Afstudeerverslag
“IPv4-connectiviteit in een IPv6 access network”

Student: Matthias van der Heide (07053940)
Instelling: Haagse Hogeschool
Versie: 2.6 - 5 juni 2011

Voorwoord

Als Technisch Informaticus in opleiding heb ik mij gespecialiseerd in het ontwikkelen van software. Het uitvoeren van een onderzoek met netwerken als hoofdonderwerp was dan ook zeker niet mijn verwachting van een geschikte afstudeeropdracht. Toch ben ik voor dit onderwerp gegaan, omdat het uitnodigde om veel te leren over technieken die aan de ene kant al twintig jaar in gebruik zijn, maar aan de andere kant continu worden vernieuwd en aangepast aan de tijd. De uitdaging waar het internet nu voor staat is groot. Om ook in de toekomst toegang tot het internet te verschaffen moeten in korte tijd vele netwerken worden aangepast. De ontwikkeling van de benodigde techniek is de afgelopen tijd in een stroomversnelling geraakt.

Na afronding van mijn opdracht kan ik tevreden stellen dat het met ‘het internet’ wel goed komt. Door een selecte groep technici bij verschillende bedrijven wordt gewerkt aan mogelijkheden om het internet toegankelijk te houden. Aan de inzet van deze mensen te merken zullen de problemen die het opraken van de internetadressen veroorzaken degelijk worden opgelost.

Mijn dank gaat uit naar Paul Boot van Bateau Knowledge voor het aanbieden van deze afstudeerstage en zijn bereidheid om mij kennis van computernetwerken bij te brengen die je niet tijdens de opleiding opdoet. Ook de leden van Stichting IPv6 Nederland hebben mij heel goed geholpen met hun kennis, tijd en contacten. Kris Nielander van We-Dare voor zijn hulp met het onderzoek naar UDP-verkeer. Een speciaal bedankje voor Joost Cassee die mij in zijn lunch en vrije tijd heeft geholpen om te reflecteren op mijn werk aan deze opdracht.

Ik wens u veel leesplezier, en hoop dat dit verslag een goed beeld schetst van mijn afstudeerwerk en de mogelijkheden die IPv6 biedt om internettoegang ook in de toekomst mogelijk te maken.

Samenvatting

In februari 2011 heeft de IANA de laatste IPv4-adresblokken uitgedeeld aan de regionale instanties die deze aan ISP's, bedrijven en organisaties toewijzen. De verwachte groei van het internet leidt hierdoor tot adrestekorten. ISP's in groeimarkten (glasvezel, mobiel) zullen als eerste worden getroffen. Hun netwerken kunnen niet zonder aanpassing blijven werken.

Bateau Knowledge ziet in het bieden van IPv4-connectiviteit over een IPv6-only netwerk de oplossing voor dit probleem. Om een techniek te vinden die IPv4-connectiviteit kan bieden in is een vergelijkend onderzoek uitgevoerd met technieken gevonden in de literatuur over IPv6 van de afgelopen jaren.

De techniek die als beste uit de bus kwam, Dual-Stack Lite, is verder onderzocht. Namens Stichting IPv6 Nederland is een testlaboratorium ingericht en met medewerking van A10 Networks en AVM zijn hun producten getest om aan te tonen of Dual-Stack Lite door providers ingezet kan worden als overgangstechniek.

Uit de tests blijkt dat met de apparatuur die momenteel beschikbaar is een werkend IPv6-netwerk kan worden gebouwd dat IPv4-connectiviteit biedt. Echter, er zijn enkele zaken die moeten worden aangepast voordat de geteste apparatuur kan worden gebruikt in een productienetwerk. Na deze aanpassingen biedt Dual-Stack Lite op IPv4-connectiviteit die qua functionaliteit en performance vergelijkbaar is met de huidige IPv4-only netwerken.

Inhoudsopgave

1. Inleiding

Dit document is het afstudeerverslag van Matthias van der Heide, en is geschreven ter afronding van de afstudeerperiode die onderdeel is van de opleiding Technische Informatica aan de Haagse Hogeschool in Delft.

In hoofdstuk twee wordt de opdracht toegelicht, zoals deze in samenspraak met de opdrachtgever is geformuleerd. Hoofdstuk drie licht de gekozen aanpak en planning toe. In de Probleemanalyse, hoofdstuk vier, wordt aan de hand van literatuur, een actorenanalyse en een analyse van het probleem beoordeeld of de oplossingsrichting die in de opdracht aangereikt wordt kan leiden tot een geldige oplossing. Vervolgens stellen we de opdracht bij. Hoofdstuk vijf verslaat het onderzoek naar kandidaatoplossingen in literatuur en op internet. Dit leidt tot de keuze voor een bepaalde techniek. Deze wordt in hoofdstuk zes door middel van het ontwerpen en bouwen van een proof of concept gevalideerd. De conclusies en aanbevelingen zijn te vinden in hoofdstuk zeven. Tot slot wordt in hoofdstuk acht het afstudeerproces geëvalueerd.

De producten van de afstudeeropdracht zijn als bijlage opgenomen. Waar van toepassing zijn secties overgenomen in dit verslag. Voor details wordt naar de bijlagen verwezen.

2. Opdracht

2.1 Opdrachtgever

Netwerkarchitect Bateau Knowledge ontwerpt voor haar klanten netwerken en onderhoudt deze. Veelal zijn de netwerken bedoeld om te communiceren met het internet (access networks). De grootte van de netwerken varieert van honderden tot duizenden hosts.

Bateau Knowledge is in de kern een eenmansbedrijf dat wordt geleidt door directeur/eigenaar Paul Boot. Voor het uitvoeren van projecten werkt hij samen met ZZP'ers. De organisatie is hierdoor klein en flexibel.

Bateau Knowledge heeft samen met vijf andere partijen de Stichting IPv6 Nederland opgericht [1]. Het doel van de stichting is om kennis op het gebied van IPv6 te delen en het bewustzijn op dit gebied te bevorderen.

2.2 Aanleiding

De invoering van Internet Protocol versie 6 (IPv6) als opvolger van het populaire Internet Protocol versie 4 (IPv4) is een opgave die de internetgemeenschap al jaren bezighoudt. Eind jaren '90 is de standaard geformaliseerd, en sinds dien wordt, hetzij mondjesmaat, netwerkapparatuur geproduceerd die in staat is om beide protocollen te ondersteunen. Het belangrijkste doel bij het ontwerp van IPv6 was het ondersteunen van de vele hosts die aan het almaar groeiende internet gekoppeld zijn. Hiervoor is het adresveld, wat IPv4 beperkt tot 4,3 miljard hosts, uitgebreid naar 128-bits. Daardoor is een ongelooflijke hoeveelheid computers met het internet te verbinden. Zo kan worden voorkomen dat er een eind komt aan de groei van het internet.

Het ministerie van Economische Zaken onderkent dit als een probleem en heeft de IPv6 Taskforce opgericht [2], met als doel om overheid en bedrijfsleven te ondersteunen bij het invoeren van IPv6. De adoptie van IPv6 is echter flink achtergebleven bij de verwachtingen, waardoor straks de situatie ontstaat dat de oude adressen eerder op zijn, dan dat iedereen overgestapt is op IPv6.

Hierdoor kunnen er geen nieuwe aansluitingen worden geleverd met een eigen IP-adres, tenzij er gebruik wordt gemaakt van IPv6-adressen. Nieuwe aansluitingen met enkel IPv6-adressen kunnen niet zonder meer verbinding maken met het 'oude' IPv4-internet.

2.3 Probleemstelling

Om te kunnen blijven groeien zullen eigenaren van accessnetwerken (internet service providers) hun netwerk moeten uitbreiden met IPv6. Dit kan door zowel IPv4 als IPv6 te gebruiken (dual stack), of door helemaal over te stappen op IPv6.

Dual-stack gaan (zowel IPv4 als IPv6 leveren) is voor de aanbieder duurder dan IPv6 alleen al vanwege de dubbele routering en de extra administratieve lasten. Daarom is het waarschijnlijk dat deze kiest voor een IPv6-only netwerk.

Klanten vragen echter om het ‘oude’ IPv4-internet te kunnen benaderen om bijvoorbeeld te surfen naar websites die niet over IPv6 beschikbaar zijn. Ook hebben ze veel apparatuur die niet werkt over IPv6, omdat deze niet vervangen kan worden, en niet geschikt te maken is door middel van een upgrade. Voorbeelden hiervan zijn spelcomputers, VoIP-telefoons, IP-camera’s en oude computers met Windows XP.

Hoe kan een provider zijn klanten zowel IPv6- als IPv4-connectiviteit bieden met een IPv6-only netwerk op een manier dat de klantenapparatuur goed blijft werken?

2.4 Opdrachtomschrijving

Realiseer een proof of concept dat verkeer van IPv4-only clients omzet naar een IPv6-only ISP-verbinding en die daarna weer naar IPv4 vertaalt en dat zo generiek mogelijk werkt. Generiek als onzichtbaar voor de IPv4-only client applicaties zodat protocollen zoals HTTP, BitTorrent en VoIP bruikbaar blijven. De oplossing moet te schalen zijn naar een netwerk van 5000 hosts.

Deelopdracht 1:
Inventariseer wat er nu op papier en in de praktijk voor handen is. Beschrijf de mogelijkheden en beperkingen van de huidige oplossingen.

Deelopdracht 2:
Maak van de meest relevante oplossing een proof of concept om theorie met de praktijk te kunnen vergelijken en beschrijf de beperkingen van de implementatie.

3. Aanpak

3.1 Methode

Voor het faseren en beheren van technische-infrastructuurprojecten zijn diverse methoden beschikbaar. Bij Bateau Knowledge wordt niet gewerkt met een vastgelegde standaard of methode.

De afstudeeropdracht stelt specifieke eisen aan de te gebruiken projectfaseringsmethode. Ten eerste is het projectteam zeer klein (twee personen, de afstudeerder en opdrachtgever) en is er slechts een externe partij (de Haagse Hogeschool) die procesmatig betrokken is. Ten tweede wordt niet een complete infrastructuur van een bestaande organisatie (her)ontworpen, maar wordt een oplossing gezocht voor een specifiek technisch onderdeel van een toekomstig netwerk. Dit maakt de opdracht minder concreet en enkele details onbekend. De projectmethode moet rekening houden met deze onzekerheden en de mogelijkheid laten om de technische aspecten te detailleren en overige aspecten bewust minder ver uit te werken.

Voor dit project zijn drie methodes bekeken:

- TOGAF
- DYA
- ASI

De belangrijkste eis voor dit project is dat de focus ligt op de ontwikkeling van een (deel van een) technische infrastructuur. Daarnaast gaat het om een kleinschalig project dat buiten een bestaande architectuur wordt uitgevoerd.

Daarom is er gekozen voor fasering volgens de ASI-methode. Deze biedt een duidelijk raamwerk voor het ontwerpen van een TI en focust niet meer dan voor dit project nodig is op invoering hiervan in een bestaande business- en technische architectuur.

Bovendien zijn de deelopdrachten goed in deze fasering te vatten. Deelopdracht één wordt ondergebracht in de architectuurfase waar alternatieve architecturen worden vergeleken. Deelopdracht twee kan volbracht worden in de ontwikkelfase, waar een toetsing van de goede werking van het ontwerp gevraagd wordt.

Een meer volledige uitleg van de kandidaatmethoden is te vinden in het Plan van Aanpak in bijlage 1.

3.2 Planning

De ASI-methode deelt het project in vier fasen in:

- Definitie
- Architectuur
- Ontwerp
- Ontwikkeling

Daarnaast is een initatiefase voor de afstudeeropdracht opgenomen, waarin de afstudeerder de tijd heeft om de opdracht te onderzoeken en eventueel bij te stellen.

Per fase vinden diverse activiteiten plaats en elke fase leidt tot een concreet resultaat. Ook moet een fase grotendeels afgesloten zijn voordat aan een volgende fase kan worden begonnen.

Op basis van de activiteiten per fase is de volgende planning opgesteld. In het Plan van Aanpak is een overzicht van deze activiteiten te vinden. Er is veel tijd voor de ontwikkelfase ingepland omdat hierin het testen plaatsvindt. Het is de verwachting dat dit de meeste tijd gaat kosten. Ook is de beschikbaarheid van apparatuur bepalend voor hoeveel tijd er getest kan worden en is daarom een langere periode gereserveerd.

De afstudeeropdracht heeft een duur van 18 weken. Alle werkzaamheden vinden in deze periode plaats. Na afsluiting van deze periode zijn drie weken beschikbaar om te schrijven aan het afstudeerverslag. De deadline van inleveren is verplaatst naar 6 juni in verband met Hemelvaart.

Stageweek	Fase	Product
1 - 3	Initiatie	Aangepaste opdracht
4 - 7	Definitie	Programma van Eisen
8 - 9	Ontwerp	Verslag Ontwerpfase
10 - 14	Ontwikkeling	Verslag Ontwikkeling Proof of concept
15 - 17	Afronding	
18	Uitloop i.v.m. opdracht bedrijfsprocessen	

4. Probleemanalyse

4.1 Literatuur

Tijdens de initiatiefase van de afstudeeropdracht zijn diverse boeken geraadpleegd om te achterhalen of het geschetste probleem daadwerkelijk optreedt. Uit het boek van Ahmed [3] blijkt dat er twee businessmodellen zijn waar ISP gebruik van maken. De eerste is het ISP-operated model, waarbij de ISP de regie heeft over het hele netwerk. Zowel de accesslaag als de daarop geboden diensten worden volledig door de ISP gecontroleerd. Het tweede model is het wholesale-model. Hierbij stelt een Network Access Provider (NAP) zijn accessnetwerk ter beschikking aan een derde, de Network Service Provider (NSP). De NSP gebruikt de verbinding die de NAP hem biedt om zijn diensten te leveren aan de klant. Een ISP kan zowel NAP als NSP zijn, ook als hij anderen toestaat op zijn netwerk. De uitvoering van de netwerken die ISP's hebben zijn gebaseerd op de keuze voor een van deze twee businessmodellen.

4.2 Marktanalyse

De markt voor internet is bijna verzadigd. Nederland heeft een van de hoogste breedbandpenetraties van de wereld. Volgens cijfers van het CBS had in 2010 94% van de Nederlanders toegang tot internet, waarbij 84% via een breedbandige verbinding. Dit betekent dat de markt voor internettoegang niet veel verder kan groeien. Wel kunnen er verschuivingen plaatsvinden tussen de verschillende toegangstechnieken zoals ADSL, kabel en glasvezel. Het is op dit punt dat veel concurrentie plaatsvindt. Kabelbedrijven betreden met telefonie en internet de markt die eerst voorbehouden was aan telecomgigant KPN.

Op basis van de bevindingen over de businessmodellen achter ISP is gekeken naar welke ISP's in Nederland actief zijn en watvoor model deze hebben. De markt voor telefonie en internettoegang wordt jaarlijks in kaart gebracht door de Onafhankelijke Post en Telecommunicatie Autoriteit, kortweg OPTA [4]. Hierbij wordt de markt voor breedband internettoegang verdeeld in twee stukken, namelijk de diensten die aangeboden worden via Unbundeled Local Loop (ULL) en Wholesale Breedband Toegang (WBT).

Bij ULL krijgt een dienstenleverancier fysieke toegang tot de aansluiting (local loop) van een consument of bedrijf. Over deze aansluiting kan hij vervolgens telefonie, ADSL/VDSL en andere diensten leveren, zonder hierbij gebonden te zijn aan het netwerk van de eigenaar van de aansluiting. Voor het gebruik van de lijn betaalt de leverancier een bedrag aan de eigenaar van het netwerk.

Momenteel is KPN de enige netwerkeigenaar die ULL-diensten aanbiedt aan derden. Afnemers zijn onder meer Online, BBned en Tele2. [5]

In het geval van Wholesale Breedband Toegang neemt een dienstenleverancier alleen virtuele toegang tot de aansluiting af van de eigenaar. De eigenaar, bijvoorbeeld KPN, fungeert dan als Network Access Provider, en levert de netwerktoegang tot de klant af op een centraal punt in het eigen netwerk. De ISP haakt op deze plek haar eigen apparatuur in het netwerk van de NAP en verbindt zicht vanaf die plek direct met de klant. Dit komt overeen met het Wholesale-model uit paragraaf 1.Dit heeft voor de ISP als voordeel dat ze geen eigen, landelijk dekkend, netwerk hoeft te beheren en toch klanten in heel het land kan bedienen. De NAP kan bovendien weer een ULL-klant zijn van een andere aanbieder. [6]

Tabel 4.1 geeft een overzicht van breedbandinternetaanbieders in Nederland en de toegangstechnieken die zij gebruiken.

Koper (ADSL/VDSL)	**Coax (kabel)**	**Glasvezel**
KPN (eigen netwerk)	Ziggo	KPN Wholesale
BBned (ULL)	UPC	BBned
Tele2 (ULL)	Delta	Reggefiber Wholesale
Online (ULL)	CAIW	

Tabel 4.1 - aanbieders verdeeld per toegangstechniek

KPN heeft als eigenaar van het telefoonnetwerk een bijna landelijke dekking op het gebied van ADSL, en breidt haar VDSL-netwerk rap uit. VDSL kan een hogere bandbreedte leveren dan ADSL maar vereist wel dat de afstand van de klant tot de centrale korter is. Hiervoor wordt de apparatuur van de centrale naar de wijk verplaatst, en verbonden door een glasvezelnetwerk. In samenwerking met Reggefiber levert KPN ook glasvezelaansluitingen.

De kabelbedrijven zijn gebonden aan de regio's waarvoor ze oorspronkelijk als gemeentebedrijf zijn opgericht. Op deze plaatsen hebben ze een monopolie, maar vanuit de OPTA is er druk om ook hen hun netwerk te laten openstellen.

De markt voor glasvezel bestaat vooral uit lokale initiatieven, omdat de investeringen voor de aanleg van glasvezel erg hoog zijn. Het aanleggen van een aansluiting kost 1000 a 2000 euro en moet in tien of meer jaar worden terugverdiend. Om aanleg mogelijk te maken wordt aan vraagbundeling gedaan op gemeenteniveau. Wijken waar voldoende interesse is worden in een project geheel 'verglaasd', en de klanten direct aangesloten. Bij dit netwerk wordt een dienstverlener (NSP) gezocht die vervolgens televisie, internet en telefonie kan leveren via dit netwerk.

In het overzicht ontbreken mobiele providers. Hoewel internet via mobiel steeds sneller wordt, valt deze sector nog niet onder de noemer breedband. Met het succes van smartphones en de groeiende investeringen in de datanetwerken is te verwachten dat mobiele operators binnenkort ook tot breedbandleveranciers gaan behoren.

4.3 Bedreigingen voor ISP's

Tijdens de loop van dit onderzoek, januari 2011, werd door de Internet Assigned Numbers Association (IANA), het laatste blok IPv4-adressen uitgedeeld aan de Regional Internet Registries (RIR's). [7] De RIR's zijn verantwoordelijk voor het toewijzen van IP-adressen aan providers, bedrijven en overheden. Deze adressen worden door de RIR in bruikleen gegeven, maar zelden teruggevorderd. Door de groei van internetbedrijven is er ook geen reden om adressen op te geven. Dit houdt in dat wanneer de voorraad van een RIR op is, organisaties geen nieuwe adressen meer kunnen aanvragen. In Azië is dit al realiteit. APNIC, de RIR aldaar, deelde op 14 april 2011 haar laatste adressen uit. [8] De RIR voor Europa is RIPE NCC. De verwachting is dat de voorraad die bij hen aanwezig is, genoeg is om tot januari 2012 de regio van voldoende IPv4-adressen te voorzien. Daarna is ook in Europa de spreekwoordelijke koek op, en kunnen bedrijven, waaronder ISP, hun bestaande producten waar IPv4 onderdeel van is niet zondermeer leveren.

Veelal wordt gesproken over een (zwarte) markt voor IPv4-adressen. De recente aankoop van IP-adressen uit de boedel van het failliete Canadese Nortel door Microsoft geeft aanleiding om te denken over een legale markt voor IP-adressen. Ook al zou deze markt ontstaan, en de adressen efficiënter herverdeeld worden, dan nog zou dit slechts genoeg zijn om enkele maanden groei van het internet te ondersteunen.

4.4 Oplossingsrichtingen

Efficiënter gebruik van IPv4-adressen

Een veelgehoorde oplossing voor het bovengenoemde adrestekort is Network Address Translation (NAT). NAT in IPv4 wordt ook wel NAT44 genoemd. Deze techniek wordt momenteel in vrijwel alle huishoudens toegepast als onderdeel van de (modem)router die daar geplaatst is. Huishoudens krijgen van hun ISP over het algemeen slechts één publiek IPv4-adres toebedeeld, hoewel veel klanten meerdere PC's en andere apparaten hebben die aan het internet verbonden zijn. Door in het interne netwerk (LAN) van de klant adressen te gebruiken die op het publieke internet niet mogen voorkomen, en deze in de router te vertalen naar het publieke IP-adres, kan dit adres door meer dan één computer worden gebruikt om toegang tot het internet te verkrijgen. Met NAT verliest de computer echter de mogelijkheid om open met hosts op het internet te praten. Uitgaande verbindingen zijn altijd mogelijk, maar inkomende verbindingen niet. Dat is een probleem voor peer-to-peer-applicaties zoals BitTorrent, Skype en maakt het onmogelijk om zelf een mail- of webserver te draaien. In de huidige routers kan dit probleem worden opgelost door handmatig uitzonderingen voor deze applicaties toe te staan. Ook zijn er protocollen zoals UPnP die dit proces automatiseren.

NAT kan worden opgeschaald door deze techniek niet alleen toe te passen op één huishouden, maar op tientallen of honderden abonnees. Deze techniek heet NAT444, wat gelezen kan worden als NAT44 + NAT44. De klant krijgt in dit scenario geen publiek IPv4-adres meer, maar slechts een privé adres uit een netwerk dat hij deelt met de andere abonnees. Als een computer een verbinding wil opzetten naar het internet vindt tweemaal de vertaalslag naar een ander adres plaats: eenmaal in de router van de klant naar de ISP, en

nogmaals in het netwerk van de ISP zelf. Dit heeft voor de ISP als voordeel dat hij met een beperkte voorraad IPv4-adressen een groter aantal klanten kan voorzien van internet. Dit brengt echter voor hem ook grotere kosten met zich mee, omdat het netwerk moet worden aangepast. Daarnaast verliest de klant in dit scenario de mogelijkheid om zelf servers te draaien, of dit door zijn applicaties te laten doen m.b.v. UPnP. Ook is het onbekend hoe applicaties die werken door één NAT zich gedragen als ze moeten werken via NAT444.

Er zitten grenzen aan het aantal gebruikers dat via één IP-adres gebruik kan maken van het internet. Via één IP-adres is het mogelijk om tegelijkertijd maximaal 65.000 TCP- en UDP-verbindingen op te zetten. TCP wordt gebruikt door HTTP, het protocol voor het World Wide Web. UDP wordt ondermeer gebruikt door applicaties zoals Skype. Metingen geven aan dat één gebruiker tot wel zeshonderd verbindingen tegelijkertijd open kan hebben staan. [9] Met een gemiddelde van 5 computers per huishouden, en een maximum van 500 gelijktijdige verbindingen per computer zou een ISP één IP-adres door 26 abonnees tegelijk kunnen laten gebruiken. Serieus onderzoek naar de grenzen van adresdeling ontbreekt echter. Wel kan aangegeven worden dat er een maximum is, en dat dit tussen de tientallen en honderden gebruikers per adres ligt.

NAT444 is daarmee een lapmiddel dat de gebruiksduur van IPv4-adressen tijdelijk oprekt, ten koste van de mogelijkheden voor de gebruiker en met extra complexiteit in het netwerk.

Overstappen op IPv6

IPv6 is gelijktijdig ontwikkeld met NAT, maar als langetermijnoplossing. Het idee was dat er voldoende tijd zou zijn om een nieuw protocol in te voeren, en op termijn het oude IPv4 uit te schakelen. Het succes van NAT heeft echter de noodzaak om IPv6 in te voeren verkleind, tot dit jaar de IPv4-adressen opraakten. Als alle computers die aan het internet verbonden zijn gebruik zouden maken van IPv6, zou het IPv4-probleem opgelost zijn. Dit is echter een utopie. Er is veel geïnvesteerd in IPv4, en dit maakt invoering van IPv6 tot een kip-ei-probleem: zolang er geen gebruikers zijn die IPv6 willen, zijn er geen aanbieders die het leveren. Zolang het niet geleverd kan worden, zal geen gebruiker er voor kiezen.

IPv4-connectiviteit naast IPv6

Deze vicieuze cirkel kan worden doorbroken door het opraken van de IPv4-adressen. De IPv6-adressen zijn wel vrij beschikbaar, en zullen naar verwachting nooit opraken. Door IPv6-connectiviteit aan te bieden kunnen providers de lasten op het IPv4-netwerk verlichten. Het 'IPv6-internet' is echter nog zeer klein. Vanwege de grote groep gebruikers op IPv4 voelen aanbieders van diensten niet de noodzaak om hun websites en applicaties ook voor IPv6 geschikt te maken. Dit is een tweede kip-ei-probleem. Het is te verwachten dat als IPv6-verbindingen beschikbaar komen, dienstenaanbieders op internet (Youtube, Facebook, Microsoft) hun websites en applicaties ook via IPv6 bereikbaar maken.

Met het oog op IPv6 als opvolger van IPv4, en het gebrek van content op het IPv6-internet is het aanbieden van verbindingen naar beide netwerken voor ISP's een mogelijk overgangsscenario. Het opraken van de IPv4-adressen maakt het echter onmogelijk om deze verbindingen native dual-stack te maken. Dat wil zeggen: zodat de gebruiker zowel een IPv4- als een IPv6-adres krijgt. Er zal moeten worden gezocht naar een manier om dit samen te doen, via één verbinding, maar zonder de dubbele kosten die aan het beheer van twee protocollen verbonden zijn.

4.5 Aanpassing opdracht

Op basis van de bevindingen van de probleemanalyse is overlegd met de opdrachtgever. Er is gezocht naar welke ISP's als eerste tegen dit probleem aanlopen. Dat zijn snelgroeiende bedrijven zoals mobiele operators, en glasvezelnetwerken die vele duizenden klanten in één keer aansluiten.

Het is niet mogelijk gebleken om een samenwerking met een (mobiele) ISP aan te gaan om voor hen dit probleem op te lossen. De oorzaak hiervan is zeer waarschijnlijk dat ISP's zelf nog niet weten hoe ze dit probleem gaan aanpakken, de technieken die hiervoor nodig zijn nog niet voldoende ontwikkeld zijn, en dat de ondersteuning van netwerkleveranciers achterblijft. Uit concurrentieoverwegingen is het voor hen slimmer om niet te praten over het probleem en hoe ze dit denken op te gaan lossen.

Daarom is de opdracht aangepast om het ontwikkeltraject te doorlopen namens een fictieve glasvezel-ISP. Op deze manier kan een voorbeeld worden genomen aan hun netwerk, de afwegingen die daarin plaats vinden en de problemen die in een dergelijk netwerk te verwachten zijn. Het netwerk van de ISP zal worden gebaseerd op informatie over bestaande ISP's die openbaar is.

5. Vergelijkend onderzoek

5.1 Literatuuronderzoek

Voor dit onderzoek is de bibliotheek van de TU Delft geraadpleegd. Er is gezocht naar boeken die IPv6 beschrijven, in het Nederlands of Engels, geschreven in de laatste 10 jaar (2000 - 2011). Daarnaast is ook gezocht naar boeken over computernetwerken waarin IPv6 is opgenomen.

Dit leverde een lijst van negen titels op.

De transitietechnieken die in deze boeken beschreven staan zijn vooral gericht op het verkrijgen van IPv6-connectiviteit in een IPv4-netwerk. Dat is niet zo verwonderlijk, aangezien IPv4 momenteel het meest gebruikt is op internet. Veelgenoemde methoden zijn 6to4, Teredo en ISATAP. Deze worden niet verder behandeld omdat ze een ander probleem oplossen dan in de opdracht is omschreven.

Een totaaloplossing die IPv4-connectiviteit over een IPv6-netwerk biedt is nergens beschreven. Wel staan in de boeken over IPv6 een aantal technieken die als bouwstenen kunnen dienen voor een oplossing. De beschrijving van deze technieken is te vinden in het document Architectuurfase. Een bewerkte versie van dit overzicht is namens Stichting IPv6 Nederland gepubliceerd op hun website. [10]

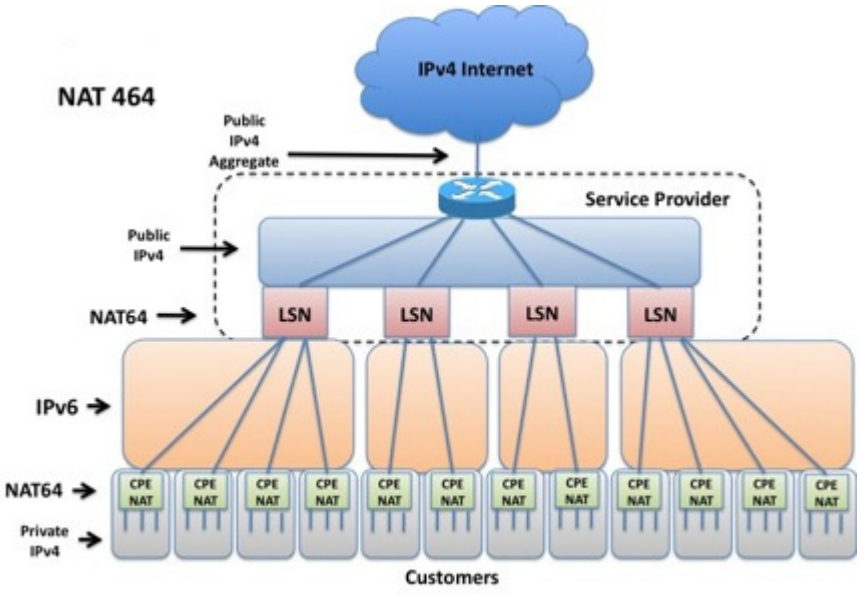
Naar aanleiding van de publicatie zijn diverse reacties en aanvullingen ontvangen. Een van de reacties betrof de techniek LISP, die volgens de voorstanders toegepast kan worden bij de overgang van IPv4 naar IPv6. Deze is in de overweging meegenomen.

