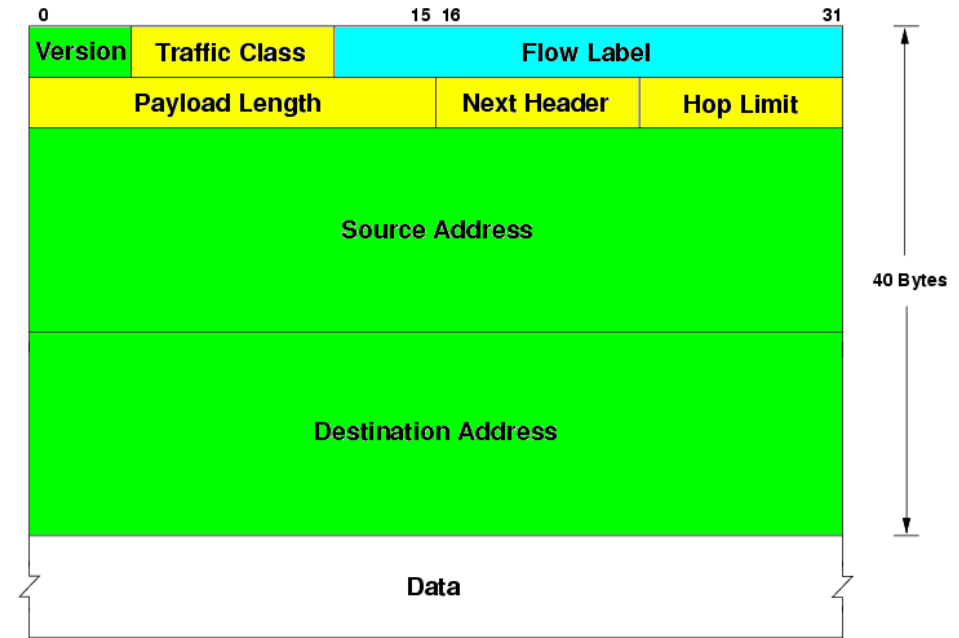
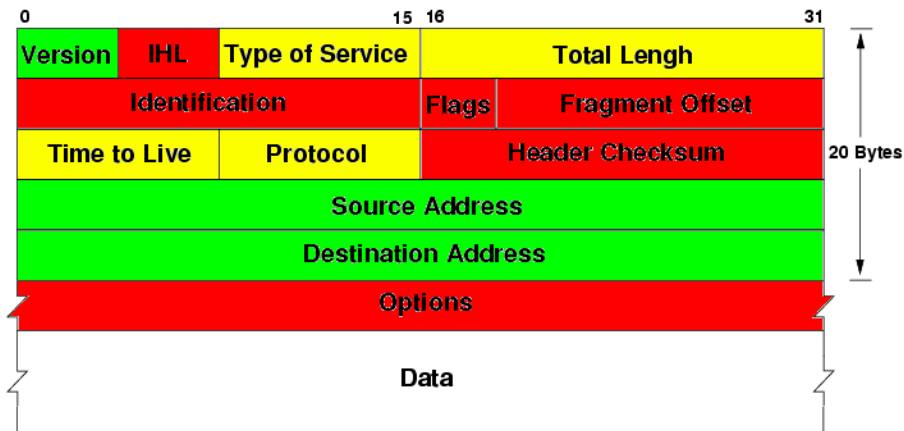
- The interface may be explicitly specified after a percent sign:

fe80::1234:abcd%eth0

Other stuff worth noting

- IPv6 routers **do not** perform fragmentation
 - Be extra careful not to break Path MTU Discovery
 - Avoid tunnels like the plague
- You (and your customers) get plenty of space, no questions asked
 - End-user assignments up to and including /48 (16K subnets sized /64) need no documentation - one size fits all
 - /32 is the minimum PA allocation size
 - Equals 16.7M /56s, or 64K /48s
 - RIPE NCC sparsely allocates to allow for growth up to a /29
 - Smaller DFZ due to less prefixes per ASN? (fingers crossed)

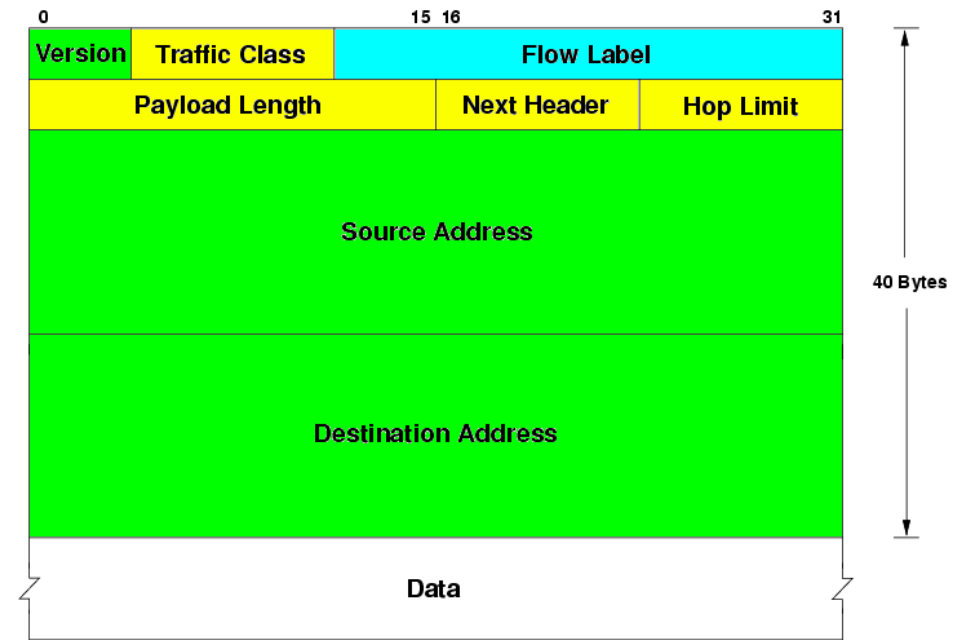
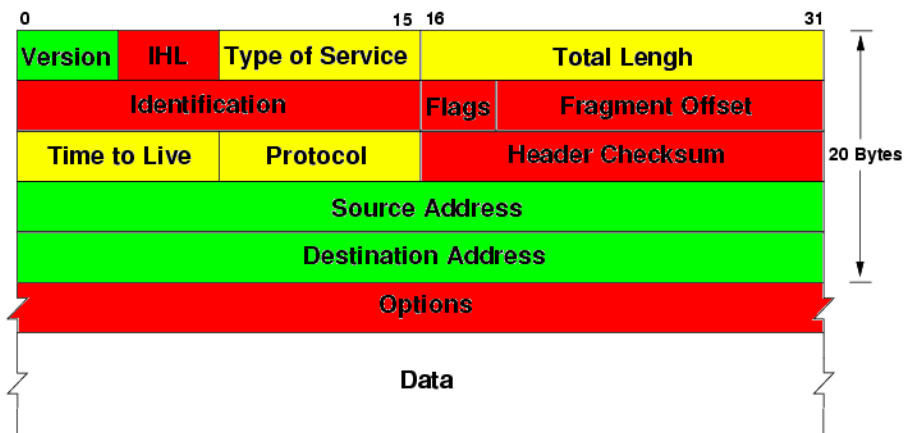
IPv4/IPv6 header comparison



Green: Unchanged fields
Yellow: Renamed/modified fields
Red: Removed in IPv6
Cyan: New in IPv6

- The **Version** and **Source/Destination Address** fields are unchanged
- The (optional) **Options** field has been removed in IPv6
- The **Internet Header Length** has been removed; an IPv6 header is always 40 bytes long
- **Header Checksum** has been removed in IPv6

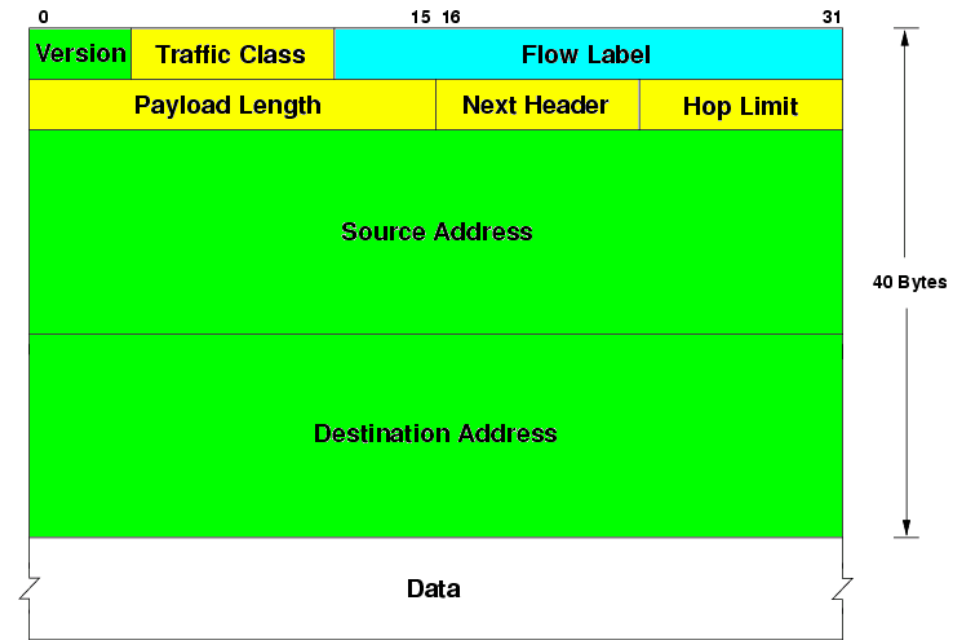
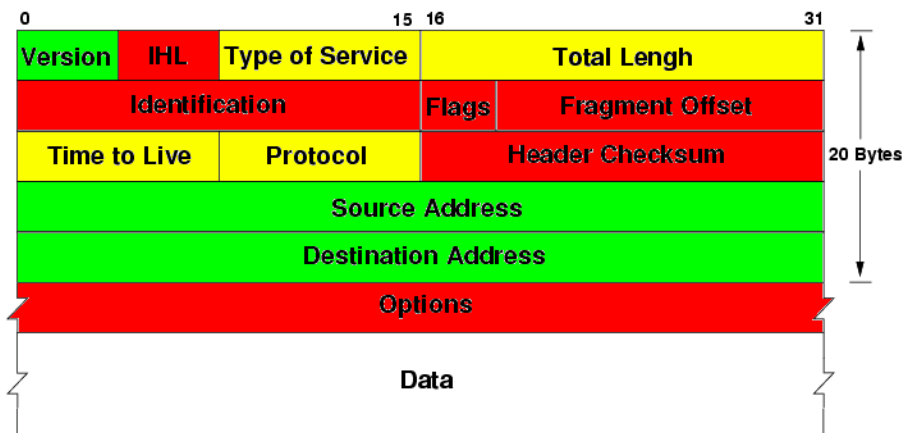
IPv4/IPv6 header comparison



Green: Unchanged fields
 Yellow: Renamed/modified fields
 Red: Removed in IPv6
 Cyan: New in IPv6

- The three fragmentation-related fields (***Identification***, ***Flags***, and ***Fragment Offset***) has been removed in IPv6
- ***Type of Service*** has been renamed to ***Traffic Class***; no change in usage (6 first bits for DiffServ/DSCP, 2 last bits for ECN)
- ***Time to Live*** has been renamed to ***Hop Limit***; no change in usage

IPv4/IPv6 header comparison



Green: Unchanged fields
 Yellow: Renamed/modified fields
 Red: Removed in IPv6
 Cyan: New in IPv6

- **Total Length** has been renamed to **Payload Length**; the value no longer includes the size of the IP header itself
- **Protocol** has been renamed to **Next Header**; the value may now also indicate the presence of an **IPv6 Extension Header**
- **Flow Label** is a new field meant to identify individual flows, for example in conjunction with ECMP load balancing

