Bachelorscriptie
Informatica

Radboud Universiteit

# De oorzaken en gevolgen van het uitstellen van de IPv6-adoptie

*Auteur:*
David Korsman
s4619579

*Inhoudelijk begeleider:*
Dr. Ir. H.P.E. Vranken
harald.vranken@ou.nl

*Tweede begeleider:*
Prof. Dr. Marko van Eekelen
marko@cs.ru.nl

27 maart 2019

**Samenvatting**

Ondanks het uitputten van de voorraad IPv4-adressen, lijkt de adoptie van de langetermijnoplossing, IPv6, nog traag te verlopen. Met het huidige tempo lijkt het tot 2024 te duren voordat de helft van het internetverkeer via IPv6 verloopt. We beschrijven verschillende factoren die de adoptie vertragen of versnellen. Zo zijn aanpassingen nodig aan bestaande netwerken en kan IPv4 worden blijven gebruikt, zelfs bij een tekort aan IPv4-adressen. Ook beschrijven we oplossingen die internetproviders kunnen toepassen om toegang te blijven bieden tot het deel van het internet dat alleen met IPv4 werkt. Een van de mogelijke oplossingen is het kopen van IPv4-adressen. Waar die optie het tekort niet oplost kan een provider meerdere klanten hetzelfde publieke IP-adres laten delen, hoewel de technieken die dat bewerkstelligen nadelen hebben.

# Inhoudsopgave


**1 Introductie** **2**

**2 Achtergrond** **3**
2.1 Geschiedenis . . . . . 3
2.2 Toewijzing van IP-adressen . . . . . 6
2.3 IPv6 . . . . . 6
2.4 DNS . . . . . 7

**3 Onderzoek** **8**
3.1 Stand van zaken . . . . . 8
3.1.1 Huidige IPv6-adoptie . . . . . 8
3.1.2 Schatting van het vervolg . . . . . 12
3.2 Vertragende en versnellende factoren . . . . . 16
3.2.1 Vertragende factoren . . . . . 16
3.2.2 Versnellende factoren . . . . . 19
3.3 Oplossingen . . . . . 20
3.3.1 IPv4-adressen kopen . . . . . 21
3.3.2 (Carrier-Grade) NAT . . . . . 21
3.3.3 NAT64 . . . . . 23

**4 Gerelateerd werk** **25**

**5 Conclusie en discussie** **26**
5.1 Conclusie . . . . . 26
5.2 Discussie . . . . . 27
5.3 Vervolgonderzoek . . . . . 27

# Hoofdstuk 1

# Introductie

Op het internet zijn unieke adressen nodig om verkeer te kunnen verzenden en ontvangen. Het hedendaagse internet is nog steeds grotendeels gebaseerd op versie 4 van het Internet Protocol (IPv4), dat sinds 1980 in gebruik is. Voor IPv4 is gekozen voor een adresformaat van 32 bits, en heeft dus een theoretisch maximum van $2^{32} = 4\,294\,967\,296$ adressen. Door de sterke groei van het internet in de afgelopen decennia is dat echter te weinig gebleken. Er bestaan verschillende oplossingen om, ondanks het tekort aan IPv4-adressen, het internet door te laten werken zoals het al jaren deed. Een van die oplossingen is het gebruikmaken van versie 6 van het Internet Protocol, IPv6. IP-adressen zijn in deze versie 128 bits lang, dus IPv6 biedt een adresruimte van $2^{128}$ adressen, wat de benodigde hoeveelheid IP-adressen met zekerheid overtreft. Het gebruik van IPv6 stijgt, maar de adoptie verloopt traag en het overgrote deel van het internetverkeer verloopt nog altijd via IPv4. Dit onderzoek richt zich op de hoofdvraag wat de oorzaken en gevolgen zijn van het uitstellen van de IPv6-adoptie. De deelvragen zijn:

- Wat is de huidige IPv6-adoptie, en wat is een redelijke schatting van het vervolg?
- Wat zijn de factoren die de adoptie beïnvloeden?
- Wat voor oplossingen bestaan er voor de uitputting van IPv4-adressen, en wat hebben die voor gevolgen?

Deze vragen zullen worden beantwoord door middel van literatuuronderzoek.

De opbouw van de rest van deze scriptie is als volgt: In hoofdstuk 2 wordt achtergrondinformatie gegeven. In hoofdstuk 3 worden de onderzoeksvragen beantwoord. Hoofdstuk 4 bespreekt gerelateerd werk. Hoofdstuk 5 omvat de conclusie en discussie, en kijkt vooruit naar mogelijk vervolgonderzoek.

# Hoofdstuk 2

# Achtergrond

## 2.1 Geschiedenis

In 1980 was de oorspronkelijke gedachte achter het internet dat iedere computer die ermee verbonden was een eigen numeriek adres zou hebben, zodat verkeer dat voor die computer bestemd was deze kon bereiken, aan de hand van alleen het adres. Zo moest er ook op internetverkeer kunnen worden geantwoord – het heeft geen zin om informatie op te vragen als die informatie niet bij de opvrager bezorgd kan worden. Elke computer, zonder onderscheid te maken tussen computers die informatie aanbieden en computers die alleen informatie opvragen, moest dus een uniek adres hebben om met andere computers te kunnen communiceren.

Het internet bestaat uit vele verschillende netwerken, de *autonomous systems* (AS), die met elkaar informatie uitwisselen over mogelijke routes naar andere delen van het internet, oftewel langs welke paden bepaalde IP-adressen bereikt kunnen worden. Via die routes kunnen in principe alle computers die zijn aangesloten op het internet verkeer naar elkaar sturen aan de hand van elkaars IP-adres. Voor IPv4 is gekozen voor een adresformaat van 32 bits, bestaande uit acht bits voor het nummer dat aan het netwerk was toegewezen (toen nog door één persoon, Jon Postel), gevolgd door drie groepen van acht bits voor lokale adressen [RFC0760; RFC0762]. Het laatste deel van een IP-adres zou dus staan voor een specifieke computer binnen dat netwerk. Aangezien de netwerken 0 en 255 gereserveerd waren zou er ruimte zijn voor 254 netwerken met elk ruimte voor $2^{24} = 16\,777\,216$ adressen.

In 1981 werd de adresruimte opnieuw ingedeeld en werd gesproken over drie klasses van netwerken, A, B en C, omdat er meer flexibiliteit nodig was in het toekennen van IP-adressen [RFC0791]. Voor netwerken die, zoals in het oude systeem, bestonden uit $2^{24}$ adressen werd nog maar de helft van de adresruimte gereserveerd, alle adressen met als eerste bit een `0`. Dit waren 127 netwerken van klasse A. Klasse B bestond uit adressen die begonnen met de bits `10`, had ruimte voor $2^{14} = 16\,384$ netwerken van $2^{16} = 65\,536$

adressen, en klasse C begon met de bits `110` en had ruimte voor $2^{21} = 2\,097\,152$ netwerken van 256 adressen. Een netwerk van klasse B kon er dus uitzien als 150.2.x.x, een netwerk van klasse C als 195.2.3.x. Zie tabel 2.1 voor de indeling van de klasses.

| | | | | |
|---|---|---|---|---|
| Klasse A | `0NNNNNNN` | `********` | `********` | `********` |
| Klasse B | `10NNNNNN` | `NNNNNNNN` | `********` | `********` |
| Klasse C | `110NNNNN` | `NNNNNNNN` | `NNNNNNNN` | `********` |

Tabel 2.1: Klassen A, B en C in binair formaat. '`N`'-bits worden gebruikt voor de identificatie van het netwerk, '`*`'-bits kunnen worden gebruikt voor alle hosts in het netwerk.

Het voordeel van het klassensysteem ten opzichte van het eerdere systeem was dat er niet slechts maximaal 254 netwerken konden bestaan, maar maximaal $127 + 2^{14} + 2^{21} = 2\,113\,663$. De meeste netwerken hadden ook significant minder adressen nodig dan $2^{24}$, dus er werd op deze manier efficiënter gebruikgemaakt van de adresruimte.

Begin jaren negentig, toen het internet nog relatief onbekend was, werd al voorspeld dat het systeem met klassen niet voldoende zou zijn om te voorkomen dat de adressen binnen een paar jaar op zouden raken [IETF90]. Daarom werd als tijdelijke oplossing afgestapt van het klassensysteem, met Classless Inter-Domain Routing (CIDR) [RFC1519]. Met CIDR kon de splitsing van het IP-adres in het gedeelte voor het netwerk en het lokale gedeelte zich tussen elke twee bits bevinden, in plaats van op drie vooraf bepaalde punten. Dit bood ook een oplossing voor de 'verspilling' van IP-adressen die nog altijd plaatsvond met het klassensysteem: een bedrijf dat bijvoorbeeld 300 IP-adressen nodig had, had niet genoeg aan een netwerk van klasse C. De enige optie zou dan klasse B zijn, wat juist veel te veel adressen bood. Met CIDR kon het netwerk dan bijvoorbeeld bestaan uit $2^9 = 512$ adressen. CIDR zou een paar jaar extra tijd creëren en in de tussentijd kon worden gewerkt aan een permanente oplossing.