Op basis van de mogelijkheden van de gevonden technieken is een shortlist van kandidaatoplossingen samengesteld. De technieken die een ISP zou kunnen toepassen zijn:

- NAT464
- DS-Lite
- A+P
- 4RD

5.2 NAT464

De techniek NAT464 is vergelijkbaar met NAT444. Tussen de klant en de provider wordt IP-verkeer vertaald. Tussen de provider en het internet vindt wederom een vertaling plaats. Bij NAT464, de middelste zes verraadt het al, is de tussenliggende adresfamilie geen IPv4 maar IPv6.



Afbeelding 5.1 - Overzicht van NAT464 © Jeff Doyle [11]

De CPE van de klant vertaalt IP-verkeer van IPv4 naar IPv6 en stuurt dit via het IPv6-netwerk naar een NAT-router (LSN) van de ISP. Deze voert de omgekeerde vertaling uit en gebruikt hiervoor publieke IPv4-adressen. Deze techniek heeft als voordeel dat het hele access netwerk van de ISP gebaseerd is op IPv6. Vertaald verkeer onderscheidt zich van ander verkeer door het gebruikte IPv6-adres.

Voordelen

- het hele access netwerk van de ISP (met uitzondering van de NAT-routers) gebruikt één protocol, IPv6
- IPv4 bevindt zich alleen aan de rand van het netwerk
- klanten kunnen het publieke internet benaderen

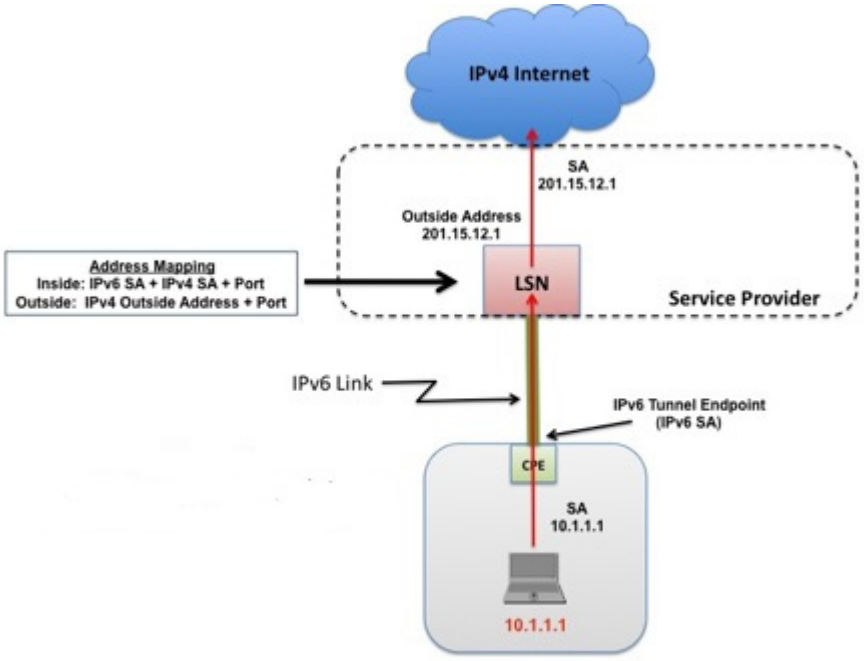
Nadelen

- er zijn geen implementaties van NAT464 bekend
- de techniek NAT464 is niet uitgewerkt in een specificatie; het is een idee
- aanpassing van de CPE is vereist
- het plaatsen van NAT-routers maakt het ISP-netwerk gecompliceerder
- het gebruik van dubbel NAT maakt het voor sommige applicaties onmogelijk om goed te werken
- het is niet mogelijk om als klant een server te draaien

In een mail van Alain Durand (lid van de IETF Softwires Working Group) bevestigt hij dat deze methode onderzocht is, maar door de Working Group is afgeschreven. In plaats daarvan is gekozen om tunneling verder te ontwikkelen. Dit heeft geresulteerd in de ontwikkeling van Dual-Stack Lite.

5.3 Dual-Stack Lite

Dual-Stack Lite lijkt in vele opzichten NAT464, maar pakt het transport van IPv4-verkeer naar de NAT-router van de provider fundamenteel anders aan. In plaats van verkeer te vertalen, wordt het getunneld in IPv6.



Afbeelding 5.2 - Werking van Dual-Stack Lite © Jeff Doyle [12]

DS-Lite lost het probleem van dubbel NAT op door de CPE aan te passen en IPv4-verkeer te tunnelen over het IPv6-netwerk van de ISP, naar de rand van het netwerk. Daar wordt NAT44 toegepast. De NAT-functionaliteit die onder controle van de klant zelf was, wordt verplaatst naar het netwerk van de ISP. Verkeer van alle computers van een bepaald aantal klanten wordt door middel van NAT vertaald naar publieke IP-adressen.

De ontwikkeling van Dual-Stack Lite gaat erg snel omdat het idee simpel is, en gebaseerd is op bestaande technieken. Hierdoor is de techniek al door enkele fabrikanten in hun producten geïmplementeerd.

Voordelen

- het hele access netwerk van de ISP (met uitzondering van de NAT-routers) gebruikt één protocol, IPv6
- IPv4 bevindt zich alleen aan de rand van het netwerk

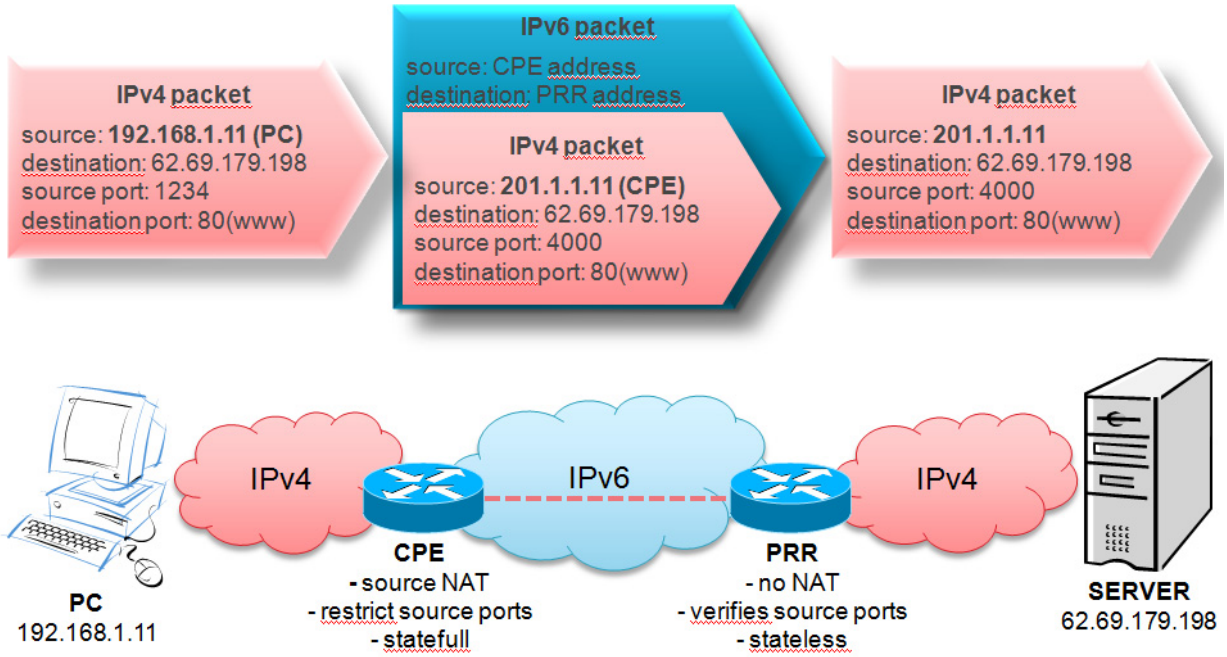
- klanten kunnen het publieke internet benaderen
- gebaseerd op bestaande technieken

Nadelen

- aanpassing van de CPE is vereist
- NAT bij de ISP maakt het netwerk gecompliceerder
- het is niet mogelijk om als klant een server te draaien

5.4 A+P

Bij Address Plus Port (A+P) wordt het adrestekort opgelost door gebruikers niet een volledig IPv4-adres ter beschikking te stellen, maar slechts een deel daarvan. Dit kan door de adressering uit te breiden naar de poortnummers. Klant A kan bijvoorbeeld gebruikmaken een IP-adres met poorten 10.000 tot 30.000 en klant B van poorten 30.000 tot 50.000 op datzelfde IP-adres. Deze techniek werkt, net als NAT444 en Dual-Stack Lite, alleen voor de protocollen TCP en UDP, die zijn gebaseerd op poortnummers.



Afbeelding 5.3 - Werking van A+P © Stichting IPv6 Nederland

De CPE van klant A weet naar welke poorten op welk publiek IP-adres hij mag vertalen. Vertaald verkeer wordt in een tunnel naar de Port Range Router (PRR) gestuurd. De PRR is een machine in het netwerk van de ISP en heeft als taak om het verkeer dat uit te tunnel komt te controleren op het gebruik van de juiste adressen en poortnummers. Vervolgens wordt het verkeer naar het publieke internet gestuurd.

Door NAT aan de gebruikerszijde tot bepaalde port ranges te beperken, kan bespaard worden op NAT in het providernetwerk. Dit heeft voordelen m.b.t. de grootte van de NAT-tabel, logging en performance. De specificatie van A+P biedt de mogelijkheid om de PRR de

functie van AFTR voor Dual-Stack Lite te laten vervullen. Hierdoor kan een mengeling van Dual-Stack Lite en A+P apparatuur worden gebruikt in hetzelfde netwerk.

Voordelen

- het hele access netwerk van de ISP (met uitzondering van de NAT-routers) gebruikt één protocol, IPv6
- IPv4 bevindt zich alleen aan de rand van het netwerk
- klanten kunnen het publieke internet benaderen
- klanten hebben controle over de NAT-functie (kunnen een server draaien)

Nadelen

- aanpassing van de CPE is vereist
- plaatsing van de PRR's maakt het netwerk gecompliceerder, doch minder dan andere technieken

5.5 4RD

IPv4 Residual Deployment (4RD) is opgesteld door Olivier Vautrin als tegenhanger van 6rd. 4RD is een manier om IPv4-verkeer te tunnelen over een IPv6-netwerk. [13] Verkeer wordt via IPv6 verzonden tussen de router van de gebruiker, de 4rd Customer Edge (CE), en de Border Router (BR) van de ISP. Als er gebruik wordt gemaakt van een algoritme om IPv4-adressen en poorten in IPv6-adressen te coderen, dan is dit stateless. De ISP hoeft dan geen gegevens bij te houden over de sessies van de gebruiker. Wordt er geen gebruik gemaakt van deze codering, dan moet voor elke verbinding een aparte NAT-sessie worden aangemaakt (stateful NAT44) en is de methode gelijk aan DS-Lite.

Sinds het verschijnen van deze draft is er niet meer aan 4rd ontwikkeld en is het werk verplaatst naar de IETF Intarea Working Group. Binnen deze groep is een nieuwe draft geschreven door Remi Depres. [14] In de draft van Remi Depres is 4RD uitgebreid met de mogelijkheid om over een ISP-netwerk te functioneren dat ofwel IPv6-only is, of RFC1918 (private IPv4 addressing) of allebei. In het geval van een RFC1918-netwerk wordt IPv6-connectiviteit geleverd door middel van 6rd, dat door Despres ontwikkeld is om IPv6 in een IPv4-netwerk te leveren [15]. Wordt zowel RFC1918-adresruimte als IPv6 gerouteerd dan vervalt het nut van 6rd en kan IPv6-verkeer direct plaatsvinden tussen hosts en het internet (end-to-end).

Er is geen eenduidige specificatie van 4rd. De hierboven beschreven documenten zijn drafts op persoonlijke titel van de auteurs, en worden dus niet (nog) door een Working Group gesteund.

Voordelen

- toepasbaar in netwerken waar IPv4 en/of IPv6 gebruikt wordt
- klanten kunnen het publieke internet benaderen

Nadelen

- ontwikkeling van de standaard wordt niet gesteund door een Working Group
- aanpassing van de CPE is vereist
- plaatsing van de Border Routers maakt het netwerk gecompliceerder, doch minder dan andere technieken

5.6 Oplossingskeuze

Voor het uitvoeren van de opdracht is het nodig om uit de vier kandidaatoplossingen één techniek te kiezen die in een proof of concept uitgewerkt kan worden. De volgende criteria zijn hiervoor van belang:

- volwassenheid van de standaard
- beschikbaarheid van implementatie

De criteria worden met punten beoordeeld. De waardering is aangegeven in de bijbehorende tabellen.

Volwassenheid van de standaard
Is de standaard stabiel? Is de specificatie voor iedereen beschikbaar en voldoende uitgewerkt? Dit zijn factoren die van belang zijn voor een goede specificatie en voor de ontwikkeling van implementaties die compatibel met elkaar zijn.

Tabel 5.1 – Puntenwaardering volwassenheid

Punten	**Voorwaarden**
5	De specificatie is openbaar (bijvoorbeeld een IETF RFC) en wordt actief ontwikkeld door de IETF of een ander instituut
4	De specificatie is een draft en wordt actief ontwikkeld door de IETF of een ander instituut met de bedoeling hier een standaard van te maken (bijvoorbeeld als onderdeel van het charter van een Working Group)
3	De specificatie is een draft en wordt actief ontwikkeld buiten het charter van een IETF Working Group of door derden
2	De specificatie is openbaar (bijvoorbeeld een IETF RFC) maar er vindt geen verdere ontwikkeling plaats
1	Er is een specificatie, maar deze is niet publiekelijk beschikbaar
0	Er is geen specificatie gepubliceerd

Beschikbaarheid van implementaties
Voor het uitwerken van een proof of concept, waarbij geen eigen software- of hardwareontwikkeling plaats vindt, is het cruciaal dat er implementaties beschikbaar. Daarom is gezocht naar leveranciers die de standaard implementeren. Dit kunnen ook

opensource-initiatieven zijn, of referentie-implementaties. Alle onderzochte technieken vereisen zowel een aanpassing aan de router van de ISP-klant (CPE) als in het netwerk van de ISP zelf. Het vermelde puntenaantal bestaat uit twee getallen, waarvan het eerste aangeeft hoeveel CPE-implementaties er gevonden zijn, en het tweede het aantal implementaties van de techniek voor in het ISP-netwerk.

Samenstelling van de eindscore
De eindscore van een techniek wordt bepaald door beschikbaarheid van implementaties. Zowel van de ISP- als de klantkant moet minstens één implementatie beschikbaar zijn. Het aantal punten voor de volwassenheid van de techniek wordt hierbij opgeteld. Op deze manier geeft een techniek die niet volwassen is, maar wel veel wordt geïmplementeerd, toch kans om onderzocht te worden. Tegelijk wordt bij een gelijk aantal implementaties voorrang gegeven aan technieken die volwassener zijn.

Score
Ten tijde van het uitvoeren van dit onderzoek zijn de websites van de IETF en diverse fabrikanten van netwerkapparatuur geraadpleegd. Ook is direct contact gezocht met diverse bedrijven om te informeren naar de beschikbaarheid van implementaties. De bevindingen zijn vermeld in de bijlage ‘Architectuurfase’.

Op basis van de eindscore wordt Dual-Stack Lite verkozen als oplossingsrichting.

Tabel 5.2 – Score oplossingsrichtingen

Techniek	**Volwassenheid**	**Beschikbaarheid**	**Eindscore**
NAT464	0	0 + 0	0
Dual-Stack Lite	4	5 + 6	15
A+P	3	1 + 1*	3
4RD	3	0 + 0	3

* de implementatie van A+P door Orange Bejing is pas na het uitvoeren van dit deel van het onderzoek bekendgemaakt, en kan dus niet worden meegeteld in de eindscore.

6. Dual-Stack Lite proof of concept

6.1 Doelstelling

De doelstelling van het proof of concept is om te bepalen of Dual-Stack Lite een geschikte oplossing is voor ISP's met een tekort aan IPv4-adressen. Hiertoe wordt een testopstelling gebouwd om te bepalen of Dual-Stack Lite in staat is om IPv4-adressen te delen op een manier die generiek is. Dat wil zeggen: transparant voor de eindgebruiker.

Dual-Stack Lite is specifiek bedoeld voor IPv6-only access netwerken waar de service provider wel toegang tot het IPv4 internet wil bieden aan hosts die niet of moeilijk te upgraden zijn. Dit is bijvoorbeeld het geval bij computers met oudere operating systems, netwerkprinters, spelcomputers en embedded devices zoals IP-camera's en IP-telefoons. Het is de verwachting dat deze devices de komende jaren allemaal nog ondersteund moeten worden.

6.2 Specificatie Dual-Stack Lite

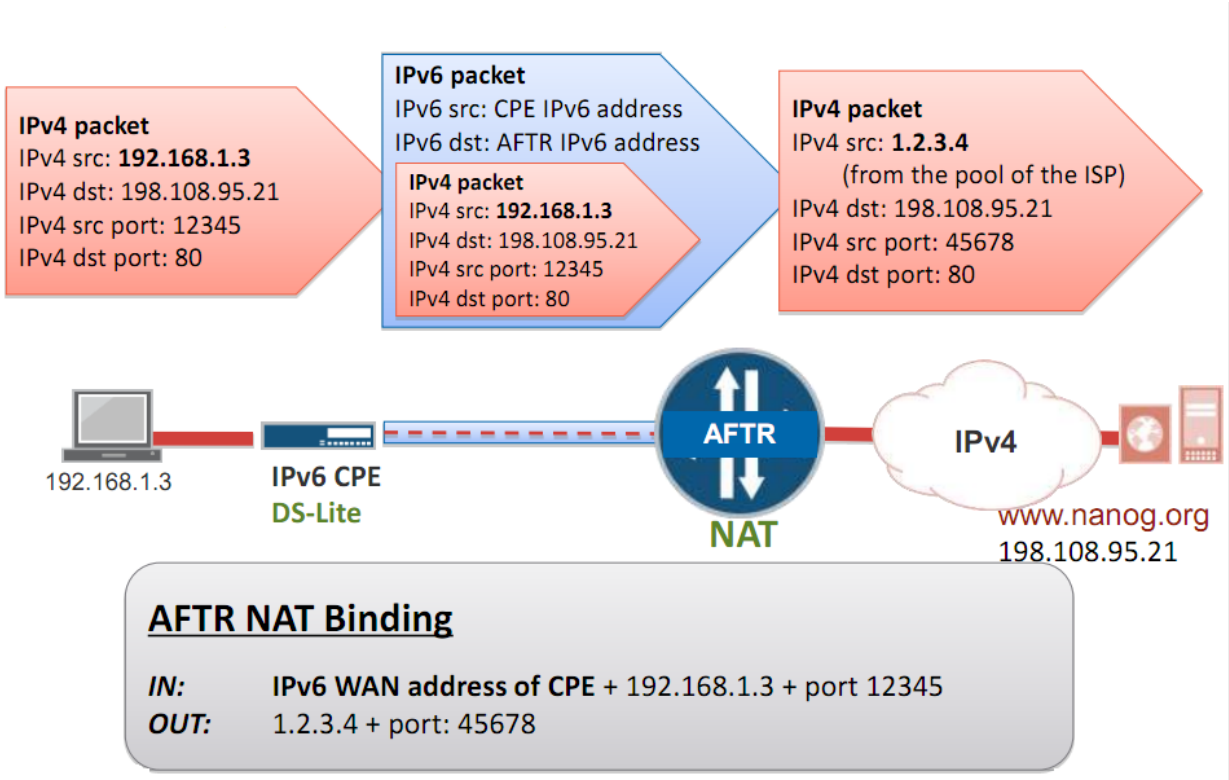
Formalisatie van de Dual-Stack Lite specificatie

De standaardisatie van Dual-Stack Lite vindt plaats binnen de Softwires Working Group van het IETF. Leden van de groep werken samen om de draft uit te werken tot een tekst die als RFC kan worden aangenomen. Op het moment dat deze RFC gepubliceerd wordt, kan de specificatie als standaard worden aangemerkt. Daarna is er de mogelijkheid om fouten in de tekst aan te geven. Voor wezenlijke veranderingen moet echter een nieuwe RFC worden gepubliceerd, die de oude vervangt. Het is daarom belangrijk om de tekst de eerste keer goed te krijgen. Het is onbekend wanneer Dual-Stack Lite wordt gepubliceerd als RFC.

Werking van Dual-Stack Lite

Het doel van Dual-Stack Lite is om IPv4-connectiviteit te bieden aan IPv4 of dual-stacked hosts achter een consumentenrouter (CPE), zonder dat deze router over een eigen, publiek, IPv4-adres beschikt. De CPE krijgt alleen een IPv6-adres van de provider. Op het interne netwerk is de CPE als DHCP-server actief voor de hosts die vragen om een IPv4-adres. Dit is een adres uit de private range, de zogenoemde RFC 1918 adresruimte. Deze adressen kunnen alleen maar in afgesloten netwerken worden gebruikt, en worden niet gerouteerd op internet. Communicatie met het IPv4-internet is alleen mogelijk als er een vorm van Network Address Translation (NAT) plaats vindt. De NAT-functie is bij Dual-Stack Lite verplaatst van de CPE naar een machine in het netwerk van de ISP. Deze heeft de naam AFTR, en voert namens de ISP-klanten NAT uit op hun IPv4-verkeer.

De IPv4-pakketten kunnen echter niet over het netwerk naar de AFTR worden verstuurd. Omdat private adresruimte door iedereen in zijn eigen netwerk gebruikt mag worden, is de kans groot dat meerdere klanten op het netwerk van de ISP dezelfde adressen gebruiken. Daarnaast gaat het hier om IPv4-pakketten. Deze zijn niet compatible met IPv6. Er moet dus een truc worden uitgehaald om deze pakketten over het IPv6-netwerk te transporteren. Hiervoor maakt Dual-Stack Lite gebruik van tunneling, of liever gezegd, encapsulation. Deze methode is vastgelegd in RFC2473. [16] De onvertaalde IPv4-pakketten van de clients, die naar hosts op het internet gericht zijn, worden door de CPE verpakt in een IPv6-pakket en doorgestuurd naar de AFTR.



Afbeelding 6.1 - Werking van Dual-Stack Lite © Alain Durand (Juniper)

De AFTR ontvangt het pakket, pakt het uit en vertaald het bronadres naar een publiek IPv4-adres. In de NAT-binding wordt naast het IPv4-adres en -poort van de client ook het IPv6-adres van de CPE geregistreerd. Als er een pakket binnenkomt op het publieke IPv4-adres van de AFTR wordt in de NAT-bindingtabel gezocht naar een regel die matcht met het publieke IP en poortnummer waar het pakket aan gericht is. De AFTR voert met behulp van het opgeslagen privé IPv4-adres destination NAT uit op het pakket, en pakt het in een IPv6-header in om het naar de CPE te sturen. De CPE pakt het pakket uit, en verstuurt het naar het IPv4-adres van de client.

Effectief wordt de IPv4 NAT-functie die normaliter in consumentenrouters is ingebouwd, verplaatst naar een centraal punt in het providernetwerk. Vanaf dit punt kan door veel meer clients een publiek IP-adres gedeeld worden dan in de huidige netwerken mogelijk is. Dit is een voordeel voor de ISP, die hierdoor met minder IPv4-adressen toekan, dan hij klanten heeft. Deze situatie zal zich na het opraken van de IPv4-adressen steeds vaker voordoen. Maar er zitten ook nadelen aan deze oplossing. Enkele nadelen zijn inherent aan het delen

van adressen tussen meerdere aansluitingen, en beschreven in een nog te publiceren RFC. [17] Voorbeelden van nadelen zijn:

- impact op performance,
- beperkt aantal poorten beschikbaar,
- beveiligingsproblemen,
- issues met fragmentatie
- introduceren van Single Points of Failure (SPOF),
- black listing is verminderd effectief

Elke techniek die een vorm van adresdeling gebruikt, dus ook Dual-Stack Lite, zal deze problemen moeten voorkomen of minimaliseren om een geschikt alternatief te zijn voor native IPv4.

Onvolkomenheden

Omdat de specificatie van Dual-Stack Lite nog in draft is, wordt hier nog volop aan geschreven. In de maanden dat de specificatie voor dit onderzoek is geraadpleegd hebben vijf revisies plaatsgevonden, die het document soms verduidelijken, maar in andere gevallen ook ambigu maken.

Zo werd een verschil gevonden in de manier waarop volgens de specificatie fragmentatie moet plaatsvinden op verkeer dat door tunneling van IPv4 in IPv6-packets te groot wordt. De draft gaat hier anders mee om dan RFC 2473, waarnaar in de specificatie verwezen wordt.

6.3 Samenstelling internetverkeer

Een van de vragen die naar aanleiding van het lezen van de specificatie voor Dual-Stack Lite naar boven kwam was: Is het zo belangrijk om fragmentatie te voorkomen? Komt fragmentatie zoveel voor, en wat is dan de impact hiervan? Is het dus wel nodig om hiertegen maatregelen te nemen?

Met behulp van Wireshark kon op een Windows 7-computer worden vastgesteld dat fragmentatie zelden voorkomt. TCP-verkeer houdt zich door het afstemmen van de Maximum Segment Size (MSS) aan de MTU van de links tussen de communicerende hosts. UDP-verkeer is over het algemeen klein, en kent daardoor weinig fragmentatie.

Voor een meer representatieve meting was verkeer nodig van meer hosts en hiervoor is samenwerking gezocht met de hosting en transit provider We Dare. Met hulp van de monitoring tools van We Dare is een profiel opgesteld van verkeer dat door een ISP wordt afgehandeld.

De eerste meting had een duur van 48 uur op twee werkdagen vond plaats op een link naar diverse bedrijven, waardoor het profiel kan afwijken van wat er in het netwerk van een consumenten-ISP wordt getransporteerd. Uit deze meting blijkt dat een groot deel van verkeer uit TCP- en UDP-verkeer bestaat, maar dat ook IPsec en GRE voorkomen. Naar verwachting is dit bij consumenten anders. De gemiddelde lengte van een UDP-pakket in

deze meting is 30446 bytes. Dit betreft de lengte van het gehele pakket, niet van de afzonderlijke fragmenten.

Bij een tweede meting van 24 uur in het core netwerk van We-Dare is alleen gekeken naar de lengte van UDP-pakketten, waarbij onderscheidt werd gemaakt op doelpoortnummer. Uit deze meting volgt dat de gemiddelde lengte van UDP-verkeer verschilt per protocol. De gevonden waarden lopen sterk uit een, maar zijn allen groter dan 1500 bytes. Hieruit kan worden geconcludeerd dat bij het testen van de netwerkprestaties rekening moet worden gehouden met UDP-packets die groter zijn dan de MTU van 1500.

Nadat de metingen waren afgerond kwam een bug in de gebruikte software naar voren. Bij het opslaan van de lengte van de IP-packets (2 bytes) worden de twee waarden omgekeerd opgeslagen. Het gevolg hiervan is dat de gerapporteerde waarde niet betrouwbaar is. De resultaten van beide metingen zijn daarom niet representatief. De conclusie dat rekening gehouden moet worden met fragmentatie van UDP is wel gebruikt om testcases op te stellen. Na uitvoeren van de tests bleek dit wel degelijk van belang om op te testen. Door deze testcase werd een ontbrekende feature in een Dual-Stack Lite-implementatie van A10 ontdekt.

6.4 Dual-Stack Lite Test Lab

Het proof of concept is uitgewerkt in een Dual-Stack Lite Test Lab. In samenwerking met de Stichting IPv6 Nederland zijn diverse leveranciers benaderd om apparatuur ter beschikking te stellen die Dual-Stack Lite implementeert.

Er is getest met de AVM Fritz!Box 6360 Cable en de A10 Networks AX2500. Beide leveranciers hebben specifieke wensen geuit voor de te testen functionaliteit. Daarnaast gelden de eisen die de fictieve glasvezel-ISP aan Dual-Stack Lite stelt. Een Linksys 160NL deed dienst als tweede CPE.

Voor deze apparatuur is een testopstelling gemaakt die de tester in staat stelt om alle verkeer dat door de apparatuur gaat te monitoren. Hierdoor kan precies bepaald worden of en waar iets fout gaat.

Op basis van de wensen van AVM, A10 Networks en het Programma van Eisen zijn testcases ontwikkeld die controleren of de apparatuur aan de gewenste specificaties voldoet.
De tests vinden plaats in een aparte, beschermde, omgeving. Voor tests waarbij communicatie met diensten op het internet noodzakelijk is wordt de testopstelling gekoppeld aan het publieke internet. Op deze manier is getest met websurfen, Skype en OpenVPN.

6.5 Testresultaten

AVM Fritz!Box 6360 Cable

De AVM is een snel modem op zowel het gebied van IPv6 als Dual-Stack Lite. Op IPv6-gebied komen de snelheden in de buurt van wat de gigabit ethernet interfaces maximaal aankunnen, bij Dual-Stack Lite is dat ongeveer de helft.

Door aanpassing van de MSS clamping functie is mogelijk een nog hogere performance te bereiken.

De router implementeert (nog) niet de DHCPv6-optie waarmee ISP's het adres van de AFTR kunnen configureren.

Linksys 160NL

De implementatie van DS-Lite door Comcast/Xavient die op de Linksys draait is minder gebruiksvriendelijk dan de AVM, maar lijkt goed te werken. Deze is niet zo uitvoerig getest omdat het een referentie-implementatie betreft.

De performance van deze router op het gebied van IPv6 blijft sterk achter bij de AVM. Hier lijkt de bandbreedte gelimiteerd door de Fast Ethernet-poorten (100 Mbps) van de 160NL. De performance via Dual-Stack Lite is niet nauwkeurig gemeten, maar lijkt sterk minder dan die van de AVM. Dit wordt waarschijnlijk veroorzaakt door beperkingen van de processor of de onderliggende software.

A10 Networks AX2500

De AFTR-functionaliteit van de AX2500 is op de meeste punten goed uitgevoerd. De performance die met behulp van de AX2500 gehaald kan worden is hoog, en geeft waarschijnlijk vooral de snelheidsbeperking van de geteste CPE aan. Het geteste NAT-gedrag voldoet aan de specificaties die Dual-Stack Lite voorschrijft.

Tussen de CPE en de AFTR ondersteunt de AX2500 echter geen fragmentatie op IPv6-niveau. De standaard vereist dit wel, en de router van AVM maakt hier ook gebruik van. Hierdoor komen grote pakketten die door de router gefragmenteerd worden niet aan op de AFTR en ontstaat een zwart gat voor dit verkeer. Een issue met port forwarding zorgt er voor dat inkomende verbindingen nooit aankomen op de CPE.