The well-known IPv6 prefixes

- `::/0` - the default route (equivalent to 0.0.0.0/0)
- `::/128` - the unspecified address (equivalent to 0.0.0.0/32)
- `::1/128` - the loopback address (equivalent to 127.0.0.1/32)
- `::/96` - IPv4-compatible addresses (e.g. `::192.0.2.1/128`). **DEPRECATED.**
- `::ffff:0:0/96` - IPv4-mapped addresses (e.g. `::ffff:192.0.2.1/128`). Used by applications using IPv6 socket APIs for communication with IPv4 hosts. Not to be seen on the wire.
- `2000::/3` - Global unicast (IANA to RIR allocations is done from this prefix)
 - `2001:0::/32` - Teredo (automatic IPv6-in-IPv4 tunneling)
 - `2001:10::/28` - ORCHID – Overlay Routable Cryptographic Hash IDentifiers
 - `2001:db8::/32` - Documentation prefix (equivalent to 192.0.2.0/24)
 - `2002::/16` - 6to4 (automatic IPv6-in-IPv4 tunneling)
- `3ffe::/16` - 6BONE v2 (experimental tunneled IPv6 overlay network). **DEPRECATED.**
- `5f00::/8` - 6BONE v1 (experimental tunneled IPv6 overlay network). **DEPRECATED.**
- `fc00::/7` - Unique Local Addresses (similar to 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16). Not to be seen on the internet.
- `fe80::/10` - Link-local unicast (equivalent to 169.254.0.0/16). Not to be seen on the internet.
- `ff00::/8` - Multicast (equivalent to 224.0.0/4). Bits 9-12 contains flags, 13-16 defines the scope. Multiple subdivisions, such as:
 - `ff02::/16` - Link-local multicast (equivalent to 224.0.0.0/24)
 - `ff05::/16` - Site-local multicast
 - `ff0e::/16` - Global multicast

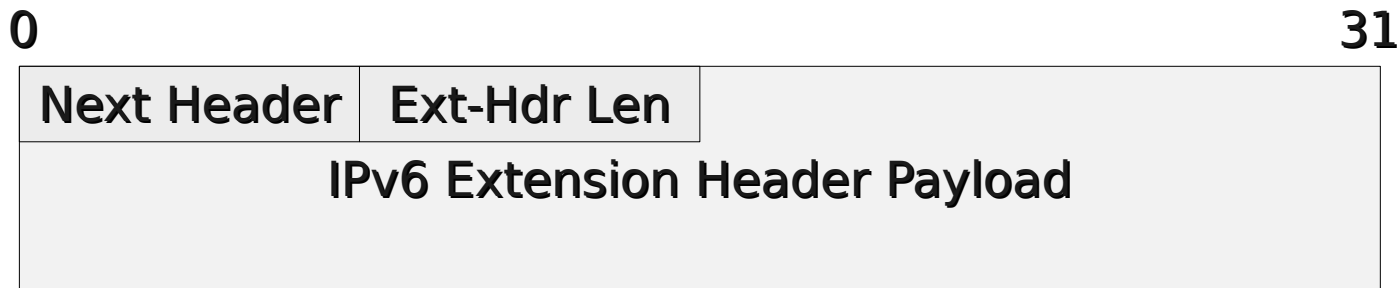
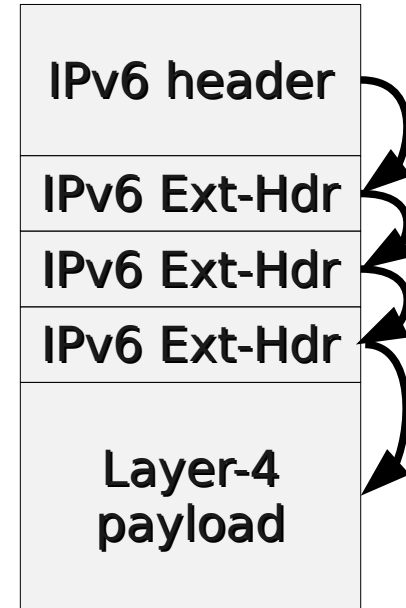
- Similar to RFC 1918 space, but vastly larger
 - fc00::/8 is reserved for future use
 - fd00::/7 is defined for local use:

fd	Randomly generated ID (40 bits)	End-user's address space (80 bits; a /48, or 64k subnets)
----	------------------------------------	--

- Bits 9-48 is defined to be random, to minimise chance of collisions
- Can provide internal addressing on a home LAN that is stable across WAN link failures and/or renumbering
- Hosts do not generally distinguish between ULAs and GUAs
 - If a host has a ULA only, plus a default route, it will generally be preferred to IPv4 for internet connectivity
 - Will result in timeouts and unhappy users

IPv6 Extension Headers

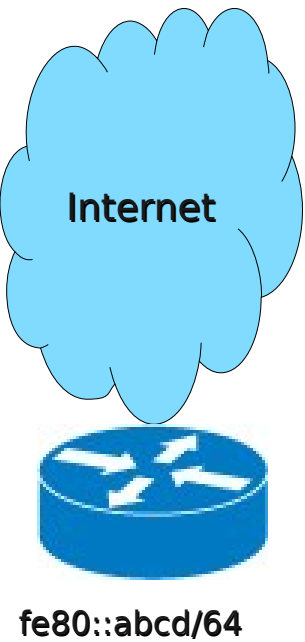
- The **Next Header** field may contain a protocol number assigned to an **IPv6 Extension Header** instead of a Layer-4 protocol like TCP or UDP
- An Extension Header always begins with another **Next Header** field, then a **Length** field, then the Extension Header Payload
- Multiple extension headers may be present, forming a chain



IPv6 extension headers, cont.

- Processed by end nodes only, except for the special **Hop-by-Hop** extension header (Next Header = 0), which **MUST** come first
 - Hop-by-Hop header is processed by the control plane – slow path
- Some examples of extension header types:
 - IPSEC (AH and ESP like in IPv4)
 - Fragmentation
 - Mobile IPv6
- Extension headers complicate filtering by intermediate nodes
 - Difficult to know where the layer-4 payload starts, as the entire extension header chain must first be followed to its end
 - Layer-4 ACLs doesn't always work when EHs are present, depending on which hardware is used

IPv6 host initialisation process



Configure Link-local addressing

The host needs to generate a unique 64-bits Interface ID for itself. The **EUI-64** algorithm is the most commonly used:

- 1) Start with the Ethernet MAC address of the host:

00:11:22:33:44:55

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

0011:22ff:fe33:4455

- 3) Flip bit 7 (the universal/local bit):

0211:22ff:fe33:4455

- 4) Prepend the well known prefix **fe80::/64**:

fe80::211:22ff:fe33:4455



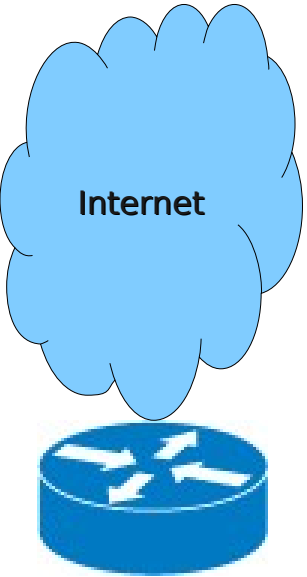
fe80::abcd/64



fe80::211:22ff:fe33:4455/64
::1/128
(state tentative)

Join multicast groups

The host must join at least two link-local multicast groups, using **Multicast Listener Discovery (MLD)** – the IPv6 replacement for IGMP:



fe80::abcd/64

- 1) **ff02::1** a.k.a. All Nodes on link, essentially IPv6 broadcast
- 2) **ff02::1:ff33:4455** a.k.a. Solicited Node, used for ICMPv6 Neighbor Discovery (the equivalent to IPv4 ARP). Combine **ff02::1:ff00:0/104** with the last 24 bits of the interface's IPv6 address.



fe80::211:22ff:fe33:4455/64
::1/128
(state tentative)

Src: :: (Unspecified)
Dst: ff02::16 (All-MLDv2-Routers)
MLDv2 Report
Interface joins groups **ff02::1** and
ff02::1:ff33:4455

Duplicate Address Detection

The host attempts to determine whether or not the chosen address is in use by sending an ICMPv6 **Neighbor Solicitation** (think ARP who-has) for the chosen address to the Solicited Node multicast address, using the **Unspecified** address as source.



fe80::abcd/64

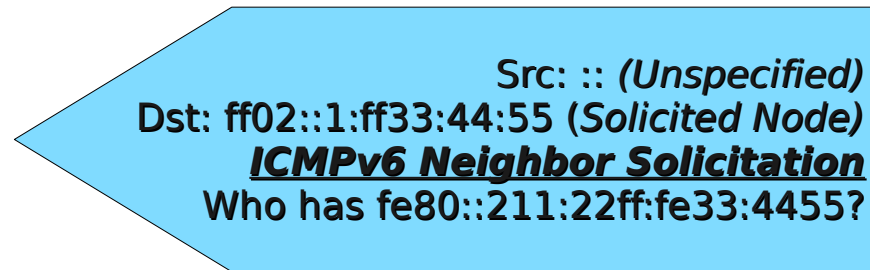
Src: :: (*Unspecified*)
Dst: ff02::1:ff33:44:55 (*Solicited Node*)
ICMPv6 Neighbor Solicitation
Who has fe80::211:22ff:fe33:4455?



fe80::211:22ff:fe33:4455/64
::1/128
(state tentative)

Duplicate Address Detection

The host attempts to determine whether or not the chosen address is in use by sending an ICMPv6 **Neighbor Solicitation** (think ARP who-has) for the chosen address to the Solicited Node multicast address, using the **Unspecified** address as source.



fe80::211:22ff:fe33:4455/64
::1/128

If no **Neighbor Advertisement** message is received in reply (to either of the multicast groups), the tentative flag is cleared and the address is ready for use.

Routing and global addressing

The host requests configuration from on-link routers by transmitting a **Router Solicitation Message**, which is replied to with a **Router Advertisement** containing config info

Src: fe80:211:22ff:fe33:4455
 Dst: ff02::2 (All Routers)
ICMPv6 Router Solicitation



fe80::abcd/64

Src: fe80::abcd
 Dst: ff02::1 (All Nodes)
ICMPv6 Router Advertisement
 Router Lifetime: 30m
 Managed flag: 0
 Other Configuration flag: 0
Prefix Information Option:
 Prefix: 2001:db8::/64
 Valid Lifetime: 24h
 Preferred Lifetime: 6h
 Autonomous flag: 1
 On-Link flag: 1
Recursive DNS Server Option:
 DNS server 1: 2001:db8::53:1
 DNS server 2: 2001:db8::53:2