Er kwam echter meer extra tijd dan die paar jaar. Er werd nog een tijdelijke oplossing ontwikkeld die ingezet zou kunnen worden als CIDR niet genoeg tijd zou winnen: Network Address Translation (NAT) [RFC3022]. Die oplossing zou wel problemen met zich meebrengen – er werd een fundamenteel principe van het IP-protocol geschonden, namelijk dat iedere computer een eigen IP-adres moest hebben. Het was dan ook bedoeld als een tijdelijke oplossing. Aanvankelijk hield NAT (Basic NAT) in dat een netwerk een kleiner aantal IP-adressen (een kleinere 'pool') kreeg dan er hosts in dat netwerk aanwezig waren. De hosts binnen het netwerk hadden (privé-)IPv4-adressen, in een bereik dat in [RFC1918] is vastgelegd. Een router in dat netwerk kon dienen als Network Address Translator, en wanneer een host in het lokale netwerk verbinding wilde maken met een host buiten dat netwerk (waarbij

het verkeer dus langs de NAT-router moet), werd een van de adressen uit de pool met globale IP-adressen gebruikt. Het aantal hosts in het netwerk dat tegelijkertijd toegang had tot externe netwerken was dus beperkt tot het aantal globale adressen dat in de pool beschikbaar was. Dit principe kan ook worden toegepast met een pool van 1 globaal IP-adres en Network Address Port Translation (NAPT)[1]. Met traditionele NAT zou dan te allen tijde slechts één computer in het netwerk tegelijkertijd toegang hebben tot externe netwerken, maar een NAT-router kan ook verbindingen van meerdere hosts via hetzelfde IP-adres laten lopen door beter bij te houden welke verbindingen gemaakt worden. Wanneer een host in het lokale netwerk verbinding wil maken met een host buiten dat netwerk, waarbij het verkeer dus langs de NAT-router moet, herschrijft deze het bron-IP-adres in pakketten naar het publieke IP-adres om te zorgen dat een eventueel antwoord weer bij zichzelf bezorgd kan worden, want de LAN-IP-adressen zijn op het internet niet routeerbaar. Antwoordpakketten zullen ook worden herschreven door dezelfde *Translator* omdat het WAN-IP-adres (het publieke IP-adres) in het lokale netwerk niet het adres is van de host die de oorspronkelijke verbinding tot stand bracht. De NAT-router moet hier dus onthouden vanaf welke IP-adressen in het LAN aanvragen worden gestuurd naar IP-adressen buitenaf. Ook moet de router een poort kiezen aan de 'buitenkant' waarvandaan deze de aanvraag verstuurt en de gebruikte poorten moeten worden onthouden – meerdere hosts achter een NAT kunnen immers aanvragen sturen naar hetzelfde IP-adres. Wanneer van buitenaf een pakket wordt beantwoord, kan de NAT aan de hand van het poortnummer bepalen naar welke lokale host het pakket moet worden doorgestuurd. De poort waarvandaan de oorspronkelijke aanvraag is verstuurd wordt in het antwoordpakket namelijk gebruikt als doelpoort. De NAT kan dus niet voor meerdere verschillende gelijktijdige aanvragen hetzelfde bronpoortnummer gebruiken.

Over het algemeen kan NAT dus worden beschouwd als *stateful*, zeker wanneer ook de poortnummers moeten worden bijgehouden – er is in dit systeem een router die een administratie moet bijhouden over elke verbinding die tot stand wordt gebracht.

Rond 2000 stagneerde de vraag naar IP-adressen en was NAT alomtegenwoordig, en de komende decennia leek de kleine adresruimte van IPv4 geen probleem meer te zullen vormen. [Ric+15] Toch bleek daarna het internet sneller te groeien dan gedacht. Steeds meer mobiele telefoons werkten met GPRS en EDGE, en later UMTS, en werden via het mobiele netwerk aangesloten op het internet. Daardoor groeide de behoefte aan IP-adressen en kwam de uitputting van IPv4 weer dichterbij. Het mobiele netwerk bood niet alleen mogelijkheden voor telefoons, maar ook voor Internet of Things-toepassingen.

[1] Ook NAPT wordt veelal aangeduid als NAT, dus we zullen in deze scriptie Basic NAT als zodanig aanduiden, en NAPT als NAT, behalve waar anders wordt aangegeven.

Er moest dus gezocht worden naar nog meer kortetermijnoplossingen. Internetproviders kunnen bijvoorbeeld ook NAT toepassen buiten de netwerken van klanten, waarmee meerdere klanten dus een IP-adres moeten delen. Dit is Carrier-Grade NAT (CGN of CGNAT), hier zal later dieper op in worden gegaan.

## 2.2 Toewijzing van IP-adressen

Adressen werden, naarmate de tijd vorderde, steeds minder informeel toegewezen. Zoals eerder genoemd werden netwerknummers aanvankelijk alleen door Jon Postel toegewezen op aanvraag. In 1988 richtte hij de Internet Assigned Numbers Authority (IANA) op. Deze was verantwoordelijk voor het toewijzen van netwerknummers, en werd nog steeds door Jon Postel beheerd. In 1998 werd de Internet Corporation for Assigned Names and Numbers (ICANN) opgericht, die de functie van IANA vanaf toen vervulde, waarmee IANA een tak van ICANN werd.

Verder werden regionale autoriteiten opgericht, de Regional Internet Registries (RIR's), die als doel hadden om netwerknummers toe te wijzen in hun eigen regio's. De eerste daarvan waren Réseaux IP Européens (RIPE) voor Europa in 1992, Asia-Pacific Network Information Centre (APNIC) in 1993, en American Registry for Internet Numbers (ARIN) in 1997. In 2002 en 2004 werden de laatste twee RIR's gevormd: Latin America and Caribbean Network Information Centre (LACNIC) en AFRINIC voor Afrika. IANA wijst blokken IP-adressen toe aan de RIR's, die vervolgens door de RIR's weer (in kleinere blokken) toegewezen kunnen worden aan Local Internet Registries (LIR's). LIR's zijn organisaties zoals internetproviders die de IP-adressen vervolgens kunnen gebruiken of toewijzen aan klanten. [Ric+15]

In februari 2011 had IANA nog vijf /8-blokken[2] over. Deze zijn gelijktijdig en gelijkmatig toegewezen aan de vijf RIR's. Zodra een RIR aan deze laatste /8 is toegekomen, wordt de voorraad van die RIR gezien als 'uitgeput', en treedt daarvoor een beleid in werking dat de toewijzing van deze IP-adressen beperkt. De voorraad van APNIC raakte als eerste uitgeput, in april 2011. In september 2012 gebeurde hetzelfde bij RIPE, in juni 2014 bij LACNIC, in september 2015 bij ARIN en tenslotte in april 2017 bij AFRINIC.[Ric+15; ARI15; AFR17]

## 2.3 IPv6

De permanente oplossing, die al in 1995 klaarlag, is IPv6 [RFC8200]. Het belangrijkste verschil tussen IPv4 en IPv6 is dat IPv6 werkt met adressen van 128 bits. De eerste 64 bits zijn in principe voor het toegewezen 'nummer' van het netwerk, hoewel dat een subnet kan zijn van een groter netwerk,

[2] Blokken waarvan het vaste netwerkdeel uit 8 bits bestaat, dus $2^{24}$ IPv4-adressen

waarvan het nummer het 'prefix' is. Een netwerkprefix kan dus bijvoorbeeld 40 bits lang zijn, waarmee 24 bits overblijven voor subnetten binnen dat netwerk. De laatste 64 bits kunnen door de hosts in het netwerk worden gekozen. Het is daarmee in de standaard niet mogelijk om, in het theoretische geval van een tekort aan IPv6-adressen, netwerken met langere prefixen te gebruiken, omdat de laatste 64 bits altijd moeten kunnen worden gekozen door de hosts zelf.

Hoewel IPv6 al geruime tijd een standaard is, is het gebruik ervan nog vrij laag, terwijl er in veel gevallen al geen IPv4-adressen meer uit te geven zijn.

IPv6 kan niet zonder meer worden gebruikt als niet zowel de bron als bestemming van een verbinding het ondersteunen. Zo kan een website die alleen een IPv4-adres heeft niet via IPv6 worden bezocht, alleen al omdat een IPv6-adres niet in een IPv4-header past - er is in de header van het IPv4-protocol ruimte voor 32-bits voor een IP-adres, en dat moet dan ook een IPv4-adres zijn.

## 2.4 DNS

Omdat IP-adressen niet gemakkelijk te onthouden zijn, is het Domain Name System (DNS) ontworpen [RFC1034; RFC1035]. Een gebruiker kan, bijvoorbeeld in een webbrowser, een domeinnaam zoals `example.com` gebruiken, waarmee via DNS een bijbehorend IP-adres kan worden opgezocht. Er bestaan onder andere `A`-records die IPv4-adressen bevatten en `AAAA`-records die IPv6-adressen bevatten. Als een domeinnaam via DNS wordt opgezocht, kunnen beide typen records worden gevonden, of een van beide. Als een DNS-client (waaronder een browser) zowel een `A`-record als een `AAAA`-record vindt en de host ondersteunt beide protocollen, kan deze kiezen via welk protocol (IPv4 of IPv6) verbinding met de website gemaakt wordt – mits dit protocol ook door het lokale netwerk wordt ondersteund. Indien slechts één van de typen DNS-records aanwezig is kan slechts één protocol worden geprobeerd, maar als dat protocol niet wordt ondersteund door de host of het lokale netwerk, zal niet zonder meer verbinding kunnen worden gemaakt.

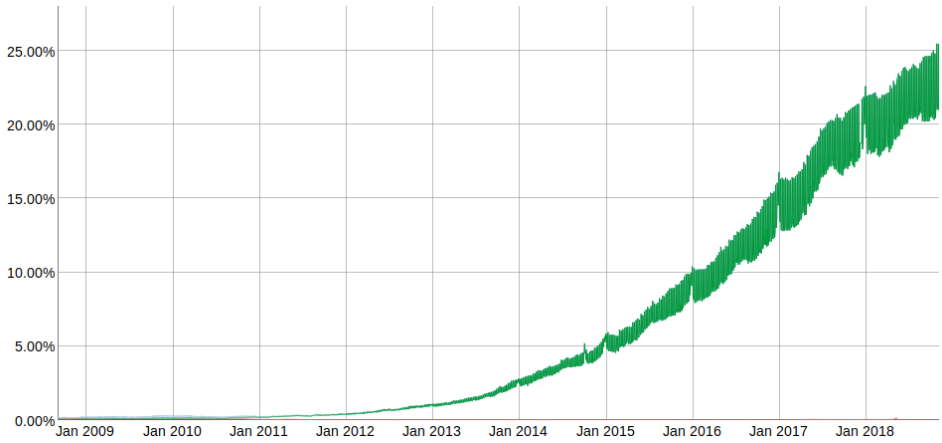
# Hoofdstuk 3

# Onderzoek

## 3.1 Stand van zaken

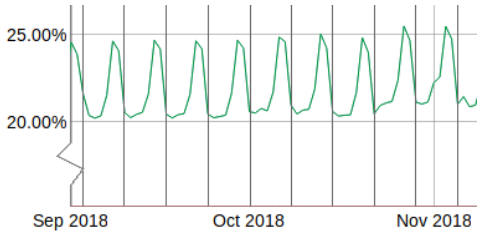
### 3.1.1 Huidige IPv6-adoptie

Er wordt door verschillende bedrijven op verschillende manieren data bijgehouden over het gebruik van IPv6 vergeleken met dat van IPv4. Google houdt sinds september 2008 bij welk percentage van haar gebruikers een internetverbinding gebruikt met toegang tot IPv6 [1] (figuur 3.1). Uit die gegevens blijkt dat dit in september 2008 nog onder 0,20% lag. Tien jaar later, in september 2018, ligt dit percentage tussen 20% en 25%.