Als deze twee beperkingen worden opgelost kan de AX2500 met vertrouwen worden getest voor gebruik in een bestaand providernetwerk.

6.6 Conclusie

Op basis van de testresultaten kan de conclusie worden getrokken dat de geteste implementaties van Dual-Stack Lite in staat zijn om IPv4-connectiviteit te bieden die vergelijkbaar is met de mogelijkheden van de huidige breedbandverbindingen.

Door een bug in de AFTR werkt port forwarding nog niet. Klanten kunnen daardoor geen servers draaien. Ook bestaat er voor grote UDP-packets het risico dat deze door de AFTR niet worden verwerkt. Naar verwachting kunnen deze problemen snel door de leveranciers worden opgelost, waarna met deze apparatuur een functionerend netwerk te bouwen is.

De geteste router heeft goede prestaties op het gebied van IPv6. De AVM is goed op het gebied van prestaties en gebruiksvriendelijkheid. Dit betekent dat met deze router een geschikte kandidaat is voor ISP’s die hun netwerk op basis van IPv6 en Dual-Stack Lite willen bouwen.

7. Conclusie en aanbevelingen

7.1 Conclusie

Het is voor Internet Service Providers met een eigen access netwerk mogelijk om IPv4-connectiviteit te leveren in een IPv6-only-netwerk. Dual-Stack Lite is momenteel een geschikte oplossing voor deze opgave. De verbinding die met deze techniek geleverd wordt is vergelijkbaar met de huidige internetverbindingen van ISP's. Leveranciers van netwerkapparatuur hebben producten beschikbaar om een ISP-netwerk aan te passen of in te richten dat voldoet aan de gestelde voorwaarden.

Wel is het noodzakelijk dat de apparatuur door de leveranciers aangepast wordt om foutsituaties te voorkomen. De tests in het testlab hebben aangetoond dat grondig testen van apparatuur noodzakelijk is om een succesvolle uitrol te verzekeren.

7.2 Aanbevelingen aan Bateau Knowledge

Op basis van de ervaringen en het onderzoek gedaan tijdens het uitvoeren van de afstudeeropdracht worden de volgende aanbevelingen gedaan:

- Raadt ISP's aan om te testen met Dual-Stack Lite als mogelijke oplossing voor hun adrestekort.

- Blijf investeren in kennis van Dual-Stack Lite, de alternatieven en IPv6 om consultancy op dit gebied te bieden.

- Houdt zicht op de alternatieven die in dit gebied worden ontwikkeld (A+P, 4RD). Dual-Stack Lite is op dit moment een goede oplossing, maar verdere ontwikkelingen van de standaarden en de komst van implementaties kunnen ervoor zorgen dat providers de focus verleggen naar concurrerende technieken.

7.3 Aanbevelingen aan Stichting IPv6 Nederland

Het werk dat voor de Stichting IPv6 Nederland verricht is in het testlab heeft geleidt tot de volgende aanbevelingen:

- Ontwikkel het Dual-Stack Lite testlab uit tot een omgeving waarin zowel leveranciers als providers terecht kunnen met hun apparatuur en vragen. De kennis die hiermee wordt opgedaan kan gebruikt worden om providers te adviseren over hun transitiestrategie. De feedback naar leveranciers verbetert hun implementaties en dus de kans van slagen van Dual-Stack Lite als transitiemechanisme. Het testlab kan Dual-Stack Lite-certificatie aanbieden met bijbehorend logo, zodat routerfabrikanten

op de verpakking van hun routers kunnen tonen dat ze hiervoor geschikt zijn.

- Testen van Dual-Stack Lite apparatuur is sterk verwant aan het testen van (modem)routers. Op dit gebied kan mogelijk samenwerking met RIPE plaatsvinden, die enkele malen per jaar het ‘CPE Survey’ op haar website publiceert. In dit onderzoek worden modemrouters beoordeeld op hun IPv6-functies.

8. Evaluatie

8.1 Opdracht

De opdracht was een uitdaging om aan te pakken en uit te voeren. De manier waarop deze is opgesteld heeft geleidt tot een project waarin veel kennis is opgedaan over oplossingen het IPv4-adrestekort het hoofd te bieden. De voorgestelde oplossingsrichting op basis van IPv6 is valide gebleken.

8.2 Aanpak

De keuze voor de ASI-methode heeft positief uitgepakt. Op basis van het ASI-rapport zijn de producten opgesteld die illustreren hoe een glasvezel-ISP zou kunnen komen tot het invoeren van Dual-Stack Lite. De informatie die tijdens dit proces is gevonden is gebruikt om tot de beantwoording van de hoofdvraag te komen.

Zoals het lot is van een planning wordt deze direct door de realiteit achterhaald. Zo ook in het geval van deze afstudeeropdracht. De definitiefase liep goed, maar tegen het eind van de architectuurfase begon de planning uit te lopen. Dit werd onder andere veroorzaakt door meer onderzoek naar transitiemechanismen.

Er is in dit gebied veel informatie, maar weinig beknopt beschreven. Ook zijn de standaarden nog volop in ontwikkeling, dus loop je geregeld achter de feiten aan.

Het schrijven aan de verslagen bleef achter bij het werk wat verzet is, waardoor de ontwikkelfase nog niet was afgesloten, terwijl er al getest werd in de ontwerpfase. Deze achterstand is tijdens de afronding van de opdracht ingelopen, maar heeft een merkbare invloed gehad op de kwaliteit van de producten.

8.3 Probleemanalyse

Het uitvoeren van de probleemanalyse heeft een positief effect gehad op de uitkomst van de opdracht. Uit de analyse blijkt dat het opraken van IP-adressen een re□el gevaar is, maar niet voor alle organisaties even nijpend is. Daarom is gekeken naar wie in de markt als eerst getroffen wordt. Op basis van de bevindingen die de analyse heeft opgeleverd is de opdracht aangepast. Daardoor is het resultaat van de opdracht beter gericht op de problemen die ISP’s in de (nabije) toekomst tegemoet treden. Dit helpt Bateau Knowledge en de Stichting IPv6 Nederland om hun activiteiten beter te sturen.

8.4 Vergelijkend onderzoek

Het onderzoeken van de mogelijke oplossingen was geen makkelijke taak. Na het verzamelen van bronnen uit boeken is op internet gezocht naar hoe deze technieken verder zijn ontwikkeld. Dit leidde tot een veelvoud aan documenten, presentaties en specificaties in ontwikkeling. Het is moeilijk om goed zicht te krijgen op de huidige stand der techniek omdat dit gebied nog zo sterk in ontwikkeling is. Geen van de gevonden oplossingen is een afgeronde specificatie met meerdere implementaties waarvan bekend is dat ze goed werken.

Van de vier kandidaatoplossingen hebben er twee een redelijke kans om binnen een jaar een standaard te worden. Het is onbekend welke van deze twee zich dan verder zal ontwikkelen. Dit hangt af van de implementaties door routerfabrikanten en welke technieken ISP's gaan invoeren. Dual-Stack Lite heeft van de twee de meeste implementaties beschikbaar.

De benchmark die gebruikt is om de oplossingsrichting te bepalen blijkt achteraf niet helemaal zuiver. Er bestaat een verband tussen de volwassenheid en het aantal implementaties van een standaard. Om algemeen geaccepteerd te worden is het nodig dat van een techniek diverse implementaties zijn, die met elkaar samenwerken. Dit bevestigt dat de techniek in kwestie goed werkt. Omdat de oplossingen in de benchmark nog niet uitontwikkeld zijn en er weinig bekend is over samenwerkende implementaties kan alsnog worden beweerd dat in dit geval deze twee onafhankelijk zijn.

8.5 Proof of concept

Uiteindelijk is getest met apparatuur van AVM en A10 Networks. Een Linksys router is door Bateau Knowledge zelf te beschikking gesteld. Met deze apparatuur is een proof of concept gebouwd om de geschiktheid te testen. Om een beter oordeel te geven over Dual-Stack Lite had de test met meer apparatuur uitgevoerd kunnen worden. Hiervoor zijn leveranciers benaderd, maar alleen A10 en AVM hebben op tijd gereageerd. Van overige leveranciers is niks vernomen. Het is onbekend waarom zij niet hebben gereageerd. Een router van Gogo6 met Dual-Stack Lite die ruimschoots op tijd besteld was, is na afloop van de tests nog niet ontvangen.

8.6 Producten

Er is veel tijd gestoken in het vergelijkend onderzoek, het proof of concept en het DS-Lite testlab. Daardoor is tijdens het project minder aandacht besteedt aan de uitwerking van de producten voor de fictieve ISP. De verslagen zijn daarom minder uitgebreid dan had gekund, was dit een opdracht voor een concrete ISP geweest. Dit vormt geen probleem aangezien de opdracht vroeg om het onderzoek en het proof of concept. Wel maakt het deze verslagen minder waardevol als product.

8.7 Competenties

Het demonstreren van competenties is een van de doelen van de afstudeeropdracht. Deze zijn door de afstudeerder gekozen en vastgelegd in afstudeerplan dat voor acceptatie van de opdracht is vastgesteld. Hoe zijn deze competenties in de uitvoer van de opdracht gedemonstreerd?

Opstellen van systeemeisen (A5)

Deze competentie komt naar voren in het Programma van Eisen in bijlage 2. Voor de fictieve ISP is bepaald aan welke eisen de uitbreiding van het netwerk moet voldoen om bepaalde diensten aan klanten te kunnen leveren.

Ontwerpen van een infrastructuur (C9)

Op basis van het Programma van Eisen is een onderzoek gestart naar mogelijke oplossingsrichtingen. De gekozen oplossing is uitgewerkt in een ontwerp voor de infrastructuur van de ISP. Dit is te vinden in het verslag van de Ontwerpfase.

8.7.3 Testen van een infrastructuur (D18)

Tijdens de afstudeeropdracht is een testlab ingericht voor Dual-Stack Lite-apparatuur. In dit lab zijn producten van twee leveranciers getest op de voorwaarden gesteld door:

- de leveranciers zelf
- de eisen van de fictieve ISP

Naar aanleiding van deze tests is verslag uitgebracht naar beide leveranciers, en de uitkomsten van de tests namens de ISP zijn verwerkt in het verslag van de Ontwikkelfase.

Bronnen

1. Website Stichting IPv6 Nederland (Stipv6)
   http://www.stipv6.nl/

2. Website IPv6 Task force
   http://www.ipv6-taskforce.nl/

3. Ahmed, “Deploying IPv6 in Broadband Access Networks”, 2009

4. Website OPTA
   http://www.opta.nl/nl/

5. OPTA, “Wat doet OPTA? - Markten - Breedband ULL”
   http://www.opta.nl/nl/wat-doet-opta/markten/breedband-ull/

6. OPTA, “Wat doet OPTA? - Markten - Breedband WBT”
   http://www.opta.nl/nl/wat-doet-opta/markten/breedband-wbt/

7. NRO, “Free Pool of IPv4 Address Space Depleted”, 3 februari 2011
   http://www.nro.net/news/ipv4-free-pool-depleted

8. APNIC, “APNIC IPv4 Address Pool Reaches Final /8”, 15 april 2011
   http://www.apnic.net/publications/news/2011/final-8

9. Alcock, Shane (University of Waikato), “Research into the Viability of Service-Provider NAT”, 5 augustus 2008
   http://www.wand.net.nz/~salcock/someisp/flow_counting/result_page.html

10. Stichting IPv6 Nederland, “Het woud van transitiemechanismen”
    http://www.stipv6.nl/transitie

11. Doyle, Jeff, “Large Scale NAT Architectures”, 30 september 2009
    http://www.networkworld.com/community/node/45776

12. Doyle, Jeff, “Understanding Dual-Stack Lite”, 22 oktober 2009
    http://www.networkworld.com/community/node/46600

13. Vautrin, O., “IPv4 Rapid Deployment on IPv6 Infrastructures (4rd)”, 29 juli 2010
    http://tools.ietf.org/html/draft-vautrin-softwire-4rd-00

14. Despres, R., “IPv4 Residual Deployment across IPv6-Service networks (4rd)”, 14 maart 2011
    http://tools.ietf.org/html/draft-despres-intarea-4rd-01

15. RFC 5969, “IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)”, augustus 2010
    http://tools.ietf.org/html/rfc5969

16. Conta, Deering, “RFC2473 - Generic Packet Tunneling in IPv6 Specification”, december 1998
    http://tools.ietf.org/html/rfc2473

17. Ford, M., “Issues with IP Address Sharing”, 3 maart 2011
http://tools.ietf.org/html/draft-ietf-intarea-shared-addressing-issues-05

18. Amoss/Minoli, “Handbook of IPv4 to IPv6 Transition”, 2008

19. Blanchet, “Migrating to IPv6”, 2006

20. Cisco, “Deploying IPv6 Networks”, 2006

21. Davies, “Understanding IPv6”, 2003

22. Goralski, “The Illustrated Network”, 2009

23. Hagen, “IPv6 Essentials”, 2002

24. Hagen, “IPv6 Essentials”, 2006

25. Stockebrand, “IPv6 In Practice”, 2007

Begrippenlijst

CPE
Afkorting voor Customer Premises Equipment. Apparatuur die door een provider geplaatst wordt bij een klant om diensten aan te kunnen bieden. Bijvoorbeeld een kabel- of ADSL-modem.

Datagram
Term voor gegevens die tussen twee hosts verstuurd worden. Vooral gebruikt in context van datalinklaag (laag 2) van het ISO-model.

Flow
Verzameling packets van eenzelfde type die zich verplaatst tussen een bron- en doelhost. Een flow kan duiden op alle verkeer dat in één richting plaatsvindt, maar de term wordt ook gebruikt om tweewegverbindingen aan te duiden. Voorbeelden van flows zijn FTP-sessies en HTTP-verbindingen.

IP-packet
Een IP-packet is een hoeveelheid data die tussen hosts wordt verstuurd, inclusief de IP-headers die hiervoor nodig zijn. De term wordt gebruikt om verkeer op ISO-laag 3, de netwerklaag, aan te duiden. IPv4- en IPv6-packets zijn voorbeelden van IP-packets.

Bijlagen

1. Plan van Aanpak
2. Programma van Eisen
3. Verslag Architectuurfase
4. Verslag Ontwerpfase
5. Verslag Ontwikkelfase
6. Testverslag A10 Networks
7. Testverslag AVM

Plan van Aanpak

Afstudeeropdracht

"IPv4-connectiviteit in een IPv6 access network"

Bateau Knowledge, Den Haag

Matthias van der Heide (07053940)

Haagse Hogeschool

Versie 1.0 – 7 februari 2011

Opdrachtomschrijving

Bedrijf en aanleiding

Netwerkarchitect Bateau Knowledge ontwerpt voor haar klanten netwerken en onderhoudt deze. Veelal zijn de netwerken bedoelt om te communiceren met het internet (access networks). De grootte van de netwerken varieert van honderden tot duizenden hosts.

Het huidige Internet Protocol versie 4 (IPv4) is door het gebruik van 32-bits adressen beperkt tot 4 miljard hosts. Naar verwachting zullen eind 2011 de laatste adressen uitgegeven zijn. Daarna is het niet meer mogelijk om nieuwe IP-adressen aan te vragen.

De oplossing is IPv6, de opvolger voor IPv4 die al meer dan tien jaar bestaat. De adoptie van IPv6 is echter flink achtergebleven bij de verwachtingen, waardoor straks de situatie ontstaat dat de oude adressen eerder op zijn, dan dat iedereen overgestapt is op IPv6. Hierdoor kunnen er geen nieuwe aansluitingen worden geleverd met een eigen IP-adres, tenzij er gebruik wordt gemaakt van IPv6-adressen.

Probleemstelling

Om te kunnen blijven groeien zullen eigenaren van accessnetwerken (internet service providers) hun netwerk moeten uitbreiden met IPv6. Dit kan door zowel IPv4 als IPv6 te gebruiken (dual stack), of door helemaal over te stappen op IPv6.

Dual-stack gaan (zowel IPv4 als IPv6 leveren) is voor de aanbieder duurder dan IPv6 alleen al vanwege de dubbele routering en de extra administratieve lasten. Daarom is het waarschijnlijk dat deze kiest voor een IPv6-only netwerk.

Klanten vragen echter om het ‘oude’ IPv4-internet te kunnen benaderen om bijvoorbeeld te surfen naar websites die niet over IPv6 beschikbaar zijn. Ook hebben ze veel apparatuur die niet werkt over IPv6, omdat deze niet vervangen kan worden, en niet geschikt te maken is door middel van een upgrade. Voorbeelden hiervan zijn spelcomputers, VoIP-telefoons, IP-camera’s en oude computers met Windows XP.

Hoe kan een provider zijn klanten zowel IPv6- als IPv4-connectiviteit bieden met een IPv6-only netwerk op een manier dat de klantenapparatuur goed blijft werken?

Doelstelling van de afstudeeropdracht

Realiseer een proof of concept die verkeer van IPv4-only clients omzet naar een IPv6-only ISP-verbinding en die daarna weer naar IPv4 vertaalt en die zo generiek mogelijk werkt. Generiek als onzichtbaar voor de IPv4-only client applicaties zodat protocollen zoals HTTP, BitTorrent en VoIP bruikbaar blijven. De oplossing moet te schalen zijn naar een netwerk van 5000 hosts.

Deelopdracht 1:
Inventariseer wat er nu op papier en in de praktijk voor handen is. Beschrijf de mogelijkheden en beperkingen van de huidige oplossingen.

Deelopdracht 2:
Maak van de meest relevante oplossing een proof of concept om theorie met de praktijk te kunnen vergelijken en beschrijf de beperkingen van de implementatie.

Resultaten

De eerste deelopdracht levert een vergelijkend onderzoek op waarin de verschillende bekende technieken opgesomd en vergeleken worden.

Op basis van dit onderzoek wordt een kandidaat bepaald. Het eindresultaat is een advies aan Bateau Knowledge welke oplossing voor de het leveren van IPv4-connectiviteit in een IPv6-netwerk zij haar klanten kan aanbevelen.

De oplossing is middels een proof of concept getoetst in een testopstelling en de voorwaarden waarin deze toegepast kan worden zijn duidelijk gemaakt.

Projectgrenzen

- Voor de opdracht wordt geen software ontwikkeld.
- Er wordt zoveel mogelijk gebruik gemaakt van bestaande technische oplossingen.

Randvoorwaarden

Voor het uitvoeren van de opdracht wordt gebruik gemaakt van de infrastructuur van Bateau Knowledge

Aanpak

Voor het faseren en beheren van technische-infrastructuurprojecten zijn diverse methoden beschikbaar. Bij Bateau Knowledge wordt niet gewerkt met een vastgelegde standaard of methode. Op basis van de eisen van het afstudeerproject en onderzochte projectmethoden wordt een keuze gemaakt voor dit specifieke project.

Projecteisen

De afstudeeropdracht stelt specifieke eisen aan de te gebruiken projectfaseringsmethode. Ten eerste is het projectteam zeer klein (twee personen, de afstudeerder en opdrachtgever) en is er slechts een externe partij (de Haagse Hogeschool) die procesmatig betrokken is. Ten tweede wordt niet een complete infrastructuur van een bestaande organisatie (her)ontworpen, maar wordt een oplossing gezocht voor een specifiek technisch onderdeel van een toekomstig netwerk. Dit maakt de opdracht minder concreet en enkele details onbekend. De projectmethode moet rekening houden met deze onzekerheden en de mogelijkheid laten om de technische aspecten te detailleren en overige aspecten bewust minder ver uit te werken.

Geselecteerde projectfaseringsmethoden

DYA

Dynamische Architectuur (DYA) is methode die door Sogeti is ontwikkeld om technische-infrastructuren te ontwikkelen en onderhouden. Het is gefocust op integratie van architectuurontwerp in de veranderprocessen van de organisatie. DYA is gebaseerd op vier processen:

- Strategische dialoog
- Architectuur services
- Ontwikkelen onder architectuur
- Ontwikkelen zonder architectuur

Deze vier kernprocessen sturen de ontwikkeling van de infrastructuur. Hierin onderscheidt DYA de business-, informatie- en technische architectuur. In het raamwerk (http://www.dya.info/Home/dya/wat_is_dya/architectuurraamwerk.jsp) worden deze nog verder uitgewerkt in onderdelen. Voor elk van deze onderdelen kan in meer of mindere mate een invulling worden gegeven in een project.

Over Dya zijn diverse Nederlandstalige boeken uitgegeven.

Bron: http://www.dya.info

ASI-rapport

Het ASI-rapport is opgesteld door het Koninklijk Instituut Van Ingenieurs (KIVI, tegenwoordig KIVI NIRIA) en het Nederlands Genootschap voor Informatica (NGI) en is bedoeld als een plan van aanpak voor architecten, ontwerpers en projectleiders van technische-infrastructuurprojecten.

Hoewel het rapport stamt uit 1995 wordt de methode actief gebruikt binnen de Technische Informatica-opleidingen van de Haagse Hogeschool.

ASI kent een gefaseerde aanpak voor het ontwerpen en ontwikkelen, maar laat invoering en exploitatie buiten beschouwing. Het onderkent dat ontwikkelen een iteratief proces is, maar dat deze iteraties binnen de fasering plaatsvinden. De vier fasen zijn:

- Definitie
- Architectuur
- Ontwerp
- Ontwikkeling

Van elke fase zijn het doel, de aanpak, de resultaten en de activiteiten beschreven. ASI is geen formele methode of een complete beschrijving, maar een handvat voor de projectleden.

Bron: ASI-rapport (verstrekt door de Haagse Hogeschool)

TOGAF / ADM

The Open Group Architecture Framework (TOGAF) is een open standaard en methode voor het ontwikkelen van architecturen in grote bedrijven. Net zoals Dya onderscheidt TOGAF de business, informatie (Data en Applicatie) en technische architectuur. Het ontwikkelen van architecturen wordt gedaan middels de Architecture Development Method (ADM). Deze focust sterk op de inpassing van de architectuur in de huidige TI en bijsturing door veranderende requirements.

Naast ADM voor de ontwikkeling van architecturen bestaat TOGAF uit het Enterprise Continuum en de Resource Base. In het Enterprise Continuum is een opslag voor architectuurdocumenten die van belang zijn voor de organisatie. De Resource Base is een verzameling van hulpmiddelen en technieken die in de ADM gebruikt kunnen worden.

Bron: website TOGAF, PDF Sogeti

Gekozen projectmethode

De belangrijkste eis voor dit project is dat de focus ligt op de ontwikkeling van een (deel van een) technische infrastructuur. Daarnaast gaat het om een kleinschalig project dat buiten een bestaande architectuur wordt uitgevoerd.

Daarom is er gekozen voor fasering volgens de ASI-methode. Deze biedt een duidelijk raamwerk voor het ontwerpen van een TI en focust niet meer dan voor dit project nodig is op invoering hiervan in een bestaande business- en technische architectuur.

Bovendien zijn de deelopdrachten goed in deze fasering te vatten. Deelopdracht één wordt ondergebracht in de architectuurfase waar alternatieve architecturen worden vergeleken. Deelopdracht twee kan volbracht worden in de ontwikkelfase, waar een toetsing van de goede werking van het ontwerp gevraagd wordt.

Planning en produkten

Inleiding

In dit hoofdstuk wordt aan de hand van de ASI-fasering uitgelegd welke werkzaamheden er plaatsvinden en wat de op te leveren produkten zijn. De deadlines voor de fases en produkten worden vermeld in de wekenplanning .

Naast de produkten voor de ASI-methode worden ook produkten voor de opleiding opgeleverd. Deze worden in de wekenplanning meegenomen.

ASI-fasering

Binnen de ASI-fasering bestaan de volgende fasen met bijbehorende activiteiten en producten.

Fase	Activiteiten	Product
Definitie	Uitvoeren van een definitiestudie Onderzoeken oplossingsrichtingen en gevolgen daarvan (voordelen/nadelen) Kiezen oplossingsrichting Detaileren eisen, wensen, uitgangspunten en randvoorwaarden	Programma van Eisen
Architectuur	Nader specificeren eisen Functionele/kwantitatieve specificatie van de infrastructuur. PvE vertalen naar platformen, diensten en netwerkverbindingen; maximaal te bieden capaciteiten en geografische distributie Opstellen alternatieve architecturen, uitwerking van de verschillen tussen de alternatieven (topologie, keuze van platform, beheerorganisatie, invoeringsscenarios, kosten/batenindicatie). Keuze voorkeursalternatief Zonodig nader uitwerken voorkeursalternatief.	Tekening netwerktopologie en –componenten, protocollen en standaarden Omschrijving invoering en migratiescenario's Kosten/batenindicatie Plan van Aanpak voor het vervolg van het project
Ontwerp	Uitwerken van de architectuur (niet tot in detail): Opstellen van het logische ontwerp Opstellen van het fysieke ontwerp - Selectie produkten/diensten/leveranciers - Bepalen principe configuratie Opstellen van ontwerp beheer/support en organisatie	Logisch ontwerp Fysiek ontwerp Beheer-/supportontwerp
Ontwikkeling	Uitwerken van de architectuur tot in detail en testen van de requirements: Uitwerking techniek - Testopstelling in laboratoriumomgeving (staging) - Beheertools (SNMP) - Fixes/workarounds (ontwerpwijzigingen, maatwerk hw/sw, richtlijnen beheer, richtlijnen gebruiker) - Acceptatietest	Ontbreekt in uitwerking rapport. Samenvatting: 'Het resultaat van de fase is een gedetailleerde voorbereiding op de technische en organisatorische invoering. Aan het eind is het ontwerp uitgewerkt tot een beproefd en te realiseren produkt.'

	Uitwerking beheer en support - Opstellen sla's - Maken handboeken Beschermde Pilot Opstellen implementatiedraaiboek	Zelf: Voor labtest: Testplan - Ontwerp testopstelling Testresultaten - Fixes/workarounds Idem voor acceptatietest als je die uitvoert SLA's Handboek beheer Handboek support Handboek gebruikers Plan pilot Verslag pilot Draaiboek implementatie

De ASI-fasering is toegespitst op het invoeren van wijzigingen aan een bestaande technische infrastructuur. Het compleet uitwerken van de invoering vereist vergaande kennis van de te wijzigen TI. De afstudeeropdracht heeft echter een scope die minder groot is.

Ten eerste wordt het project gedaan voor een vooralsnog onbekende, toekomstige klant. De precieze invulling van de TI van de klant is daardoor onbekend, en kan niet volledig worden uitgewerkt. Dat betekent onder meer dat bepaalde operationele eisen onbekend zijn en dat er geen helder beeld is van de huidige logische en fysieke indeling van het netwerk.

Deze informatie kan ook niet makkelijk worden ingewonnen zonder met het bedrijf zelf te overleggen, bijvoorbeeld door het lezen van publicaties en informatie op websites. Wel kan dit een globaal beeld geven van de mogelijke huidige infrastructuur van het bedrijf.

Ten tweede spitst de opdracht zich toe op een specifiek onderdeel van het netwerk van een Internet Service Provider. Het is waarschijnlijk dat een ISP meerdere van deze netwerken heeft, opgebouwd uit verschillende componenten. Daarom zal waar mogelijk een keuze worden gemaakt om de uitwerking toe te spitsen op een specifiek deel van het netwerk, en niet de TI als geheel.

Door te kiezen voor een deel van het netwerk, en doordat de specifieke TI van de klant onbekend is, verliezen een aantal van de activiteiten en producten uit de ASI-fasering hun directe waarde.

Met name activiteiten in de Ontwikkelfase die gericht zijn op invoering van het ontwerp in de concrete, huidige TI zijn niet van toepassing voor dit project. Documenten zoals het Programma van Eisen en de uitwerking van de huidige infrastructuur bevatten hiaten die met medewerking van een concrete klant ingevuld hadden kunnen worden.

Definitie

In de definitiefase wordt een beeld gevormd van het doel van het project, en de omgeving waarin deze plaatsvindt. Welk probleem wordt er opgelost? In welke organisatie speelt dit probleem? Wat is de oorzaak? Op welke manieren kan het probleem opgelost worden?

Activiteiten

- Uitvoeren van een definitiestudie
- Onderzoeken oplossingsrichtingen en gevolgen daarvan
- Kiezen van een oplossingsrichting
- Detailleren eisen, wensen, uitgangspunten en randvoorwaarden van de oplossing

Producten

Het product van de definitiefase is het Programma van Eisen. Het bevat de resultaten van de definitiestudie en het onderzoek naar de oplossingsrichtingen. Op basis van een vergelijking is een keuze gemaakt voor een bepaalde richting. Eisen, wensen, uitgangspunten en randvoorwaarden zijn verder uitgewerkt.

Architectuur

Tijdens de architectuurfase worden binnen de gekozen oplossingsrichting diverse alternatieven onderzocht. Hun architectuur wordt globaal uitgewerkt en de verschillen tussen de scenario's worden aangeduid. Daarna wordt een keuze gemaakt voor een uit te voeren scenario.

Activiteiten

- Nader specificeren eisen
- Functionele/kwantitatieve specificatie van de infrastructuur. PvE vertalen naar platformen, diensten en netwerkverbindingen; maximaal te bieden capaciteiten en geografische distributie
- Opstellen alternatieve architecturen, uitwerking van de verschillen tussen de alternatieven (topologie, keuze van platform, beheerorganisatie, invoeringsscenario's, kosten/batenindicatie).
- Keuze voorkeursalternatief

Producten

- Tekening netwerktopologie en –componenten, protocollen en standaarden
- Uitwerking alternatieve architecturen en keuze voorkeursalternatief
- Plan van Aanpak voor het vervolg van het project

Ontwerp

Het voorkeursalternatief wordt concreter gemaakt. Hier wordt de vertaalslag gemaakt van een globaal concept naar een specifieker logisch en fysiek ontwerp voor de aan te passen TI. Omdat er nog een keuze gemaakt moet worden voor leverancier ligt het niveau van detail op de grens waar dit nog hetzelfde is voor alle leveranciers. Bijvoorbeeld wel een IP-adresplan, maar geen configuratiebestanden. Leveranciers worden benaderd om, op basis van het ontwerp, offertes te leveren voor de aan te schaffen apparatuur en diensten. Ook wordt er gekeken hoe de aanpassingen opgenomen moet worden in de beheerorganisatie.

Activiteiten

- Uitwerken van de architectuur (niet tot in detail)
    - Opstellen van het logische ontwerp

    - Opstellen van het fysieke ontwerp
- Selectie produkten/diensten/leveranciers
- Bepalen principe configuratie
- Opstellen van ontwerp beheer/support en organisatie

Producten

- Logisch ontwerp
- Fysiek ontwerp
- Beheer-/supportontwerp

Ontwikkeling

Het ontwerp wordt uitgewerkt tot een geïntegreerd geheel. Producten van verschillende leveranciers worden getest om te bepalen of ze voldoen aan de requirements. De onderdelen uit ASI die betrekking hebben op invoering (acceptatietests, pilot, uitwerking SLA's, implementatiehandboeken) worden niet uitgevoerd.