fe80::211:22ff:fe33:4455/64
 ::1/128

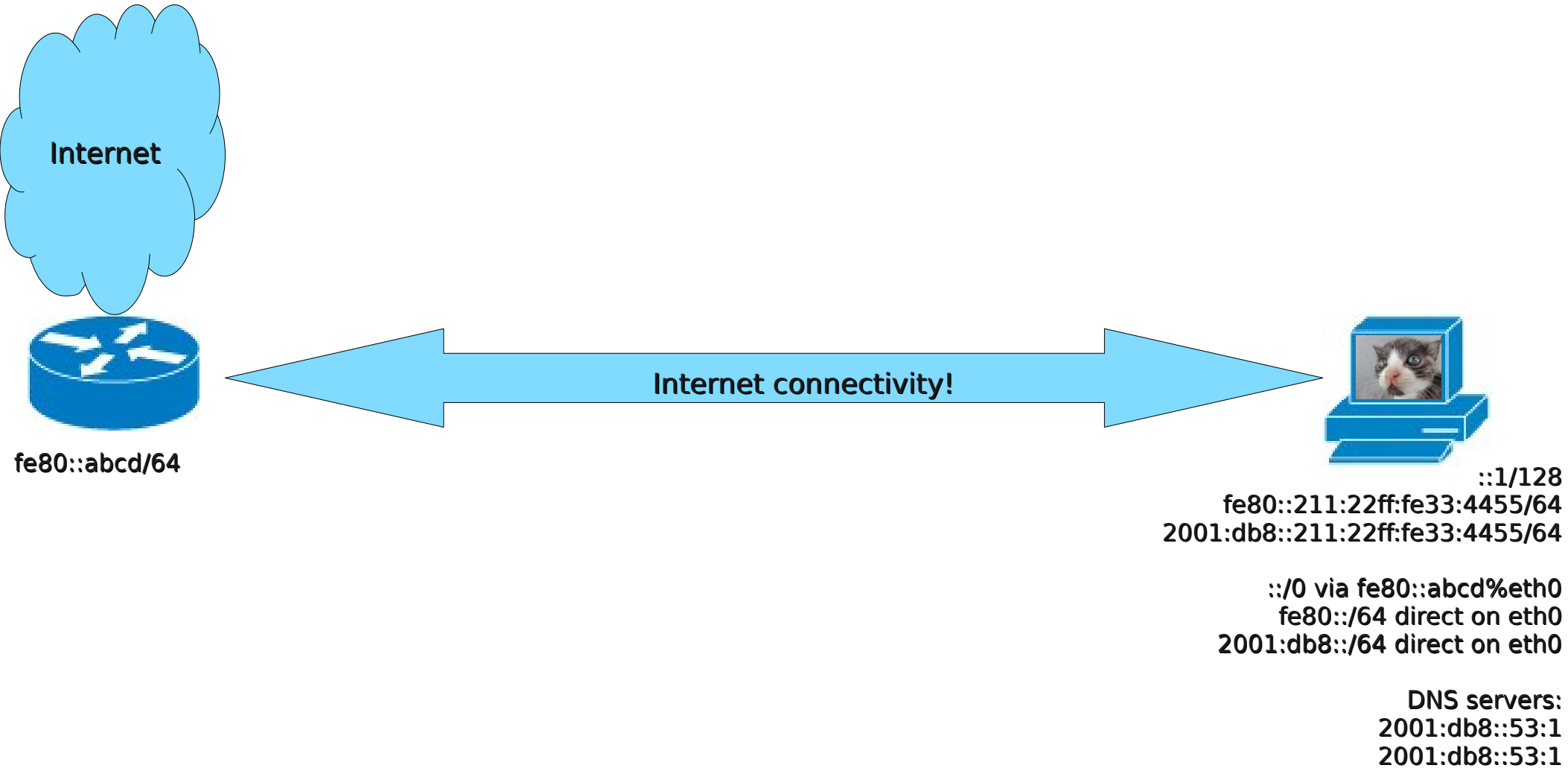
Processing Router Advertisements

- If **Router Lifetime** is >0 , add the router to the list of default routers. This is the **only way** a regular IPv6 node can learn the default router!
- If the **Managed** flag is set, the host should attempt to obtain an address lease and other config (e.g. DNS/NTP servers) from **DHCPv6**
 - Otherwise, if the **OtherConfig** flag is set, consult **DHCPv6** for config (DNS/NTP/..) but do not request a lease (*Information-Only DHCPv6*).
- For all **Prefix Information Options** with the **Autonomous** flag set:
 - Generate a 64-bit interface ID (e.g. using *EUI-64*), combine with specified prefix, configure on interface, and perform **DAD**
 - If **On-Link** is set, add a direct route to the prefix on the interface. (If unset, all traffic between hosts must pass through the router.)
- Add any **Recursive DNS Servers** to the OS' stub resolver configuration

Processing Router Advertisements

- Various timers starts counting down after RA processing:
 - When **Router Lifetime** reaches 0, the default route is removed
 - When an address' **Preferred Lifetime** reaches 0, it is no longer a candidate for establishing new connections with
 - When an address' **Valid Lifetime** reaches 0, it is removed from the interface
 - DNS servers also have lifetimes
- There's no shortage of other RA options, such as:
 - More-Specific Routes (with lifetimes)
 - DNS Domain Search List (with lifetimes)
 - Router preference
 - Hardware address of advertising router (so that ND can be skipped)

We're done! (Well, for now...)

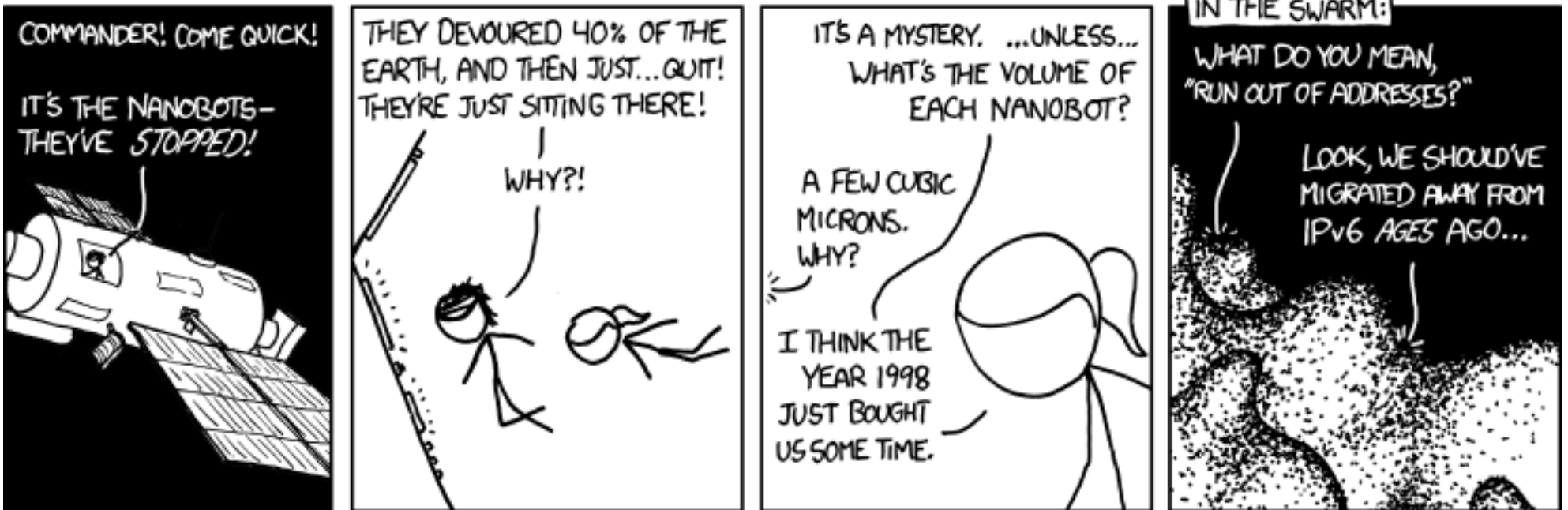


Hosts never stop listening for RAs

- Routers will periodically transmit unsolicited RAs in order to:
 - Reset lifetime counters
 - Perform link renumbering or reconfiguration:
 - Send old prefix/configuration with timers set to 0
 - Send new prefix/configuration with normal timers
 - Remove itself from the default router list, e.g.:
 - Before shutting down/maintenance
 - After its WAN/upstream connectivity has failed

Operating System support

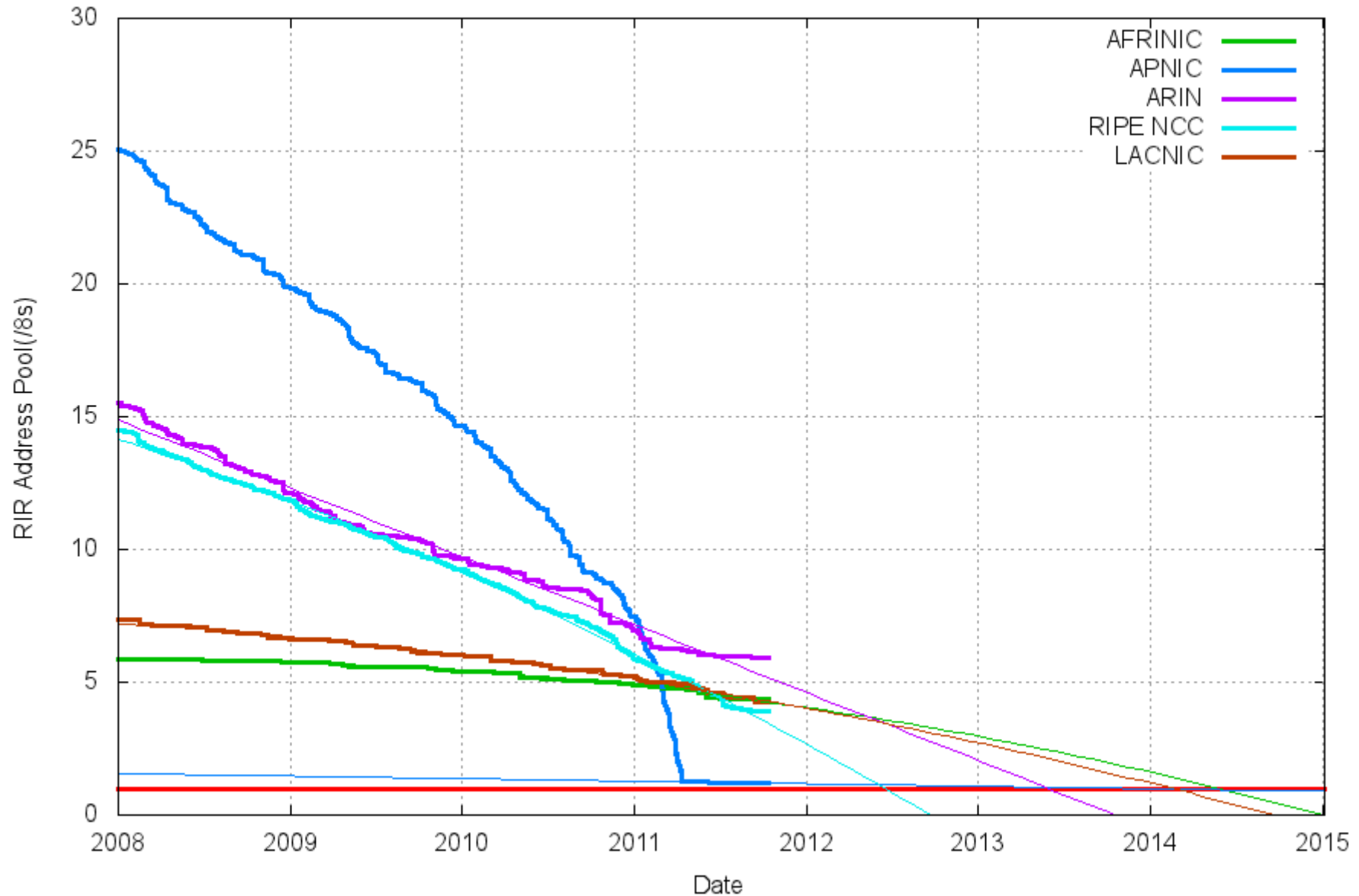
- All modern OS-es support acquiring an IPv6 address and the default route from RAs
 - Called **SLAAC** (Stateless Address AutoConfiguration)
 - DHCPv6 is less supported (Windows and recent OS X/iOS/Linux)
- DNS servers are more tricky
 - Windows only supports DHCPv6, not RDNSS option in RA
 - OS X 10.6 only support RDNSS, no DHCPv6
 - OS X 10.7 and very recent Linux and iOS supports both
 - For maximum host compatibility, advertise DNS servers in both places; RDNSS in RA and DHCPv6 (set the OtherConfig RA flag to 1)
- No shortage of implementations that do not properly support IPv6
 - Android, Windows pre Vista, OS X pre 10.6, PS3, Xbox,
 - Your customers are going to continue to demand IPv4!