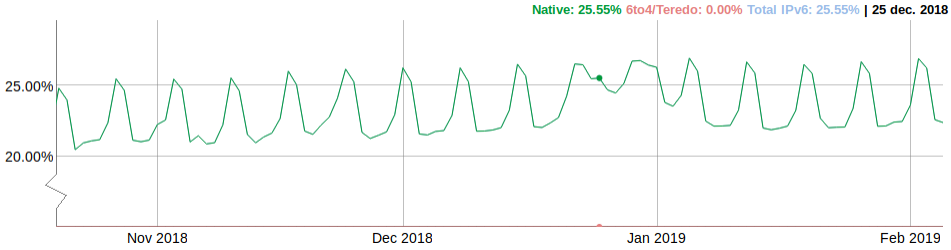



Figuur 3.1: De wereldwijde IPv6-adoptie per dag volgens Google, van september 2008 tot 6 november 2018

[1] https://www.google.com/intl/en/ipv6/statistics.html




Figuur 3.2: Een uitvergroot deel van de grafiek van Google. De maandagen zijn met verticale lijnen aangegeven.




Figuur 3.3: 25 december 2018 in de grafiek van Google, met een stip gemarkeerd, bij wijze van voorbeeld dat het gebruikelijke patroon tijdens kerst wordt onderbroken

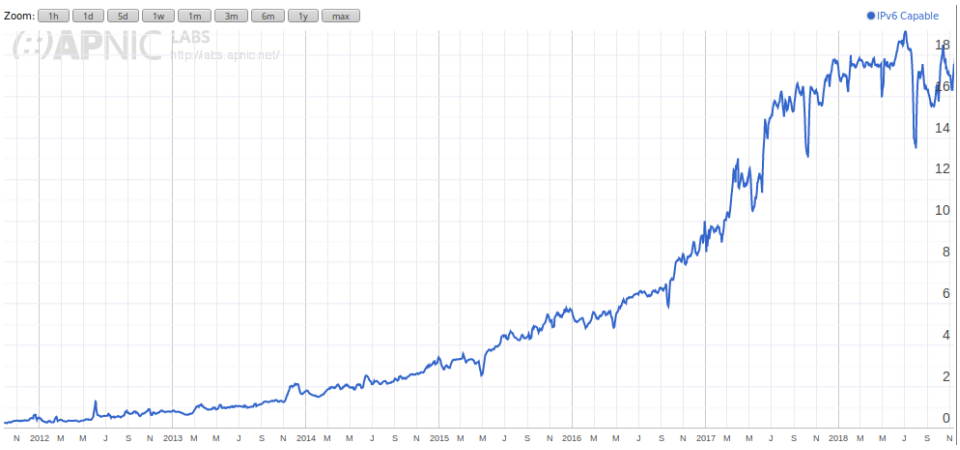
Google verzamelt deze data door aan een kleine hoeveelheid van de Google-zoekresultaten een JavaScript-fragment toe te voegen dat de browser een bepaalde pagina doet opvragen. In 10% van de gevallen is deze pagina *IPv4-only* door een domein te gebruiken waarvoor alleen een `A`-DNS-record kan worden gevonden, in de overige 90% is deze over zowel IPv4 als IPv6 te bereiken, en kan gekeken worden of de browser verbinding maakt via IPv6 of niet. Deze methode gebruikt Google om te bepalen hoe groot de beschikbaarheid van IPv6 is bij haar gebruikers. Hierbij moet worden opgemerkt dat het IPv6-gebruik per land verschilt, en Google zal niet in elk land evenveel gebruikt worden (Google is bijvoorbeeld geblokkeerd in China[2]) dus deze data geeft een indicatie en geen exact beeld. [Col+10]

In de grafiek van Google valt een aantal dingen op. De grafiek schommelt op en neer (zichtbaar in figuur 3.2), waarbij er nagenoeg altijd een duidelijke piek is in het percentage IPv6-gebruikers op zaterdagen en zondagen. Verder wordt dit patroon elk jaar rond 25 december verbroken en zijn de lokale minima in de dagen erna aanzienlijk minder laag dan in de weken ervoor en erna (figuur 3.3). Een waarschijnlijke interpretatie van deze patronen is dat een hoger percentage van de internetverbindingen thuis IPv6 heeft dan de internetverbindingen "op het werk", omdat werknemers in weekeindes en tijdens kerst waarschijnlijk thuis zullen zijn. In de publicatie van Google

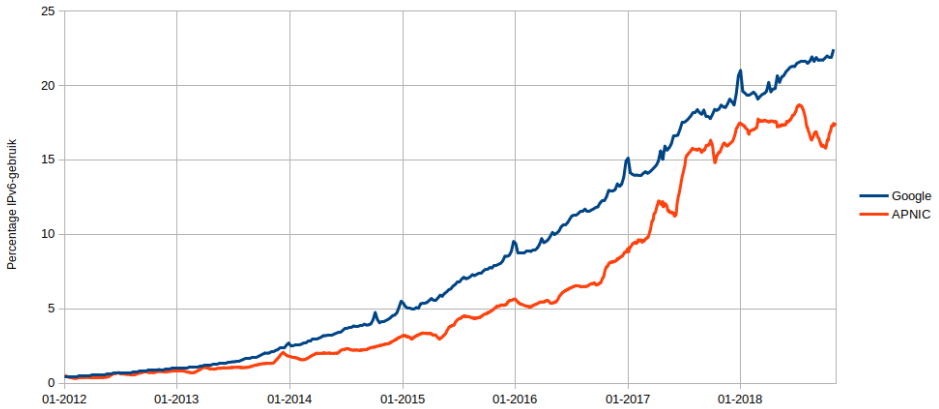
[2] `https://transparencyreport.google.com/traffic/overview`

[Col+10] wordt die conclusie ook al getrokken aan de hand van de data van september 2009. De schommelingen kunnen worden weggenomen door per week het gemiddelde te nemen. Vanaf nu zullen we dus meetintervallen van een week hanteren.

APNIC, de Regional Internet Registry van Azië en Oceanië, verzamelt zelf ook data over het gebruik van IPv6 wereldwijd [3] (figuur 3.4). Deze data wordt verzameld door JavaScript-fragmenten en Flash-applets toe te voegen aan advertenties of op pagina's zelf, die, soortgelijk aan de methode van Google, de browser bepaalde pagina's doet opvragen. [Hus13]




Figuur 3.4: Het wereldwijde percentage IPv6-adoptie per week volgens APNIC Labs, van september 2011 tot 15 november 2018




Figuur 3.5: De data van Google en APNIC per week

De data van APNIC geeft aan dat IPv6 wereldwijd in slechts rond 17% van de verbindingen gebruikt wordt, wat dus lager uitvalt dan de data van

[3] `https://stats.labs.apnic.net/ipv6`

Google. In figuur 3.5 kunnen de datasets van Google als APNIC worden vergeleken. Te zien is dat de data van Google significant hoger uitvalt dan de data van APNIC. De data van Google valt dus te beschrijven als optimistischer.

Er zitten grenzen aan zowel de methode van Google als die van APNIC. Zo baseert Google haar data op de gebruikers die gebruikmaken van de websites van Google. Dit geeft een goede schatting voor de landen waar Google in voldoende mate gebruikt wordt, maar dat is niet in alle landen het geval. Wat dit verschil mogelijk zou verklaren is dat China slechts rond 3% IPv6-adoptie heeft (zowel in de data van Google als APNIC), veel lager dan het wereldwijde gemiddelde. Zoals eerder opgemerkt is de zoekmachine van Google geblokkeerd in China, terwijl de IPv6-data van Google is gebaseerd op zoekopdrachten in de zoekmachine [Col+10]. APNIC publiceert, in tegenstelling tot Google, uit hoeveel datapunten de metingen bestaan. Op dit moment zijn er wereldwijd in totaal 201 349 226 datapunten, waarvan 14 444 126 uit China (7%). Daarmee is China het derde land ter wereld met de meeste datapunten, na India en de Verenigde Staten. De hypothese is dat het verschil tussen APNIC en Google als volgt te verklaren is: het wereldwijd gemiddelde van APNIC wordt omlaag getrokken doordat meer dan 7% van de data uit China komt, waar het gemiddelde rond 3% ligt, en in het gemiddelde van Google is China bijna niet vertegenwoordigd, waardoor dat percentage hoger uitvalt.

Om die theorie te toetsen kan worden berekend wat het wereldwijde percentage volgens APNIC zou zijn als alle datapunten uit China daarin niet zouden worden meegeteld, om te proberen het percentage van Google te benaderen (de data van Google zou door de blokkade immers voor 0% uit China moeten komen). De verhouding tussen het aantal datapunten uit China en uit de hele wereld in de data van APNIC is $\frac{14\,444\,126}{201\,349\,226}$, dit noemen we $W_C$. Het wereldwijde percentage datapunten dat "IPv6 Capable" is, is volgens APNIC $17,68\%$, voor China is dit $3,06\%$. Dit noemen we respectievelijk $P_A$ en $P_C$.

Het wereldwijde percentage zonder de datapunten van China kan dan als volgt worden gedefinieerd:

$$\frac{P_A - P_C \times W_C}{1 - W_C}$$

Oftewel:

$$\frac{17{,}68 - 3{,}06(\frac{14\,444\,126}{201\,349\,226})}{1 - \frac{14\,444\,126}{201\,349\,226}} = 18{,}81\%$$

Het verwijderen van de data van China brengt het percentage van APNIC wel iets dichter bij dat van Google, maar 18,81% ligt nog steeds dichter bij de oorspronkelijke 17,68% dan bij zelfs de laagste waarden die Google meet, die al maanden boven de 20% liggen. China kan dus een van de redenen zijn dat de data afwijkt, maar het is niet de enige reden.