Activiteiten

- Testopstelling in laboratoriumomgeving (staging)
- Beheertools (SNMP)
- Fixes/workarounds (ontwerpwijzigingen, maatwerk hw/sw, richtlijnen beheer, richtlijnen gebruiker)

Producten

- Testplan
- Testverslag

Wekenplanning

Op basis van de activiteiten per fase is de volgende planning opgesteld. Er is veel tijd voor de ontwikkelfase ingepland omdat hierin het testen plaatsvindt. Het is de verwachting dat dit de meeste tijd gaat kosten. Ook is de beschikbaarheid van apparatuur bepalend voor hoeveel tijd er getest kan worden en is daarom een langere periode gereserveerd.

Stageweek	Kalenderweek	Datum	Fase
1	2	Maandag 10 januari	Initiatie
2	3		
3	4		
4	5		Definitie
5	6		Architectuur
6	7		
7	8		
8	9		Ontwerp
9	10		
10	11		Ontwikkeling
11	12		
12	13		
13	14		

14	15		
15	16		Afronding
16	17		
17	18	Vrijdag 6 mei	Afronding; einde afstudeerperiode
18	19	Vrijdag 13 mei	Uitloop i.v.m. opdracht bedrijfsprocessen
	20		
	21		
	22	Vrijdag 3 juni	Inleveren afstudeerdossier
	23		
	24		
	25	20 juni t/m 6 juli	Examenzittingen
	26		
	27		

Projectinrichting

Projectorganisatie

Binnen het afstudeerproject zijn er drie rollen: de afstudeerder, de opdrachtgever en de begeleider.

Afstudeerder

De afstudeerder is Matthias van der Heide. Zijn taak is het uitvoeren van de afstudeeropdracht zoals beschreven in dit document. De afstudeerder werkt op werkdagen (ma-vr) aan zijn opdracht en is die dagen beschikbaar.

Opdrachtgever

De opdrachtgever is Bateau Knowledge, vertegenwoordigd door Paul Boot. Hij begeleidt de afstudeerder en beantwoord vragen over de opdracht. De opdrachtgever is minstens één dag per week voor overleg beschikbaar en voor verdere afspraken beschikbaar afhankelijk van werk binnen of buiten kantoor.

Begeleider

De begeleider namens de Haagse Hogeschool is Pieter Burghouwt. Indien hij niet beschikbaar is wordt zijn rol overgenomen door Adri Pronk. De begeleider is voor de afstudeerder aanspreekpunt in geval van inhoudelijke en procesmatige vragen over het afstuderen. De begeleider beoordeelt het afstudeerdossier en of dit voldoende is voor een afstudeerzitting.

Pieter Burghouwt

Beschikbaarheid:	maandag tot en met vrijdag
Telefoon:	015-2606288
E-mail:	p.burghouwt@hhs.nl

Adri Pronk

Beschikbaarheid:	maandag, dinsdag, woensdag, vrijdag
Telefoon:	015-2606279
E-mail:	A.G.P.Pronk@hhs.nl

Informatievoorziening

Centraal punt voor communicatie is de afstudeerder. Hij verzorgt het maken van afspraken tussen de verschillende projectdeelnemers en het doorsturen van documenten.

Documenten worden gedeeld via e-mail of Google Docs. Informatie over het bedrijfsnetwerk kan gevonden worden in de bedrijfswiki.

Er vindt wekelijks overleg plaats tussen de opdrachtgever en de afstudeerder. Van dit overleg worden door de afstudeerder notulen gemaakt.

Faciliteiten

Bateau Knowledge richt een werkplek in voor de afstudeerder op het kantoor in Den Haag. De afstudeerder werkt op een eigen laptop. Internettoegang en stroom is geregeld.

Voor het testen kan gebruik gemaakt worden van het netwerk van Bateau Knowledge. In overleg met de opdrachtgever wordt de configuratie van het netwerk hiervoor aangepast.

Programma van Eisen

Afstudeeropdracht

“**IPv4-connectiviteit in een IPv6 access network**”

Bateau Knowledge, Den Haag

Matthias van der Heide (07053940)

Haagse Hogeschool

Versie 2.0 –6 juni 2011

Inhoudsopgave

1. Inleiding

Het doel van dit document is een overzicht te geven van de eisen aan de aanpassing van het IPv4-accessnetwerk van een grote Internet Service Provider (ISP). Het netwerk wordt ontwikkeld in opdracht van Bateau Knowledge en is een verkenning van de mogelijkheden die er zijn om IPv4 connectiviteit te bieden aan consumenten als het netwerk is gebaseerd op IPv6 en er een tekort is aan IPv4-adressen.

Hoofdstuk twee beschrijft de bestaande situatie bij de (fictieve) ISP. Daarna wordt in hoofdstuk drie de het huidige netwerk van de provider uitgelegd. De reden voor de aanpassing wordt gegeven in hoofdstuk vier. Hoofdstuk vijf, ‘Eisen’, beschrijft welke functionele en operationele eisen gesteld worden aan de oplossing.

2. Opdracht

2.1. Aanleiding

Netwerkarchitect Bateau Knowledge voorziet dat met het opraken van de IPv4-adressen er in de toekomst een probleem ontstaat voor ISP's. De huidige netwerken zijn gebouwd met als basisprincipe dat elke klant (minstens) een IPv4-adres krijgt toegewezen. Providers die voldoende adressen ter beschikking hebben kunnen hun netwerk voorlopig blijven gebruiken, maar als een ISP wil groeien en dus meer klanten wil aansluiten kan deze in de problemen raken.

Daarom laat Bateau een onderzoek uitvoeren naar IPv6-transitietechnieken. Dit onderzoek wordt uitgevoerd door Matthias van der Heide, student Technische Informatica aan de Haagse Hogeschool.

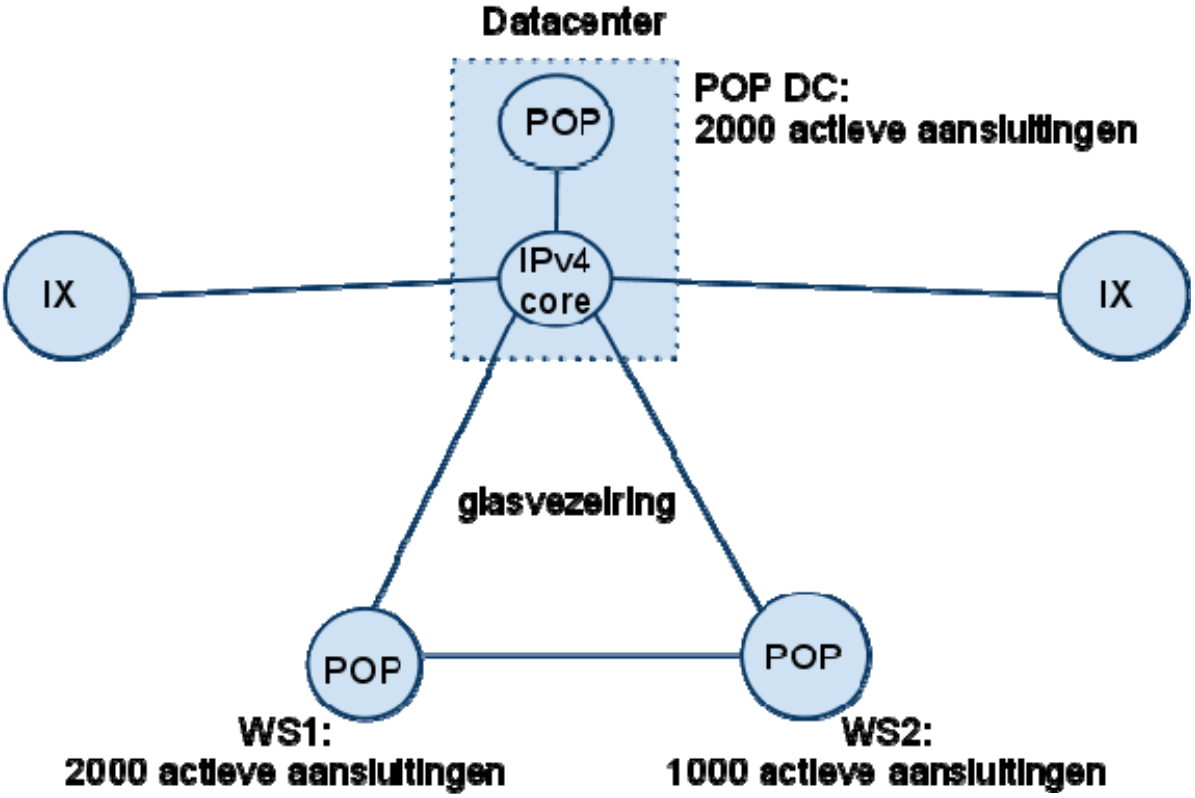
Binnen dit onderzoek wordt een deel van het netwerk ontwikkeld voor een fictieve klant, een regionale ISP met een glasvezelnetwerk. Dit biedt een context voor het onderzoek en op deze manier wordt er niet alleen rekening gehouden met de (technisch) functionele eisen maar ook met de operationele eisen die gelden vanuit een ISP.

2.2 Doel van het project

Dit project heeft tot doel om een manier te vinden waarop een provider met een (aankomend) tekort aan IPv4-adressen zijn klanten toch toegang tot het IPv4-internet kan bieden. Hierbij wordt gekeken naar de mogelijkheden om dit met behulp van IPv6 te realiseren, en in de toekomst ook IPv6-connectiviteit te leveren.

3. Bestaande situatie

Het netwerk van de ISP bevindt zich in een middelgrote stedelijke gemeente met 36.000 inwoners. Samen vormen zij 16.900 huishoudens. De gemeente bestaat uit vier wijken. Drie van de wijken zijn al voorzien van glasvezel. Hiervoor zijn drie Points of Presence (POP's) gebouwd, waar de kale glasvezels van klanten uitkomen. In een POP, ook wel wijkstation genoemd, worden de glasvezels van klanten die een aansluiting afnemen gekoppeld aan de netwerkapparatuur van de provider.



Afbeelding 3.1 - huidige situatie

De POP's worden op hun beurt door middel van glasvezel verbonden met de core routers van het ISP-netwerk, die geplaatst zijn in het datacenter. De core wordt aangesloten op de rest van het internet via glasvezelverbindingen naar Internet Exchanges (IX) waar met andere partijen peering plaatsvindt. Verkeer dat niet via peering kan worden afgeleverd wordt via een leverancier van transitverbindingen afgehandeld. Het netwerk zoals het fysiek is opgebouwd is te zien in afbeelding 1.

De huidige situatie in het accessnetwerk is als volgt:

Lokatie POP	Aansluitingen	Actief	Percentage actief
Datacenter	6.000	2.000	33%
Wijkstation 1	4.200	2.000	48%
Wijkstation 2	3.600	1.000	28%
Totaal	**13.800**	**5.000**	**36%**

Tabel 3.1 – Aansluitingen (aangelegd/actief) per POP

Bij de aanleg van het netwerk is er van uit gegaan dat een derde van de aangesloten huizen internettoegang zou afnemen, zo'n 4600 adressen. Hiervoor zijn in de loop van de tijd drie IPv4-adresblokken van /21 aangevraagd, met een gezamenlijke grootte van 6144 adressen. Deze blokken zijn in kleinere stukken opgedeeld en aan de lokaties en klanten uitgedeeld. Een deel van de adressen wordt gebruikt om de infrastructuur te adresseren. Deze bestaat uit servers voor DNS, DHCP, en andere diensten. Ook alle routers en switches hebben IP-adressen.

Klanten worden aangesloten in groepen van 250 aansluitingen, waardoor voor elke groep klanten een IPv4-subnet met een grootte van /24 (254 adressen) nodig is. De groepen klanten worden in een gezamenlijk VLAN geplaatst, waarbij dit VLAN overeenkomt met één IP-subnet. De router krijgt het hoogste adres in dit subnet. De overige adressen worden door een DHCP-server uitgedeeld aan de routers van klanten. Beveiligingsmaatregelen op netwerklaag twee (layer 2) zorgen dat verkeer tussen klanten alleen via de routers van de ISP kan plaatsvinden, en niet onderling tussen klanten in een VLAN.

Gebruikersgroep	Aantal /24-blokken
Infrastructuur	1
Kantoor ISP	1
Klanten DC	8
Klanten WS1	8
Klanten WS2	4
Totaal in gebruik	**22**
Totaal beschikbaar	**24 (3 maal een /21)**

Tabel 3.2 – Adresverbruik per groep

De huidige opzet gebruikt 22 adresblokken van /24, ofwel 5632 van de 6144 beschikbare IP-adressen. Hierdoor is slechts 10% adresruimte over voor uitbreiding van het netwerk en aansluiten van nieuwe klanten. Met deze twee blokken kan een groei van 500 klanten worden opgevangen.

4. Uitbreiding naar de vierde wijk

Wijk nummer vier, waar in eerste instantie geen glasvezel is aangelegd, is gerenoveerd en het is gebleken dat hier voldoende animo is voor internetdiensten om een derde wijkstation te bouwen en alle huizen te voorzien van glasvezel. Hiermee is de gehele gemeente ‘verglaasd’.

De nieuwe situatie zou er dan ongeveer zo uitzien:

Lokatie POP	Aansluitingen	Actief	Percentage actief
Datacenter	6.000	2.000	33%
Wijkstation 1	4.200	2.000	48%
Wijkstation 2	3.600	1.000	28%
Wijkstation 3	3.100	1.200	39%
Totaal	**13.800**	**6.200**	**44%**

Tabel 4.1 – Aansluitingen (aangelegd/actief) per POP, nieuwe situatie

Dit stelt de ISP voor een direct probleem: het (verwachtte) totale aantal aansluitingen komt boven huidige IPv4-adresvoorraad uit. Door het opraken van de adressen kan niet voldoende adresruimte bij RIPE aangevraagd worden.

Wel kan een /22 (1024 hosts) worden aangevraagd bij RIPE om de nood te ledigen. Dit subnet is afkomstig uit een voorraad die is gereserveerd om als laatste redmiddel te worden gebruikt. Voorwaarde voor het ontvangen van dit blok is dat de ISP al IPv6-adresruimte heeft aangevraagd. De toelichting en procedures voor het aanvragen van zijn te vinden in de document RIPE509 [1].

Ook met deze extra ruimte komt de ISP slechts op 7168 adressen uit. Dat is mogelijk net voldoende voor de aanleg van wijkstation 3, maar er bestaat een reëel gevaar dat bij groei van het aantal geactiveerde aansluitingen er een tekort aan IPv4-adressen optreedt.

De ISP besluit daarom een onderzoek te starten naar methoden om IPv4-adressen efficiënter te gebruiken. Een mogelijke oplossing is om meerdere klanten gebruik te laten maken van een IP-adres.

Ook is er de wens om in de toekomst IPv6-connectiviteit te bieden aan klanten, als daar om gevraagd wordt.

Daarnaast is er de wens om het netwerk simpel te houden door een protocol te gebruiken. De ISP vreest dat het routeren van twee protocollen dubbele kosten met zich mee brengt.

5. Eisen

5.1 Inleiding

In dit hoofdstuk worden de eisen die de ISP stelt aan de te leveren dienst vastgesteld. Er wordt een definitie gegeven van internettoegang om de functionele eisen duidelijk te maken. Daarna worden de operationele eisen uitgewerkt. In het ontwerpen van netwerken ligt de focus sterk op operationele eisen, omdat deze de kaders van de mogelijke oplossingen stellen.

5.2 Functionele eisen

“De te ontwikkelen technische infrastructuur biedt breedbandige internettoegang aan alle huishoudens in de gemeente.”

Om deze eis te verduidelijken worden de gebruikte begrippen toegelicht:

Internettoegang
We kunnen internettoegang definieren als een netwerkverbinding die inkomend en uitgaand verkeer volgens het Internet Protocol toestaat. De verbinding stelt de klant in staat om te communiceren met alle actieve systemen op het internet.

De meestgebruikte versie van het Internet Protocol is IPv4. Bijna alle systemen op het internet werken met dit protocol. De adressen voor dit protocol raken echter op. Daarom is een nieuw protocol ontwikkeld, IPv6, dat veel meer systemen kan adresseren.

Voor deze opdracht wordt uitgegaan van het IPv4-internet, dat wil zeggen dat er van wordt uitgegaan dat de ISP zijn klanten een verbinding naar IPv4-systemen op het internet aanbiedt. Als de ISP ook verbinding naar IPv6-systemen wil bieden, dan moet dit ook mogelijk zijn.

Breedband
Breedbandinternet is internettoegang met een hoge doorvoercapaciteit (ook wel snelheid genoemd). De capaciteit wordt gemeten in bits per seconde en kan voor downstream (verkeer van internet naar de gebruiker) anders zijn dan voor upstream (verkeer van de gebruiker richting het internet). De huidige breedbandverbindingen hebben een downstreamcapaciteit tussen de 1 en 120 megabit per seconde en een upstreamcapaciteit van 0,5 tot 10 megabit.

De OECD definieert breedbandtoegang als een verbinding met een downstreamcapaciteit groter dan 256 kilobit per seconde (0,25 Mbps). [2]

Bij veel verbindingen is de doorvoercapaciteit niet gegarandeerd, maar wordt een overboekingsfactor gehanteerd. De daadwerkelijke capaciteit hangt dan bijvoorbeeld af van de drukte op het netwerk van de aanbieder. Verbindingen met een lage overboekingsfactor (minder dan 1:20) worden vooral aan bedrijven geleverd, vanwege de gegarandeerde capaciteit. Verbindingen met een hoge overboekingsfactor zijn goedkoper en komen daarom bij consumenten veel voor. [3]

Huishoudens in de gemeente
De 16.900 privehuishoudens binnen de gemeentegrenzen.

5.3 Operationele eisen

Beschikbaarheid
De internettoegang is in principe altijd (24 uur per dagen, 7 dagen per week) beschikbaar, maar een downtime van enkele uren per maand voor de eindgebruiker is acceptabel. In de overeenkomst met de gebruiker wordt vastgelegd hoeveel uren dit maximaal kan zijn, en of hier een vergoeding tegenover staat. Alleen het niet-werken van de toegangsverbinding zelf wordt als downtime gerekend. Netwerkproblemen buiten het netwerk van de provider vallen niet onder downtime.

De provider neemt maatregelen om te voorkomen dat derden de beschikbaarheid negatief beinvloeden (hackers, DDoS)

Eis:

- de geleverde verbinding heeft maximaal een downtime van 1 uur per maand (uptime van ca. 99,8%)

Capaciteit
Het netwerk moet in staat zijn om de eindgebruiker de snelheid te bieden die in het contract is vastgelegd. Voor upstream en downstream gelden dezelfde bandbreedtes. Daarnaast geldt voor de verbindingen een overboekingsfactor, waardoor de totale te leveren capaciteit vele malen lager is (1/10e) dan de som van capaciteiten van de eindgebruikers. In de meeste gevallen (bij rust op het netwerk) is de volledige verbinding beschikbaar, maar bij drukte wordt 1/10e van de afgenomen capaciteit gegarandeerd. Gemiddeld ligt de gebruikte capaciteit van de verbindingen nog onder deze gegarandeerde capaciteit.

Voor de opdracht gaan we uit van een symmetrische breedbandverbinding met een capaciteit van 100 Mbit downstream en 100 Mbit upstream met een overboekingsfactor van 1:10. Dat wil zeggen: binnen het access netwerk wordt 1/10e van de geadverteerde snelheid gegarandeerd. Een andere manier om dit te verwoorden is: een aansluiting deelt met maximaal tien anderen dezelfde verbindingscapaciteit naar het internet.

De verwachting is dat het aantal actieve aansluitingen na de aanleg van wijkstation 3 groeit van 5.000 naar 6.200. De benodige capaciteit is uit te rekenen met de volgende formule:

aantal aansluitingen * verbindingscapaciteit * overboekingsfactor = maximaal benodigde totale capaciteit

Dit betekent dat de huidige capaciteit moet worden uitgebreid van 50 Gbit naar 62 Gbit. In de toekomst moet een capaciteit van 169 Gbit mogelijk zijn, bij activering van alle aansluitingen.

Berekeningen:
5.000 * 100 mbit * (1:10) = 50 Gbit
6.200 * 100 mbit * (1:10) = 62 Gbit
16.900 * 100 mbit * (1:10) = 169 Gbit

Eis:

- de capaciteit wordt uitgebreid naar 62 Gbit
- de capaciteit kan (op termijn) uitgebreid worden naar 169 Gbit als dit nodig is

Performance (prestatie)
Netwerkprestaties worden over het algemeen uitgedrukt in de volgende termen:

- packet loss (aantal niet-afgeleverde pakketten in een tijdspanne)
- latency (tijd tussen het verzenden van een pakket en het afleveren ervan)
- jitter (variatie in de latency; lage jitter is belangrijk voor diensten zoals videostreaming, audiostreaming en VoIP)

Idealiter worden deze geminimaliseerd naar nul. Dit is echter onhaalbaar, om diverse redenen.

Vanwege capaciteitsproblemen kan altijd packet loss optreden. Ook als er voldoende capaciteit is kan een pakket nogsteeds verloren gaan buiten het netwerk van de ISP. Er wordt dus gestreeft naar een zo laag mogelijke packet loss binnen het netwerk, maar packet loss kan nooit geheel voorkomen worden. Bovendien is TCP/IP gebouwd om hier mee om te gaan. Een bovengrens van 1% is acceptabel. [4]

Latency is de tijd die een packet er over doet om van de bron naar het doel te komen. Binnen computernetwerken wordt vaak de term RTT, round-trip time gebruikt. Dit is de tijd die een packet nodig heeft om naar het doel en terug te komen. RTT is dus ruwweg twee keer de latency. De tijd die nodig is om een bepaalde afstand door glasvezel (of welk ander medium dan ook) af te leggen is gelimiteerd door de snelheid van het licht. Voor glasvezelverbindingen binnen Nederland is een RTT naar een server in Nederland maximaal ca. 100 ms, net als de RTT naar Google wereldwijd. [5] Met sommige servers in Nederland kan 10 ms worden gehaald (blijkt uit snelheidstesten vanuit Bateau Knowledge).

Jitter wordt veroorzaakt door verschillen in transmissietijden van packets. Dit kan veroorzaakt worden door opstoppingen (congestion), drukke routers en omleidingen in het internet. Hoge jitter heeft gevolgen voor streams, zoals audiochat en videobellen. Packets die te laat binnenkomen moeten als verloren beschouwd worden. Hierdoor verslechteren de prestaties van deze toepassingen.

Het invoeren van IPv6 als tweede netwerkprotocol leidt onvermijdelijk tot verschillen in de prestaties van IPv6 en IPv4. Voor dit project mogen de prestaties ruim van elkaar afwijken. Een factor twee is het maximum, in het voordeel van IPv4 of IPv6.

Eis:

- packet loss binnen het netwerk van de ISP moet minimaal zijn, met een bovengrens van 1%
- latency binnen het netwerk is maximaal 10 ms
- jitter is maximaal gelijk aan de latency
- de waarden voor voor IPv4 versus IPv6 schelen maximaal een factor 2

Schaalbaarheid

Indien er meer gebruik wordt gemaakt van het systeem dan verwacht (meer gebruikers, meer dataverkeer) dan moet het systeem hier mee om kunnen gaan. Momenteel wordt ongeveer 35% van de aansluitingen daadwerkelijk gebruikt. Maar in potentie kan dit groeien naar 100%. Ook wordt door het toenemende dataverbruik (downloaden, streaming video en audio) verwacht dat de te leveren capaciteit per gebruiker toe zal nemen. De beschikbare hoeveelheid IPv4-adressen is beperkt op 7.168, ongeacht het aantal nieuwe aansluitingen.

Eis:

- het netwerk kan groeien naar zijn maximale capaciteit van 16.900 aansluitingen

Veiligheid

Het moet niet mogelijk zijn voor iemand om zich voor te doen als een andere klant.
Het is aan de klant zelf om te bepalen welk verkeer hij wel en niet toestaat op zijn netwerk.
Derden mogen geen controle krijgen over het netwerk van de provider.

Duurzaamheid

Duurzaamheid is op verschillende manieren belangrijk. Bijvoorbeeld:

- apparatuur moet na gebruik milieuvriendelijk worden verwijderd (geen gevaarlijke stoffen, recyclen)
- het netwerk mag niet teveel energie verbruiken. Behalve dat dit slecht is voor het milieu kost het veel geld.

Bronnen

1. RIPE, “RIPE509 - IPv4 Address Allocation and Assignment Policies for the RIPE NCC Service Region”, paragraaf 5.6, 18 januari 2011
   http://www.ripe.net/ripe/docs/ripe-509#----use-of-last----for-pa-allocations

2. OECD, “OECD Broadband Subscriber Criteria 2010”, geraadpleegd februari 2011
   http://www.oecd.org/document/46/0,3746,en_2649_34225_39575598_1_1_1_1,00.html

3. OPTA, “Wat doet OPTA - Breedband WBT”, geraadpleegd februari 2011
   http://www.opta.nl/nl/wat-doet-opta/markten/breedband-wbt/

4. The Chromium Project, “SPDY: An experimental protocol for a faster web”, geraadpleegd februari 2011
   http://www.chromium.org/spdy/spdy-whitepaper

5. Mike Belshe, “More Bandwith Doesn’t Matter (Much)”, april 2010
   http://docs.google.com/a/chromium.org/viewer?a=v&pid=sites&srcid=Y2hyb21pdW0ub3JnfGRldnxneDoxMzcyOWI1N2I4YzI3NzE2

Verslag Architectuurfase

Afstudeeropdracht

“**IPv4-connectiviteit in een IPv6 access network**”

Bateau Knowledge, Den Haag

Matthias van der Heide (07053940)

Haagse Hogeschool

Versie 1.0 – 6 juni 2011

Inhoudsopgave

1. Inleiding

Dit verslag van de Architectuurfase brengt in kaart hoe op basis van de opdracht een onderzoek is gestart naar mogelijke oplossingsrichtingen.

Eerst wordt in hoofdstuk drie uitgelegd hoe het literatuuronderzoek is uitgevoerd, en hoe op basis van deze en andere bronnen een selectie van kandidaattechnieken is gemaakt. In de benchmark in hoofdstuk vier wordt van deze technieken de meest geschikte oplossing gekozen.

In hoofdstuk vijf, ‘Architectuuraanpassing’, wordt een aangepast ontwerp van de architectuur ontwikkeld op basis van de nieuwgekozen oplossing.

2. Doelstelling

Het doel van de architectuurfase is om de alternatieve oplossingsrichtingen voor de aanpassing van het netwerk te inventariseren, en er een te kiezen die het meest geschikt is. De opdrachtomschrijving van de afstudeeropdracht, en de eisen uit het Programma van Eisen zijn hierbij leidend.

In de opdracht wordt gevraagd om een techniek te vinden die het mogelijk maakt IPv4-connectiviteit te bieden via een IPv6-netwerk op een manier die zo generiek mogelijk is.

Het Programma van Eisen stelt dat de geboden connectiviteit op vergelijkbaar niveau moet zijn als dat momenteel haalbaar is. Vanwege een tekort aan adressen moeten deze met enkele aansluitingen gedeeld worden. Het accessnetwerk moet uiteindelijk IPv6-only worden.

3. Literatuuronderzoek

3.1 Onderzoeksopdracht

Om de opdracht te kunnen uitvoeren moet een antwoord worden gevonden op de vraag: “Welke techniek maakt het mogelijk om generieke wijze verkeer van IPv4-only clients om te zetten naar een IPv6-only ISP-verbinding en die daarna weer naar IPv4 te vertalen?”

Voor dit onderzoek is de bibliotheek van de TU Delft geraadpleegd. Er is gezocht naar boeken die IPv6 beschrijven, in het Nederlands of Engels, geschreven in de laatste 10 jaar (2001 - 2011). Daarnaast is ook gezocht naar boeken over computernetwerken waarin IPv6 is opgenomen.

3.2 Literatuur

Op basis van bovenstaande criteria is een selectie van negen boeken gevonden. Deze zijn allen gelezen. De lijst met titels is te vinden in de bronnenlijst.

De transitietechnieken die in deze boeken beschreven staan zijn vooral gericht op het verkrijgen van IPv6-connectiviteit in een bestaand IPv4-netwerk. Dat is niet zo verwonderlijk, aangezien IPv4 momenteel het meest gebruikt is op internet. Voorbeelden van technieken zijn 6to4, Teredo en ISATAP. Deze zijn primair bedoeld voor een scenario waarbij IPv6 in een IPv4-netwerk wordt uitgerold. Dit is de tegenovergestelde situatie van die beschreven in de onderzoeksopdracht. Daarom worden ze hier niet verder behandeld.

De belangrijkste bron voor standaarden voor het internet is de Internet Engineering Taskforce (IETF). Binnen de Working Groups van de IETF worden Request For Comment-documenten opgesteld (RFC’s) . Voordat een RFC wordt goedgekeurd heeft het document een ‘draft status’, wat betekent dat de tekst nog in ontwikkeling is. Na discussie binnen de werkgroep en op de mailinglijst wordt na goedkeuring van het IETF een draft als RFC gepubliceerd.

In tabel 1 zijn de transitiemechanismen die in gebruikte boeken worden beschreven gesorteerd op RFC-nummer. Dit geeft een indicatie van de leeftijd van de techniek (lagere nummers zijn ouder). Daarnaast is de status van de RFC vermeld. Een bewerkte versie van dit overzicht is namens ‘Stichting IPv6 Nederland’ gepubliceerd hun website. [1]

Tabel 3.1 - Gevonden transitiemechanismen

Techniek	IETF RFC/draft	Status
IP Encapsulation within IP	2003	Proposed Standard
Generic Packet Tunneling in IPv6 Specification	2473	Proposed Standard

Stateless IP/ICMP Translation Algorithm (SIIT)	2765	Proposed Standard
Network Address Translation - Protocol Translation (NAT-PT)	2766	Historic
Generic Routing Encapsulation (GRE)	2784	Proposed Standard
An IPv6-to-IPv4 Transport Relay Translator (TRT)	3142	Informational
Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)	4798	Proposed Standard
IPv6 Tunnel Broker with the Tunnel Setup Protocol (TSP)	5572	Experimental
NAT64/DNS64	6144, 6145, 6146	RFC
Dual-Stack Lite (DS-Lite)	draft-ietf-softwire-dual-stack-lite-07	draft
Dual Stack IPv6 Dominant Transition Mechanism (DSTM)	draft-bound-dstm-exp-04	draft

3.3 Internetbronnen

Met behulp van tabel 3.1 is op internet gezocht naar meer informatie over de genoemde technieken. Dit bracht aan het licht dat werk aan deze standaarden de afgelopen jaren is voortgezet in diverse werkgroepen van het IETF, en dat daar een veelvoud aan ideeën is over hoe de overgang van een IPv4- naar een IPv6-internet kan worden gemaakt.