IPv4 depletion

IPv4 depletion timeline

RIR IPv4 Address Run-Down Model



IPv4 depletion facts

- IANA ran out of IPv4 addresses in February, APNIC in April
- 64M unallocated addresses remain in the RIPE region (Oct 10)
- RIPE's current allocation horizon is three months' usage
- The «allocations from the last /8» policy (ripe-509 5.6) states:
 - Each LIR may allocate a single /22 only - 1024 addresses
 - ...provided the applicant also has an IPv6 allocation
- An IPv4 address market will likely show up
 - Prices and supply highly uncertain
 - Nortel/Microsoft deal: 666K addrs for US\$7.5 -> US\$11.50/addr
- Deployment of IPv6 will NOT stop IPv4 depletion from happening
 - But it may give you more tools to better deal with it

It takes two to tango

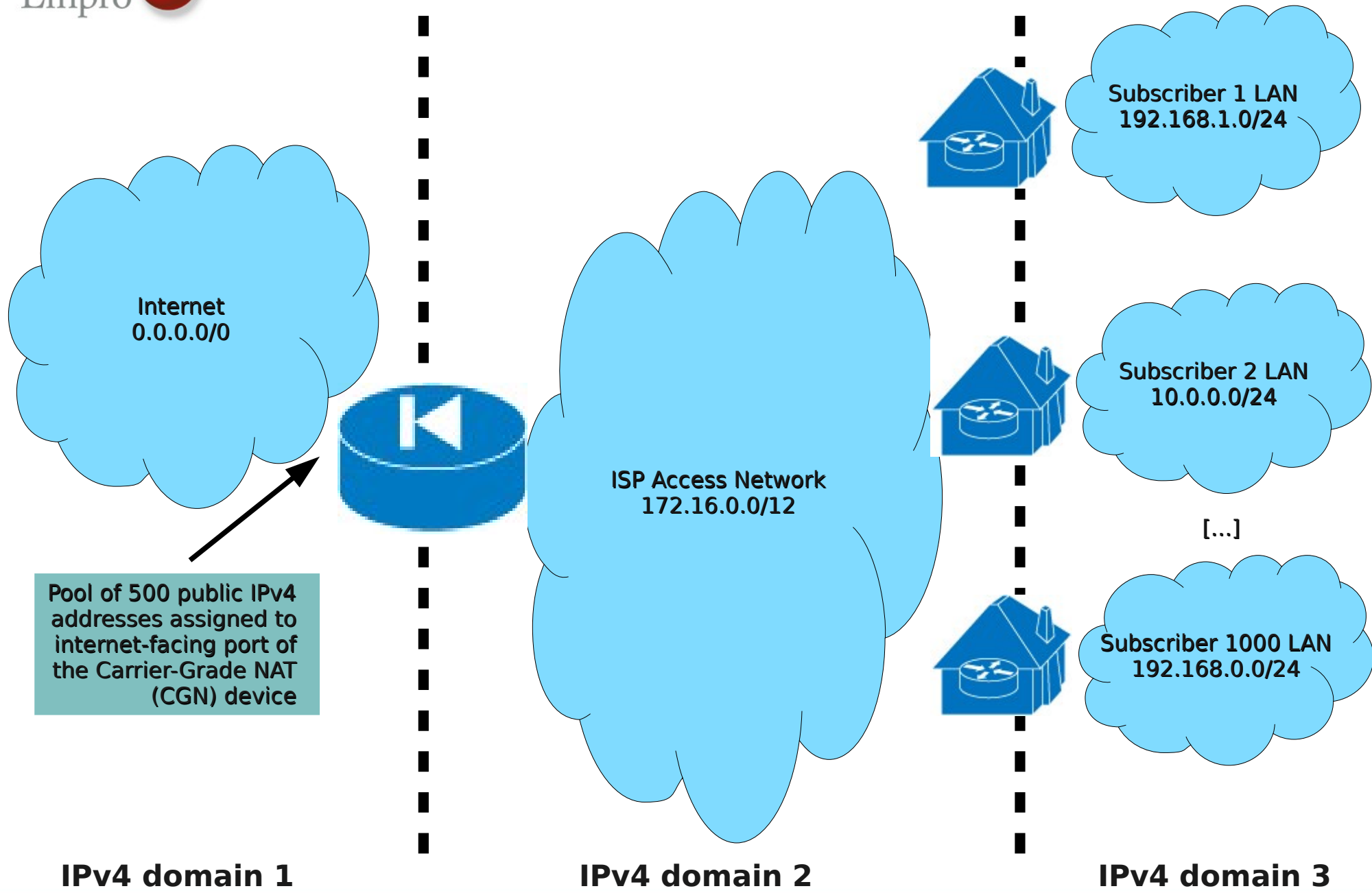
- Do you:
 - ...have plenty of IPv4 addresses?
 - ...no growth in your customer base?
 - ...no customer demand for IPv6?
 - ...not see the point of bothering with IPv6 at all?
- Then consider whether the above is true for everyone your customers would like to communicate with over the internet
- Also consider what you will do the day:
 - IPv6 deployment on the internet is reaching critical mass, and
 - Pear, Inc. releases a new iThing that requires IPv6?

ISP deployment scenarios

NAT444

Also known as: Carrier Grade NAT (CGN) or Large Scale NAT (LSN)

NAT444: three distinct IPv4 domains separated by NATs



NAT444 highlights

- Handles IPv4 depletion
- Relatively easy to implement
 - Only one new box to purchase
 - Offerings available from numerous vendors
 - The CPEs don't have to be touched at all
 - Doesn't require major changes to access network
 - Freeing up public IPv4 addresses by renumbering
 - Deployment of central NAT box(es)

NAT444 disadvantages

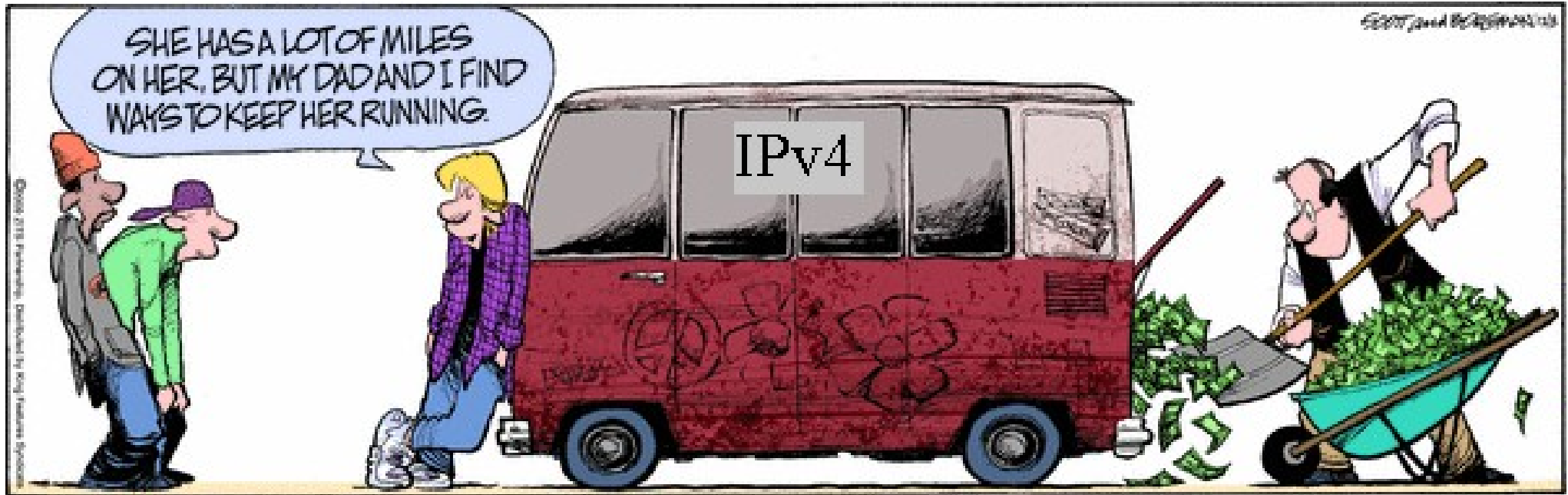
- Stateful, flow-tracking devices
 - Much more complex (and thus expensive) than routers
 - Limited by flows/s rather than raw pps/bps throughput
 - Potential DoS target (also non-malicious, e.g., BitTorrent)
- Existing methods of NAT traversal (UPnP, NAT-PMP) break
 - Affects BitTorrent, video streaming, SIP, online gaming, etc.
 - See [I-D.donley-nat444-impacts](#)
 - Static port forwards by customer
- Abuse handling and compliance with legal regulations is tricky
 - Which subscriber sent spam from public IPv4 address X?

NAT444 disadvantages

- How to number the access network
 - RFC1918 space?
 - May conflict with users' LAN prefixes
 - Multiple overlapping realms using provider's public addresses?
 - Wastes precious IPv4 resources, adds significant complexity
 - Squatting on currently unadvertised space?
 - Risk of future reassignment and use on the internet
 - I-D.weil-shared-transition-space-request
 - Aims to make IANA/ARIN allocate a /10 for this exact purpose
 - Non-RFC1918 might make the CPE assume direct internet access
 - Particularly damaging for CPEs with 6to4 built in

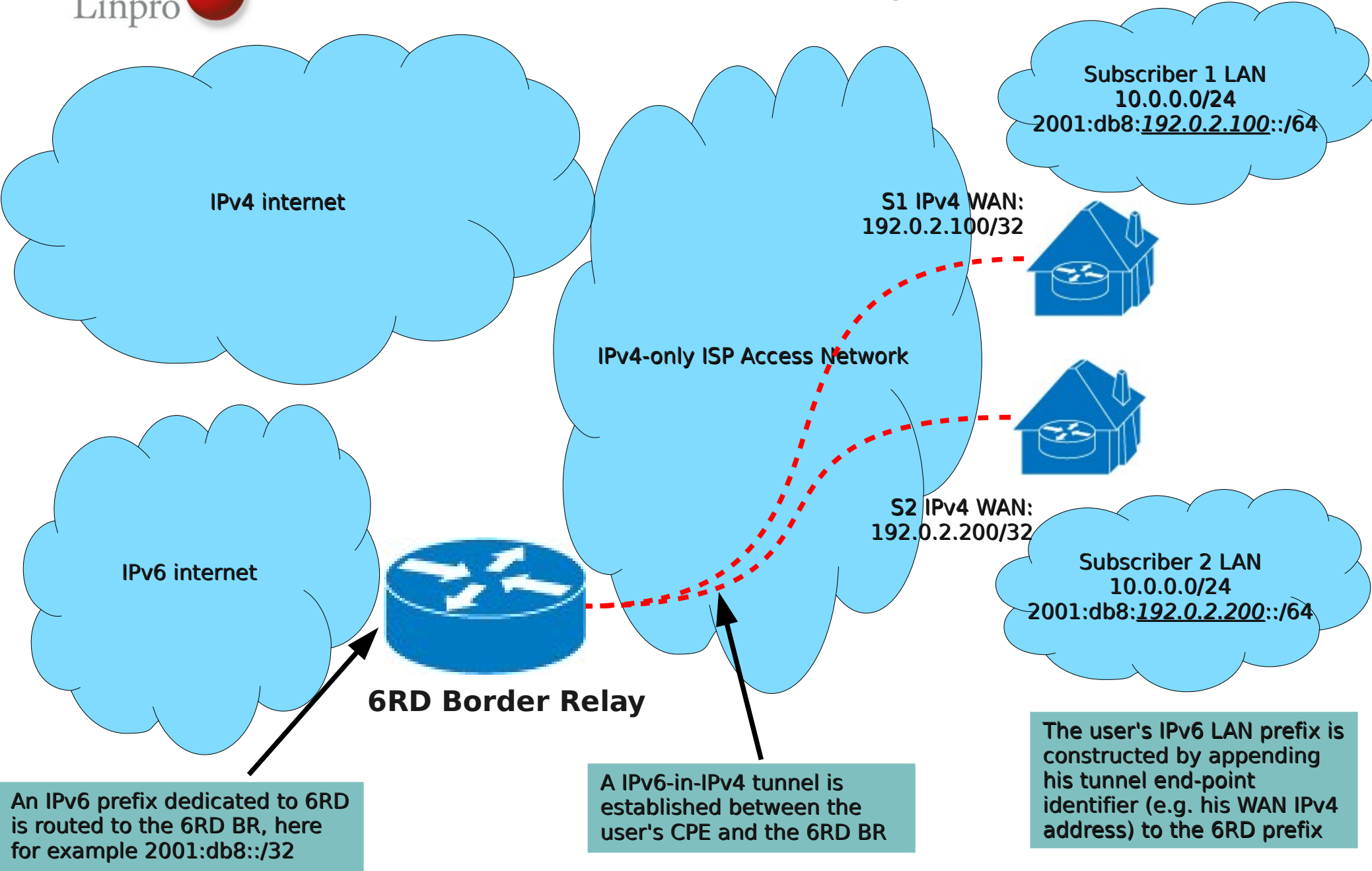
NAT444 disadvantages

- Traffic must be flow bidirectionally across the NAT device
 - 1) Backhaul traffic to and from central NAT site(s), or
 - 2) Distribute NAT devices across all or most PoPs
- NAT devices must be scaled up as overall traffic levels increase



