

Samenvatting Infrastructure

Noeël Moeskops

October 30, 2017

Contents

1	Computer Networks and the Internet	3
1.1	What is the Internet	3
1.1.1	A Nuts-and-Bolts description	3
1.1.2	A service Description	3
1.1.3	What is a protocol	3
1.2	The Network Edge	4
1.2.1	Access Network	4
1.2.2	Physical Media	4
1.3	The Network Core	4
1.3.1	Packet Switching	4
1.3.2	Circuit Switching	4
1.3.3	A Network of Networks	5
1.4	Delay, Loss And Throughput in Packet-Switched Networks . . .	5
1.5	Protocol Layers and Their Service Models	6
1.5.1	Layerd Architecture	6
1.5.2	Encapsulation	6
1.6	Network Under Attack	7
2	Application Layer	8
2.1	Principles of Network Applications	8
2.1.1	Network Application Architectures	8
2.1.2	Processes Communicating	8
2.2	The Web and HTTP	8
2.2.1	Overview of HTTP	9
2.2.2	Non-Persistent and Persistent Connections	9
2.2.3	HTTP Message Format	9
2.2.4	User-Server Interaction: Cookies	11
2.2.5	Web Caching	11
2.3	Electronic Mail in the Internet	12
2.3.1	SMTP	12
2.3.2	Comparison with HTTP	12
2.3.3	Mail Message Formats	12
2.3.4	Mail Access Protocols	12
2.4	DNS-The Internet's Directory Service	12
2.4.1	Services Provided by DNS	12
2.4.2	Overview of How DNS Works	13
2.4.3	DNS Records and Messages	14
2.5	Peer-to-Peer File Distribution	14

1 Computer Networks and the Internet

1.1 What is the Internet

1.1.1 A Nuts-and-Bolts description

Het internet bestaat uit verschillende lagen de **kern** en de **hosts/end systems**. Eind systemen zijn apparaten die de **application layer** protocol verstaan. Hier uit kan je verstaan dat het laptops, mobiels, computers, tv's, game consoles etc. Deze zijn met elkaar verbonden door middel van **communicaton links en packets switches** (coax cable, radio etc.) Al deze dingen hebben verschillende **transmission rate**. Deze worden altijd afgebeeld in *bits/s. Meestal zie je mb/s of kb/s. Om van bit naar byte te gaan moet je het getal gedeelt door 8 doen.

De meest voorkomende vorm van packet switchers zijn **routers en link-layer switches**. routers worden meer in de netwerk core gebruikt. Het pad wat een pakketje neemt van van host A naar host B te komen word een **route of path** genoemd.

Een **Internet Service Provider (ISP)** geeft je een ip adres en verbind je met het internet. Je hebt hier verschillende lagen in. Je hebt lokale ISP (laag 3) laag 2 ISPs en in de core zitten laag 1 ISP.

Alle applicaties op je computer die verbinding willen maken met een andere eind-systeem maken gebruik van een **application protocol** voorbeelden hiervan zijn http, https, ftp, skype, nfs etc. De belangrijkste protocolen voor het internet zijn **Transmission Control Protocol (TCP)** en **Internet Protocol (IP)**. Het internet model wat nu in gebruik is heet daarom ook het **TCP/IP** model.

Internet standaarden zijn super belangrijk om te kunnen communiseren tussen systemen. Deze worden vaak gemaakt door het **Internet Engineering Task Force (IETF)**. Openbare vrijstandaarden noem je **requests for comments (RFCs)** daar mag iedereen aan werken; implementeren en wijzigen. Of het word goed gekeurd is een ander verhaal.

1.1.2 A service Description

distributed applications zijn programma's die verdeelt zijn onder meer mensen dus games, sociale media, streaming services omdat ze meer system ding doen ofzo (pag 33.).

Een eindsysteem bind zichzelf aan een socket. een webserver bijvoorbeeld poort 80. Dit noem je een **socket interface**.