De volgende technieken werden nog niet in boeken, maar alleen binnen de werkgroepen van het IETF gevonden:

- A+P
- A+P Lite
- 4rd
- Dual-Stack Extra Lite

Discussies en blogs op internet vestigden de aandacht op twee andere technieken: NAT444 en NAT464. NAT444 wordt ook wel Large-Scale NAT (LSN) of Carrier-grade NAT (CGN) genoemd. De artikelen van Jeff Doyle bieden een bijzonder leesbare uitleg van de deze twee transitietechnieken.

Naar aanleiding van de publicatie van tabel 1 in een artikel op de site van Stichting IPv6 Nederland is de techniek LISP bekeken.

Een gesprek met Iljitsch van Beijnum gaf duidelijkheid over de stand van ontwikkeling, het werk binnen de IETF en het verschil tussen NAT en tunneling als transitiestrategie. Van Beijnum is mede-auteur van de RFC's over NAT64 en schrijver van de boeken "Running IPv6" (Apress) en "BGP" (O'reilly).

3.4 Techniekbeschrijvingen

6PE

Sommige ISP's hebben hun netwerk gebaseerd op Multiprotocol Label Switching (MPLS). Bij MPLS worden IPv4-pakketten op de router aan de rand van het netwerk voorzien van een label. Dit label geeft aan welk pad, een zogenaamd Label Switched Path (LSP), dit pakket moet volgen naar de eindbestemming. De eindbestemming is de MPLS-router in het netwerk die het dichtst bij het doel is. De labels zitten tussen de informatie voor de datalinklaag (bijvoorbeeld de Ethernet-header) en de IP-header. MPLS wordt daarom ook wel een laag 2,5 protocol genoemd. Bij de methode 6PE (IPv6 Provider Edge) wordt IPv6-connectiviteit geboden door op de Provider Edge routers ontvangen IPv6-packets op gelijke wijze te voorzien van een label, en op basis daarvan door het netwerk te routeren. Deze techniek is alleen van toepassing als het netwerk al is gebaseerd op MPLS. In haar boek legt Silvia Hagen dat als volgt uit: "The fact that MPLS can be used to transport IPv6 packets over IPv4 does not mean you should implement MPLS for this purpose. If you do not have an MPLS infrastructure in place, other tunneling mechanisms may be better suited to reach your goal. But if you already have MPLS, it's a great starting point." [2]. 6PE is dus geen handige oplossing om IPv6 in te voeren als het netwerk nog niet op MPLS gebaseerd is. Het biedt geen mogelijkheden om IPv4-verbindingen over IPv6 te leveren. De protocollen bestaan naast elkaar, en maken gebruik van MPLS als onderliggende techniek. Ook IPv4-adresdeling is met 6PE niet mogelijk.

A+P

Bij Address Plus Port (A+P) wordt het adrestekort opgelost door gebruikers niet een volledig IPv4-adres ter beschikking te stellen, maar slechts een deel daarvan. Dit kan door de adressering uit te breiden naar de poortnummers. Een 32-bits IPv4-adres wordt dan effectief 16-bits langer. Klant A kan bijvoorbeeld gebruikmaken een IP-adres met poorten 10.000 tot 30.000 en klant B van poorten 30.000 tot 50.000 op datzelfde IP-adres. Deze techniek werkt, net als NAT444 en Dual-Stack Lite, alleen voor de protocollen TCP en UDP, die zijn gebaseerd op poortnummers.

De CPE van klant A weet naar welke poorten op welk publiek IP-adres hij mag vertalen. Vertaald verkeer wordt in een tunnel naar de Port Range Router (PRR) gestuurd. De PRR is een machine in het netwerk van de ISP en heeft als taak om het verkeer dat uit te tunnel komt te controleren op het gebruik van de juiste adressen en poortnummers. Vervolgens wordt het verkeer naar het publieke internet gestuurd.

Door NAT aan de gebruikerszijde tot bepaalde port ranges te beperken, kan bespaard worden op NAT in het providernetwerk. Dit heeft voordelen m.b.t. de grootte van de NAT-tabel, logging en performance. De specificatie van A+P biedt de mogelijkheid om de PRR de functie van AFTR voor Dual-Stack Lite te laten vervullen. Hierdoor kan een mengeling van Dual-Stack Lite en A+P apparatuur worden gebruikt in hetzelfde netwerk. [3]

Dual-stack Lite

Dual-stack Lite is een overgangsmechanisme waarbij automatische tunnels in IPv6-(access)netwerken connectiviteit bieden voor IPv4-hosts. Het systeem bestaat uit twee elementen: de Basic Bridging BroadBand element (B4) op de router aan de klantzijde (CPE) en de Address Family Transitional Router (AFTR) in het netwerk van de ISP.

De B4 dient als default gateway voor de IPv4-hosts in het netwerk achter de CPE. Deze tunnelt door middel van encapsulation in IPv6 het verkeer naar een AFTR. De AFTR voert vervolgens NAT-vertaling uit op het getunnelde IPv4-pakket en routeert het naar het internet. [4]

A+P Lite

A+P Lite [5] is een combinatie van DS-Lite en A+P. Er wordt een koppeling gemaakt tussen het IPv6-adres van de klant en de het IPv4-adres en de verzameling poorten waarvan de klant gebruik mag maken. Dit moet zorgen voor het verminderen van toestandsinformatie (state) in het netwerk. Het voorkomen of verkleinen van de hoeveelheid state in het netwerk wordt als een groot goed gezien, omdat dit het netwerk simpeler en daarmee betrouwbaarder maakt.

4RD

IPv4 Residual Deployment (4RD) is opgesteld door Olivier Vautrin als tegenhanger van 6rd. 4RD is een manier om IPv4-verkeer te tunnelen over een IPv6-netwerk. [6] Verkeer wordt via IPv6 verzonden tussen de router van de gebruiker, de 4rd Customer Edge (CE), en de Border Router (BR) van de ISP. Als er gebruik wordt gemaakt van een algoritme om IPv4-adressen en poorten in IPv6-adressen te coderen, dan is dit stateless. De ISP hoeft dan geen gegevens bij te houden over de sessies van de gebruiker. Wordt er geen gebruik gemaakt van deze codering, dan moet voor elke verbinding een aparte NAT-sessie worden aangemaakt (stateful NAT44) en is de methode gelijk aan DS-Lite.

Sinds het verschijnen van deze draft is er niet meer aan 4rd ontwikkeld en is het werk verplaatst naar de IETF Intarea Working Group. Binnen deze groep is een nieuwe draft geschreven door Remi Depres. [7] In de draft van Remi Depres is 4RD uitgebreid met de mogelijkheid om over een ISP-netwerk te functioneren dat ofwel IPv6-only is, of RFC1918 (private IPv4 addressing) of allebei. In het geval van een RFC1918-netwerk wordt IPv6-connectiviteit geleverd door middel van 6rd, dat door Despres ontwikkeld is om IPv6 in een IPv4-netwerk te leveren [8]. Wordt zowel RFC1918-adresruimte als IPv6 gerouteerd dan vervalt het nut van 6rd en kan IPv6-verkeer direct plaatsvinden tussen hosts en het internet (end-to-end).

Er is geen eenduidige specificatie van 4rd. De hierboven beschreven documenten zijn drafts op persoonlijke titel van de auteurs, en worden dus niet (nog) door een Working Group gesteund.

Dual-Stack Extra Lite

Dual-Stack Extra Lite is een uitbreiding op Dual-Stack Lite. [9] Deze methode gaat er van uit dat als achter een router vele devices hangen die elk op een point-to-point link zitten, en waarvan de broadcast domains (subnets) niet overlappen, zoals bij consumentennetwerken, dat het dan niet nodig is om alle devices een uniek IP-adres te geven. Het apparaat kan worden geïdentificeerd door de point-to-point-link. Daarom kan elk apparaat hetzelfde, private, adres gebruiken, zonder dat dit problemen oplevert. Dit is bijvoorbeeld het geval als DSL-modems via PPP over Ethernet (PPPoE) of PPP over ATM (PPPoA) een verbinding maken. Een vergelijkbare situatie bestaat bij mobiele aanbieders, waarbij dataverbindingen ook bestaan uit tunnels naar een centraal punt.

Bij Dual-Stack Lite wordt voor alle clients achter een aansluiting het privé IPv4-adres wordt bewaard in de NAT-mapping. Bij Dual-Stack Extra Lite hoeft deze informatie niet extra te worden opgeslagen, omdat altijd hetzelfde private IP-adres wordt gebruikt voor de vertaling. Dit bespaart op het geheugengebruik van NAT-apparatuur.

DSTM

Het ‘Dual-Stack Transition Mechanism’, of DSTM, is een eerste transitiemechanismes bedacht om IPv4-hosts te verbinden over een IPv6-netwerk. Hosts die IPv4-connectiviteit willen, benaderen een centrale server en verzoeken het tijdelijke gebruik (leasen) van een publiek IPv4-adres.

De draftspecificatie van DSTM [10] schrijft geen specifieke tunnelingtechniek voor, en is daarmee niet concreet genoeg voor een implementatie. DSTM is onder andere ontwikkeld door Alain Durand, die later Dual-Stack Lite bedacht. Doorontwikkeling van DSTM is gestaakt in 2002. Er is geen werk gaande om van DSTM een internetstandaard te maken. Wel is er een werkende implementatie gemaakt door Freenet6, het latere Gogo6. [11]

Generic Packet Tunneling in IPv6 Specification (RFC 2473)

Om het mogelijk te maken om IPv6- en IPv4-verkeer te tunnelen over IPv6 is RFC 2473 ontwikkeld. [12] Deze RFC beschrijft hoe, door simpelweg IPv6-headers voor een IPv4- of IPv6-packet te plaatsen, dit verkeer over een IPv6-netwerk kan worden geleid. RFC2473 is de IPv6-variant van IP-in-IP-encapsulation zoals beschreven in RFC 2003 [13].

De IPv6-tunnel bestaat uit een begin- en een eindpunt, maar legt niet vast of en hoe deze elkaar kennen. Het opzetten van deze tunnels moet dus aan een andere techniek worden overgelaten. Een pakket dat de tunnel betreedt wordt voorzien van een IPv6-header van 40 bytes, met als Next Header-waarde het IP-type van het getunnelde pakket. Door de extra header kan een packet te groot worden voor de MTU van het onderliggende IPv6-netwerk. Indien mogelijk wordt het packet gefragmenteerd in IPv6 en in delen naar het eindpunt gestuurd. Deze voegt de delen samen, en routeert het oorspronkelijke packet verder.

Generic Routing Encapsulation (GRE)

Een techniek door Cisco ontworpen om netwerkprotocollen generiek te tunnelen over andere netwerkprotocollen. Tunnelinterfaces die GRE gebruiken zijn hoofdzakelijk stateless, wat betekent dat een host of router verkeer kan tunnelen richting een andere host of router zonder dat deze hier van op de hoogte is. Hierdoor is onbekend of een link up of down is, anders dan door te controleren of de doelhost van de tunnel bereikbaar is. [14]

GRE kan gebruikt worden om IPv4-verkeer te tunnelen in IPv6-pakketten, maar ontbeert mogelijkheden om IP-adressen te delen.

IP Encapsulation within IP (IP-in-IP)

IP-in-IP is in 1996, voor de standaardisatie van IPv6, bedacht om source routing mogelijk te maken voor Mobile IP-toepassingen. Mobile IP kan worden toegepast als een apparaat verhuist van zijn eigen naar een ander netwerk. Bijvoorbeeld als je je telefoon of laptop meeneemt naar je werk. De router thuis kan in dit geval het verkeer voor je laptop naar het netwerk van je werkgever sturen. Dit principe heet source routing, en is geïmplementeerd in IPv4. De ondersteuning voor source routing is om redenen van veiligheid en complexiteit erg slecht. IP-in-IP is bedacht om dit probleem op te lossen. Het kapselt de packets in, door er een IPv4-header voor te plakken. Deze packets worden naar het adres van je laptop in het netwerk van je werkgever gestuurd. [13]

LISP

Het Locator Identifier Seperation Protocol, afgekort LISP, heeft als doel om het internet weer schaalbaar te maken door onderscheid te maken tussen ‘wie’ en ‘waar’ in de routeringstabellen. Momenteel zijn de tabellen voor routers in de default-free zone (DFZ) erg groot, en ze groeien nog steeds. In LISP wordt het aantal regels in de tabellen verminderd door IP-adressen te gebruiken om hosts te identificeren (‘wie’ stuurt iets) en te routeren op basis van locators (‘waar’ moet het pakket heen). Hierdoor zou meer aggregatie van kunnen plaatsvinden in de routeringstabellen, zodat deze kleiner worden. Routers die LISP ondersteunen versturen verkeer naar elkaar via een tunnel. De Ingress Tunnel Router (ITR) zoekt op basis van de identifier van het pakket (het IP-doeladres) de bijbehorende locator (ook een IP-adres) op. Vervolgens wordt het pakket ingepakt in een UDP-pakket en naar de Egress Tunnel Router (ETR) verstuurt. Een verzameling ITR- en ETR-routers vormt samen een LISP-domein. [15]

LISP-tunnel zijn in principe adresfamilie-onafhankelijk. Daardoor kan een tunnel over IPv4 zowel IPv4- als IPv6-verkeer transporteren. Hierdoor kan een netwerk dat alleen via IPv4 te bereiken is ook IPv6-verkeer accepteren. Dit is de reden dat LISP als transitiemechanisme wordt geopperd.

NAT444

Een mogelijkheid om de levensduur van IPv4-adressen op te rekken is door een extra laag NAT-apparatuur aan te brengen. Huidige klanten van ISP's hebben al een NAT-functie (adresdeling door meerdere pc's) in hun router. Dit wordt ook wel NAT44 genoemd. De 44 betekent dat er vertaling plaatsvindt van een IPv4- naar een ander IPv4-adres. Bij NAT444 plaatst de provider zijn klanten ook nog eens achter een NAT-device. In het netwerk tussen de klant en de provider wordt dan RFC1918-adresruimte gebruikt (adressen uit het 10.0.0.0- of 192.168.0.0-bereik). Pakketten gericht naar hosts op het publieke internet worden nogmaals vertaald. Vandaar de drie vieren in NAT444. [16]

NAT444 is ook bekend onder de namen Large-Scale NAT (LSN) en Carrier-Grade NAT (CGN).

Een gevaar van dubbel-NAT is dat sommige applicaties die rekenen op de aanwezigheid van een NAT, niet meer werken als er twee NAT-apparaten in het pad naar het internet zitten. Daarnaast zijn er problemen te verwachten met routers die slecht onderscheid kunnen maken tussen de adresruimte die in het LAN van de klant wordt gebruikt, en dezelfde adresruimte in het netwerk van de ISP.

Om deze problemen te voorkomen is een specificatie opgesteld voor Large-Scale NAT. Deze moet er voor zorgen dat bovenstaande problemen geminimaliseerd worden. [17]

Hoewel NAT444 gebruikt kan worden om IPv4-adressen effectiever te gebruiken, wordt het gezien als een lapmiddel die de groeiende vraag naar IPv4-adressen vertraagd, maar niet kan doen dalen. Hierdoor is het geen geschikte oplossing op de lange termijn.

NAT464

Vergelijkbaar met NAT444 is de vorm NAT464. Hierbij wordt het netwerk van de provider geheel in IPv6 uitgevoerd. Verbindingen van IPv4-clients naar hosts op het internet worden door de CPE vertaald naar IPv6 door middel van NAT46. Wanneer de packets aan de rand van het providernetwerk komen worden ze door de routers van de provider weer vertaald naar IPv4. IPv6-packets met IPv4-data zijn onherkenbaar van normaal IPv6-verkeer, op het adres na. Hierdoor kan de provider deze pakketten naar een centrale locatie routeren en terugvertalen naar IPv4. [18]

Hoewel NAT464 als voordeel heeft dat het netwerk van de provider compleet IPv6 kan worden gemaakt, blijven de gevaren van dubbel-NAT.

Er zijn geen werkende implementaties van NAT464 bekend. Werk aan NAT464 in de context van de IETF Softwires Working Group is na stemming gestaakt, laat Alain Durand per e-mail weten.

NAT-PT

Network Address Translation - Protocol Translation (RFC 2766) [19] is een vroege poging om vertaling van IPv4 naar IPv6 te doen. De techniek wordt meerdere malen beschreven in de boeken, maar is door de internetgemeenschap (IETF) als verouderd verklaart, en ongeschikt bevonden als transitiemechanisme voor het gehele internet. Hun bezwaren zijn samengevat in RFC 4966, “Reasons to Move NAT-PT to Historic status”. [20] Redenen zijn onder andere de beperkingen opgelegd door het gebruik van een Application Level Gateway voor DNS, problemen met fragmentatie, multicast en IPv4-adressen die voorkomen in de data van de IP-pakketten.

Een aangepaste versie van NAT-PT, genaamd NAT64+DNS64 maakt een comeback als transitiemechanisme voor mobiele netwerken en netwerken met IPv6-only hosts. [21]

NAT64/DNS64

NAT64 is een standaard opgesteld om het verouderde NAT-PT (zie hierboven) een nieuwe kans te geven. De techniek is opnieuw beschreven en op sommige punten aangepast. NAT64 is bedoeld om IPv6-only hosts toegang te bieden tot het IPv4-internet. Een IPv6-client die verbinding wil maken met een server op het internet vraagt het IPv6-adres van deze server op door een verzoek naar de DNS-server te sturen. Op de DNS-server is een DNS64 Application Level Gateway (ALG) actief, die nagaat of voor deze server een IPv6-adres beschikbaar is (AAAA record). Is dat niet het geval, en is alleen een IPv4-adres bekend, dan genereert de ALG op basis van het IPv4-adres en een door de provider gekozen IPv6-prefix een nieuw IPv6-adres, dat als antwoord naar de client wordt gestuurd.

De client maakt vervolgens een IPv6-verbinding met dit adres, wat uitkomt op een router die NAT64 ondersteunt. Op basis van de gegevens in het IPv6-adres voert deze de vertaling van IPv6 naar IPv4 uit. [21]

Voor de werking van NAT64 en DNS64 is extra apparatuur in het netwerk nodig. NAT64 kan IPv6-only hosts toegang tot het IPv4-internet bieden. Daarom is het geschikt voor gesloten netwerken, waar de hosts beheerd zijn door de provider, of modern genoeg zijn om alleen via IPv6 te werken. Dat geldt bijvoorbeeld voor mobiele netwerken, waarin de provider bepaalt welke toestellen worden verkocht.

Stateless IP/ICMP Translation (SIIT)

Aangezien IPv4 en IPv6 grotendeels dezelfde functionaliteit bieden (transport van data tussen hosts) bevatten de IP-packets grotendeels dezelfde informatie. RFC 2765 beschrijft een algoritme voor een Protocol Translator. [22] Dit apparaat kan een vertaalslag maken van IPv4-packets naar IPv6-packets en omgekeerd. SIIT is dus geen transitiemechanisme op zich, maar een onmisbaar onderdeel van een grotere techniek. Mechanismen zoals stateless NAT64 en NAT-PT maken gebruik van (een aangepaste versie van) SIIT.

TRT

De Transport Relay Translator (RFC 3142) [23] is een onderdeel in het netwerk dat bidirectionele transportverbindingen (TCP of UDP) omzet van IPv6 naar IPv4. Het apparaat is het eindpunt van de IPv6-verbinding die bijvoorbeeld over TCP loopt en zet een IPv4-verbinding op naar het beoogde eindpunt. Tijdens de duur van de verbinding zet het de gegevens uit de packets over van de ene adresfamilie naar de andere. Hier vindt geen vertaling plaats, maar wordt de data op transportniveau (laag 4) overgezet van IPv6 naar IPv4 en vice versa. TRT kan daarmee omschreven worden als een laag-4-NAT. Het is bedoeld om gebruikt te worden in een IPv6-only netwerk, als manier om verbindingen naar IPv4-only hosts te ondersteunen. Het is alleen geschikt voor tweewegcommunicatie en is stateful, zoals NAT44 en stateful NAT64. Ook heeft het aparte voorzieningen nodig om NAT-onvriendelijke protocollen toe te staan.

TSP

Het Tunnel Setup Protocol (TSP) is geen compleet transitiemechanisme maar een XML-protocol dat clients kunnen gebruiken om tunnels op te zetten. [24][25] Het protocol is generiek, waardoor het tunnels voor alle combinaties van IPv4, IPv6 en andere (netwerk)protocollen kan opzetten. Het kan samen met een tunnelingmechanisme gebruikt worden om een systeem te bouwen dat automatische tunnels voor IPv4-over-IPv6 maakt. TSP-clients zouden in CPE-routers ingebouwd kunnen worden om klanten een simpele interface te geven om tunnels te configureren.

Een implementatie van TSP is verkrijgbaar bij Gogo6. Deze maakt gebruik van L2TP als onderliggend tunnelingmechanisme. [26]

4. Benchmark

4.1 Overzicht van oplossingsrichtingen

Als we de transitiemechanismen indelen op basis van de onderliggende techniek dan komen we tot de volgende ruwe indeling:

- **Dual-stack**
  Netwerk waarin zowel IPv4 als IPv6 gerouteerd wordt.

- **Tunnelen (tunneling/encapsulation)**
  Het inkapselen (encapsulation) van IPv4-verkeer in IPv6-pakketten. Sommigen maken onderscheid tussen inkapseling (encapsulation), wat stateless kan zijn, en tunnelen (tunneling) waarbij aan beide eindpunten de toestand (state) van de tunnel wordt bijgehouden.

- **Vertaling (translation)**
  Vertalen van datapakketen van de ene adresfamilie naar de andere.

Daarnaast zijn er technieken die een combinatie vormen van deze concepten. DS-Lite is bijvoorbeeld encapsulation gecombineerd met vertaling in het accessnetwerk. Ook zijn er technieken die voortbouwen op onderdelen van andere technieken, zoals NAT64/DNS64 gebruik maakt van SIIT met een DNS Application Level Gateway.

4.2 Eliminatie

Op basis van de mogelijkheden die de technieken bieden is een aantal scenario's bedacht, dat in de vergelijking tegen elkaar af is gezet. Niet alle gevonden technieken kunnen zelfstandig gebruikt worden of als onderdeel van een bedacht scenario. Om die reden zijn de volgende technieken niet meegenomen in de afweging:

- 6PE
- GRE
- NAT-PT (verouderd)
- SIIT
- TRT
- GPT
- IP-in-IP (RFC 2003)

De lijst in tabel 2 geeft een overzicht van technieken die gebruikt kunnen worden. De Voor de volledigheid noemen we ook de non-oplossingen IPv4-only, IPv6-only en Dual-Stack mee, om de andere oplossingen hiermee te kunnen vergelijken.

Voor het kiezen van geschikte technieken zijn drie criteria van belang:

- Dual-stack LAN mogelijk; met deze techniek kunnen zowel IPv4- als IPv6-capable hosts in het klantnetwerk bediend worden.
- IPv4-adresdeling tussen aansluitingen: de techniek stelt de ISP in staat om de beperkte voorraad IPv4-adressen te gebruiken met meer klanten dan hij adressen heeft
- IPv6-only accessnetwerk; de techniek werkt in een ISP-netwerk dat alleen gebruik maakt van IPv6

Tabel 4.2 - mogelijkheden van transitiemechanismen

Scenario	**Dual-stack klantnetwerk**	**IPv4-adresdeling tussen aansluitingen**	**IPv6-only accessnetwerk**	**Beperkingen**
IPv4 only	nee	nee	nee	
IPv6 only	nee	n.v.t. *	ja	* geen IPv4-verbinding aanwezig
Dual-stack	ja	nee	nee	
NAT444 + IPv6	ja	ja	nee	
A+P	ja	ja	ja	
NAT64/DNS64	nee	ja	ja	
NAT464	ja	ja	ja	
DS-Lite	ja	ja	ja	
4RD	ja	ja	ja	
IPv6-only met TSP client in CPE	ja	nee	ja	
DSTM	ja	nee	ja	
LISP	ja	nee	ja	

Op basis van bovenstaande afweging komen we op vier technieken die aan alle eisen voldoen:

- NAT464
- DS-Lite
- A+P
- 4RD

4.2 Selectie

Voor het uitvoeren van de opdracht is het nodig om één techniek te kiezen die in een proof of concept uitgewerkt kan worden. De volgende criteria zijn hiervoor van belang:

- volwassenheid van de standaard
- beschikbaarheid van implementatie

De criteria worden met punten beoordeeld. De waardering is aangegeven in de bijbehorende tabellen.

Volwassenheid van de standaard
Is de standaard stabiel? Is de specificatie voor iedereen beschikbaar en voldoende uitgewerkt? Dit zijn factoren die van belang zijn voor een goede specificatie en voor de ontwikkeling van implementaties die compatibel met elkaar zijn.

Punten	Voorwaarden
5	De specificatie is openbaar (bijvoorbeeld een IETF RFC) en wordt actief ontwikkeld door de IETF of een ander instituut
4	De specificatie is een draft en wordt actief ontwikkeld door de IETF of een ander instituut met de bedoeling hier een standaard van te maken (bijvoorbeeld als onderdeel van het charter van een Working Group)
3	De specificatie is een draft en wordt actief ontwikkeld buiten het charter van een IETF Working Group of door derden
2	De specificatie is openbaar (bijvoorbeeld een IETF RFC) maar er vindt geen verdere ontwikkeling plaats
1	Er is een specificatie, maar deze is niet publiekelijk beschikbaar
0	Er is geen specificatie gepubliceerd

Beschikbaarheid van implementaties
Voor het uitwerken van een proof of concept, waarbij geen eigen software- of hardwareontwikkeling plaats vindt, is het cruciaal dat er implementaties beschikbaar. Daarom is gezocht naar leveranciers die de standaard implementeren. Dit kunnen ook opensource-initiatieven zijn, of referentie-implementaties. Alle onderzochte technieken vereisen zowel een aanpassing aan de router van de ISP-klant (CPE) als in het netwerk van de ISP zelf. Het vermelde puntenaantal bestaat uit twee getallen, waarvan het eerste aangeeft hoeveel CPE-implementaties er gevonden zijn, en het tweede het aantal implementaties van de techniek voor in het ISP-netwerk.

Samenstelling van de eindscore

De eindscore van een techniek wordt bepaald door beschikbaarheid van implementaties. Zowel van de ISP- als de klantkant moet minstens één implementatie beschikbaar zijn. Het aantal punten voor de volwassenheid van de techniek wordt hierbij opgeteld. Op deze manier geeft een techniek die niet volwassen is, maar wel veel wordt geïmplementeerd, toch kans om onderzocht te worden. Tegelijk wordt bij een gelijk aantal implementaties voorrang gegeven aan technieken die volwassener zijn.

4.3 Score

Ten tijde van het uitvoeren van dit onderzoek zijn de websites van de IETF en diverse fabrikanten van netwerkapparatuur geraadpleegd. Ook is direct contact gezocht met diverse bedrijven om te informeren naar de beschikbaarheid van implementaties. De bevindingen zijn vermeld in de bijlage ‘Implementaties van Dual-Stack Lite’.

Techniek	Volwassenheid	Beschikbaarheid	Eindscore
NAT464	0	0 + 0	0
Dual-Stack Lite	4	5 + 6	15
A+P	3	1 + 1*	3
4RD	3	0 + 0	3

* de implementatie van A+P door Orange Bejing is pas na het uitvoeren van dit deel van het onderzoek bekendgemaakt, en kan dus niet worden meegeteld in de eindscore.

4.4 Conclusie

Op basis van de beschikbaarheid van implementaties, en de volwassenheid van de standaard, is gekozen voor de techniek Dual-Stack Lite. Deze heeft verreweg de meeste beschikbare implementaties en is ook onderweg om een standaard te worden binnen de IETF.

5. Architectuuraanpassing

Nu de keus gemaakt is om Dual-Stack Lite toe te passen in het providernetwerk, kan worden nagedacht over de aanpassing van de architectuur. Een gedeelte van het netwerk moet IPv6 gaan ondersteunen, om communicatie tussen de CPE's van de klanten en de AFTR mogelijk te maken. Routers van klanten moeten omgeruild voor exemplaren die Dual-Stack Lite ondersteunen. En een of meerdere servers die AFTR-functionaliteit bieden moeten in het netwerk worden opgenomen. Hiervoor zijn diverse voorbeeldarchitecturen bedacht, die we toelichten met de voor- en nadelen. Uiteindelijk wordt hier de beste uit gekozen, die zal worden uitgewerkt in de ontwikkelfase.

5.1 Aanpassing klantapparatuur

In een netwerk dat is opgebouwd op basis van IPv4 kan niet zomaar worden overgeschakeld naar Dual-Stack Lite en IPv6. Een belangrijk onderdeel dat aangepast moet worden, is de router (CPE) van de klant. Bestaande CPE's zijn zelden of nooit geschikt voor IPv6. Ondersteuning voor nieuwere technieken, zoals Dual-Stack Lite, is geheel afwezig.

Updates

De aanpassingen die nodig zijn om Dual-Stack Lite te ondersteunen in een router bevinden zich allemaal in software. DS-Lite is gebouwd op basis van bestaande IPv4- en IPv6-technieken en vereist geen nieuwe hardware, mits deze voldoende capaciteit heeft om IPv4-verkeer te tunnelen over IPv6. In netwerken waar de ISP controle heeft over de CPE is het mogelijk om de software op afstand aan te passen. Als de leverancier van de CPE een firmware met ondersteuning voor Dual-Stack Lite ter beschikking stelt is dit de snelste en goedkoopste manier om apparatuur bij klanten te upgraden. De ontwikkeling van een aangepaste firmware vergt een investering van de leverancier, die hij mogelijk doorbelast aan de ISP.