IPv6 Rapid Deployment - 6RD

6RD: IPv6 tunneled across an IPv4-only access network



6RD highlights

- Easy to implement, no changes to access/core network required
 - The 6RD BR can be placed anywhere in the network
- Stateless operation
 - The 6RD prefix and the 6RD BR address may be anycasted
- Production quality code and implementations exist
 - Supported in the standard Linux kernel
 - Very similar to 6to4 which is implemented in many CPEs already, not difficult to add the missing 6RD pieces to the firmware
 - Standardisation is complete, RFCs are published
- Free (a French ISP) has run 6RD in production for several years, with well over a million active subscribers

6RD disadvantages

- Does not in any way help with IPv4 depletion
 - However, it may be combined with NAT444
- Does not provide any IPv4 exit strategy (quite the opposite)
- Requires support in CPE
- Uses tunneling
 - Encap/decap overhead
 - Reduced MTU for IPv6 flows
- Requires end-user to have relatively stable IPv4 addressing

6RD in detail

- The CPE learns the IPv4 address of the 6RD BR and the delegated IPv6 prefix through a DHCPv4 option (212)
- In the simplest form, all the 32 bits of the end user's IPv4 address is encoded in the delegated IPv6 prefix
 - Means customer will only get a /64 if the 6RD prefix is a /32:

6RD prefix (32 bits)	End-user's IPv4 address (32 bits)	End-user's assigned address space (64 bits – a single subnet)
2001:db8:	192.0.2.100 (c000:0264)	::

- You will likely be able to receive a larger prefix from the RIPE NCC if you plan on using 6RD (for example, Free got a /26)

Conserving bits

- Identical leading bits from the IPv4 addresses when constructing the end-users' tunnel endpoint identifiers may be stripped
- For example, assume all your 6RD users are in **192.0.2.0/24**:

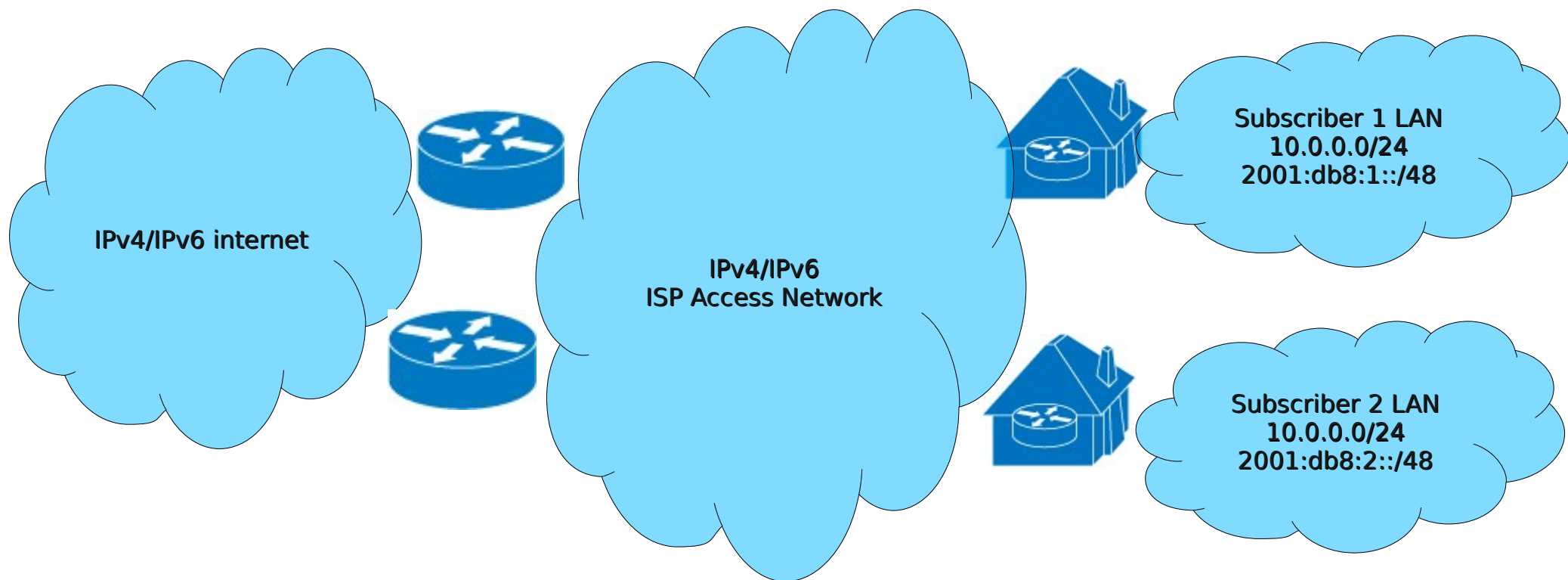
6RD prefix (40 bits)	EU ID 8 bits	End-user's assigned address space (80 bits; 2001:db8:0064::/48, or 64k subnets)
2001:db8:00	192.0.2.100 (64)	::

- The number of bits to strip is controlled by the 6RD DHCPv4 option
- With several IPv4 prefixes, you can create multiple 6RD domains using a different 6RD prefixes. For example, with up to 16 IPv4 prefixes, none larger than a /16, you could do:

6RD allocation (36 bits)	6RD DomID 4 bits	EU ID 16 bits	End-user's assigned address space (72 bits; a /56, or 256 subnets)
			<i>(6RD prefix = 6RD allocation + 6RD Domain ID)</i>

Dual Stack

Dual Stack: Run native IPv6 in parallel with IPv4, everywhere



Dual Stack

- IPv6 and IPv4 are equal citizens, same performance characteristics
 - Technically superior solution (for the end user)
- IPv6 may be gradually be deployed on existing infrastructure without touching IPv4
- CPEs and access network components must all support IPv6
- The CPE learns its IPv6 prefix using **DHCPv6 Prefix Delegation**
 - This means DHCPv6 is involved in routing (unlike DHCPv4)
- Provides an IPv4 exit strategy
 - The IPv6 service has no dependency on IPv4

DHCPv6 Prefix Delegation

PE router
fe80::1

CPE router
fe80::2

2. DHCPv6 server allocates a free IPv6 prefix (here 2001:db8::/48) from a pool, and configures inserts route:
2001:db8::/48 via fe80::2%foo

ICMPv6 RA
(May or may not configure a global IPv6 WAN address on CPE with SLAAC or Managed=1)

DHCPv6-PD request

DHCPv6-PD response
Prefix: 2001:db8::/48
DNS servers: [...]

1. Configures route:
::/0 via fe80::1%WAN

3. Configures routes:
2001:db8::/48 discard
2001:db8::/64 dev LAN

Configures RA/DHCPv6 service on LAN interface to use 2001:db8::/64

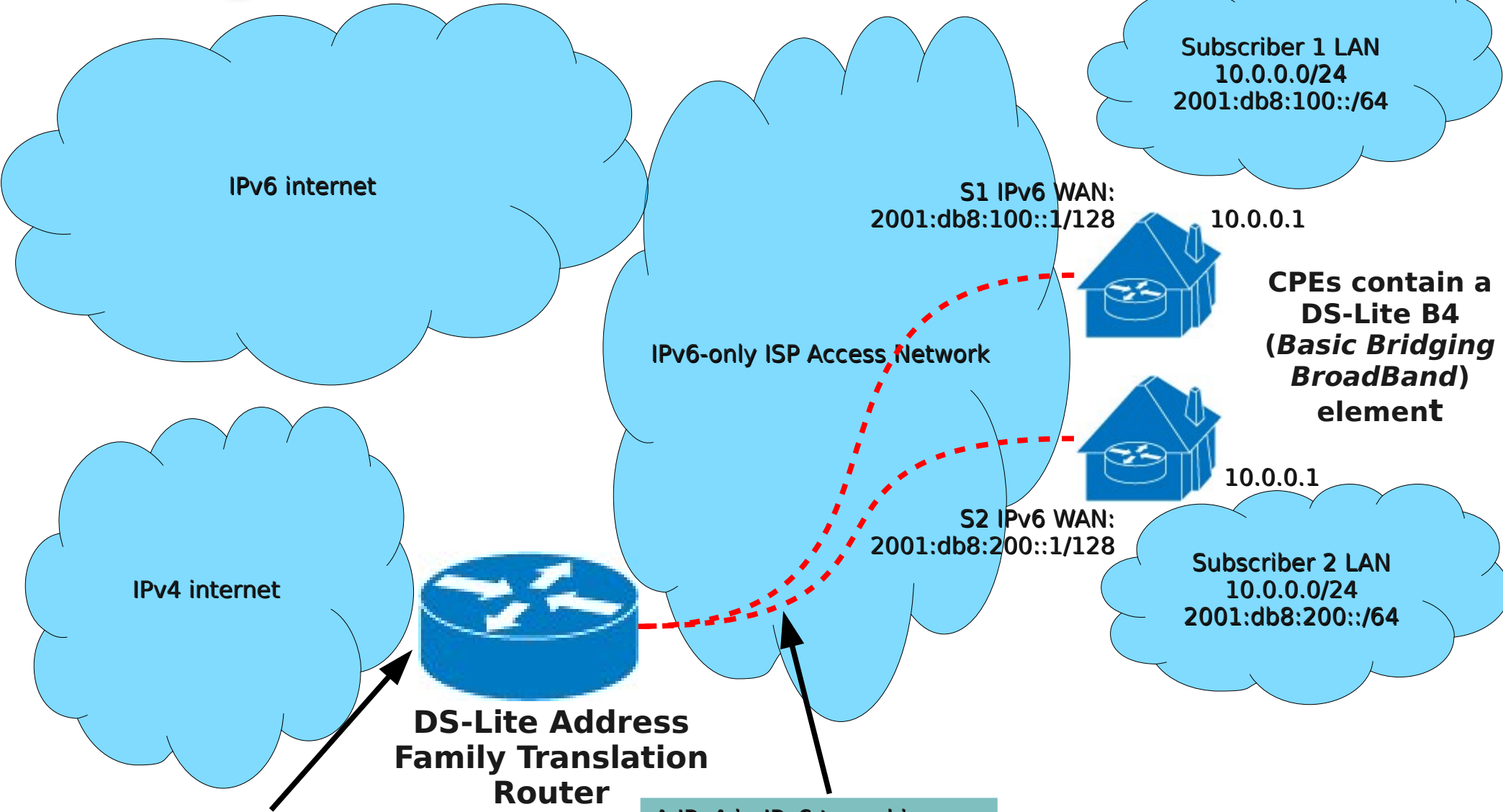
Optionally serves any DHCPv6-PD clients on LAN using prefixes from 2001:db8::/48 pool

Dual Stack disadvantages

- Double the amounts of stacks, double the amount of work...
 - Monitoring, ACLs, configuration, troubleshooting, accounting, etc.
- Doesn't help with IPv4 depletion
 - Combine with NAT444
- May require large/expensive replacements in the access network
- Missing IPv6 support in parts of access network might force you to have incongruent layer 3 topologies for IPv4 and IPv6
- Achieving same level of security as in IPv4 with N:1 Subscriber:VLAN models are challenging at best (more on that later)

Dual Stack Lite (DS-Lite)

DS-Lite: IPv4 tunneled across an IPv6-only access network, terminating in a central NAT44 (+ native IPv6)