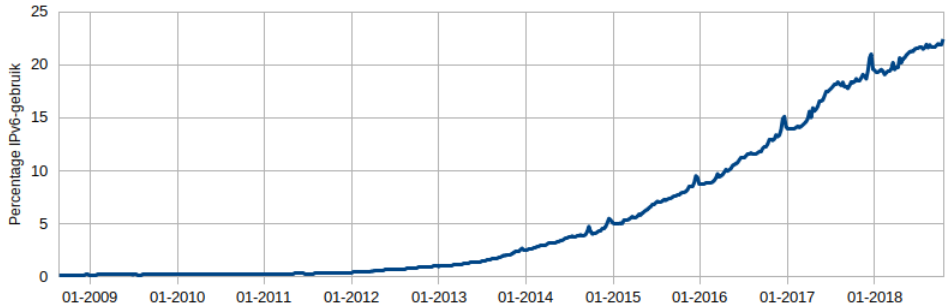
Mogelijke andere redenen zijn te vinden in de grenzen van de meetmethodes van Google en APNIC, waardoor mogelijk een bias in de meetresultaten ontstaat. Zowel de methoden van Google als APNIC werken niet als de browser de scripts die de meting uitvoeren blokkeert. Groepen internetgebruikers die een advertentieblokkeerder of NoScript gebruiken kunnen wellicht niet worden meegenomen in de telling. Ook kan het zijn dat bepaalde groepen internetgebruikers bijvoorbeeld de websites van Google mijden. Het is denkbaar dat er ook groepen gebruikers zijn die bewust kiezen voor een internetprovider die IPv6 aanbiedt, en tegelijkertijd bij een van de groepen horen die moeilijk door Google en APNIC te meten zijn.

Over het algemeen is het nauwkeurig meten van het gebruik van IPv6 niet makkelijk, omdat het internet niet gecentraliseerd is. Een mogelijke andere manier om aan gegevens te komen is om het te vragen aan internetproviders. In 2016 zijn Richter et al. door middel van een enquête van 75 internetproviders te weten gekomen wat hun voortgang was met betrekking tot IPv6 [Ric+16]. Daaruit bleek dat 32% van de providers IPv6 had uitgerold voor de meeste of alle klanten, en nog 35% voor een deel van de klanten. Slechts 11% van de providers had geen plannen om IPv6 uit te rollen. Deze resultaten lijken af te wijken van de datasets van zowel Google als APNIC, maar het uitrollen van IPv6 door een provider hoeft niet te betekenen dat er ook gebruik van wordt gemaakt. Ook hebben Richter et al. expliciet niet beweerd dat hun steekproef statistisch onbevooroordeeld is.

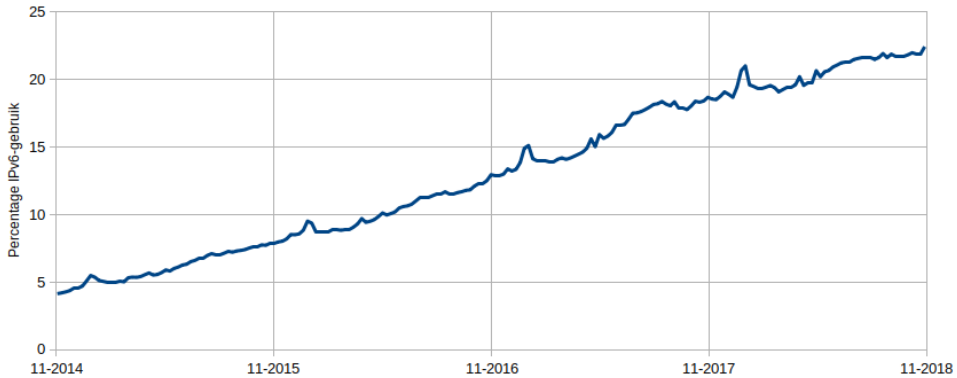
Vervolgonderzoek zou kunnen bestaan uit het meten van IPv6-ondersteuning op websites (bijvoorbeeld de populairste websites) en het bepalen van een theoretische bovengrens van IPv6-adoptie door te kijken hoeveel blokken IPv6-adressen tot nu toe zijn uitgegeven door IANA en de RIR's. Hier zal verder op worden ingegaan in sectie 5.3.

### 3.1.2 Schatting van het vervolg

Het is interessant om op basis van de beschikbare historische gegevens te voorspellen hoe de IPv6-adoptie zich verder zal ontwikkelen. We zullen een schatting maken op basis van de data van Google. Op basis van deze data kan een optimistische schatting gemaakt worden, met andere woorden: we willen benaderen op welk moment de IPv6-adoptie op zijn vroegst een bepaalde hoogste bereikt. Om dit makkelijker te maken zullen we de gegevens per week bekijken in plaats van per dag, om de schommelingen tussen doordeweekse dagen en weekeindes weg te werken (figuur 3.6 en 3.7). De data begint op 4 september 2008, dit zien we als de start van de eerste week. Voor elke week wordt het gemiddelde van de dagen in die week genomen.




Figuur 3.6: De data van Google per week, van 4 september 2008 tot 31 oktober 2018

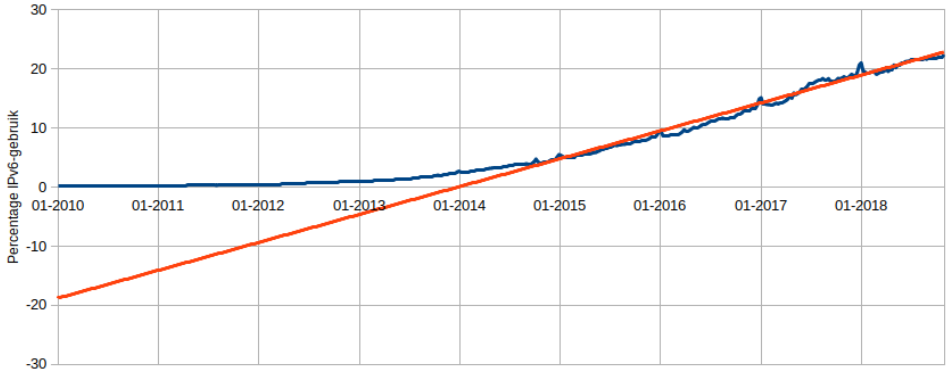



Figuur 3.7: De data van Google per week, van 30 oktober 2014 tot 31 oktober 2018

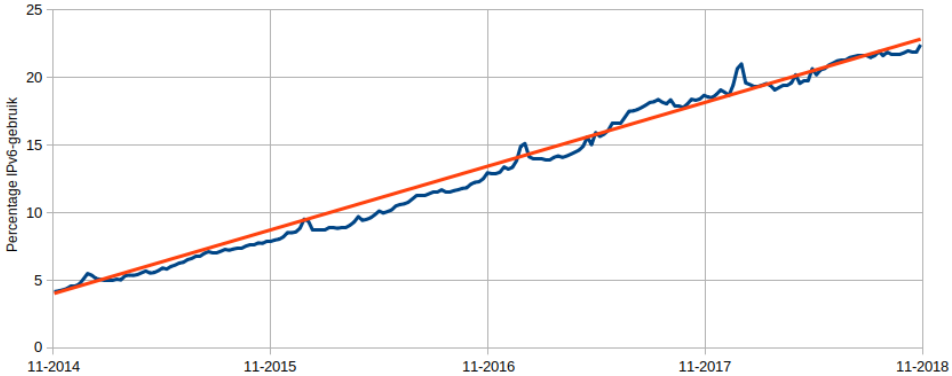
We willen graag de grafiek extrapoleren, daarvoor benaderen we de gegevens van de laatste 104 weken (dus ongeveer twee jaar, de periode van 3 november 2016 tot 31 oktober 2018) met een lineaire functie. Deze functie is met behulp van de kleinste-kwadratenmethode berekend als volgt:

$$y = -25{,}00 + 0{,}09w$$

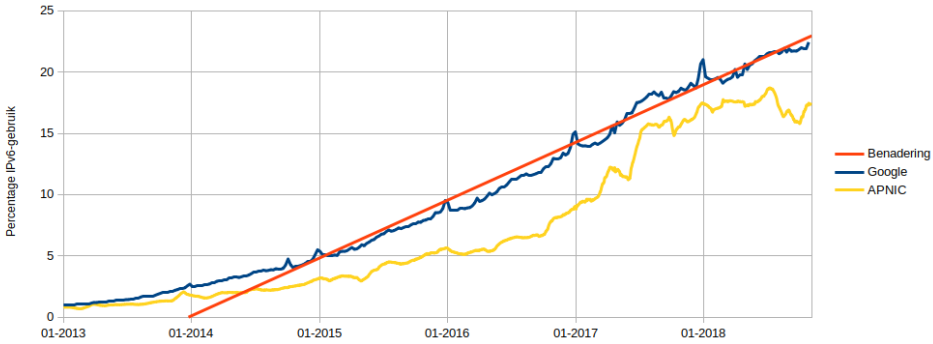
Hier is $w$ het weeknummer, dat begint bij de week die begint op 4 september 2008, wat week 0 is. In figuren 3.8 en 3.9 is deze functie weergegeven. Tevens worden in figuur 3.10 de datasets van zowel Google als APNIC weergegeven, samen met de benadering van de dataset van Google.




Figuur 3.8: De data van Google per week, van 1 januari 2010 tot 31 oktober 2018, met de lineaire benadering

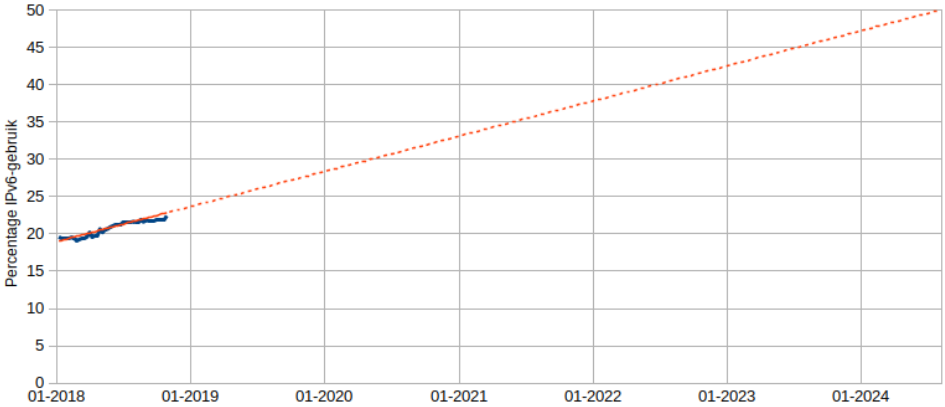



Figuur 3.9: De data van Google per week, van 30 oktober 2014 tot 31 oktober 2018, met de lineaire benadering




Figuur 3.10: De data van Google en APNIC met de lineaire benadering van de data van Google

Vanaf eind 2014 lijkt de benaderde lijn redelijk samen te vallen met de werkelijke data van Google. Nu kan de lijn worden doorgetrokken tot in de toekomst.