1.1.3 What is a protocol

Een protocol is een van te voren afgesproken manier om met elkaar te praten. Dit hebben computers nodig om succesvol met elkaar te kunnen communiseren. Elk protocol heeft zo haar eigen doel.

1.2 The Network Edge

Het internet bestaat uit **clients** en **host** een client is een systeem wat iets opvraagt van een host. En een host is een systeem wat informatie geeft aan een client op aanvraag. Een **data center** kan bestaan uit meerdere **end-systems/host** kan oplopen in de tonnen.

1.2.1 Access Network

Om met het internet te verbinden heb je natuurlijk ook een kabel nodig. Een **Digital subscriber line (DSL)** is z'n soort kabel. ookal word DSL voor meer dan 85% gebruikt kan je veel sneller gaan met **fiber to the home (FTTH)**

1.2.2 Physical Media

fysieke media vallen in twee categorieën: **guided media** en **unguided media**. guided media is bedraad en unguided media in draad-loos.

Unshielded twisted pair (UTP) word vaak gebruikt binnen een gebouw. transfer rates gaan van 10 Mbps tot 10Gbps. Coax cabels zijn een guided **shared medium**, op een shared medium kunnen meerdere systemen aangesloten zijn met maar 1 kabel. terwijl de data niet shared is.

In communicatie worden er twee verschillende soorten satelite gebruikt. **Geostationary satellites** deze blijven op een plek boven de aarde zweven op 36.000 KM. En **low-eath orbiting (LEO) satallites** deze zijn een stuk dichterbij en staan ook nooit stil.

1.3 The Network Core

1.3.1 Packet Switching

Wanneer je een bericht stuurt over het internet word hij in kleine stukjes gehakt, dit noem je **packets** dit pakketje gaat door allemaal **packet switches** (routers en link-layer switches).

De meeste packetswitchers zijn **store-and-forward transmission** dit betekent dat ze eerst wachten totdat alle stukjes van een pakket binnen is voordat hij het door verstuurt. Voor iedere link aan een router heb je ook een **output buffer/output queue** dit is een stukje ruimte waar de router pakketjes stopt voordat hij ze versuurt. Een pakketje kan dus een queuing delay oplopen. Wanneer een queue vol zit kan het pakketje worden gedropt en ontstaat er dus **packet loss**.

Er zijn ook verschillende **routing protocols** Voor routers om met elkaar te praten.

1.3.2 Circuit Switching

word niet behandeld. (pag. 55)

1.3.3 A Network of Networks

Omdat er meerdere lagen van ISPs bestaan word degene waar jij in contact mee komt de access ISP genoemd en de bovenste laag de global ISP. Je zou ook kunnen stellen dat de access ISP een **customer** is van global ISP en dat hij weer de **provider** is.

Een **PoP** is een verzameling van routers op een lokatie (op een niet access ISP level) die word gebruikt om meerdere access ISP te verbinden naar local of global ISP. Een super switch eigenlijk maar dan van routers. ISPs kunnen ook direct met elkaar contact leggen zonder hulp van een local of global ISP dit bespaart kosten en noem je **peer**.

Een random bedrijf kan ook een **Internet Exchange Point (IXP)** opzetten. Dit is wel letterlijk een mega switch waar iedereen verbinding mee kan maken om zo internet te delen. Er bestaan ook **content-provider networks** Dit zijn mega bedrijven zoals Google, facebook, Amazon etc. Die als tier 1 (global) ISP spelen.

1.4 Delay, Loss And Throughput in Packet-Switched Networks

Wanneer een pakketje bij een node is kan het vertraging oplopen. Dit kan **nodal processing delay, gueueing dealy, transmission delay of propagation delay** zijn. Al deze dingen samen noemen we **total nodal delay**.

processing delay is de tijd dat het duurt voor een node om naar de header van een pakketje te kijken. en om te kijken waar het naartoe moet.

queuing delay is de tijd dat het pakketje in de rij moet wachten tot dat het verstuurt kan worden.

transmission delay Dit is de tijd dat het duurt om alle pakketjes op de link te pushen. Dit is niet de tijd van een node naar de andere.

propagation delay Dit is de tijd dat het duurt om een pakketje van een node naar de anderen te sturen.

traffic intensity word bepaald door $\text{paketsgrote} \times \text{aantalpakketjes} / \text{transmissionrate}$ (La/R). Als een queue van een node is kan hij het pakketje **droppen** dan is de data verloren, hij kan er ook voor kiezen om een ander pakketje te droppen.