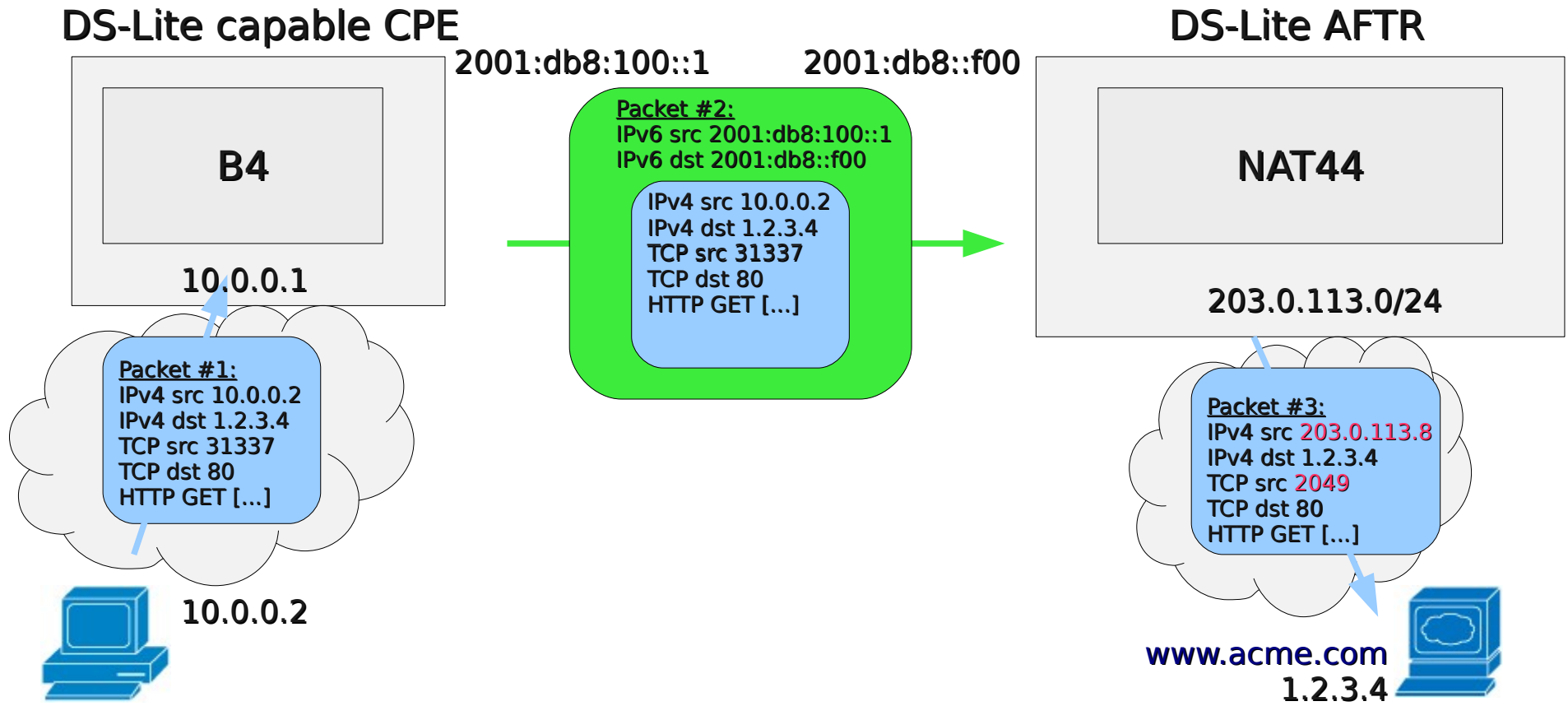
One or more public IPv4 prefixes are routed to the AFTR, used by a NAT44 function

A IPv4-in-IPv6 tunnel is established between the user's CPEs B4 element and the AFTR

Dual Stack Lite key points

- Single-stack access network, less operational overhead
 - Native IPv6 - works the same as IPv6 in standard Dual Stack
 - Maximises IPv4 address usage efficiency
- Provides an IPv4 exit strategy
 - AFTR traffic will diminish as IPv6 is deployed on the internet
- The AFTR(s) may be placed anywhere in the network
- The DS-Lite CPE (B4) element..
 - ..learns the address of the AFTR from a DHCPv6 option
 - ..does **NOT** perform NAT44 – single-level NAT44 end-to-end
 - ..provides DHCPv4 and DNS service for IPv4 LAN hosts

DS-Lite NAT44 element



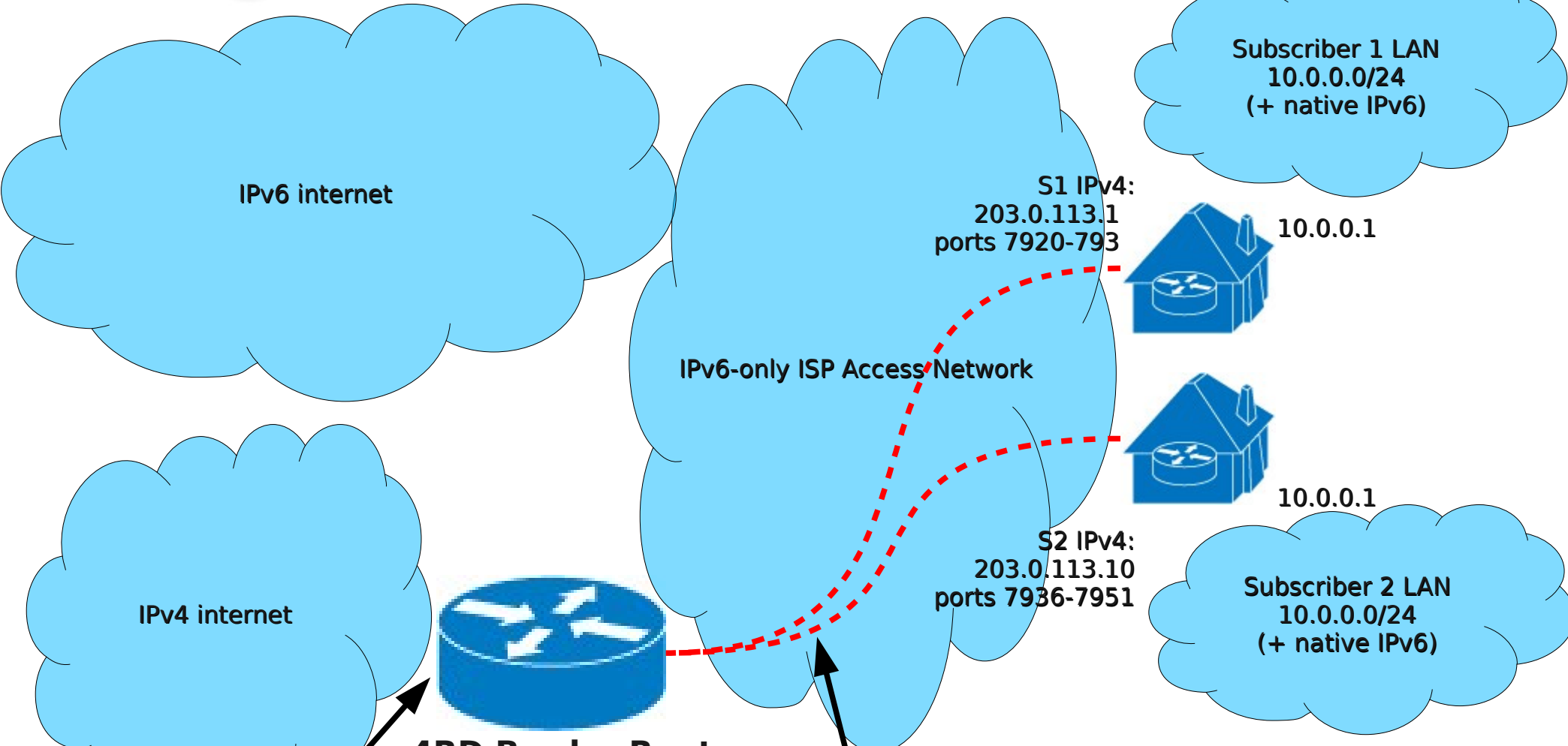
NAT44 mapping on AFTR:
2001:db8:100::1/10.0.0.2/TCP/31337 <---> 203.0.113.8/TCP/2049

DS-Lite disadvantages

- The AFTR's NAT44 is a stateful flow-tracking device
 - Performance and cost concerns, same as with CGN
- IPv4 traffic is tunnelled
 - Reduced MTU
 - Encap/decap overhead
- The DS-Lite standard isn't fully mature at the moment
 - Draft status in the IETF
 - Not many CPEs to choose from
 - No significant production-level deployments to date, trials only

IPv4 Residual Deployment (4RD)

4RD: IPv4 tunneled across an IPv6-only access network, borrowing bits from TCP/UDP port space for routing (+ native IPv6)



One or more public IPv4 prefixes are routed to the BR (here e.g. 203.0.113.0/24)

A IPv4-in-IPv6 tunnel is established between the user's CPEs 4RD BR

The user's CPE performs IPv4 NA(P)T. Several subscribers may share the same IPv4 address - if so, each have a limited unique set of ports available to him

4RD highlights

- Similar use-case as DS-Lite's
 - Native IPv6-only SP access network + automatic IPv4 tunnels
 - Single-level NAT44 between end user and the internet
 - Helps deal with IPv4 depletion
- A 4RD CPE may be delegated a dedicated IPv4 prefix, a dedicated IPv4 address, or a shared IPv4 address
 - You could charge customers extra for «whole» IPv4 addresses
 - Prepare early for depletion, start sharing only when you have to
- The 6RD BR is a stateless device, does not contain a NAT44 element
 - May be load balanced and anycasted, similar to a 4RD BR
 - Scales with traffic, not sessions/subscribers (unlike DS-Lite and CGN)
- IPv6 address of 6RD BR is discovered using a DHCPv6 option