Figuur 3.11: De lineaire benadering, vanaf januari 2018 tot augustus 2024

| | | | |
|---|---|---|---|
| 25% | 04-2019 | 65% | 10-2027 |
| 30% | 05-2020 | 70% | 10-2028 |
| 35% | 05-2021 | 75% | 11-2029 |
| 40% | 06-2022 | 80% | 12-2030 |
| 45% | 07-2023 | 85% | 12-2031 |
| 50% | 07-2024 | 90% | 01-2033 |
| 55% | 08-2025 | 95% | 02-2034 |
| 60% | 09-2026 | 100% | 03-2035 |

Tabel 3.1: Maanden waarin bepaalde gemiddelde percentages ongeveer zullen worden bereikt, volgens de benaderde lineaire functie

Figuur 3.11 en tabel 3.1 tonen de geëxtrapoleerde datums waarop naar verwachting het gemiddelde percentage bepaalde waarden bereikt. Met het huidige tempo zal het dus tot 2024 duren voordat het IPv6-gebruik tot 50% is gestegen, en tot 2035 voor 100%. Het is waarschijnlijk dat de werkelijkheid van deze voorspelling zal gaan afwijken. Zodra de 50% is overschreden, betekent dat dat verbindingen die alleen IPv4 ondersteunen in de minderheid beginnen te komen. Op een gegeven moment zou het percentage IPv6-ondersteunende aansluitingen hoog genoeg kunnen worden dat het voor bijvoorbeeld websites en aanbieders van online-diensten aantrekkelijk wordt om alleen nog IPv6 te ondersteunen, wat weer zou kunnen leiden tot een extra impuls voor achterblijvende providers om ook IPv6 aan te

gaan bieden. Dit zou daarmee een voorbeeld zijn van de wet van Metcalfe: de waarde van IPv6 zou kwadratisch kunnen stijgen naarmate het gebruik ervan stijgt. Ook kan het gebeuren dat het tekort aan IPv4-adressen de komende jaren dermate dringend wordt dat er door providers toch extra haast gemaakt gaat worden met IPv6.

## 3.2 Vertragende en versnellende factoren

### 3.2.1 Vertragende factoren

De overstap van IPv4 naar IPv6 duurt dus langer dan gedacht, maar waar wordt dat eigenlijk door veroorzaakt?

**Bestaande netwerken** Bestaande netwerken en apparatuur zijn door de jaren heen ingericht op IPv4. IPv4 en IPv6 zijn niet met elkaar compatibel, waardoor netwerkbeheerders vaak zelf actie moeten ondernemen om al hun reeds werkende systemen om te zetten. Een voorbeeld daarvan is YouTube, waar slechtere prestaties werden gemeten via IPv6 dan via IPv4 (augustus 2014 – juni 2017) omdat Google Global Caches minder goed beschikbaar waren over IPv6 [Baj+17]. Ook Facebook moest, ondanks de motivatie om over te stappen op IPv6, veel problemen overwinnen die te maken hadden met verouderde software en code die op IPv4 was afgestemd [Fac14].

IPv6 moet softwarematig ook worden ondersteund door de apparaten die op een IPv6-netwerk worden aangesloten voordat IPv6 door dat apparaat daadwerkelijk net zo goed gebruikt kan worden als IPv4. Zo is het formaat van de header veranderd in IPv6 [RFC8200], en software kan er nog steeds van uitgaan dat IP-adressen altijd 32 bits lang zijn en uit vier octets bestaan tussen 0 en 255, waardoor het niet mogelijk is om IPv6-adressen in te vullen.

Als gebruikte software, zoals besturingssystemen van computers of routers, in een netwerk alleen geschikt is voor IPv4, en er geen update beschikbaar is om IPv6-ondersteuning toe te voegen, is dat al een belemmering voor een omschakeling naar IPv6, omdat er eerst een update moet worden gecreëerd voor die software. Dit kan onder meer inhouden dat ondersteuning voor het IPv6-protocol en de configuratie daarvan moet worden toegevoegd, de implementatie van DNS ook `AAAA`-records moet ondersteunen en de interface moet worden aangepast om IPv6-adressen als invoer te accepteren. Als het creëren van een update niet mogelijk is zal de software of het apparaat in het geheel moeten worden vervangen. Waar IPv6 al wel door het gehele netwerk volledig wordt ondersteund, kan de configuratie nog bestaan uit "hard-coded" IPv4-adressen. Denk aan IP-adressen van databaseservers en

proxyservers, webservers die alleen geconfigureerd zijn om te luisteren via een IPv4-adres, firewalls die al het binnenkomend verkeer blokkeren behalve bepaalde poorten, maar onbedoeld alleen via IPv4, of juist alleen verbindingen vanaf bepaalde bekende IPv4-adressen accepteren. Hoe complexer een netwerk, hoe meer van dit soort configuratie opnieuw bekeken moet worden.

**IPv4-'life support' bestaat** Er bestaan mechanismen (zoals Carrier-Grade NAT) om, ondanks het tekort aan IPv4-adressen, het leven van IPv4 te verlengen zonder dat IPv6 hoeft te worden ondersteund. Daar wordt dan ook gebruik van gemaakt, omdat dat vaak gemakkelijker is dan meteen al rekening te moeten houden met de mogelijke problemen die bij het IPv6-geschikt maken van een netwerk komen kijken. [Ško+14; Ric+16]

In de jaren 90 dreigde al een tekort aan IPv4-adressen [IETF90]. Het ontwerp van Network Address Translation werd gepubliceerd in 1994 [RFC1631; RFC3022], als kortetermijnoplossing. In 1995 werd het eerste ontwerp van de beoogde langetermijnoplossing gepubliceerd, IPv6 [RFC1883]. IPv6 werd ontworpen met de verwachting dat het op korte termijn gebruikt zou gaan worden, NAT was immers een tijdelijke oplossing om wat ademruimte te creëren. In plaats daarvan heeft het nog een lange tijd geduurd voor IPv6 van de grond kwam en is NAT op een permanente oplossing gaan lijken.

**Kip-en-eiprobleem** Die 'life support'-mechanismen stellen niet alleen de transitie uit, ze zijn ook nodig zolang een groot deel van de rest van het internet nog alleen te bereiken is via IPv4; laat staan *alleen* via IPv6. IPv6 is daarmee een mooi voorbeeld van het kip-en-eiprobleem: het is voor de aanbieders van internetverbindingen voor overwegend clientgebruikers (zoals internetproviders voor particulieren en bedrijven) nog niet heel interessant om IPv6 aan te bieden, omdat er nog bijna geen websites en diensten zijn die niet ook gewoon via IPv4 te bereiken zijn, en voor websites en diensten is het niet heel interessant omdat iedere internetgebruiker toch nog een IPv4-verbinding heeft.

**Coördinatie is moeilijk** Het kip-en-eiprobleem is des te groter omdat het internet niet meer slechts vier universiteiten met elkaar verbindt, ondertussen is er een groot aantal apparaten op het internet aangesloten, wat de transitie niet makkelijk maakt [TKA08]. Het is niet meer denkbaar dat alle organisaties die zijn aangesloten op het internet in een vergadering afspreken om van IPv4 af te stappen en alle netwerken en systemen voor een bepaalde datum geschikt te maken voor IPv6. Het internet is van nature een gedecentraliseerd netwerk. Er is geen internetautoriteit die IPv4 wereldwijd kan uitschakelen. Wat waarschijnlijk het dichtst in de buurt zou komen is wetgeving die providers

verplicht om een IPv6-verbinding aan te bieden, maar dat is nog geen coördinatie.

**"Het werkt toch?"** Aangezien zowel internetproviders als beheerders van websites en diensten nog niet echt de keuze hebben om te stoppen met IPv4, is de verwachting dat het IPv4-internet voorlopig nog volledig zal blijven werken. Ook NAT en Carrier-Grade NAT zijn bedoeld om alledaags internetgebruik mogelijk te maken. Eenieder die Google, YouTube, Facebook, en andere populaire websites bezoekt hoeft zich geen zorgen te maken dat het tekort aan IPv4-adressen ervoor zorgt dat die websites ontoegankelijk zullen worden, dat is namelijk in niemands belang – prioriteit is dat alles voor gebruikers blijft werken zoals het altijd al werkt. Bedrijven als Google en Facebook zouden geld mislopen als de meerderheid van de gebruikers hun sites niet meer kon bezoeken, en providers lopen het risico op veel vragen en klachten van klanten omdat websites niet meer werken, en dat klanten over zullen stappen op een concurrent.

Voor de gemiddelde internetgebruiker is niet direct een voordeel te halen uit IPv6. Er is voor de meeste internetgebruikers geen reden om contact op te nemen met de provider omdat ze behoefte hebben aan een eigenschap van IPv6 – zolang de websites die ze bezoeken maar goed werken. In principe zou de gemiddelde internetgebruiker ook niets moeten merken wanneer de provider ondersteuning voor IPv6 inschakelt; de meeste apparaten en bijna alle besturingssystemen ondersteunen al IPv6, de browser zal vaker aan de hand van `AAAA`-DNS-records een server vinden in plaats van `A`-records. Dit gebeurt allemaal 'achter de schermen'. De gebruiker kan echter wel tegen problemen aanlopen die niet voorkwamen in een 'IPv4-only'-omgeving. Een website kan bijvoorbeeld verkeerd geconfigureerd zijn en alleen goed werken over IPv4. Als de gebruiker merkt dat de site bij anderen wel werkt, lijkt het alsof het probleem niet wordt veroorzaakt door de website zelf, maar door de eigen provider. Als de gebruiker ontdekt dat het probleem verholpen kan worden door IPv6 uit te schakelen (of wordt aangeraden dit te doen), wordt IPv6 gezien als de oorzaak van het probleem, en eenmaal uitgeschakeld is het waarschijnlijk dat er niet meer naar wordt omgekeken. Datzelfde kan gebeuren wanneer de provider IPv6-ondersteuning toevoegt en dit niet heel soepel verloopt: de klant schakelt IPv6 uit en vervolgens raakt deze actie in de vergetelheid.

**Niet elke internetprovider heeft een tekort** Er zijn providers die vroeger heel veel IPv4-adressen toegewezen hebben gekregen, soms veel meer dan ze tot op de dag van vandaag nodig hebben. Die hebben er dus nog meer dan genoeg en voelen niet altijd de noodzaak om haast te maken met de transitie naar IPv6. [Ric+16]

**Handel** IPv4-adressen kunnen worden doorverkocht [Ric+15; Ric+16]. Daarmee is er in mindere mate een tekort aan IPv4-adressen, want er is nog een aanbod.