Vervanging

Het is te verwachten dat sommige typen routers niet voldoen aan de hardwarematige eisen om IPv6 en Dual-Stack Lite te ondersteunen. IPv4-routers zijn niet ontwikkeld met uitbreidingsmogelijkheden in het achterhoofd. De hardware is vaak toegespitst om zijn taak met een zo klein (en dus goedkoop) mogelijk geheugen en afdoende processor te vervullen. Ook kan het zijn dat een leverancier van CPE's liever het nieuwste model verkoopt, in plaats van het gratis of tegen betaling aanpassen van de firmware van een ouder model dat is uitontwikkeld. Is hij hier wel toe bereid, dan kan de ISP de kosten voor aanpassing nog te hoog vinden. En de mogelijkheid bestaat dat de huidige leverancier deze aanpassing niet kan of wil leveren, bijvoorbeeld omdat ze geen reden zien om hier in te investeren, op wat

voor manier dan ook. Vervanging van de CPE is dan de enige optie. Mogelijk moet een nieuwe leverancier worden gevonden, die bereidt is om samen met de ISP een Dual-Stack Lite-router te ontwikkelen.

Investeringen in CPE vormen een aanzienlijk deel van de investeringen in het netwerk. Providers rekenen hun klanten een bedrag van enkele tientjes voor de aanschaf van een modemrouter (ADSL) of rekenen tot 50 euro borg (kabel). Dit betekent dat vervanging van 6000 CPE's een kostenpost van 300.000 euro kan zijn. Daar zijn de kosten voor instellen, verzending en ondersteuning nog niet bij gerekend. Nieuwe klanten accepteren mogelijk dat dit bedrag bij hen in rekening wordt gebracht bij het afsluiten van een abonnement. Bestaande klanten zullen echter niet bereid zijn om te betalen voor vervanging van apparatuur die in hun ogen prima werkt. In dit geval kan vervanging daarom het beste plaatsvinden bij verhuizing, wijziging van abonnement of defect raken van de oude router.

5.2 Plaatsing AFTR

Op basis van onderzoek door de University of Waikato kan de conclusie worden getrokken dat een klant maximaal 500 poorten nodig heeft om normaal gebruik te maken van het internet. [27] Wel moeten we er rekening mee houden dat dit mogelijk geldt voor één host en in het geval een klant meerdere hosts wil aansluiten, dit vermenigvuldigd moet worden. Ook bestaat er een groot verschil tussen veelverbruikers en klanten die weinig gebruik maken van internet, maar toch een breedbandverbinding hebben. We gaan er van uit dat een klant maximaal 5 computers aansluit op zijn internetaansluiting, en dat deze elk maximaal 500 poorten gebruiken om verbindingen naar het internet op te zetten.

Eén IP-adres kan maximaal 64.000 verbindingen tegelijk faciliteren, omdat het poortnummer van TCP en UDP een lengte van 16 bits heeft. Hiermee zouden maximaal 25 klanten een IP-adres kunnen delen. Dit is een bovengrens, maar dit aantal kan ook minder zijn. Sterker nog, op basis van het huidige productaanbod van de ISP, de adresvoorraad na uitbreiding (7.000) en de maximale aantallen aansluitingen (16.900) zou een IP adres maximaal door drie klanten gedeeld hoeven te worden. Als alleen rekening wordt gehouden de aanleg van wijkstation 3, zou dit getal uitkomen op vier. Dat betekent dat adresdeling al kan plaatsvinden op het laagste niveau, dat van het wijkstation.

We gaan er van uit dat netwerkleveranciers apparatuur ontwikkelen die voor 1000 publieke IPv4-adressen de AFTR-functionaliteit levert. Dit is een conservatieve schatting, op basis van specificaties van Large-Scale NAT-implementaties. Hiermee kunnen in het netwerk van de ISP 4.000 klanten worden bediend.

Architectuur 1

Bij adresdeling op wijkniveau wordt de AFTR in het wijkstation geplaatst en gekoppeld met de router. De AFTR handelt het verkeer af voor 1200 klanten (start) met de mogelijkheid om door te groeien naar 3100 aansluitingen. Het netwerk tussen de AFTR en de klanten is IPv6-only. Het corenetwerk en de verbinding met het wijkstation zijn IPv4-only.

Deze opzet heeft als nadeel dat alleen het accessnetwerk in wijkstation 3 geschikt is voor IPv6. Omdat IPv6 op het corenetwerk en de tussenliggende verbinding ontbreekt, kan geen IPv6-connectiviteit geleverd worden.

Architectuur 2

Door de verbinding tussen wijkstation 3 en het corenetwerk ook te voorzien van IPv6 kan aan klanten IPv6-connectiviteit worden geboden. Het is dan wel noodzakelijk dat het corenetwerk dual-stack (IPv4 + IPv6) wordt gemaakt, en dat voor IPv6-verkeer een contract wordt afgesloten met een transitprovider. Op deze manier voldoet het wijkstation aan de gestelde eis om IPv4- en IPv6-connectiviteit te leveren.

Architectuur 3

Op het moment dat een AFTR uitvalt ontstaat voor alle gebruikers die op dat moment een verbinding via Dual-Stack Lite hebben een probleem. Onderdeel van de AFTR is de NAT-sessietabel. In deze tabel staat welke klant gebruik maakt van welk IP-adres en welke poorten. Bij uitval van de AFTR gaan deze gegevens verloren. Alle verbindingen zijn dan verbroken, en moeten (na herstel van het probleem) opnieuw worden opgebouwd.

Dit probleem kan worden verkleind door de AFTR redundant uit te voeren. De plaatsing van een tweede AFTR voorkomt dat een defect aan één apparaat alle IPv4-verbindingen verbreekt. De apparaten synchroniseren hun NAT-sessietabel, en in geval van een storing neemt de backup-AFTR de taak van de primaire AFTR over. De kans dat beide AFTR's tegelijk uitvallen door storing is zeer klein. Klanten kunnen in dat geval terugvallen op IPv6-verbindingen.

Het redundant uitvoeren van de AFTR verdubbelt de investering per locatie, en dus de investering per klant.

Architectuur 4

Hoewel architectuur 3 voldoet aan de eis om IPv4- en IPv6-connectiviteit te leveren, is de schaalbaarheid van deze oplossing beperkt. Als ook andere wijkstations worden aangepast om Dual-Stack Lite en IPv6 te ondersteunen, moet in elk station twee AFTR's geplaatst worden. Dit leidt tot een grote overcapaciteit, omdat in deze stations maar 1000 tot 2000 aansluitingen actief zijn. Als de gebruikers in deze wijken één voor één overstappen van native IPv4 naar Dual-Stack Lite, zal het gebruik van deze apparatuur in het begin minimaal zijn.

Door de AFTR's in het corenetwerk te plaatsen kan het gebruik veel efficiënter. In eerste instantie worden de twee AFTR's voor wijkstation 3 in het datacenter geplaatst. Deze opstelling heeft voldoende capaciteit voor de eerste 1200 klanten die van Dual-Stack Lite gebruik maken. De overcapaciteit (2800 klanten) kan gebruikt worden om nieuwe en bestaande klanten te migreren naar Dual-Stack Lite en IPv6. Als de grenzen bereikt zijn, wordt de capaciteit uitgebreid door het bijplaatsen van AFTR's. De investering is op deze manier gekoppeld aan het gebruik van Dual-Stack Lite, niet aan het aantal wijkstations dat geschikt wordt gemaakt.

Daarnaast biedt deze variant het voordeel dat op termijn alle wijkstations kunnen worden overgezet naar een IPv6-only netwerk. Hiervoor moeten alle routers van klanten geschikt zijn voor Dual-Stack Lite. Met de overzetting naar IPv6-only is het doel om in het accessnetwerk maar één protocol te routeren behaald.

6. Conclusie

Op basis van de gestelde eisen komt Dual-Stack Lite als best geschikte oplossingsrichting naar voren. Deze techniek is het meest volwassen en uitgewerkt van de geboden opties en biedt de meeste implementaties om te evalueren.

Architectuur 4 biedt van de voorgestelde aanpassingen de beste match met de wensen van de ISP en de laagste kosten voor invoering. Daarom wordt deze in de Ontwerpfase verder uitgewerkt.

Bronnen

1. Stichting IPv6 Nederland, “Het woud van transitiemechanismen”
   http://www.stipv6.nl/transitie
2. Hagen, Silvia, “IPv6 Essentials”, 2002
3. Bush, R., “The A+P Approach to the IPv4 Address Shortage”, 17 februari 2011
   http://tools.ietf.org/html/draft-ymbk-aplusp-09, geraadpleegd 24-2-2011
4. Durand, “Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion”, 11 augustus 2010
   http://tools.ietf.org/html/draft-ietf-softwire-dual-stack-lite-06
5. Xiaoyu, Z., “Aplusp Lite - A light weight aplusp approach”, 26 mei 2010
   http://tools.ietf.org/html/draft-xiaoyu-aplusp-lite-00
6. Vautrin, “IPv4 Rapid Deployment on IPv6 Infrastructures (4rd)”, 29 juli 2010
   http://tools.ietf.org/html/draft-vautrin-softwire-4rd-00
7. Despres, R., “IPv4 Residual Deployment across IPv6-Service networks (4rd)”, 14 maart 2011
   http://tools.ietf.org/html/draft-despres-intarea-4rd-01
8. Despres, R., “IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)”, januari 2010
   http://tools.ietf.org/html/rfc5569
9. Arkko, “Scalable Operation of Address Translators with Per-Interface Bindings”, 4 februari 2011
   http://tools.ietf.org/search/draft-arkko-dual-stack-extra-lite-05
10. Bound, “Dual Stack Transition Mechanism (DSTM)”
    http://tools.ietf.org/html/draft-ietf-ngtrans-dstm-08
11. Gogo6, “Freenet6 DSTM”
    http://gogonet.gogo6.com/page/freenet6-dstm
12. Deering, “Generic Packet Tunneling in IPv6”, december 1998
    http://tools.ietf.org/html/rfc2473
13. Perkins, “IP Encapsulation within IP”, oktober 1996
    http://tools.ietf.org/html/rfc2003
14. Farinacci, “Generic Routing Encapsulation (GRE)”, maart 2000
    http://tools.ietf.org/html/rfc2784
15. Website LISP
    http://www.lisp4.net
16. Doyle, Jeff, “Large Scale NAT Architectures”, 30 september 2009
    http://www.networkworld.com/community/node/45776
17. Perreault, “Common requirements for IP address sharing schemes”, 12 maart 2011
    http://tools.ietf.org/html/draft-ietf-behave-lsn-requirements-01

18. Doyle, Jeff, “Large Scale NAT Architectures”, 30 september 2009
http://www.networkworld.com/community/node/45776

19. RFC 2766, “Network Address Translation - Protocol Translation (NAT-PT)”, februari 2000
http://tools.ietf.org/html/rfc2766

20. RFC 4966, “Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status”, juli 2007
http://tools.ietf.org/html/rfc4966

21. Bagnulo, “Stateful NAT64”, april 2011
http://tools.ietf.org/html/rfc6146

22. Nordmark, “Stateless IP/ICMP Translation Algorithm (SIIT)”, februari 2000
http://tools.ietf.org/html/rfc2765

23. RFC 3142, “An IPv6-to-IPv4 Transport Relay Translator”, juni 2000
http://tools.ietf.org/html/rfc3142

24. Blanchet, “IPv6 Tunnel Broker with the Tunnel Setup Protocol (TSP)”, februari 2010
http://tools.ietf.org/html/rfc5572

25. Blanchet, ”Migrating to IPv6”, 2006

26. Gogo6, “gogoCPE”
http://gogoware.gogo6.com/4105/description.asp?product_id=180

27. Alcock, Shane (University of Waikato), “Research into the Viability of Service-Provider NAT”, augustus 2008
http://www.wand.net.nz/~salcock/someisp/flow_counting/result_page.html

28. Ahmed, “Deploying IPv6 in Broadband Access Networks”, 2009

29. Amoss/Minoli, “Handbook of IPv4 to IPv6 Transition”, 2008

30. Cisco, “Deploying IPv6 Networks”, 2006

31. Davies, “Understanding IPv6”, 2003

32. Goralski, “The Illustrated Network”, 2009

33. Hagen, “IPv6 Essentials”, 2006

34. Stockebrand, “IPv6 In Practice”, 2007

Bijlage – Implementaties van Dual-Stack Lite

Producten die Dual-Stack Lite B4 implementeren

Fabrikant	Product	Bron
D-Link	DIR-655 HW version B1 with special firmware	Hans_Liu@dlink.com.tw
Gogo6	Gogo CPE DS-Lite	http://gogoware.gogo6.com/4105/description.asp?product_id=180
AVM	Fritz!BOX 6360 Cable	http://www.avm.de/en/news/artikel/2010/Cebit_2010_6630.html
Comcast/Xavient	DS-Lite client for OpenWRT on Linksys 160NL and WRT54GL	http://www.comcast6.net/6rd-config.php#config3
Zyxel	onbekend	http://www.zyxel.nl/web/news_news.php?sqno=444

Producten die Dual-Stack Lite AFTR implementeren

Fabrikant	Product	Bron
Gogo6	Gogo Server	http://gogoware.gogo6.com/4105/description.asp?product_id=178
A10	AX-series	http://www.a10networks.com/news/2010/100222-LSN_SDLite.php
Juniper	M-series	http://www.juniper.net/techpubs/en_US/junos10.4/topics/example/ipv6-access-ds-lite-configuring.html
	T-series	
	MX-series	
ISC	AFTR (for Linux)	http://www.isc.org/software/aftr
Cisco	CRS-1	http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6553/brochure_c02-560497_ns1017_Networking_Solutions_Brochure.html
Alcatel	naam onbekend	AVM

Verslag Ontwerpfase

Afstudeeropdracht

“**IPv4-connectiviteit in een IPv6 access network**”

Bateau Knowledge, Den Haag

Matthias van der Heide (07053940)

Haagse Hogeschool

Versie 1.0 – 6 juni 2011

Inhoudsopgave

1. Inleiding

Dit verslag van de Ontwerpfase geeft een indruk van hoe het netwerk van de glasvezel-ISP er na de aanpassing uit komt te zien. Het ontwerp wordt toegelicht op logisch en fysiek niveau. Er wordt getracht een beeld te vormen van de impact van de invoering op de processen binnen de ISP.

2. Logisch ontwerp

Adresplan

Door de aanleg van wijkstation drie komt het aantal klanten uit boven de beschikbare IPv4-adresvoorraad. Het aansluiten van 1200 nieuwe klanten op de oude manier zou 5 adresblokken van /24 groot (250 klanten) vergen. Deze adresruimte is er niet. Uit het oude adresplan zijn nog slechts twee blokken van /24 beschikbaar.

Gebruikersgroep	**Aantal /24-blokken**
Infrastructuur	1
Kantoor ISP	1
Klanten DC	8
Klanten WS1	8
Klanten WS2	4
Totaal in gebruik	**22**
Totaal beschikbaar	**24 (3 maal een /21)**

Tabel 1 – oude adresplan

Door bij RIPE een IPv6-adresblok aan te vragen kan de ISP aanspraak maken op een /22 IPv4-adresblok uit de laatste voorraad. Deze is speciaal gereserveerd voor overgangsscenario's. De minimum grootte voor een IPv6-adresblok is een /32, wat overeenkomt met 65536 subnets van /48. Hiermee kunnen ruim zestigduizend klanten een eigen netwerk bouwen met elk weer zestigduizend subnets. In eerste instantie wordt de adresruimte benut om het corenetwerk van IPv6-adressen te voorzien. Hiervoor is één /48 toereikend.

Gebruikersgroep	**Aantal /24-blokken (IPv4)**	**Aantal /48 subnets (IPv6)**
Infrastructuur	1	1
Kantoor ISP	1	-
Klanten DC	8	-
Klanten WS1	8	-
Klanten WS2	4	-
Klanten WS3	5	-
Totaal in gebruik	**27**	**1**
Totaal beschikbaar	**28 (3 maal een /21 + één maal een /22)**	**65536**

Tabel 2 – nieuwe adresplan

Zelfs met dit extra blok adressen blijft het adresplan erg krap. Er is slechts één blok IPv4-adressen over voor eventuele uitbreiding. Dat is erg weinig voor een ISP met een groeiend aantal abonnees. Daarom wordt er voor gekozen om wijkstation 3 alleen door middel van IPv6 aan te sluiten, en met behulp van Dual-Stack Lite IPv4-connectiviteit te bieden. De

AFTR's, de apparatuur voor DS-Lite, worden centraal in het netwerk geplaatst, en de klanten van wijkstation drie krijgen geen eigen IPv4-adressen meer. In plaats daarvan worden hun adressen gebruikt in de adrespool van de AFTR's. Deze pools zijn kleiner dan het aantal klanten dat hiermee bedient wordt. Doordat IPv4-adressen met behulp van Dual-Stack Lite worden gedeeld tussen klanten, kan de provider met minder adressen toe dan nodig is bij native IPv4.

Met behulp van een simpele berekening is na te gaan hoe groot de adrespools moeten zijn. Dit wordt bepaald door de verhouding tussen klanten en beschikbare adressen (ondergrens) en het maximaal aantal verbindingen dat een klant via Dual-Stack Lite maakt (bovengrens).

De ondergrens wordt berekend door het maximale aantal actieve aansluitingen (13.800) te delen door het maximaal beschikbare aantal IPv4-adressen voor klanten (26 * 250). Dit geeft een verhouding van 13.800:6.500 of afgerond naar boven: 3 aansluitingen per IP-adres. In de toekomst, bij volledig gebruik van het netwerk, zullen minimaal drie klanten gebruik maken van één IPv4-adres.

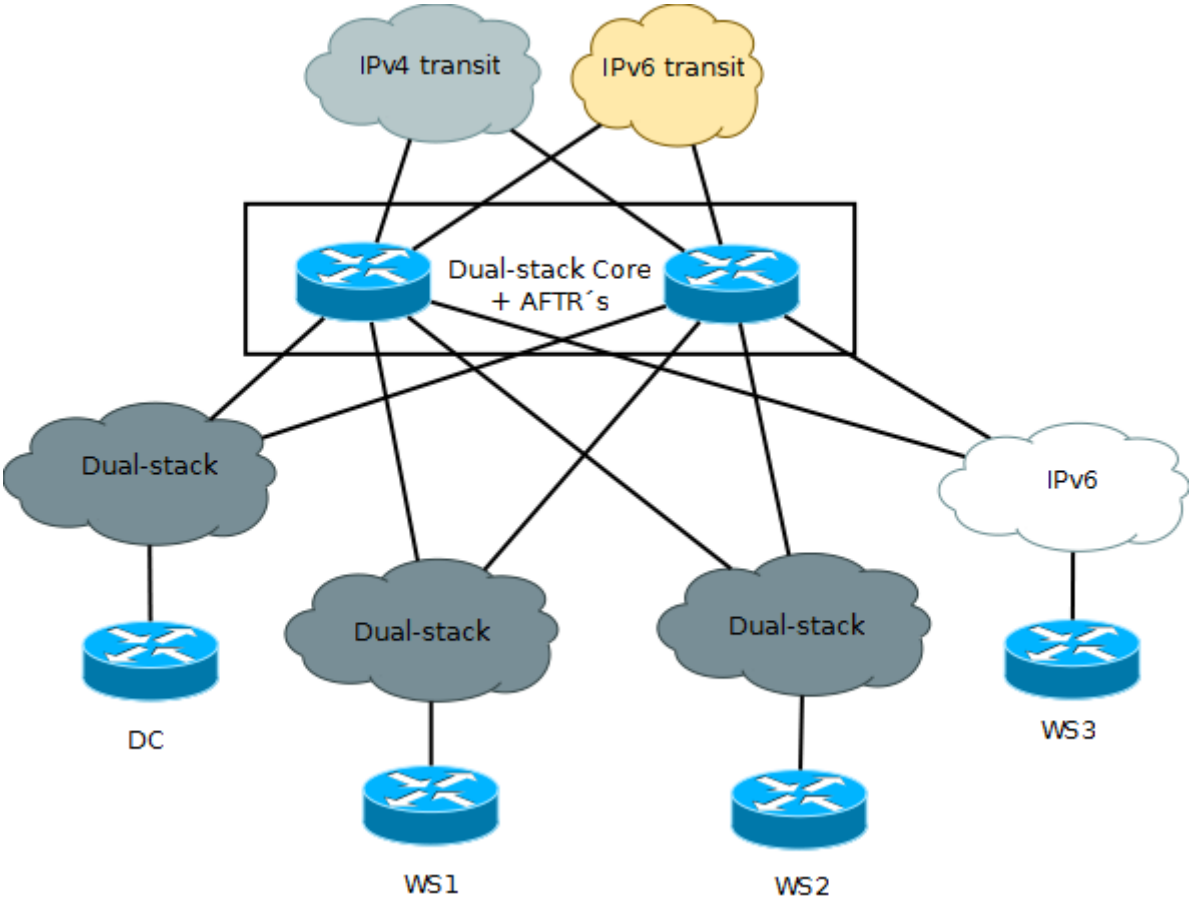
De bovengrens is een technische beperking van adresdeling. Verbindingen van klanten lopen via de TCP- en UDP-poortnummers. Hiervan heeft elk IP-adres er ongeveer 64.000. Maar elke host heeft maximaal rond de 500 poorten nodig om goed te werken. Als klanten tot vijf computers mogen aansluiten, betekent dit dat één klant al 2.500 poorten gebruikt. Dan kunnen maximaal 25 klanten gebruik maken van één IP-adres.

De adrespool moet dus een voldoende groot zijn om tussen de 3 en 25 klanten te bedienen. De adrespool voor wijkstation drie is 1000 adressen groot. Hiermee kunnen tussen de 3.000 en 25.000 klanten worden bedient. Ruim voldoende voor de toekomst!

Gebruikersgroep	**Aantal /24-blokken (IPv4)**	**Aantal /48 subnets (IPv6)**
Infrastructuur	1	1
Kantoor ISP	1	-
Klanten DC	8	-
Klanten WS1	8	-
Klanten WS2	4	-
Klanten WS3	-	5
AFTR adrespool	4 (/22 van RIPE)	-
Totaal in gebruik	**26**	**6**
Totaal beschikbaar	**28**	**65536**

Tabel 3 – nieuwe adresplan, dual-stack

Netwerkoverzicht

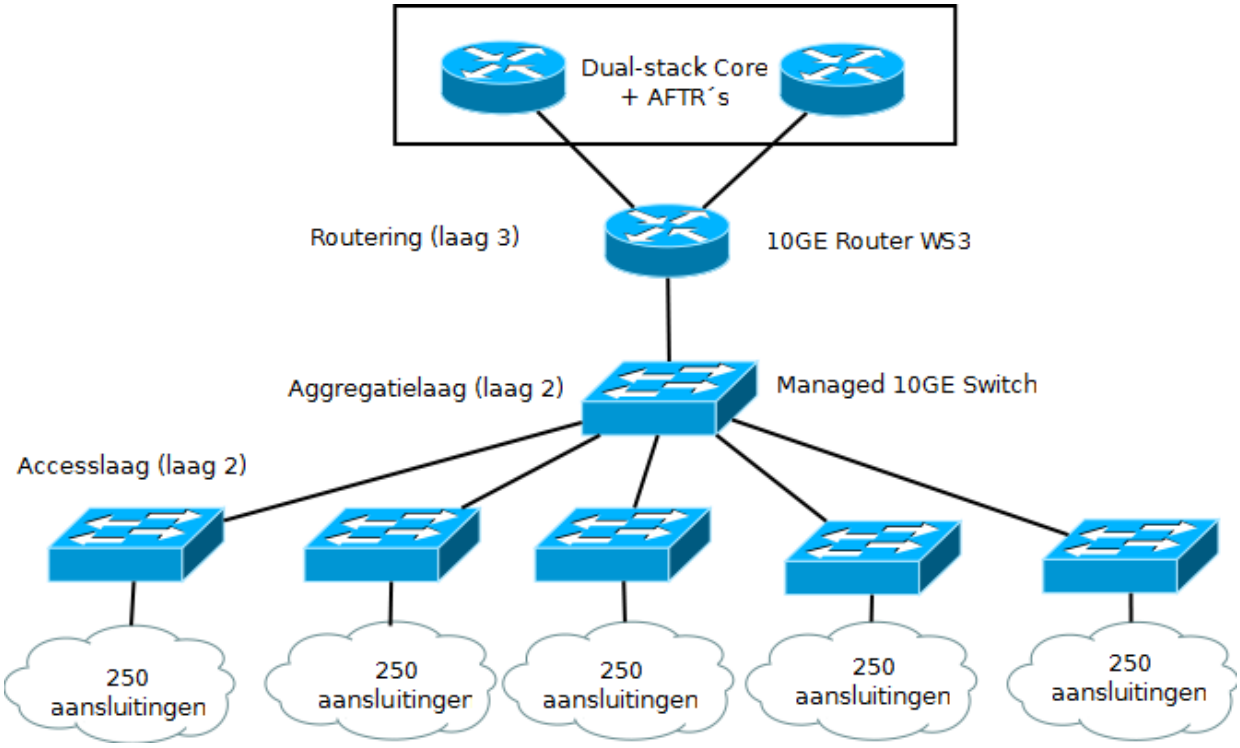


Afbeelding 5 - totaaloverzicht op logisch niveau

Als dit geheel wordt samengevoegd, wordt het het beeld van het ISP-netwerk op logisch niveau duidelijk. De AFTR's zijn centraal geplaatst in de buurt van de core routers, zodat ze hun diensten kunnen leveren aan alle POP's die aangesloten zijn. Wijkstation 3 is al helemaal over op IPv6, wat aangegeven wordt met een witte wolk. De overige POP's zijn voorlopig dual-stack om de mix van klanten (bestaand IPv4 en Dual-Stack Lite) te kunnen bedienen. Op termijn is het de bedoeling dat deze allemaal IPv6-only worden, en alleen de core nog IPv4 hoeft te routeren.

3. Fysiek ontwerp wijkstation 3

De fysieke opbouw van wijkstation drie is gebaseerd op bestaande glasvezelinfrastructuur. Aansluitingen van klanten komen uit op een switch die de aansluiting in een VLAN met maximaal 250 klanten plaatst. Communicatie tussen klanten wordt door middel van filters geblokkeerd. Alleen verkeer van en naar router WS3 is mogelijk.



Het tagged verkeer van de access switch komt uit op een 10 GE switch in de aggregatielaag. Het is de taak van deze switch om het verkeer te bundelen richting de router. Door plaatsing van deze tussenlaag kan de capaciteit van een POP worden uitgebreid op laag 2, zonder dat onderhoud nodig is aan de router. Wel moet van tevoren de link tussen de router en de aggregatielaag voldoende groot zijn ingeschat. Het is niet ondenkbaar dat deze uit meerdere, redundante, hogesnelheidslinks bestaat.

4. Organisatie

Met de invoering van IPv6 en Dual-stack Lite in het netwerk van de provider moet ook veel aangepast worden aan de organisatie.

Aan de volgende zaken moet onder meer gedacht worden bij het plannen van de uitrol:

Netwerkbeheer
- Trainen netwerk engineers
- Inrichting van managementtools
- Aanpassen van alle processen waar IP-adressen in voorkomen

Support
- Opleiden supportmedewerkers

Inkoop
- IPv6-transit afnemen

Verkoop
- Updaten bestelprocedures voor nieuwe aansluitingen

5. Conclusie

De invoering van Dual-Stack Lite verhoogt de complexiteit van het netwerk. Op termijn kan hopelijk in het access-deel van het netwerk IPv4 worden uitgeschakeld. Tot die tijd bestaat een gemengde omgeving. Invoering van Dual-Stack Lite raakt alle afdelingen binnen het bedrijf.

Verslag Ontwikkelfase

Afstudeeropdracht

“**IPv4-connectiviteit in een IPv6 access network**”

Bateau Knowledge, Den Haag

Matthias van der Heide (07053940)

Haagse Hogeschool

Versie 1.0 – 6 juni 2011

Inhoudsopgave

Inleiding

Dit document is een verslag van de ontwikkeling van het proof of concept voor Dual-Stack Lite. Het beschrijft de te testen functies, de testopstelling die hiervoor is opgezet en de tests die zijn ontworpen.

In hoofdstuk één wordt de doelstelling van het proof of concept gegeven. In hoofdstuk twee wordt in grote lijnen de te testen functionaliteit geschetst. Hoofdstuk drie behandelt de gebouwde testopstelling. Hoofdstuk vier gaat in op de tools die gebruikt zijn voor het ontwikkelen van tests. De testcases zijn uitgewerkt in hoofdstuk vijf. In hoofdstuk zes wordt de geteste apparatuur opgesomd. Hoofdstuk zeven geeft de testresultaten. Hoofdstuk acht bevat de conclusie van de tests met het proof of concept.

1. Doelstelling

Het doel van de ontwikkelfase is om het ontwerp dat in de ontwerpfase is opgesteld te valideren. Dit betreft de aanpassing van het providernetwerk om Dual-Stack Lite te implementeren. De testopstelling zal dienen als het proof of concept zoals dat in de afstudeeropdracht, deelopdracht twee, is bedoeld.

Daarnaast is door Stichting IPv6 Nederland verzocht om tests uit te voeren namens het Stipv6 Dual-Stack Lite testlab.

Door de twee opdrachten te combineren kan Stipv6 helpen om fabrikanten te benaderen om hun medewerking te verlenen. De kennis en contacten die hieruit voortvloeien kunnen worden gebruikt om ISP’s en leveranciers van netwerkapparatuur te helpen met het implementeren van Dual-Stack Lite.

2. Te testen functionaliteit

Voor zowel de AFTR als de CPE is een aantal tests bedacht die de werking van de onderdelen verifieert. Hierbij worden alleen de onderdelen die van belang zijn voor DS-Lite bekeken. Bijvoorbeeld: er wordt aangenomen dat een CPE die DS-Lite implementeert een werkende, volwaardige implementatie van de IPv6-netwerkstack heeft.

2.1 Interoperabiliteit

Voor een succesvolle implementatie is het noodzakelijk dat apparatuur, die onafhankelijk van elkaar is ontwikkeld, goed met elkaar samenwerkt. Hieruit blijkt dat de specificatie eenduidig is, en dat deze uniform is uitgewerkt door verschillende fabrikanten.