IPv4 address sharing in 4RD

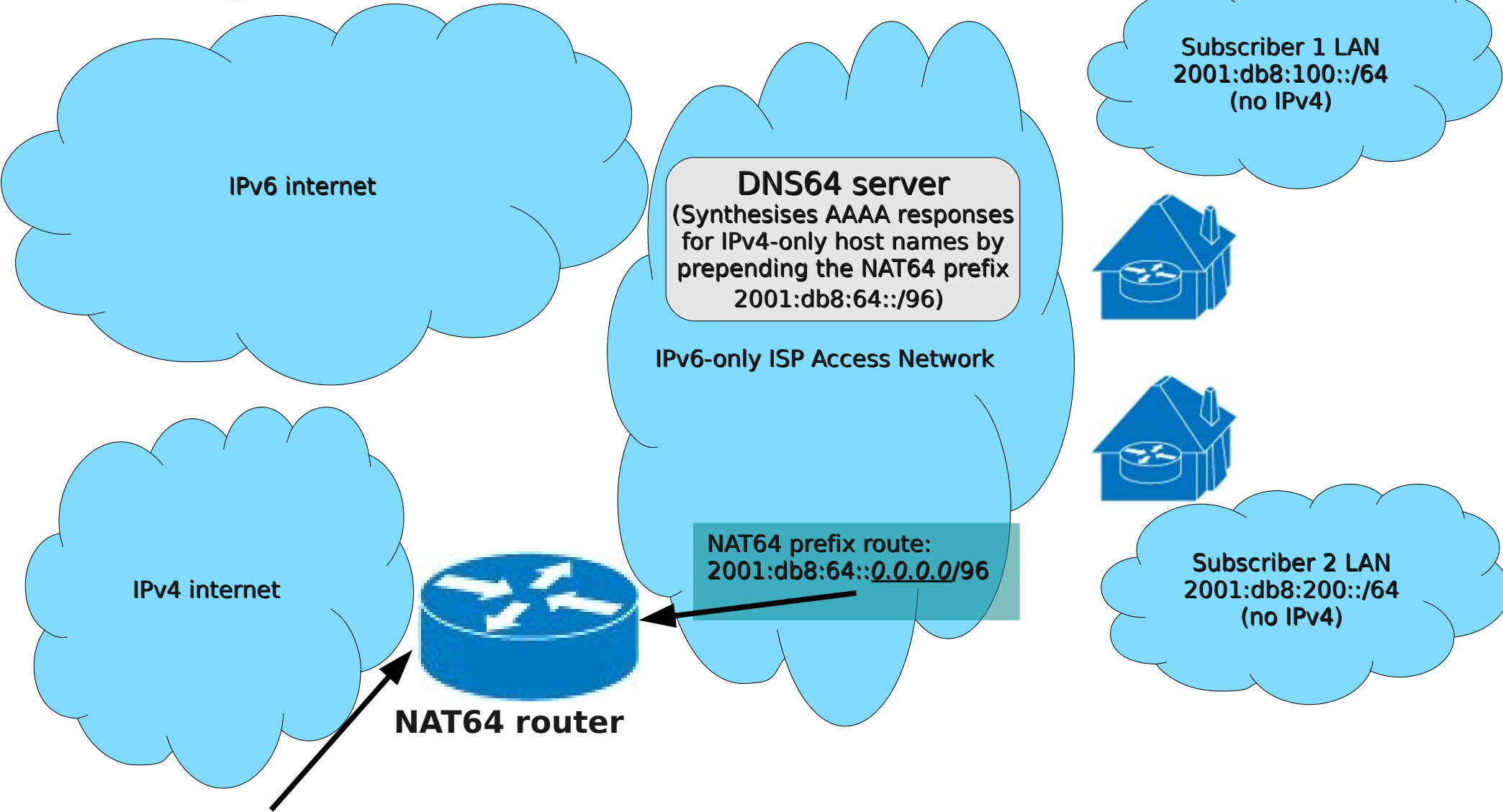
- The CPE's IPv4 prefix, address, or address+port-set is derived from its IPv6 address according to a set of mapping rules
 - 4RD-internal P2P traffic flows directly, does not have to pass the BR
- When sharing addresses, the CPE has a limited set of ports available
 - Ports 0-4095 are «more valuable» and therefore never assigned
 - 30k ports (2 users/address), 15k ports (4 u/a),, 1 port (60k u/a)
 - The CPE must translate the source ports of outbound connections
- End user is in control of the NAT44 function
 - May forward (assigned) ports as he wishes
 - Excessive connection initiation rates will only overload the user's CPE (just like with today's common NAT44-in-CPE deployments)

4RD disadvantages

- Tunnelling – overhead related to encap/decap and MTU
- No dynamic allocation of source ports possible (unlike DS-Lite/NAT444)
 - Different users require different number of source ports
 - Easiest solution: largest size fits all
 - Or: guesstimate individual users' requirements and allocate different-sized port ranges accordingly
 - Smaller number of users per IPv4 address possible
- Immature standard (more so than DS-Lite)
 - Early draft stage in the IETF
 - No implementations exist (beyond proof of concept quality)

NAT64/DNS64

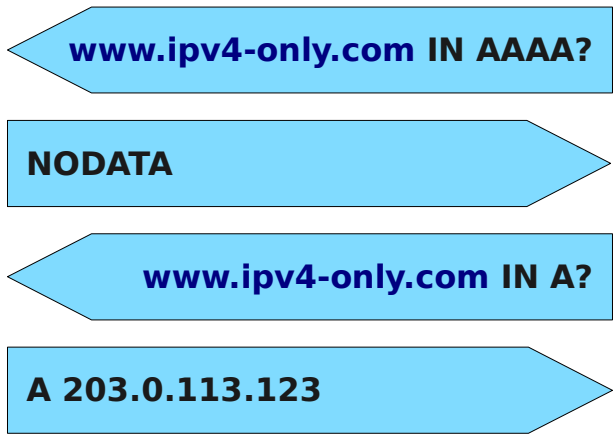
NAT64/DNS64: Map the IPv4 internet into an IPv6 prefix, translate between them, and rewrite DNS responses accordingly (+ native IPv6)



One or more public IPv4 prefixes are routed to the NAT64 router

NAT64/DNS64 packet flow

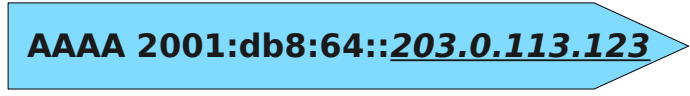
AuthNS
for
ipv4-only.
com



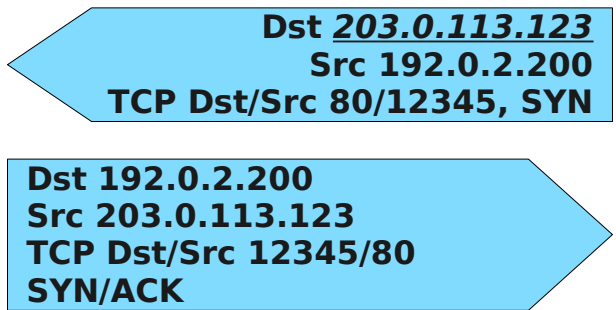
DNS64



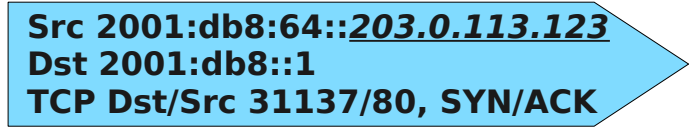
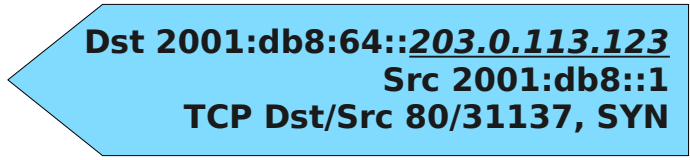
End
User
2001:db8::1



www.
ipv4-only.
com



NAT64
(NAT pool:
192.0.2/24)



NAT64 mappings:

- Static: 2001:db8:64::[32 last bits]/ANY
- Dynamic: 2001:db8::1/TCP/31337

- > [32 last bits]/ANY
- <---> 203.0.113.123/TCP/12345

NAT64/DNS64

- No shared state between NAT64 and DNS64 components
 - They must be configured with the same NAT64 prefix
- NAT64 is stateful, similar scaling issues as a DS-Lite AFTR
- May be placed anywhere in the network, unlike a NAT444 CGN
- Does not require any special CPE/host support
- IPv4-only clients/devices are not supported at all
 - Suitable only in controlled environments (enterprise, mobile, ...)
- Does not work with protocols/applications that do not (always) use DNS
 - Skype, `http://192.0.2.123/foo.html`, FTP, ... (ALGs may exist though)
- Mature standard, published RFCs
 - Production-quality implementations exist (hardware and opensource)
 - Pilot deployments by e.g. T-Mobile USA

Other proposals, and IPv4 address sharing concerns

Lurking on the IETF lists

- There's many proposals floating around in the IETF, for example:
 - dIVI – stateless NAT464
 - 6to4-PMT – 6to4 in CPE + stateless SP-operated NAT66
 - And many, many more...
- Port Control Protocol (PCP) – generic protocol for opening inbound ports
 - Replacment UPnP and NAT-PMP
 - Aims for compatibility with all types of NAT (both SP and CPE)
 - A CPE may proxy PCP/UPnP/NAT-PMP to PCP-capable SP NATs
 - Still at draft stage in the IETF

IPv4 address sharing will suck

- Data retention directive compliance w/on-demand source port allocation
- Abuse handling, DNS-blacklisting ... who's the culprit?
- Users can't expect to get specific ports forwarded to them
 - No more home web/ssh/etc servers
 - Protocols that require specific source ports break (e.g. IKE)
 - UPnP v1 breaks as it can only request specific ports
- Port-less protocols will generally break (e.g. ESP)
- Geo-location will break
- Solutions using tunnelling may require PMTUD or fragmentation
 - Raise the MTU in the SP network to allow for inner MTU ≥ 1500
- RFC 6269 «Issues with IP address sharing» is a must-read

A few observations on IPv6 security

To firewall or not to firewall...

- Commonly heard arguments FOR:
 - A host/device receiving inbound unsolicited connections is a sitting duck for attackers or internet worms
 - CPE-based NAT44 prevents inbound connections by default, so there is an implied expectation that this will be continued in IPv6
- Commonly heard arguments AGAINST:
 - Restoration of the end-to-end principle, encouraging innovation
 - NAT44 drops inbound connections out of necessity, it isn't and never was intended to be a security-enhancing feature
 - The sitting ducks were Win98 and the like, do not in support IPv6 in any case – modern OSes that support IPv6 are hardened by default
 - People are mobile and connect to airport and café networks anyway
 - Some ISPs don't provide CPEs and thus no NAT44 - no problem

Issues with 64-bits subnets

- Packet looping on point-to-point links that do not use ICMPv6 ND

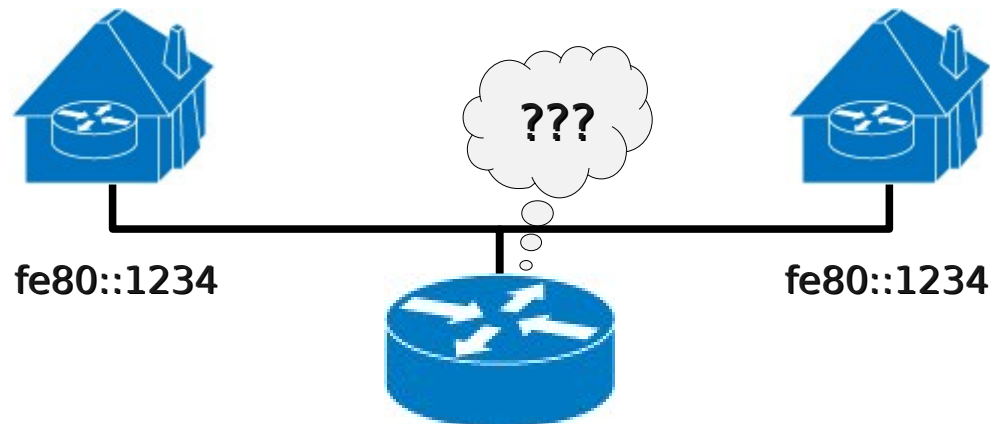


- >100x traffic amplification; <100Mb needed to fill 10Gb
- Use /127s instead, or filter all unused addresses on the link
- ICMPv6 ND exhaustion DoS
 - Sweeping a /64 results in incomplete/failed neigh entries on router
 - Make sure your preferred vendor handles this gracefully
 - /127s can be used within SP networks, or filter unused addresses

- An EUI-64 address discloses the end users' Ethernet MAC address
 - Allows for tracking a user as he moves between networks
- With SLAAC the host may pick any address instead of (or in addition to) using EUI-64, provided DAD for the selected address completes
 - RFC 4941 Privacy Extensions for SLAAC exploits this by randomly generating new temporary addresses once in a while
 - Default on in Windows and some Linux distros
 - Disliked by network admins in enterprises and similar environments

Layer 2 security in N:1 SP networks

- DHCPv6's Option 37 (equivalent to DHCPv4 Option 82) may only be inserted by layer-3 routers, not layer-2 bridges
 - RFC 6221 fixes this, but no implementations yet as far as I know
- DHCPv6/ND Snooping + IPv6 Source Guard? Practically non-existent
 - In any case you can't know the link-local WAN address of the CE
 - Nor any SLAAC-assigned global address



Layer 2 security in N:1 SP networks

- CE-to-CE attacks include:
 - Rogue RA – customer transmitting Router Advertisements to others
 - MS Windows w/Internet Connection Sharing **LOVES** to do this
 - RA Guard prevents this, available on some switches
 - DAD DoS – Evil customer preventing others from completing DAD
- Private VLANs and off-link prefixes may help, but not against spoofing
- SeND (Secure Neighbor Discovery) protects against ICMPv6 ND attacks
 - Requires certificates to be distributed with CPE
- Other solutions:
 - PPPoE (yuk!)
 - 1:1 VLAN model...