### 3.2.2 Versnellende factoren

Toch is er ook een aantal redenen waarom tegenwoordig steeds meer verkeer via IPv6 verloopt.

**IPv4 en IPv6 kunnen naast elkaar bestaan** Het is niet zo dat het gehele internet tijdelijk moet worden uitgeschakeld om te kunnen overstappen van IPv4 op IPv6. Daar zou verstrekkende coördinatie voor nodig zijn geweest, tot een hoogte die waarschijnlijk niet was bereikt (zie 3.2.1). Het internet had dan waarschijnlijk nog steeds voor 100% op IPv4 gedraaid, terwijl het alleen maar doorgroeit.

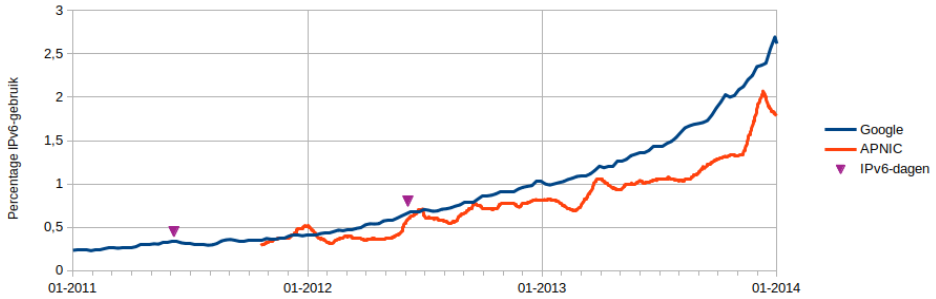
IPv4 en IPv6 kunnen dus wel naast elkaar gebruikt worden, en één systeem kan dat tegelijkertijd (dual-stack). Als hetzelfde netwerk zowel IPv4 als IPv6 kan gebruiken, is het werk voor dat netwerk voorlopig voor het grootste deel klaar; websites kunnen worden bezocht over zowel IPv4 als IPv6, en wanneer het daar tijd voor is hoeft alleen nog maar IPv4 te worden verwijderd.

**Nadelen van IPv4 en voordelen van IPv6** Het is moeilijk voor te stellen dat het gebruik van IPv6 de afgelopen jaren zo zou zijn gestegen als er nog genoeg IPv4-adressen zouden zijn. De kleine adresruimte is daarmee een nadeel van IPv4 dat IPv6 oplost [Zho+06; Ško+14; RFC8200]. Door het tekort aan IPv4-adressen moet NAT worden gebruikt wat ook weer nadelen heeft: netwerken worden er complexer van, het belemmert end-to-end-verbindingen en flexibiliteit, en voor sommige toepassingen werkt het helemaal niet [Zho+06; TKA08]. Microsoft [Kea17] en Facebook [Fac14] hebben aangegeven gebruik te zullen maken van netwerken die intern alleen IPv6 gebruiken, omdat ze zelfs bijna alle privé-IPv4-adressen gedefinieerd in [RFC1918] al gebruikt hebben, en omdat het hun netwerken minder complex zal maken.

IPv6 heeft nog meer voordelen ten opzichte van IPv4. Zo kan IPv6 Internet of Things-apparaten beter faciliteren. Naast de grotere adresruimte, die van pas komt wanneer een groot aantal apparaten via IP bereikbaar moeten zijn, is het ook mogelijk om IPv6-pakketten te versturen over IEEE 802.15.4-netwerken (voor apparaten die met weinig stroomverbruik werken) [RFC4944]. Verder is bij de ontwikkeling van IPv4 minder rekening gehouden met security dan bij de ontwikkeling van IPv6. Zo is IPsec standaard ingebouwd in IPv6, en wordt aanbevolen ('`SHOULD`') dat dit door IPv6-nodes wordt ondersteund,

terwijl IPsec voor IPv4-implementaties slechts optioneel is [RFC4301; RFC8504]

**IPv6 wordt gepromoot** In 2011 en 2012 zijn door de Internet Society een *World IPv6 Day* en een *World IPv6 Launch* georganiseerd. Op *World IPv6 Day* (8 juni 2011) konden deelnemende bedrijven en andere websitebeheerders hun websites beschikbaar maken over IPv6 bij wijze van proef, en de *World IPv6 Launch* (6 juni 2012) was bedoeld om organisaties zoals internetproviders te motiveren om ondersteuning voor IPv6 permanent toe te voegen aan hun producten en diensten. Ten tijde van de *World IPv6 Day* en *World IPv6 Launch* was het percentage IPv6-gebruik nog onder de 1% (zie figuur 3.12). Te zien is dat deze dagen weinig rechtstreekse invloed lijken te hebben gehad, maar het IPv6-gebruik van toen staat in schril contrast met dat van tegenwoordig (zie sectie 3.1). [WIPv6L; Baj+17]




Figuur 3.12: Het IPv6-gebruik volgens Google en APNIC rond *World IPv6 Day* en *World IPv6 Launch*. *World IPv6 Day* en *World IPv6 Launch* zijn aangegeven.

## 3.3 Oplossingen

Het tekort aan IPv4-adressen betekent dat er een grotere vraag is dan een aanbod – anders is er geen tekort. De druk daarvan is naar verwachting het grootst bij de bedrijven die financieel afhankelijk zijn van hun toegekende IPv4-adressen en deze uitgeven aan klanten, zoals providers. Als de druk niet verlicht wordt, dat wil zeggen, de uiteindelijke 'uitgever' heeft geen IP-adressen meer om uit te geven en er wordt niks aan gedaan, zal deze abrupt geen nieuwe klanten meer aan kunnen sluiten, wat er dus toe leidt dat nieuwe klanten redenen hebben om een andere provider te zoeken. Er bestaan natuurlijk wel methoden om de druk van het gebrek aan IPv4-adressen te verlichten, zonder over te hoeven stappen op netwerken die alleen IPv6 ondersteunen en geen contact kunnen leggen met een aanzienlijk deel van het internet dat alleen maar IPv4 ondersteunt.

### 3.3.1 IPv4-adressen kopen

Zoals eerder genoemd is er een markt voor IPv4-adressen waar deze kunnen worden overgekocht. Als een LIR een bepaalde hoeveelheid IPv4-adressen nodig heeft en het geen probleem vindt om te betalen voor die benodigde hoeveelheid, dan is dit een heel voor de hand liggende oplossing. Het is ook een heel 'comfortabele' kortetermijnoplossing, want netwerken hoeven er niet voor te worden aangepast, en eindgebruikers kunnen niet ineens minder met hun internetverbinding doen dan ze eerder al konden, zoals servers opzetten die via IPv4 bereikt kunnen worden. Sterker nog, als het gaat om IP-adressen voor servers is er vaak geen andere optie, omdat de gemiddelde webserver een IPv4-adres moet hebben; NAT biedt vooral uitkomst wanneer een netwerk niet toegankelijk hoeft te zijn vanaf het internet.

Bij de meeste RIR's bestaat de mogelijkheid voor LIR's om adresruimte aan elkaar over te dragen. De RIR kan dan administratief de toewijzing veranderen. Ook is het in sommige gevallen mogelijk om adresruimte over te dragen naar een LIR die lid is van een andere RIR. [Ric+15]

Een nadeel van het toepassen van deze methode is dat het de aanbieder geld kost, en dit kan worden doorberekend aan klanten, waardoor de dienst ook voor klanten onaantrekkelijker wordt om af te nemen. Nieuwe spelers op de markt zullen vooraf moeten investeren in de benodigde adresruimte, en een nieuwe speler kan zich die kosten wellicht niet altijd veroorloven. Verder bereidt het aanbieders niet voor op de toekomst van het internet.

### 3.3.2 (Carrier-Grade) NAT

Network Address Translation [RFC3022] is een bewezen techniek die al sinds de jaren 90 wordt toegepast om het gebruik van IP-adressen te beperken. De techniek zoals deze zich heeft bewezen werkt door aan iedere eindgebruiker (dus ieder huishouden) slechts één IPv4-adres toe te wijzen, dat wordt gebruikt voor de 'Network Address Translator', wat normaal gesproken de eerste router in het huis is. Zoals eerder genoemd kan NAT worden beschouwd als *stateful*, wat het internet complexer maakt.

NAT werpt een probleem op wanneer een host die zich erachter bevindt van buitenaf te benaderen moet zijn, met andere woorden, wanneer iemand 'thuis' een server wil draaien. Dit is niet zonder meer mogelijk, want hosts binnen het lokale netwerk hebben geen publiek IP-adres. Dit probleem kan worden opgelost door *port forwarding* in te stellen op de NAT-router (waarbij binnenkomend verkeer naar een bepaalde poort wordt doorgestuurd naar een vooraf gedefinieerde host) of al het binnenkomend verkeer naar een bepaalde host te laten doorsturen (een *demilitarized zone host*).

Met NAT is een provider dus altijd slechts één publiek (routeerbaar) IPv4-adres per klant kwijt geweest. Dit is wat het tekort jarenlang in feite heeft beperkt. Als een provider echter ook niet genoeg IP-adressen op voor-

raad heeft om elke klant van een publiek IP-adres te voorzien, kan NAT ook worden toegepast buiten de netwerken van klanten, waarmee meerdere klanten dus een IP-adres moeten delen. Dit verlicht de druk van de provider, maar heeft veel nadelen. Zo kunnen gebruikers vrijwel zeker niet meer *port forwarden*, waardoor bijvoorbeeld beveiligingscamera's en servers voor spellen niet meer van buitenaf kunnen worden benaderd. De vertaling van IP-adressen gebeurt namelijk op een router die niet beheerd kan worden door klanten zelf, en het is vanuit het perspectief van de provider gezien niet praktisch om een uitzondering te maken voor het instellen van port forwarding-regels [Ško+14].