Functionaliteit die getest wordt:

- RFC 2473-tunneling
- MSS clamping
- Gecombineerde performance van B4 en AFTR
- NAT-gedrag van de AFTR

2.2 B4

De CPE heeft de volgende functies die onafhankelijk van de AFTR getest kunnen worden:

- IPv6-prestaties
- Automatische WAN-configuratie (DHCPv6)
- Handmatige configuratie (WAN+LAN)
- DHCPv4-server (LAN)
- Fragmentatie

2.3 AFTR

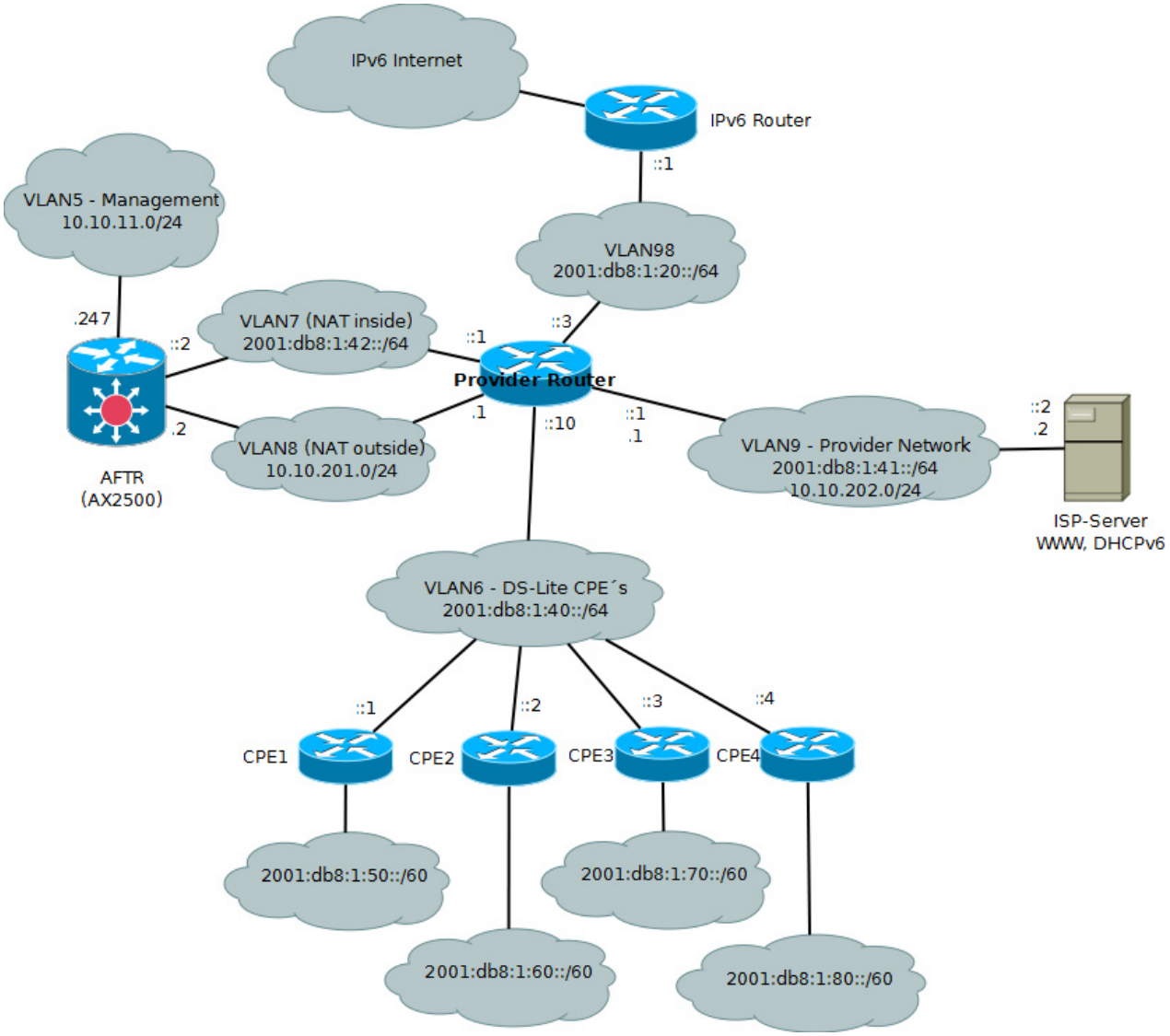
Aan de kant van de AFTR is het ondermeer belangrijk om de volgende zaken te testen:

- Port forwardings
- Configuratieopties

3. Testopstelling

Voor het testen van de verschillende onderdelen zijn afwijkende testopstellingen nodig. Deze zijn gebaseerd op de onderstaande architectuur.

Bij het ontwerp (afbeelding 1) is er van uitgegaan dat alle verkeer door één router (Provider Router) geleidt wordt. Hierdoor is het mogelijk om alle subnets van de opstelling te monitoren en te kijken hoe verkeer gerouteerd wordt door het IPv4- en IPv6-netwerk.



Afbeelding 1 - ontwerp van de testopstelling

4. Testtools

Er wordt gewerkt met een standaard koper 100 mbit ethernetnetwerk, bestaand uit bekabeling, switches en hosts. Als het voor de performance noodzakelijk is om meer bandbreedte ter beschikking te hebben, kan een gigabit switch en gigabit NIC worden gebruikt op de punten waar dat is vereist. Servers die geen hoge load hoeven te verwerken kunnen op een VMware machine worden geplaatst die in het lab beschikbaar is.

Daarnaast kan gebruik gemaakt worden van diverse tools om de apparatuur te testen. Een overzicht is hieronder gegeven.

4.1 Wireshark

Met behulp van Wireshark kan netwerkverkeer worden afgevangen, geanalyseerd en opgeslagen. Met behulp van filters kan een selectie van het verkeer worden uitgelicht en onderzocht. De toegevoegde waarde van Wireshark zit in de vele analysers die standaard worden meegeleverd. Met behulp van deze plugins kunnen problemen met specifieke protocollen worden onderzocht. Zo is met de TCP Protocol Analyser precies te volgen welke gegevens over en weer worden verstuurd gedurende een TCP-verbinding.

Wireshark maakt voor opslag gebruik van het het PCAP-formaat, dat ook door tcpdump wordt gebruikt.

De website van Wireshark is te vinden op http://www.wireshark.org/.

4.2 Tcpdump

De command-line utility tcpdump is een Zwitsers zakmes voor Linux-gebruikers met netwerkproblemen. Het vervult via de prompt veel van de taken waar anders Wireshark voor gebruikt zou worden. Onder meer de volgende functies worden ondersteunt:

- Tonen van verkeer op een specifieke netwerkinterface (sniffing)
- Filteren van getoond verkeer op basis van type verkeer, bron- of doeladres
- Opslaan van traces

Ook analyseert tcpdump het verkeer, en geeft het meldingen als bijvoorbeeld checksums niet in orde zijn. Tcpdump is standaard onderdeel van vele Linux-distributies en op http://www.tcpdump.org/.

4.3 Iperf

De tool Iperf is een opensource programma voor onder meer Linux en Windows. Het bepaalt de bandbreedte van een verbinding door TCP- of UDP-verbindingen op te zetten. Hiervoor kent Iperf een client- en een servermodus. De client genereert verkeer, en stuurt dit naar de opgegeven server. Op de server draait Iperf in server-modus en vangt daar alle verkeer op. Beide processen rapporteren vervolgens de gemeten bandbreedte en testduur.

Voor de tests in deze opdracht is de bandbreedte bepaald door vijf testruns van een minuut te doorlopen. De gemeten waarden zijn vervolgens gemiddeld.

Omdat Iperf de bandbreedte meet van client naar server, is het voor tweewegmetingen nodig om de rol van client en server om te draaien. Indien het te meten apparaat een firewall of NAT-functie heeft, moet hiervoor een firewalluitzondering of port forwarding voor worden aangemaakt.

Iperf is te downloaden op http://sourceforge.net/projects/iperf/.

4.4 MGEN

De Multi-Generator (MGEN) is een tool ontwikkeld door de Amerikaanse marine om netwerken te testen. Het kan gebruikt worden om TCP- en UDP-verkeer te genereren volgens een precies getimed script. Aan de hand van loggings kan worden bepaald of het verkeer correct ontvangen is.

De scripts bestaan uit combinaties van de commando's ON, OFF, LISTEN en IGNORE. Met ON en OFF kan het versturen van een stroom packets (flow) worden bedient. Met behulp van parameters kan het type, de packetgrootte, de frequentie en de duur van verzending worden bepaald. De commando's LISTEN en IGNORE geven aan wanneer de ontvanger een bepaalde flow verwacht te ontvangen.

MGEN is zeer geschikt om normale netwerkflows te genereren. Door de parameters te variëren kan een breed scenario aan verkeer gesimuleerd worden. MGEN is onder meer gebruikt voor het testen van MSS clamping en het bepalen van de fragmentatieparameters van de geteste systemen.

MGEN is te vinden op de website van het Navy Research Lab op http://cs.itd.nrl.navy.mil/work/mgen/index.php.

4.5 Scapy

Scapy is een library voor de populaire programmeertaal Python, die de gebruiker in staat stelt om netwerkpackets te genereren. In tegenstelling tot veel andere programma's kan de gebruiker in Scapy praktisch elke bit in het packet zelf bepalen. In normale gevallen werkt Scapy met default-waarden, zodat niet letterlijk elk bit hoeft te worden ingegeven.

Daarnaast biedt Scapy enkele handige functies voor het verzenden van packets. Zo kan de respons op verzonden packets bewaard worden bij het packet zelf, waardoor het makkelijk is om code te schrijven die gebaseerd is op dit principe, zoals de ping-utility.

Deze commando's gaan er echter van uit dat de respons wordt ontvangen op dezelfde interface als waar deze op verstuurd wordt. Voor het testen van Dual-Stack Lite is het nodig om te controleren of het versturen van een packet naar één interface van een router, leidt tot het versturen van een packet op een andere interface.

Daarom is een aangepaste versie van ping gemaakt die het mogelijk maakt om met behulp van een computer met twee netwerkkaarten, te testen of het versturen van een packet op één interface er voor zorgt dat een bepaald packet wordt ontvangen op de ander. De broncode is te vinden in listing 1.

Listing 1 - Broncode van BKping.py

```
# BK Ping, by Bateau Knowledge
# Authors: Joost Cassee, Matthias van der Heide
# (C) Bateau Knowledge, 2011

print "BK Ping"

from scapy.all import *
from threading import Timer
from unittest import TestCase,main

class PingTestCase(TestCase):
  def test_simple_ping(self):
      request = Ether() / IP(dst="192.168.1.1") / ICMP(type="echo-request")
      delay(sendp, request, iface="eth2")
      replies = sniff(count=1, filter="icmp and src 192.168.1.1", timeout=5, iface="eth2")
      self.assertTrue(replies)

def delay(func, *args, **kwargs):
  Timer(0.1, func, args, kwargs).start()

if __name__ == "__main__":
  main()
```

Dit prototype maakt het mogelijk om aan te geven op welke interface een respons verwacht wordt. Hiervoor is het alleen nodig om de ‘iface’ parameter aan te passen. In Linux kan de naam van de interface gebruikt worden, in Windows wordt de Linux-naamgeving nagebootst en is het nodig om met trial-and-error de juiste naam te vinden.
De ‘request’-regel maakt duidelijk hoe simpel het is om met Scapy IP-packets te genereren. In dit geval wordt een Ethernet-datagram aangemaakt, waarvan source en destination MAC-adres door Scapy automatisch worden bepaald. Hierin komt een IPv4-packet met als ‘dst’-parameter het IP-adres van de doelhost. Het ‘vullen’ van packets gaat in Scapy met de schuine streep (/), wat het erg simpel maakt om gegevens ‘in te pakken’ in andere typen packets. Wederom worden de overige velden door Scapy ingevuld. Vervolgens wordt het IP-packet gevuld met een ‘ICMP Echo Request’. Hiervan hoeft alleen het type aan Scapy te worden doorgegeven.

Op basis van deze ene regel genereert Scapy het complete datagram dat naar de doelhost verstuurd wordt. Daarna luisteren we vijf seconden lang voor antwoorden, en geven aan of deze wel of niet ontvangen zijn.

De getoonde werkwijze laat zien dat het simpel is om packets te versturen en deze compleet naar je hand te zetten. Dit is ideaal voor het genereren van gekreupelde packets, die bijvoorbeeld tot een foutmelding van de doelhost moeten leiden. Scapy ondersteunt ook IPv6, waardoor het gebruikt kan worden om de testomgeving te automatiseren.

De broncode van Scapy en uitvoerige tutorials zijn te vinden op http://www.secdev.org/projects/scapy/.

5. Testcases

De testcases zijn genummerd en worden apart uitgewerkt. De testcases worden gegroepeerd op het aspect wat getest wordt.

Sommige cases testen specifiek een onderdeel van de hele opstelling (CPE, AFTR). De meerderheid van de cases test de opstelling als geheel (end to end).

B4-tests (B)

B1 De CPE fungeert als DHCPv4- en DNS-server voor het LAN
B2 De CPE krijgt handmatig het IPv6-adres van de AFTR ingesteld
B3 De CPE krijgt handmatig de FQDN van de AFTR ingesteld
B4 De CPE krijgt via DHCPv6 de FQDN van de AFTR ingesteld
B5 Bepaalde de downloadperformance van de CPE over IPv6

Interoperabiliteit (I)

I1 Clients kunnen hosts op het internet pingen
I2 Bepaal de downloadperformance van het systeem over DS-Lite
I3 Bepaal de uploadperformance van het systeem over DS-Lite
I4 Controleer of MSS clamping correct werkt
I5 Toets het NAT-gedrag van de AFTR
I6 Fragmentatie van grote pakketten

AFTR-tests (A)

A1 Stel een port forwarding in

A1 - Stel een port forwarding in

Uitgangssituatie

- De AFTR is geconfigureerd voor het testnetwerk
- De CPE is ingesteld om de AFTR te gebruiken
- Client 1 is aangesloten op het LAN van de CPE

Acties

1. Monitor het CPE-subnet aan de AFTR-kant
2. Stel op de AFTR een port forwarding in voor client 1
3. Zet vanaf de testserver een verbinding op met deze poort

Verwachte resultaat

- De packets van de testserver naar client 1 zijn zichtbaar in het CPE-subnet

B1 - De CPE fungeert als DHCPv4- en DNS-server voor het LAN

Uitgangssituatie

- De CPE is geconfigureerd als Dual-Stack Lite client
- De CPE is geconfigureerd met het adres van een AFTR

Acties

4. Sluit een client aan op de LAN-zijde van
5. Bepaal met Wireshark of de CPE zich als DHCP-server adverteert

Verwachte resultaat

- De CPE functioneert als DHCP-server en reageert op aanvragen voor IPv4-adressen en de nameserver

B2 - De CPE krijgt handmatig het IPv6-adres van de AFTR ingesteld

Uitgangssituatie

- De CPE is opgestart
- Er is een client verbonden met de WAN-zijde
- Er is een client verbonden met de LAN-zijde

Acties

1. Benader de CPE via de webinterface
2. Stel het IPv6-adres van de AFTR in
3. Sniff met Wireshark de WAN-interface van de CPE
4. Start op een client op een LAN-interface een verbinding naar buiten

Verwachte resultaat

- De CPE stuurt getunneld IPv4-verkeer naar het adres van de AFTR

B3 - De CPE krijgt handmatig de FQDN van de AFTR ingesteld

Uitgangssituatie

- De CPE is opgestart
- Er is een client verbonden met de WAN-zijde
- Er is een client verbonden met de LAN-zijde

Acties

1. Benader de CPE via de webinterface
2. Stel het de DNS-naam van de AFTR in
3. Sniff met Wireshark de WAN-interface van de CPE
4. Start op een client op een LAN-interface een verbinding naar buiten

Verwachte resultaat

- De CPE vraagt via DNS het adres van de AFTR op
- De CPE stuurt getunneld IPv4-verkeer naar het adres van de AFTR

B4 - De CPE krijgt via DHCPv6 de FQDN van de AFTR ingesteld

Uitgangssituatie

- De CPE is opgestart
- De DHCPv6-server is ingesteld om het adres van de AFTR te geven
- Verkeer op de WAN-zijde van de CPE wordt gemonitord
- Er is een client aangesloten op de LAN-zijde

Acties

1. Benader de webinterface van de CPE
2. Stel deze in op automatische configuratie
3. Start indien nodig de CPE opnieuw op
4. Observeer het verkeer op de WAN-zijde
5. Maak een verbinding naar een IPv4-hosts op het internet vanaf de client op het LAN

Verwachte resultaat

- De CPE vraagt via DHCPv6 om optie 64, OPTION_AFTR_NAME
- De DHCPv6-server reageert
- De CPE vraagt via DNS het adres van de AFTR op
- De CPE tunnelt het IPv4-verkeer naar de AFTR

B5 - Bepaalde de downloadperformance van de CPE over IPv6

Uitgangssituatie

- De CPE is ingesteld als IPv6-router
- Aan de LAN-zijde is een client verbonden
- Via de WAN-zijde is de testserver bereikbaar

Acties

1. Doe een ping6 naar de testserver
2. Voer wget uit om een groot bestand te downloaden via IPv6 (>500 MB)
3. Noteer de downloadsnelheid

Verwachte resultaat

- De testserver is bereikbaar
- Het bestand wordt succesvol gedownload.

I1 - Clients kunnen hosts op het internet pingen

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van de CPE
- De CPE is geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het testnetwerk te gebruiken

Acties

1. Voer op client 1 een ping4 uit naar de testserver.

Verwachte resultaat

- De ping wordt beantwoordt door de testserver

I2 - Bepaal de downloadperformance van het systeem over DS-Lite

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van de CPE
- De CPE is geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het testnetwerk te gebruiken

Acties

1. Start op client1 een wget van een groot bestand op de testserver
2. Noteer de downloadsnelheid.
3. Maak een portforwarding aan op de AFTR om client1 te benaderen (test A1)
4. Start iperf op client 1 in servermodus
5. Start vanaf de server een iperf client naar client 1, duratie 60 seconden
6. Herhaal stap 4 vijf maal.
7. Noteer het gemiddelde van de gevonden snelheden

Verwachte resultaat

- De download is succesvol
- Er kan een verbinding van de testserver naar client1 gemaakt worden

I3 - Bepaal de uploadperformance van het systeem over DS-Lite

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van de CPE
- De CPE is geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het testnetwerk te gebruiken

Acties

1. Start iperf opde testserver in servermodus
2. Start vanaf client1 een iperf client naar de testserver, duratie 60 seconden
3. Herhaal stap 2 vijf maal.
4. Noteer het gemiddelde van de gevonden snelheden

Verwachte resultaat

- Er kan een verbinding van client1 naar de testserver gemaakt worden

I4 - Controleer of MSS clamping correct werkt

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van de CPE
- De CPE is geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het testnetwerk te gebruiken

Acties

1. Start op de testserver het MGEN-script ‘tcp_test_server.mgn’
2. Start op client 1 het MGEN-script ‘tcp4_test.mgn’
3. Log met Wireshark en TCP-dump op alle subnets de TCP-verbinding
4. Verwerk de gerapporteerde MSS (twee richtingen) in een tabel

Verwachte resultaat

- De CPE voert MSS-clamping van 40 bytes uit (MSS op WAN-zijde is 40 kleiner dan op LAN-zijde)
- De AFTR voert in omgekeerde richting MSS-clamping van 40-bytes uit (MSS tussen de AFTR en CPE is 40 kleiner dan tussen server en AFTR)

I5 - Toets het NAT-gedrag van de AFTR

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van CPE1
- Client 2 is gekoppeld aan het LAN van CPE2
- De CPE's zijn geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het publieke internet te gebruiken

Acties

1. Voer op beide clients Wireshark uit
2. Start op beide clients Skype en log in
3. Zet een gesprek op tussen de gebruikers op beide clients
4. Observeer de publieke IP-adressen en poortnummers die worden gebruikt

Verwachte resultaat

- De clients ontvangen beide verkeer van het externe IPv4-adres van de AFTR

I6 - Fragmentatie van grote pakketten

Uitgangssituatie

- Client 1 is gekoppeld aan het LAN van de CPE
- De CPE is geconfigureerd om de AFTR te gebruiken
- De AFTR is geconfigureerd om het testnetwerk te gebruiken

Acties

1. Start op de testserver het MGEN-script 'udp_test_server.mgn'
2. Start op client 1 het MGEN-script 'udp4_test.mgn'
3. Log met Wireshark en TCP-dump op alle subnets de UDP-verbinding

Verwachte resultaat

- Packets kleiner en gelijk aan 1500 bytes komen door
- Packets groter dan 1460 bytes worden gefragmenteerd tussen de CPE en de AFTR

6. Geteste apparatuur

Op uitnodiging van de Stichting IPv6 Nederland hebben twee bedrijven positief gereageerd. AVM heeft op Cebit 2011 aangekondigd Dual-Stack Lite te ondersteunen in haar consumentenrouters. A10 Networks heeft Dual-Stack Lite sinds eind 2010 werkend geimplementeerd in haar AX-serie van producten. Met behulp van de firmware die door Comcast/Xavient is ontwikkeld is een Linksys WRT54GL uit het lab geschikt gemaakt om als B4 dienst te doen.

6.1 AVM Fritz!Box 6360 Cable

Het kabelmodem van AVM kan gebruikt worden als een normale router waarbij LAN-poort 1 als WAN-poort dienst doet. Er is een experimentele firmware die Dual-Stack Lite implementeert.

6.2 Linksys 160NL

De Linksys 160NL met OpenWRT-firmware van Xavient kan dienst doen als Dual-Stack Lite B4-element. Deze beschikt over Fast Ethernet-poorten voor LAN (4 stuks) en WAN (één poort).

6.3 A10 Networks AX2500

De AX2500 is van huis uit een server load-balancer (SLB). Een van de functies die het apparaat standaar ondersteunt is NAT44 (IPv4 NAPT). Met een aangepaste firmware is deze ook over IPv6 te gebruiken en ondersteunt dan Dual-Stack Lite.

7. Testresultaten

Alle B4-tests zijn uitgevoerd op beide modems.

Test	AVM 6360	Linksys 160NL	Opmerking
B1	geslaagd	geslaagd	
B2	geslaagd	geslaagd	
B3	gefaald	gefaald	
B4	gefaald	gefaald	beide modems ondersteunen OPTION_AFTR_NAME niet
B5	+/- 750 mbps	+/- 90 mbps	

Interoperabiliteitstests zijn alleen uitgevoerd met de AVM 6360 als CPE1 en de A10 AX2500 als AFTR. Met uitzondering van I5, waar de Linksys 160NL als CPE2 is geconfigureerd.

Test	Uitslag	Opmerking
I1	geslaagd	
I2	80 mbps	Omdat een port forwarding niet werkte is alleen getest met wget. De 80 mbps is omgerekend van de 8 MB/s bandbreedte die wget aangaf
I3	35,4 mbps	
I4	geslaagd	AX2500: verlaagt de MSS correct met 40 bytes AVM 6360: gebruikt een veilige waarde van 1200 bytes
I5	geslaagd	
I6	gefaald	packets boven de 1460 bytes komen niet aan

De AFTR-test is uitgevoerd op de A10 AX2500.

Test	Uitslag	Opmerking
A1	gefaald	de packets van AFTR naar CPE hebben een leeg IPv6-bronadres en komen daarom niet aan

8. Conclusie

De testresulaten geven aan dat op basis van de geteste apparatuur een Dual-Stack Lite netwerk is te bouwen dat IPv4-connectiviteit biedt in een IPv6-netwerk.

In de AX2500 van A10 Networks zijn issues gevonden met het afhandelen van in IPv6-gefragmenteerde packets tussen de CPE en de AFTR, en ging er iets mis met het forwarden van poorten.

De referentie-implementatie van Comcast op de Linksys 160NL werkt, maar is niet uitvoerig getest. Er is voorrang gegeven aan de tests van de AVM 6360. De IPv6 performance is in orde voor een breedbandrouter, maar lijkt gelimiteerd door de Fast Ethernet poorten.

De AVM Fritz!Box 6360 Cable is een snelle router voor IPv6 en bereikt ook op Dual-Stack Lite goede resultaten. De snelheden gemeten snelheden van 80 mbps download en 35,4 mbps upload benaderen de gevraagde 100 mbps symmetrisch. Mogelijk kan door het aanpassen van MSS-clamping op de 6360 een nog hogere snelheid gehaald worden.

Test report
A10 Networks AX2500

Dual-Stack Lite capabilities

Published by:	Stichting IPv6 Nederland (The Netherlands IPv6 Foundation)
Author:	Matthias van der Heide
Date:	June 3, 2011
Version:	2011.03

Summary

In this report we show how the A10 Networks AX2500 performs in tests of its Dual-Stack Lite capabilities.

Although the AX2500 performed well regarding performance and NAT behaviour, issues have been found in the handling of IPv6 fragmented packets from CPEs and port forwarding.

When these issues get resolved, the AX2500 can serve as a good example of a working implementation of the Dual-Stack Lite AFTR element and be deployed by ISPs to solve their IPv4 exhaustion problem.

Introduction

This test report is the result of work performed by the Dual-Stack Lite Test Lab of Stichting IPv6 Nederland (The Netherlands IPv6 Foundation; Stipv6). The lab has been established to investigate the viability of Dual-Stack Lite as an IPv6 transition mechanism for Internet Service Providers.

The work was carried out by Matthias van der Heide as part of his graduation for the title Bachelor of ICT at the Haagse Hogeschool (The Hague University of Applied Sciences).

We thank Mischa Peters at A10 Networks for granting access to an AX2500 for the duration of the tests.

Scope

The IETF draft specification of Dual-Stack Lite details the functionality of an Address Family Transition Router (AFTR) as part of the Dual-Stack Lite transition mechanism. The aim of this technique is to allow broadband internet customer to connect to the internet without having been assigned a full IPv4 address to their home gateway or Customer Premises Equipment (CPE). The CPE is extended with a virtual interface, called Basic Bridging BroadBand element (B4), which tunnels IPv4 traffic to the AFTR over IPv6.

Dual-Stack Lite is not a finished standard. Rather it is constantly being updated by its authors based on new insights and comments from the IETF Softwires Working Group.

For the purpose of this test we adhere to the draft specification that was current during the start of this project, on April 1st, 2011. This is version 7 of the draft, dated March 3rd, 2011. The latest specification is version 11 and was published May 28, 2011. The new document changed some requirements from ‘should’ to ‘must’ status, but does not introduce new behavior.

The Dual-Stack Lite specification makes use of techniques defined in RFCs. Those RFCs are listed here and are considered part of the specification:

- Generic Packet Tunneling in IPv6 Specification (RFC2473)
- Network Address Translation (NAT) Behavioral Requirements for Unicast UDP (RFC4787)
- NAT Behavioral Requirements for TCP (RFC5382)
  NAT Behavioral Requirements for ICMP (RFC5508)

Links:

- Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion - http://tools.ietf.org/html/draft-ietf-softwire-dual-stack-lite-07
- Generic Packet Tunneling in IPv6 Specification (RFC2473) http://tools.ietf.org/html/rfc2473
- Network Address Translation (NAT) Behavioral Requirements for Unicast UDP (RFC4787)
  http://tools.ietf.org/html/rfc4787
- NAT Behavioral Requirements for TCP (RFC5382)
  http://tools.ietf.org/html/rfc5382
- NAT Behavioral Requirements for ICMP (RFC5508)
  http://tools.ietf.org/html/rfc5508

Device Under Test

The device tested is an A10 Networks AX2500 Load-balancer. It features 8 copper Gigabit Ethernet ports and 4 Small Form-factor Pluggable (SFP) tranceiver slots. This allows the AX2500 to be configured with up to 12 x 1 Gbps ports for high-performance duties.

The supplied device was running the 64-bit Advanced Core OS (ACOS) version 2.6.4, build 62 (Mar-17-2011,04:19)

Test objective

The purpose of the tests is to establish whether the AFTR implementation of the AX2500 is capable of performing Dual-Stack Lite.

The focus is on the ability of the AFTR to perform correct encapsulation and decapsulation of IPv4 traffic, NAT behaviour and performance.

Test setup

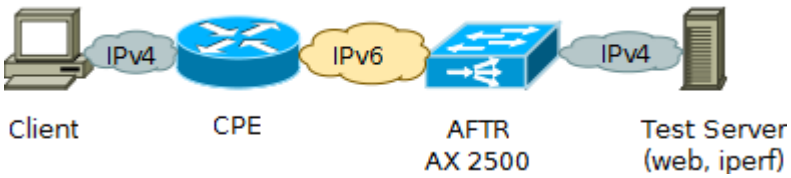
For the purpose of these tests a setup was built to perform tests both under lab conditions and with real world connectivity.

The network is depicted in diagram 1 on page 8. It consists of gigabit ethernet switches, configured with several VLANs to create separated Layer 2 ethernet networks. For hosts we used real and virtualized clients, a virtualized Ubuntu server, and a virtualized Vyatta 6.2 router in the center acting as the provider router. Having all traffic pass through the provider router enables us to monitor all traffic flows, both IPv4 and IPv6, in a central location.
The public IPv4 and IPv6 addresses that were used during tests have been replaced by documentation-friendly versions.

Tests were run using one of the two following configurations:

Private configuration

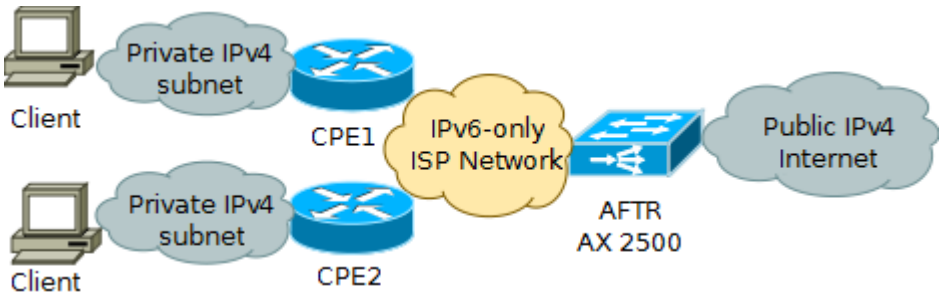
The private configuration allows us to test the AFTR under lab conditions. All traffic is flowing through the provider router and can be monitored, captured and analysed. The design is meant to eliminate factors that are beyond our control. Tests are performed between the client and the test server, without interference from other traffic.



The AFTR is assigned an address in the private IPv4 address range (RFC1918). See listing 1 for the complete configuration.