My closet



WAN Interface Edit

General

Name :

Type :

Mode :

WANServiceType :

IPv6/IPv4 Mode:

VLAN

Enable VLAN :

Enter 802.1P Priority [0-7] :

Enter 802.1Q VLAN ID [1-4094] : (3900 ~ 3905 are reserved.)

MTU

MTU

WAN Interface Edit

DNS Server

Obtain DNS info Automatically

Use the following Static DNS IP Address

IPv6 Address

Obtain IPv6 Address / Prefix Automatically

Enable Non-temporary Addresses

Enable Prefix Delegation

Static IPv6 Address

IPv6 DNS Server

Obtain IPv6 DNS info Automatically

Use the following Static DNS IPv6 Address

LAN IPv6 address Setup

- Delegate prefix from WAN
- Static Configuration

VdsIWAN1/ptm0.3900 (2001:840:3033:20::/60) ▼

ULA IPv6 address Setup

ULA IPv6 Address :

fd00::1

LAN IPv6 Address assign Setup

- Stateless + DNS send by RADVD (DHCPv6 server disable)
- Stateless + DNS send by DHCPv6 (DHCPv6 server enable)
- Stateful + DHCPv6 Server (DHCPv6 server enable)
- Stateful + DHCPv6 Relay (DHCPv6 Relay enable)

DHCPv6 Configuration

DHCPv6 State :

DHCPv6 Server

IPv6 DNS Values

IPv6 DNS Server 1

From ISP ▼

2001:840:100::

IPv6 DNS Server 2

From ISP ▼

2001:840:200::

IPv6 DNS Server 3

From ISP ▼

IPv6 Router Advertisement State

RADVD:

Enable

Apply

Cancel

Firewall

General Services Access Control DoS **IPv6 General** IPv6 Services

The firewall blocks unauthorized accesses to your IPv6 network.

IPv6 Firewall Enable Disable

Packet Direction	Default Action
WAN to LAN	Drop
LAN to WAN	Permit

Apply Cancel

Firewall

General Services Access Control DoS IPv6 General **IPv6 Services**

If you want to stop certain Internet services (e.g. instant messaging or downloading files from FTP sites), enable LAN-to-WAN Services Blocking and add them to Blocked Services.

Packet Direction : WAN to LAN ▾

Add new rule

#	Status	Source IP Address	Destination IP Address	Services	Action	Modify
1	💡	Any IP	2001:840:3033:20:208:a1ff:fe...	SSH	Permit	✎ 🗑
2	💡	Any IP	2001:840:3033:20:208:a1ff:fe...	TCP/UDP(62524~62524)	Permit	✎ 🗑

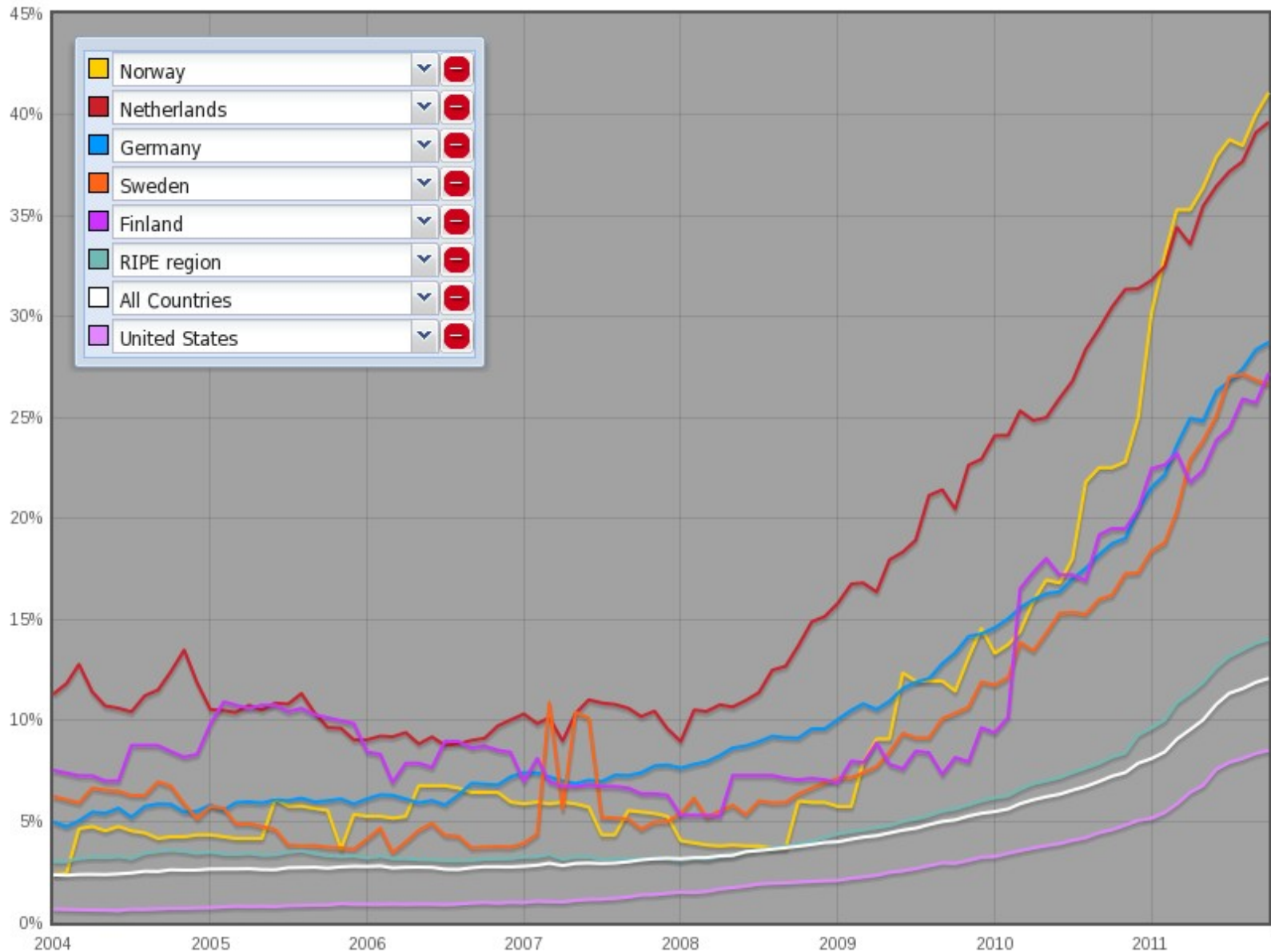
Clear All

ZyXEL P-2812HNU-F3

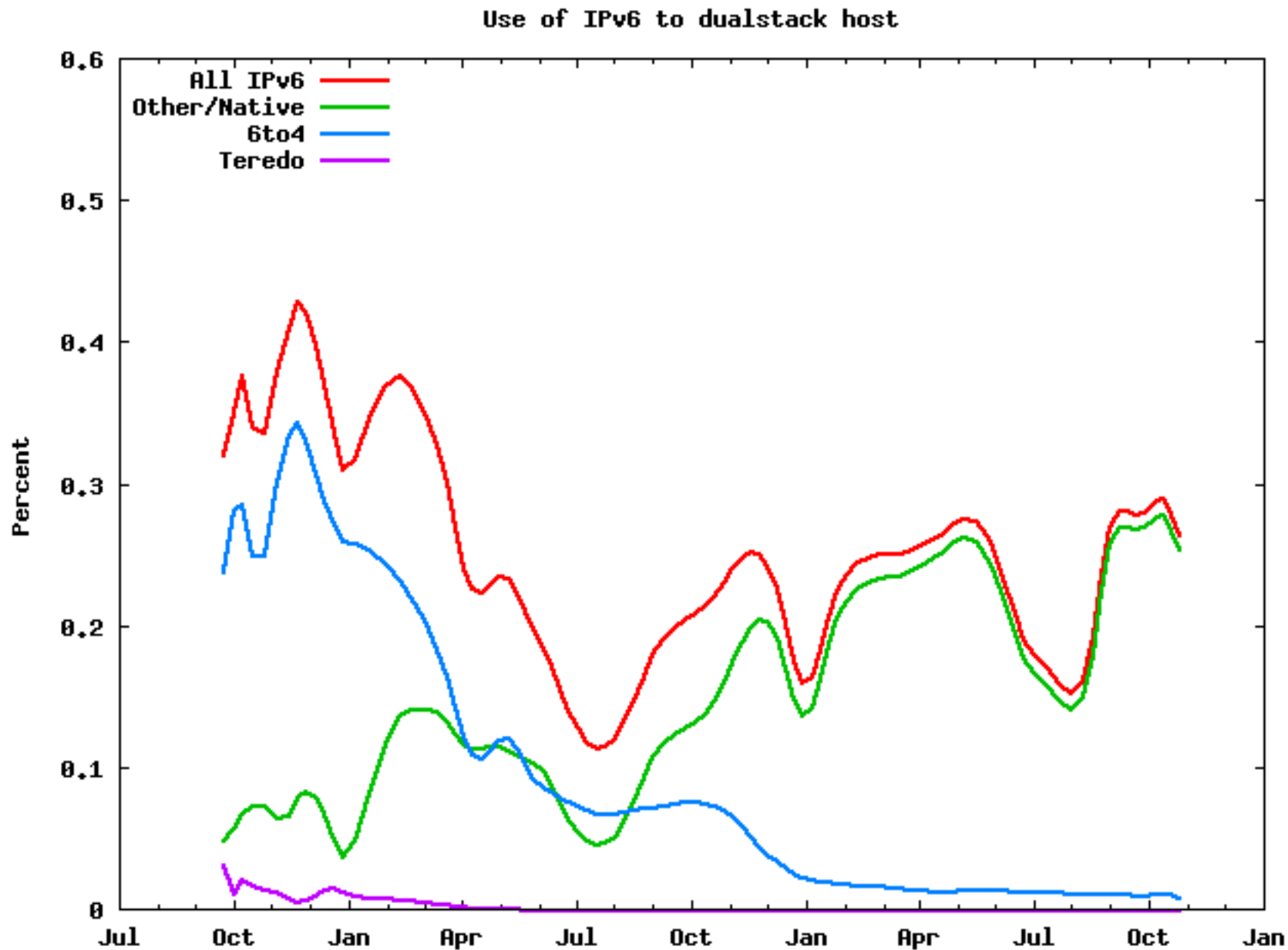
System Info

Device Information	
Host Name:	router
Model Name:	P-2812HNU-F3
MAC Address:	c8:6c:87:ab:da:5f
Firmware Version:	V3.11(TUS.0)
WAN 1 Information:	(VDSL WAN 1)
- Mode:	IPoE
- IP Address:	77.40.195.250
- IP Subnet Mask:	255.255.255.192
- Global IPv6 Address:	2001:840:330:3148:ca6c:87ff:feab:da5...
- Link-Local IPv6 Address:	fe80::ca6c:87ff:feab:da58/64
LAN Information:	
- IP Address:	192.168.1.1
- IP Subnet Mask:	255.255.255.0
- DHCP Server:	Server
- Global IPv6 Address:	2001:840:3033:20:ca6c:87ff:feab:da5f/60
- Link-Local IPv6 Address:	fe80::ca6c:87ff:feab:da5f/64
- DHCPv6 Server:	Server

ISPs announcing IPv6 prefixes



End-users in Norway with IPv6





[Forsiden](#) [IPv6 konferansen](#) [Konferanseprogram](#) [Samarbeidspartnere](#) [Presentasjoner og foredrag](#) [Om IPv6 forum](#)

IPv6 konferanse i Stavanger 21 og 22 november

Posted on [21.09.2011](#) by [Torgeir Waterhouse](#)

Neste IPv6 konferanse arrangeres i Stavanger 21 og 22 november – sett av datoene nå!

Tema for konferansen blir IPv6 og: innkjøp, investering og innovasjon.

Programmet ferdigstilles nå, og vil bli publisert her så snart det er klart, annen informasjon om konferansen publiseres fortløpende her.

Alle spørsmål og henvendelser sendes til: post@ip6forum.no



Sign up for our mailing list.

First Name :

Questions?