Ook is er een mogelijk probleem met reputatie van IP-adressen, zo kan één klant zich bijvoorbeeld misdragen op Wikipedia en het gedeelde IP-adres geblokkeerd laten worden, of op spam-blacklists laten komen [Ric+16]. Providers kunnen wettelijk verplicht zijn om klanten aan de hand van een IP-adres en een tijdstip te kunnen identificeren [Ric+16; Ško+14]. Als elke klant unieke IP-adressen krijgt toegewezen is daar makkelijk aan te voldoen, zelfs als toewijzingen soms veranderen en gebruikte IP-adressen naar andere klanten gaan [Ško+14]. Een provider hoeft namelijk alleen te 'onthouden' wie op welk moment welk IP-adres toegewezen heeft gekregen. Voor elke klant dus een eenvoudig lijstje met IP-adressen en begin- en eindtijden, of andersom: voor elk IP-adres een eenvoudig lijstje met klanten en begin- en eindtijden. Bij Carrier-Grade NAT (CGN of CGNAT) is het iets ingewikkelder, aangezien per definitie meerdere klanten tegelijkertijd één IP-adres gebruiken. Als het bewijs voor eventueel misbruik alleen bestaat uit het IP-adres van een internetgebruiker en een tijdstempel, is dat dus onvoldoende om te weten om welke klant het gaat. Er is pas voldoende informatie als de provider opslaat welke klant wanneer gebruikmaakt van welke poort, en het genoemde bewijs ook het bronpoortnummer omvat. Is dat niet bekend, maar wel het IP-adres en poortnummer van de bestemming van de verbinding, dan moet de provider die informatie zelf hebben opgeslagen om de traceerbaarheid nog enigzins te kunnen waarborgen. Dat kan privacybezwaren oproepen, omdat het simpel gezegd gelijkstaat aan het bijhouden van een logboek welke websites door welke klant bezocht worden. Ook maakt dit snelle, betrouwbare en grote opslagsystemen nodig om de logboeken bij te kunnen houden. [Ško+14] Voor mensen die liever niet onder de loep van inlichtingendiensten komen kan CGN zowel voor- als nadelen hebben. Aan de ene kant biedt een gedeeld IP-adres een mate van anonimiteit, een verbinding kan tot stand gebracht zijn door één van misschien wel 128 'huishoudens' [Ric+16]. Aan de andere kant zal de kans groter zijn dat een opsporingsdienst geïnteresseerd is in iemand anders met hetzelfde publieke IP-adres, waardoor veel meer klanten onder het vergrootglas gehouden zullen worden dan anders nodig zou zijn.

Zoals gezegd zorgt NAT voor extra complexiteit in een netwerk. Met CGN wordt dit effect versterkt, aangezien het normaal is om de huishou-

delijke NAT in stand te houden bij CGN. In de meestvoorkomende vorm van CGN krijgt ieder lokaal netwerk nog altijd één IPv4-adres, dat is echter geen publiek routeerbaar IP-adres.

Er is ook een vorm mogelijk van CGN waarbij geen lokaal IPv4-adres gebruikt wordt, maar waarbij IPv4-pakketten in IPv6 worden ingepakt. Dit wordt gebruikt bij Dual-Stack Lite (DS-Lite) [RFC6333]. Verder kan ook de techniek Address plus Port (A+P) gebruikt worden [RFC6346]. Dit is grotendeels hetzelfde als wat eerder is besproken: meerdere netwerken delen hetzelfde IP-adres, met als verschil dat een uitgaande verbinding niet elke willekeurige poort zou kunnen gebruiken, maar dat elk netwerk van tevoren een bepaalde hoeveelheid poorten toegewezen krijgt. Dit zorgt ervoor dat A+P minder *stateful* is dan dubbele NAT – de CGN hoeft immers niet per verbinding bij te houden welke poortnummers gebruikt worden – maar mogelijke problemen met de reputatie van IP-adressen en wettelijke verplichtingen tot identificatie aan de hand van een IP-adres en een tijdstip blijven.

Het is moeilijk te meten in hoeverre CGN wordt gebruikt. Aan een publiek IP-adres is namelijk niet te zien door hoeveel netwerken dit IP-adres wordt gedeeld. Toch zijn in 2016 door de werking van BitTorrent DHT en de dienst Netalyzr meer dan 500 unieke implementaties van CGN gevonden. [Ric+16]

### 3.3.3 NAT64

NAT64 is een methode die ervoor zorgt dat al het verkeer tussen een netwerk en de rest van het internet via IPv6 verloopt, terwijl IPv4-hosts nog bereikbaar blijven. Het maakt gebruik van het feit dat de meeste verbindingen over het internet niet tot stand komen voordat een DNS-server is geraadpleegd om te weten wat het IP-adres van een server is. Een gemiddelde website wordt bezocht via een domeinnaam, en de bijbehorende IP-adressen worden dan via DNS opgevraagd. Als via DNS een `AAAA`-record terugkomt, kan geprobeerd worden de website via het IPv6-adres te bezoeken. Als een website echter alleen is geconfigureerd voor IPv4, zal er hoogstwaarschijnlijk ook alleen een `A`-record terugkomen, en moet de verbinding met de website wel via IPv4 worden gemaakt. NAT64 zorgt er op een creatieve manier voor dat er verbinding kan worden gemaakt met dit IPv4-adres: de DNS-server (in dit geval een DNS64-server [RFC6147]) geeft niet het oorspronkelijke IPv4-adres terug, maar maakt zelf een IPv6-adres dat het IPv4-adres representeert. In feite behoort dit IPv6-adres tot een netwerk van de NAT64-gateway, wat het IPv4-adres uit het gemaakte IPv6-adres kan aflezen en dan de verbinding kan opzetten met dat IPv4-adres zoals de bedoeling was. De NAT64-gateway vertaalt dus IPv6 naar IPv4 en omgekeerd, nadat de DNS-server heeft 'gelogen' over het IP-adres om het verkeer daarheen te sturen. [Ško+14; RFC6146]

Een beperking van NAT64 is dat het geen effect heeft wanneer IPv4-adressen rechtstreeks worden gebruikt in plaats van een domeinnaam.

Als NAT64 in een netwerk wordt gebruikt betekent dat in ieder geval dat dat netwerk al helemaal klaar moet zijn voor IPv6, en dat deze methode dus op zich niet zorgt voor een vertraging van de IPv6-adoptie.

Er wordt in de RFC over NAT64 wel een aantal *Security Considerations* genoemd. Zo is NAT64 stateful, wat onder andere het risico met zich meebrengt dat een DoS-aanval het geheugen van de NAT64 vult met informatie over verbindingen die een bepaalde tijd bewaard moeten blijven. [RFC6146]

# Hoofdstuk 4

# Gerelateerd werk

Richter et al. hebben in 2016 het gebruik van een van de genoemde oplossingen, Carrier-Grade NAT, geanalyseerd [Ric+16].

Eerder, in 2015, heeft een deel van dezelfde groep onderzoekers (Richter et al.) de uitputting van IPv4-adressen en bijbehorende problemen beschreven [Ric+15].

Škoberne et al. hebben mechanismen voor het delen van IPv4-adressen ingedeeld in negen klassen [Ško+14].

Er zijn meerdere onderzoeken gedaan naar de prestaties van zowel IPv6 als mechanismen zoals CGN. Zo hebben Dhamdhere et al. gevonden dat de prestaties van IPv6 en IPv4 in de regel vergelijkbaar moeten zijn [Dha+12], en hebben Ohara et al. gevonden dat het gebruik van CGN in mobiele netwerken geen significante invloed heeft op de prestaties [Oha+14].

Ook zijn er eerder onderzoeken gedaan naar IPv6-adoptie. Zo hebben Dhamdhere et al. in 2012 aan de hand van Border Gateway Protocol-data de IPv6-adoptie geanalyseerd [Dha+12].

# Hoofdstuk 5

# Conclusie en discussie

## 5.1 Conclusie

In dit onderzoek hebben we onderzocht wat de oorzaken en gevolgen zijn van het uitstellen van de IPv6-adoptie.

De oorzaken zijn te vinden in een aantal vertragende factoren, waaronder:

- Bestaande netwerken en apparatuur zijn door de jaren heen ingericht op IPv4, en IPv4 en IPv6 zijn niet met elkaar compatibel.
- Er bestaan 'life-support'-mechanismen om, ondanks het tekort aan IPv4-adressen, het leven van IPv4 te verlengen zonder dat IPv6 hoeft te worden ondersteund.
- IPv6 is een voorbeeld van het kip-en-eiprobleem.
- Het internet is groot en gedecentraliseerd, waardoor coördinatie moeilijk is.
- IPv4 kan blijven werken, en voor de gemiddelde internetgebruiker is niet direct een voordeel te halen uit IPv6.
- Niet elke internetprovider heeft een tekort aan IPv4-adressen.
- IPv4-adressen kunnen worden doorverkocht, waardoor er in mindere mate een tekort aan IPv4-adressen bestaat.

Voor de gevolgen van het uitstellen van de IPv6-adoptie kan worden gekeken naar de gevolgen van de methoden die worden gebruikt om het leven van IPv4 te verlengen. Voor nieuwe spelers op de markt geldt dat ze wellicht moeten investeren in adresruimte die ze zich niet kunnen veroorloven, omdat IPv4-adresruimte soms een vereiste is om een dienst op te kunnen starten. NAT heeft al gevolgen gehad voor de werking van het internet, namelijk dat clients die zich achter een NAT bevinden niet zonder bijzondere

configuratie ook de rol van server kunnen aannemen. Doordat IPv6 nog niet in voldoende mate gebruikt wordt kan het, als verdere maatregel, nodig zijn om NAT te gebruiken om meerdere netwerken, die nu al NAT gebruiken, ook gezamelijk één IP-adres te laten delen. Dat kan ervoor zorgen dat de eerder genoemde configuratiemogelijkheden om zich toch een server achter een NAT te laten bevinden verdwijnen. Dit geldt ook voor NAT64, waar in het geheel geen sprake meer is van een IPv4-verbinding tot aan een netwerk dat ervan gebruik maakt.

## 5.2 Discussie

Aangezien het internet een gedecentraliseerd en open netwerk is, is het moeilijk om nauwkeurige metingen van het wereldwijde gebruik van het internet te verrichten en daarmee iets over het internet als geheel te zeggen. Een bedrijf als Google kan gemakkelijk bepaalde metingen uitvoeren op de gebruikers van haar diensten, maar voordat de data iets kan zeggen over het gehele internet zouden die gegevens representatief moeten zijn voor alle internetgebruikers, in alle landen ter wereld. Aangezien zelfs Google niet door honderd procent van de gebruikers in alle landen gebruikt wordt – in sommige landen kunnen andere bedrijven populairder zijn – kunnen bijna geen metingen worden verricht die volledig representatief zijn. Dergelijke metingen zijn uiteraard ook beperkt tot wat door de meter gemeten kan worden. Een website kan bijvoorbeeld niet direct zien door hoeveel gebruikers een bepaald IP-adres gedeeld wordt, omdat technieken zoals NAT en CGN gebaseerd zijn op de bestaande protocollen van het internet en compatibel moeten zijn met apparaten die uitgaan van deze bestaande protocollen. Dat wil zeggen dat er geen ruimte (noch noodzaak) is om IPv4-pakketten die door een NAT of CGN worden doorgestuurd herkenbaar te maken als pakketten waarvan het IP-adres vertaald wordt.