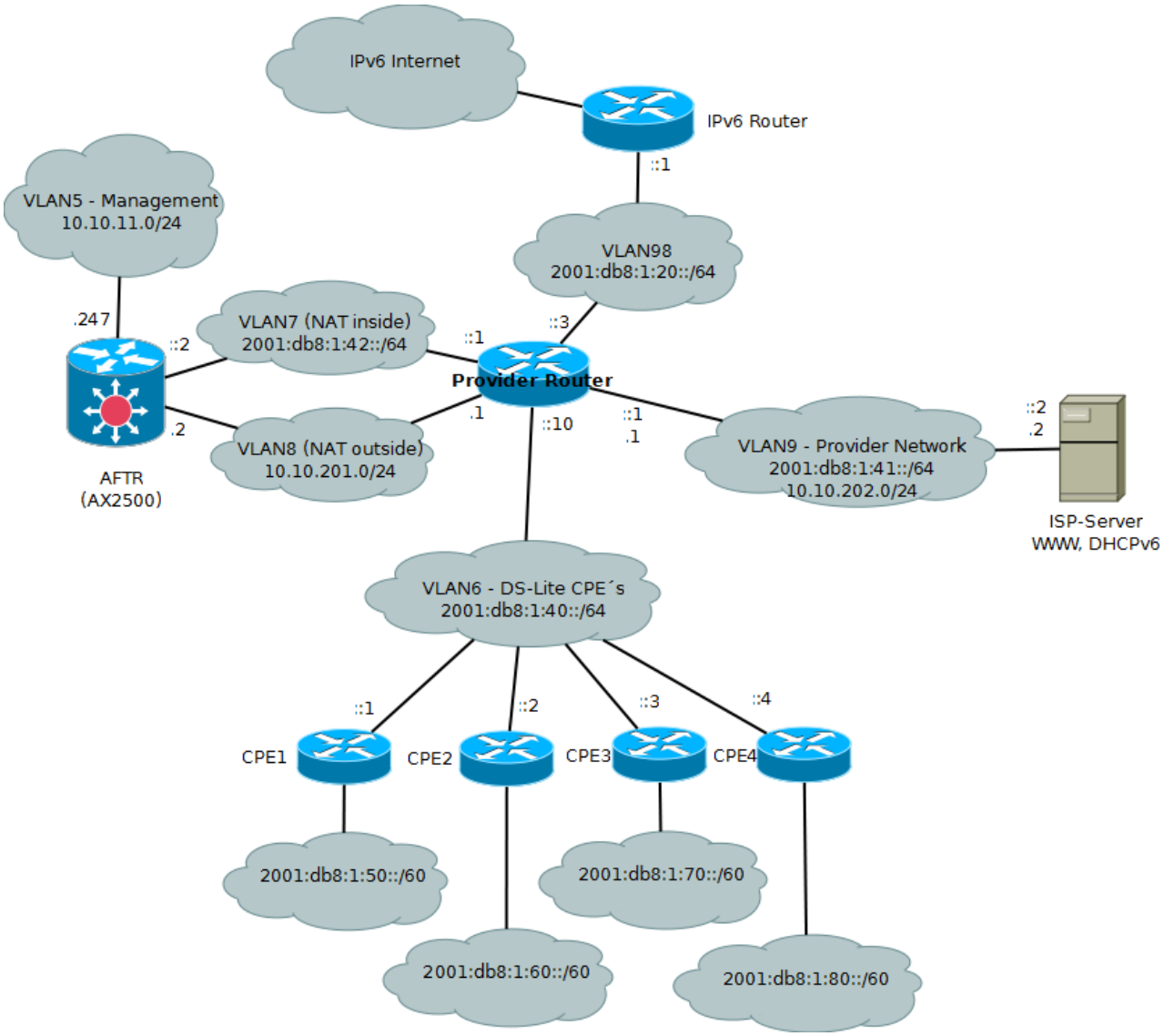
Public configuration

In this configuration the outside interface of the AFTR is directly connected to a public IPv4 subnet, instead of the ISP's IPv4-network (VLAN8). Because the public IPv4 subnet is not connected to the Provider Router, this traffic is not being monitored and performance of the AFTR now depends in part on the performance of the ISP uplink at the test lab and performance of the hosts accessed on the internet.



The AFTR is assigned an address in the public IPv4 range in use by the test lab. See listing 2 for the complete configuration.

Diagram 1 - Network topology (layer 3)



Listing 1 - Private configuration file (shortened)

```
hostname AX2500
!
clock timezone Europe/Amsterdam
!
class-list dslite_class
2001:db8:1:40::/64 lsn-lid 1
!
interface management
ip address 10.10.11.247 255.255.255.0
ip default-gateway 10.10.11.1
!
interface ethernet 1
ip address 10.10.200.2 255.255.255.0
ipv6 address 2001:db8:1:42::2/64
ipv6 nat inside
name "CPE"
!
interface ethernet 2
ip address 10.10.201.2 255.255.255.0
ip nat outside
name "PUB"
!
ip route 10.10.202.0 /24 10.10.201.1
ip route 0.0.0.0 /0 10.10.201.1
!
ipv6 route 2001:db8:1:41::/64 2001:db8:1:42::1
ipv6 route ::/0 2001:db8:1:42::1
!
ds-lite inside source class-list dslite_class
!
ip nat pool dslite_pool 10.10.201.3 10.10.201.3 netmask /24  lsn
!
lsn-lid 1
source-nat-pool dslite_pool
!
enable-buff-debug
!
end
```

Listing 2 - Public configuration file (shortened)

```
hostname AX2500
!
clock timezone Europe/Amsterdam
!
class-list dslite_class
2001:db8:1:40::/64 lsn-lid 1
!
interface management
ip address 10.10.11.247 255.255.255.0
ip default-gateway 10.10.11.1
!
interface ethernet 1
ip address 10.10.200.2 255.255.255.0
ipv6 address 2001:db8:1:42::2/64
ipv6 nat inside
name "CPE"
!
interface ethernet 2
ip address 192.0.2.192 255.255.255.192
ip nat outside
name "PUB"
!
ip route 0.0.0.0 /0 192.0.2.254
!
ipv6 route 2001:db8:1:41::/64 2001:db8:1:42::1
ipv6 route ::/0 2001:db8:1:42::1
!
ds-lite inside source class-list dslite_class
!
ip nat pool dslite_pool 192.0.2.193 192.0.2.193 netmask /26  lsn
!
lsn-lid 1
source-nat-pool dslite_pool
!
enable-buff-debug
!
end
```

Results

Summary

In general, the AFTR was found to perform very well. It enabled clients behind the CPEs to access the internet in the same way they would behind a regular NAT, and with good enough performance. The AFTR allows clients behind CPEs to sent TCP, UDP and ICMP traffic through the NAT to outside IPv4 addresses. An issue was detected in the handling of IPv6 fragments from CPEs. See the ‘Fragmentation’ section for more details. Also the AFTR does not handle port forwardings correctly. See ‘Port forwardings’ for a description of the issue.

Performance

The performance of the AFTR has been tested using real and virtual clients behind two different CPEs, both in private and public configuration. Only one client was active at the same time. As such, this is not a test of the maximum throughput, active sessions or supported clients, but rather a test of the combined performance of the particular CPE and the AX2500.

Tests with a Linksys WRT160NL obtained a maximum downstream speed of 3 Mbps in private, and 2,5 Mbps in public configuration. This is a rather disappointing result, which may be caused by hardware limitations of the WRT160NL and/or the implementation in software by Xavient.

The AVM Fritz!Box 6360 Cable performed much better at approximately 80 Mbps upstream and 50 Mbps downstream.

In both cases the CPE is believed to be the limiting factor, as the AX2500 showed low CPU load, unchanged memory usage and unsaturated network interfaces.

There is no reason to think the AX2500 can not handle a full load of 1 Gbps throughput.

NAT behaviour for UDP

To allow applications to function transparently through a NAT, RFC 4787 states 14 requirements. We tested some of these using Skype, placing a call between two clients behind CPE1 and CPE2. This causes a bi-directional flow of UDP packets between client 1, CPE1, the AFTR, CPE2 and client 2.

The test was successful. The clients, both behind the AFTR NAT, established a working Skype call without the use of a relay on the public internet. Traces show the following requirements are met:

- REQ-1: The AFTR implements “Endpoint-independent Mapping” behaviour.
- REQ-3: The AFTR does not do port overloading on port assignment.
- REQ-5: A NAT UDP mapping timer MUST NOT expire in less than two minutes.
- REQ-6: The NAT mapping Refresh Direction MUST have a “NAT Outbound refresh behaviour” of “True”.
- REQ-6a: The NAT mapping Refresh Direction MAY have a “NAT Inbound refresh behaviour” of “True”.
- REQ-9: A NAT MUST support “Hairpinning”
- REQ-9a: A NAT Hairpinning behavior MUST be "External source IP address and port".

Other requirements or subrequirements (5a, 5b, 5c) were not tested.

ICMP Application Level Gateway

To allow hosts behind a NAT to succesfully perform ping requests to hosts on the internet, the AX2500 features an ICMP Application Level Gateway (ALG). The DS-Lite specification demands that AFTRs must implement this, according to RFC5508, “NAT Behavioral Requirements for ICMP”.

The AX2500 was found to correctly forward ping requests and corresponding incoming replies both to local test servers and public hosts.

This satisfies requirement REQ-1 of RFC5508. Other requirements have not been tested.

Fragmentation

Encapsulation of IPv4 packets in IPv6 packets decreases the effective capacity of a packet with 40 bytes. This means some packets sent between the B4 (the Dual-Stack Lite interface in the CPE) and AFTR are to large to fit on the link. Their size, after encapsulation, exceeds the MTU of the underlying link. In Dual-Stack Lite this problem is solved by fragmenting the packets after encapsulation. The details of this procedure are provided in RFC2473, “Generic Packet Tunneling in IPv6 Specification”.

The AX2500 was found NOT supporting fragmentation according to RFC2473 on the incoming interface. Rather it only supports fragmentation in IPv4. Any packets fragmented in IPv6 will be dropped by the AFTR. This means the CPE has to fragment packets in IPv4 before encapsulation in IPv6.

In tests this was simulated by decreasing the MTU of the CPE to 1460 bytes manually. This forces the CPE to fragment packets in IPv4 instead of IPv6, so they won't get dropped by the the AFTR.

This scenario is shown in the following trace (output of tcpdump), made on the outside interface of the AFTR. A host behind CPE1 sends 5 consecutive UDP packets of increasing size to the test server. The UDP payloads are of size 1400, 1420, 1430, 1432 and 1450. The size of the fourth packet has been carefully chosen so the resulting packet is of a size equal to the Tunnel MTU (1460 bytes). It should not be fragmented. Packet 5 should fragmented in IPv4 by the CPE, and the fragments forwarded or reassembled by the AFTR.

```
12:23:46.489877 IP (tos 0x0, ttl 63, id 0, offset 0, flags [DF], proto UDP (17), length 1428)
        10.10.201.3.9000 > 10.10.202.2.9001: [udp sum ok] UDP, length 1400
12:23:47.475060 IP (tos 0x0, ttl 63, id 0, offset 0, flags [DF], proto UDP (17), length 1448)
        10.10.201.3.9000 > 10.10.202.2.9001: [udp sum ok] UDP, length 1420
12:23:48.475047 IP (tos 0x0, ttl 63, id 0, offset 0, flags [DF], proto UDP (17), length 1458)
        10.10.201.3.9000 > 10.10.202.2.9001: [udp sum ok] UDP, length 1430
12:23:49.475074 IP (tos 0x0, ttl 63, id 0, offset 0, flags [DF], proto UDP (17), length 1460)
        10.10.201.3.9000 > 10.10.202.2.9001: [udp sum ok] UDP, length 1432
12:23:50.477743 IP (tos 0x0, ttl 63, id 56933, offset 0, flags [none], proto UDP (17),
length 1478)
        10.10.201.3.9000 > 10.10.202.2.9001: [udp sum ok] UDP, length 1450
```

Output of tcpdump on the outside interface of the AFTR

The AX2500 was found to perform defragmentation of the IPv4 fragments before sending them out. This results in the AFTR issuing almost the exact same IPv4 packet (shown in bold) on the outgoing interface as would have been sent when IPv6 fragmentation would have been supported. The difference being, that the packet now has an Fragment ID that is not zero, because it has previously been fragmented in IPv4.

Fragmentation in the opposite direction (outside to inside) is support in both IPv4 and IPv6 (default). This conforms to the Dual-Stack Lite specification. Only IPv6 fragmentation was tested, and found to be working correctly.

MSS clamping

To avoid fragmentation of encapsulated TCP packets between the CPE and the AFTR, the TCP Maximum Segment Size (MSS) parameter can be adjusted to influence the size of packets sent during the TCP connection. This is called MSS clamping or MSS rewrite.

The AX2500 was tested in a network using Gigabit Ethernet links with an MTU of 1500 bytes. The TCP MSS was monitored on a bi-directional connection made from a client behind CPE1 to the test webserver.

Traces show that the AX2500 decreased the MSS of the connection in both directions with 40 bytes (the size of the Tunnel IPv6 Header), which is an effective way to avoid fragmentation of TCP packets.

The settings for MSS clamping can be altered using the ACOS command line and provide well defined settings and defaults.

Port forwarding

Because Dual-Stack Lite moves the NAT from the CPE to the core if the ISP network, the customer loses the ability to open ports for incoming connections. This can be mitigated by allowing customers to add static mappings to the AFTR NAT table. The AX2500 AFTR implements command line options to insert static mappings in the NAT table, making this a manual action that can only be performed by the network operator.

We tested port forwarding by adding mappings in two different port ranges (443 and 22000-23000) from the outside IPv4 address of the AFTR to the private IPv4 address of client 1 behind CPE1.

```
AX2500>show ds-lite port-reservations
LSN Port Reservations
Tunnel Src IPv6 Address                 Inside Address   Start    End      NAT Address      Start    End
-------------------------------------------------------------------------------------------------------
-----------------
2001:610:6bb:40::1                      192.168.178.22   443      443      10.10.201.3      443      443
2001:610:6bb:40::1                      192.168.178.22   22000    23000    10.10.201.3      22000    23000
Total Static Port Reservations: 2
```

Attempting to set up a TCP connection from the test server to client 1 results in the following packet being output on the inside interface of the AFTR:

13:54:09.322081 IP6 (hlim 64, next-header IPIP (4) payload length: 60) **:: >** 2001:610:6bb:40::1: IP (tos 0x0, ttl 62, id 33651, offset 0, flags [DF], proto TCP (6), length 60)

10.10.202.2.45333 > 192.168.178.22.https: Flags [S], cksum 0x3cab (correct), seq 780541849, win 5840, options [mss 1420,sackOK,TS val 269130986 ecr 0,nop,wscale 4], length 0

The IPv4 packet has been correctly translated to go to the private address of client 1. It has been encapsulated in IPv6, and sent towards CPE1. However, the IPv6 source address is the unspecified address (**:: or 0:0:0:0:0:0:0:0**). Packets without a valid source address will not be routed by the rest of the network, so this packet will never reach CPE1. This seems to be a bug in the implementation of the softwire (RFC2473).

Recommendations

- We suggest that the AFTR implements IPv6 fragmentation of packets coming from CPEs, according to the Dual-Stack Lite specification. Because CPEs exist that by default rely on this functionality, ISP customers might experience connection difficulties. It is expected that ISPs will request this feature in the future when they deploy these CPEs.

- The issue with port forwarding is probably a bug and should be investigated and fixed if necessary.

- Documentation is needed of the AFTR functionality and command line settings for operators and testers to understand the capabilities and incapabilities of the AX2500 implementation of Dual-Stack Lite.

- We encourage A10 Networks to cooperate with vendors of more CPEs to test their implementation and to provide feedback to the community (IETF Softwires Working Group) on issues found and best practices. The Stichting IPv6 Nederland (The Netherlands IPv6 Foundation) can assist A10 with more independent interoperability testing.

Afterword

Stichting IPv6 Nederland would like to thank A10 Networks, and especially Mischa Peters, for cooperating in this test.

We hope our findings prove valuable in your efforts to further improve the Dual-Stack Lite functionality in the AX series of products.

Test report
AVM Fritz!Box 6360 Cable

Dual-Stack Lite capabilities

Published by:	Stichting IPv6 Nederland (The Netherlands IPv6 Foundation)
Author:	Matthias van der Heide
Date:	June 5, 2011
Version:	2011.02

Introduction

This test report is the result of work performed by the Dual-Stack Lite Test Lab of Stichting IPv6 Nederland (The Netherlands IPv6 Foundation; Stipv6). The lab has been established to investigate the viability of Dual-Stack Lite as an IPv6 transition mechanism for Internet Service Providers.

The work was carried out by Matthias van der Heide as part of his graduation for the title Bachelor of ICT at the Haagse Hogeschool (The Hague University of Applied Sciences).

We thank Eric van Uden at AVM for granting access to a Fritz!Box 6360 Cable for the duration of the tests.

Summary

In this report we show how the Fritz!Box 6360 Cable performs in tests of its Dual-Stack Lite capabilities.

The 6360 performs very well in performance tests. Speeds up to 80 mbps down are achieved. Upload speeds are about half, at 35.4 mbps. The TCP MSS clamping setting could be adjusted to increase performance even more.

IPv6 throughput is almost 10 times the Dual-Stack Lite performance. Download speeds of 750 mbps are available if the interfaces are in full gigabit ethernet mode. Upload speed is just a little lower at 698 mbps.

The AVM Fritz!Box 6360 Cable is a router that is ready for IPv6 and Dual-Stack Lite.

Scope

The IETF draft specification of Dual-Stack Lite details the functionality of a Basic Bridging Broadband element (B4) as part of the Dual-Stack Lite transition mechanism. The aim of this technique is to allow broadband internet customer to connect to the internet without having been assigned a full IPv4 address to their home gateway or Customer Premises Equipment (CPE).

Dual-Stack Lite is not a finished standard. Rather it is constantly being updated by its authors based on new insights and comments from the IETF Softwires Working Group.

For the purpose of this test we adhere to the draft specification that was current during the start of this project, on April 1st, 2011. This is version 7 of the draft, dated March 3rd, 2011. The latest specification is version 11 and was published May 28, 2011. The new document changed some requirements from ‘should’ to ‘must’ status, but does not introduce new behavior.

The Dual-Stack Lite specification makes use of techniques defined in RFCs and IETF drafts. Those documents are listed here and are considered part of the specification:

- Generic Packet Tunneling in IPv6 Specification (RFC2473)
- DHCPv6 Option for Dual-Stack Lite

Links:

- Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion - http://tools.ietf.org/html/draft-ietf-softwire-dual-stack-lite-07
- Generic Packet Tunneling in IPv6 Specification (RFC2473) http://tools.ietf.org/html/rfc2473
- DHCPv6 Option for Dual-Stack Lite http://tools.ietf.org/html/draft-ietf-softwire-ds-lite-tunnel-option-09

Device Under Test

The device tested is an AVM Fritz!Box 6360 Cable. This full featured EuroDOCSIS cable modem can connect ISDN phones, DECT phones, USB devices and up to 4 computers directly. It has 4 gigabit ethernet ports. For these tests, the router was configured to act as an ethernet router, using port LAN1 as the provider uplink, and ports LAN2, 3 and LAN4 for clients. Other options were disabled.

The 6360 we tested came with firmware version 85.05.04-19714.

Test objective

The purpose of the tests is to establish the forwarding performance of the 6360 using Dual-Stack Lite and to test interoperability with AFTR devices.

Test setup

For the purpose of these tests a setup was built to perform tests both under lab conditions and with real world connectivity.

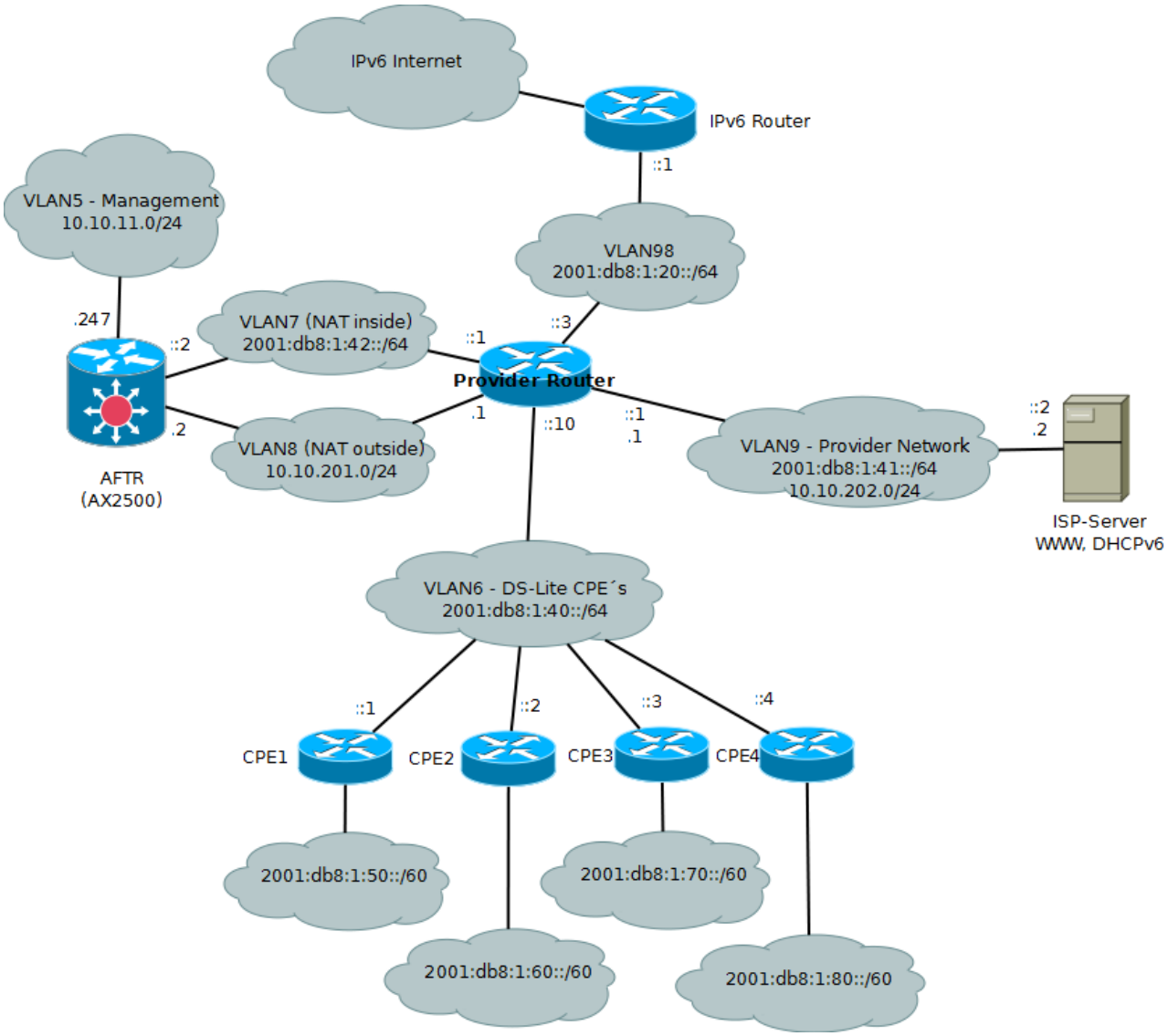
The network is depicted in diagram 1 on page 7. It consists of gigabit ethernet switches, configured with several VLANs to create separated Layer 2 ethernet networks. For hosts we used real and virtualized clients, a virtualized Ubuntu server, and a virtualized Vyatta 6.2 router in the center acting as the provider router. Having all traffic pass through the provider router enables us to monitor all traffic flows, both IPv4 and IPv6, in a central location. The public IPv4 and IPv6 addresses that were used during tests have been replaced by documentation-friendly versions.

The 6360 was configured as CPE1, delegating an IPv6 /60 subnet to it's connected hosts and acting as a B4 element for Dual-Stack Lite IPv4 connectivity.

A Linksys 160NL with the Comcast/Xavient implementation of Dual-Stack Lite running on OpenWRT was configured as CPE2, for the purpose of customer-to-customer connection tests.

The role of Dual-Stack Lite AFTR was performed by an A10 Networks AX2500.

Diagram 1 - Network topology (layer 3)



Results

Manual configuration

The web interface of the 6360 is very user friendly. It operates both over IPv4 and IPv6. It allows the user to connectivity using native IPv4, IPv6 or Dual-Stack Lite.

Automatic configuration

When the 6360 is configured to ask the DHCPv6 server for an AFTR address, it does so using an Option Request Option with the value 54. This option is reserved for RFC 5678, Mobility Services Discovery.

While several known implementations (all using the OpenWRT code) use this value to ask for configuration of an AFTR address, it is reserved for a different usage, and should not be used outside closed environments.

A new DHCPv6 Option has been described in an IETF draft, and is referenced by the Dual-Stack Lite specification. This option allows CPE to request an AFTR domain name from the ISP's DNS server, which will resolve to the IPv6 address of the AFTR.

DNS forwarding

DNS requests from IPv4 hosts in the LAN get forwarded over IPv6 to the configured DNS server. This is the correct behaviour according to the Dual-Stack Lite specification.

Performance

Bandwidth tests using iperf confirmed that an upload speed of 35 mbps is possible. Unfortunately, because of an issue with port forwarding in the A10 AX2500, download speed can not be tested using iperf. Using wget in V6-mode, a thoughput of approximately 80 mbps was achieved.

IPv6 Performance

Performance using IPv6 has been tested. To gain maximum results, the gigabit interfaces had to be switched from 'green mode' (100mbit) to full gigabit ethernet. Tests were executed using iperf in IPv6 mode.

Download performance:	750 mbps
Upload performance:	698 mbps

Fragmentation

Tunneling packets through an IPv4-in-IPv6 tunnel as with Dual-Stack Lite leads to fragmentation of larger packets. In Dual-Stack Lite this has to be done using IPv6 fragmentation. We have tested this using UDP traffic and found the 6360 to fragment the packets correctly, splitting them in two roughly same-sized parts.

There is an issue with the A10 Networks AX2500, which doesn't accept IPv6-fragmented packets from CPEs. This causes packets that are too large for the tunnel, and have to be fragmented in IPv6, to be dropped. If the client isn't aware of this, attempts to retransmit these packets are useless.

To allow fragments through, the client computer was configured with a smaller MTU. This causes it to fragment in IPv4, before packets arrive at the CPE. The resulting IPv4 packets are smaller than the MTU of the Dual-Stack Lite tunnel. The 6360 forwarded these packets to the AFTR, never having to fragment them in IPv6. In this configuration, the network is stable and any packet originating from the client always get delivered to the AFTR.

Because the CPE is not aware of this limitation on the AFTR, it can not be expected to prevent sending IPv6-fragmented packets. However, if customers ask for it, AVM could implement measures to circumvent the problem, by:

1. performing fragmentation in IPv4
2. using DHCPv4 to lower the MTU of clients in the LAN

The first solution violates the Dual-Stack Lite specification, but is not expected to introduce errors. It solves the issue in the router itself, without affecting the configuration of the clients. The second solution is to configure all clients to use a lower MTU. It is unknown whether all DHCPv4 clients implement this configuration option, and adjust their behaviour accordingly.

MSS clamping

To avoid fragmentation of encapsulated TCP packets between the CPE and the AFTR, the TCP Maximum Segment Size (MSS) parameter can be adjusted to influence the size of packets sent during the TCP connection. This is called MSS clamping or MSS rewrite.

The 6360 was tested in a network using Gigabit Ethernet links with an MTU of 1500 bytes. The TCP MSS was monitored on a bi-directional connection made from a client behind CPE1 to the test webserver.

Traces show that the 6360 decreased the MSS of the connection in both directions to 1200 bytes, which is a safe default to prevent fragmentation of TCP packets. The number is derived from the IPv6 minimum MTU of 1280 bytes. Sending packets of this size guarantees delivery, without the danger of fragmentation. The MSS value of 1200 is obtained by subtracting the IPv6 tunnel header size (40 bytes), IPv4 header (20 bytes) and TCP header (20 bytes). So for 1200 bytes of TCP data, 80 bytes of overhead are needed. Half of it is caused by the tunneling mechanism (the IPv6 header) and half of it by regular IPv4 and TCP headers.

However, in a setup like the one tested, much higher MSS values are possible. An MSS

value of 1200, with 80 bytes overhead, is 93,75% efficient. Increasing the MSS to fit the Ethernet link MTU of 1500 results in an MSS of 1420, at a 94,67% efficiency. Links supporting a larger MTU (9000 bytes on Ethernet using jumbograms) are uncommon, but could benefit even more from this. At a 9K MTU, the ratio of headers to data is so low, efficiency peaks at 99,11%.

In reality the MTU of the path between CPE and AFTR is expected to be 1500 or slightly less. This can be caused by the access network itself (for instance PPPoE) or tunnels inside the provider network.

RFC2473 states that the Path MTU of a tunnel should be dynamic. The MSS rewrite value could be based on this Path MTU to gain a much higher efficiency and hence, higher performance for TCP connections. Because the path from CPE to AFTR in a provider network is usually stable, the MTU value shouldn't change unless perhaps a different AFTR is configured, because of failover or change of network topology.

IPv6 Port forwarding

A minor usability issue was encountered during testing of IPv6 port forwarding. This is unrelated to Dual-Stack Lite functionality.

The webinterface allows users to open up IPv6 hosts inside their network to connections coming from the public internet. This allows clients on the internet to access servers on the netwerk of the user using IPv6 connections.

The input form allows the user to select a known host inside the network, or specify the host part of the IPv6 address.

When attempting to edit the IPv6 address, groups of characters that have the value zero (0000) are left out. This leaves one or more fields of the address empty. Submitting the form correctly warns the user of the flawed address.

Recommendations

- The value chosen for MSS clamping (1200 bytes) is a safe default, but prohibits a better performance. We recommend making it depended on the Path MTU to the AFTR element, to obtain the best possible performance and efficiency.

- The ease the configuration of the 6360 for use in a Dual-Stack Lite setup we encourage AVM to implement the DHCPv6 Dual-Stack Lite Tunnel Option that is currently under development by the IETF.

- There is an interoperability issue with the A10 Networks AX2500 AFTR which causes large UDP-packets to be dropped when sent from the CPE to AFTR using IPv6 fragments. This is caused by the AX2500 which does not implement that part of the Dual-Stack Lite spec. If the feature is not implemented by A10, AVM can take measures to prevent sending IPv6 fragments.

- The interface for setting up IPv6 port forwardings appears to have a bug in it, that makes it hard to edit the IPv6 address correctly. This issue should be easy to solve.

- We recommend that AVM work together suppliers of AFTR devices, such as A10, Cisco and Juniper, to test their implementation of Dual-Stack Lite and to provide feedback to the community (IETF Softwires Working Group) on issues found and best practices. The Stichting IPv6 Nederland (The Netherlands IPv6 Foundation) can assist AVM with more independent interoperability testing.

Afterword

Stichting IPv6 Nederland would like to thank AVM, and especially Eric van Uden, for cooperating in this test.

We hope our findings will help to further improve the Dual-Stack Lite functionality in your products.