Desondanks geven bestaande metingen wel een indicatie dat de IPv6-adoptie traag verloopt. Dat is mede omdat men over het algemeen met IPv4 nog in voldoende mate uit de voeten kan, vooral dankzij de methoden die het mogelijk maken om zoveel mogelijk netwerken van zo min mogelijk IP-adressen gebruik te laten maken, ondanks het feit dat die methoden ook nadelen hebben.

## 5.3 Vervolgonderzoek

Zoals eerder genoemd zou vervolgonderzoek de IPv6-ondersteuning van de populairste websites kunnen bepalen. Er kan dan bijvoorbeeld worden gekeken hoeveel procent van de populairste websites volgens Alexa[1] een `AAAA`-DNS-record heeft dat naar behoren werkt. Dat hebben we in deze scriptie

[1] `https://www.alexa.com/topsites`

niet gedaan, omdat de gratis toegankelijke lijst van Alexa nog slechts 50 sites omvat, en dat alleen iets zegt over websites en niet over internetgebruikers. Er zouden ook voor langere tijd metingen moeten worden verricht om het tijdsverloop te bestuderen.

Ook zou vervolgonderzoek een theoretische bovengrens kunnen bepalen van de IPv6-adoptie door te kijken hoeveel blokken IPv6-adressen tot nu toe zijn uitgegeven door IANA en de RIR's. Dit onderzoek zou echter ook een beperking hebben: aangezien de IPv6-adresruimte groot is kunnen grote blokken IP-adressen worden toegewezen aan providers, blokken die een veelvoud zijn van de totale IPv4-adresruimte. Wel zouden in theorie deze IPv6-toewijzingen kunnen worden vergeleken met de bestaande IPv4-toewijzingen, om te kijken hoeveel providers IPv4-adressen hebben maar geen IPv6-adressen.

# Bibliografie

## Academisch


[Baj+17] Vaibhav Bajpai e.a. "Measuring YouTube Content Delivery over IPv6". In: *SIGCOMM Comput. Commun. Rev.* 47.5 (okt 2017), p. 2–11. ISSN: 0146-4833. DOI: 10.1145/3155055.3155057. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/3155055.3155057.

[Dha+12] Amogh Dhamdhere e.a. "Measuring the Deployment of IPv6: Topology, Routing and Performance". In: *Proceedings of the 2012 Internet Measurement Conference*. IMC '12. Boston, Massachusetts, USA: ACM, 2012, p. 537–550. ISBN: 978-1-4503-1705-4. DOI: 10.1145/2398776.2398832. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/2398776.2398832.

[Oha+14] Yasuhiro Ohara e.a. "On the Impact of Mobile Network Delays on Connection Establishment Performance of a Carrier Grade NAT Device". In: *Proceedings of the AINTEC 2014 on Asian Internet Engineering Conference*. AINTEC '14. Chiang Mai, Thailand: ACM, 2014, 1:1–1:8. ISBN: 978-1-4503-3251-4. DOI: 10.1145/2684793.2684794. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/2684793.2684794.

[Ric+16] Philipp Richter e.a. "A Multi-perspective Analysis of Carrier-Grade NAT Deployment". In: *Proceedings of the 2016 Internet Measurement Conference*. IMC '16. Santa Monica, California, USA: ACM, 2016, p. 215–229. ISBN: 978-1-4503-4526-2. DOI: 10.1145/2987443.2987474. URL: http://doi.acm.org/10.1145/2987443.2987474.

[Ško+14] Nejc Škoberne e.a. "IPv4 Address Sharing Mechanism Classification and Tradeoff Analysis". In: *IEEE/ACM Trans. Netw.* 22.2 (apr 2014), p. 391–404. ISSN: 1063-6692. DOI: 10.1109/TNET.2013.2256147. URL: http://dx.doi.org.ru.idm.oclc.org/10.1109/TNET.2013.2256147.

[TKA08] Kuljaree Tantayakul, Sinchai Kamolphiwong en Touchai Angchuan. "IPv6@HOME". In: *Proceedings of the International Conference on Mobile Technology, Applications, and Systems*. Mobility '08. Yilan, Taiwan: ACM, 2008, 92:1–92:6. ISBN: 978-1-60558-089-0. DOI: 10.1145/1506270.1506383. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/1506270.1506383.

[Zho+06] Xiaoming Zhou e.a. "Estimation of Perceived Quality of Service for Applications on IPv6 Networks". In: *Proceedings of the ACM International Workshop on Performance Monitoring, Measurement, and Evaluation of Heterogeneous Wireless and Wired Networks*. PM2HW2N '06. Terromolinos, Spain: ACM, 2006, p. 74–81. ISBN: 1-59593-502-9. DOI: 10.1145/1163653.1163667. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/1163653.1163667.

## Technisch

[Col+10] Lorenzo Colitti e.a. "Evaluating IPv6 adoption in the Internet". In: *PAM 2010*. 2010. URL: http://www.pam2010.ethz.ch/papers/full-length/15.pdf.

[RFC0760] J. Postel. *DoD standard Internet Protocol*. RFC 760. RFC Editor, jan 1980.

[RFC0762] J. Postel. *Assigned numbers*. RFC 762. RFC Editor, jan 1980.

[RFC0791] Jon Postel. *Internet Protocol*. STD 5. RFC Editor, sep 1981. URL: http://www.rfc-editor.org/rfc/rfc791.txt.

[RFC1034] P. Mockapetris. *Domain names - concepts and facilities*. STD 13. RFC Editor, nov 1987. URL: http://www.rfc-editor.org/rfc/rfc1034.txt.

[RFC1035] P. Mockapetris. *Domain names - implementation and specification*. STD 13. RFC Editor, nov 1987. URL: http://www.rfc-editor.org/rfc/rfc1035.txt.

[RFC1519] Vince Fuller e.a. *Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy*. RFC 1519. RFC Editor, sep 1993. URL: http://www.rfc-editor.org/rfc/rfc1519.txt.

[RFC1631] Kjeld Borch Egevang en Paul Francis. *The IP Network Address Translator (NAT)*. RFC 1631. RFC Editor, mei 1994. URL: http://www.rfc-editor.org/rfc/rfc1631.txt.

[RFC1883] Stephen E. Deering en Robert M. Hinden. *Internet Protocol, Version 6 (IPv6) Specification*. RFC 1883. RFC Editor, dec 1995. URL: http://www.rfc-editor.org/rfc/rfc1883.txt.

[RFC1918] Yakov Rekhter e.a. *Address Allocation for Private Internets*. BCP 5. RFC Editor, feb 1996. URL: `http://www.rfc-editor.org/rfc/rfc1918.txt`.

[RFC3022] P. Srisuresh en K. Egevang. *Traditional IP Network Address Translator (Traditional NAT)*. RFC 3022. RFC Editor, jan 2001.

[RFC4301] S. Kent en K. Seo. *Security Architecture for the Internet Protocol*. RFC 4301. RFC Editor, dec 2005. URL: `http://www.rfc-editor.org/rfc/rfc4301.txt`.

[RFC4944] G. Montenegro e.a. *Transmission of IPv6 Packets over IEEE 802.15.4 Networks*. RFC 4944. RFC Editor, sep 2007. URL: `http://www.rfc-editor.org/rfc/rfc4944.txt`.

[RFC6146] M. Bagnulo, P. Matthews en I. van Beijnum. *Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers*. RFC 6146. RFC Editor, apr 2011.

[RFC6147] M. Bagnulo e.a. *DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers*. RFC 6147. RFC Editor, apr 2011.

[RFC6333] A. Durand e.a. *Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion*. RFC 6333. RFC Editor, aug 2011. URL: `http://www.rfc-editor.org/rfc/rfc6333.txt`.

[RFC6346] R. Bush. *The Address plus Port (A+P) Approach to the IPv4 Address Shortage*. RFC 6346. RFC Editor, aug 2011.

[RFC8200] S. Deering en R. Hinden. *Internet Protocol, Version 6 (IPv6) Specification*. STD 86. RFC Editor, jul 2017.

[RFC8504] T. Chown, J. Loughney en T. Winters. *IPv6 Node Requirements*. BCP 220. RFC Editor, jan 2019.

## Overig

[AFR17] AFRINIC. *AFRINIC Enters IPv4 Exhaustion Phase 1*. 2017. URL: `https://afrinic.net/library/news/2053-afrinic-enters-ipv4-exhaustion-phase-1` (bezocht op 16-03-2019).

[ARI15] ARIN. *ARIN IPv4 Free Pool Reaches Zero*. 2015. URL: `https://www.arin.net/vault/announcements/2015/20150924.html` (bezocht op 16-03-2019).

[Fac14] Facebook. *The Road To IPv6*. 2014. URL: `https://www.internetsociety.org/wp-content/uploads/2014/04/WorldIPv6Congress-IPv6_LH-v2.pdf` (bezocht op 16-03-2019).

[Hus13] Geoff Huston. *APNIC Labs IPv6 Measurement System*. 2013. URL: https://labs.apnic.net/?p=348 (bezocht op 16-03-2019).

[IETF90] *Proceedings of the Eighteenth Internet Engineering Task Force*. (University of British Columbia, Vancouver, 30 jul–3 aug 1990). https://www.ietf.org/proceedings/18.pdf, pagina's 59–61. IETF, 1990. URL: https://www.ietf.org/proceedings/18.pdf.

[Kea17] Marcus Keane. *IPv6-only at Microsoft*. 2017. URL: https://blog.apnic.net/2017/01/19/ipv6-only-at-microsoft/ (bezocht op 20-03-2019).

[Ric+15] Philipp Richter e.a. "A Primer on IPv4 Scarcity". In: *SIGCOMM Comput. Commun. Rev.* 45.2 (apr 2015), p. 21–31. ISSN: 0146-4833. DOI: 10.1145/2766330.2766335. URL: http://doi.acm.org.ru.idm.oclc.org/10.1145/2766330.2766335.

[WIPv6L] *FAQ — World IPv6 Launch*. URL: https://www.worldipv6launch.org/faq/ (bezocht op 20-03-2019).