Instantaneous throughput is de throughput in een moment. **average throughput** is de gemiddelde throughput.

1.5 Protocol Layers and Their Service Models

1.5.1 Layerd Architecture

Protocol Layering Het idee van protocol layering is dat je al je dingen in andere lagen zet. Iedere laag heeft dan end-points waarmee het een service biedt aan de laag erboven. en een end-point waarmee het service vraagt aan de laag eronder. De inhoud van de layer kan veranderen als de end-points maar hetzelfde blijven. Dit zorgt voor een schaalbaar systeem.

Applicatie layer Dit is de layer die drijft op een eind-systeem protocollen zoals HTTP, FPS, Skype, IMAP etc. dit gaat oneindig door. data wordt hier een **message** genoemd

Transport Layer deze laag zorgt ervoor dat de data van de applicatie laag uit je computer gaat en dat inkomende data naar de juiste applicatie worden geleid. Deze laag kent maar twee protocollen: **TCP & UDP**. Data wordt hier een **segment** genoemd

Network Layer Deze laag zorgt voor de navigatie van node naar node. Dit kent ook maar een aantal protocollen: **IP, en not iets?** //TODO: protocollen. Data wordt hier een **datagram** genoemd.

Link layer dit is de laag wat letterlijk de data van node A naar node B brengt. Het gebruikt hiervoor de physical layer. Dit doet het met het **PPP** protocol. Data op deze laag worden **frames** genoemd.

physical layer dit is letterlijk een ethernet kabel of wi-fi verbinding.

1.5.2 Encapsulation

Wanneer je data verstuurt over het internet doe je aan encapsulation. Een beetje als een russische pop met meerdere lagen. Een Message wordt in een segment gestopt. waarbij de transport layer alleen kijkt naar de header van de message. Hij voegt dan wat nuttige data toe aan de segment header en geeft het dan aan de network laag. die maakt er een datagram van en kijkt weer naar de header van de segment. Zo gaat het door tot aan de Link laag. Wanneer de Frame bij een switch komt wordt hij niet uitgepakt omdat deze ook alleen gebruik maakt van de link layer. Een router werkt echter op de netwerk laag, het pakketje wordt dan uitgepakt naar een datagram wanneer het bij een router is en weer ingepakt als de router het weer verstuurt. Wanneer het bij zijn eindbestemming aankomt wordt het weer helemaal uitgepakt tot een Message.

1.6 Network Under Attack

slechte software word **malware** genoemd. Je hebt hierin meerde varianten: **Viruses** die je zelf op je computer werkt en **worms** die een via een lek in je systeem komen. Veel malware in **self-replicating** het verspreid zichzelf.

deniel-of-service (DoS) is dat je heel veel verkeer stuurt naar een host om zo zijn toegang tot het internet te blokeren. Dit kan op drie manieren *vulnerability attack* is een aanval speciviek gemaakt om een lek te gebruiken om toegang te blokeren. *bandwidth flooding* Met deze aanval stuur/vraag je mega veel en vaak data naar/van een host. *connection flooding* is dat je alle sockets van een host bezet houd door een (half) open connectie te behouden. Je kan ook een **distributed DoS (DDos)** uitvoeren doormiddel van een **botnet**.

Je kan ook op een lokaal netwerk **packets sniffen** dan inspecteer je alle pakketjes op een netwerk. Je kan ook je **ip spoofen** dan doe je net alsof je een andere IP hebt.

2 Application Layer

Dit hoofdstuk gaat over de Application Layer. waar eigenlijk al je netwerk programma's opdraaien.

2.1 Principles of Network Applications

Het principe van netwerk layers is dat je je maar meestal zorgen hoeft te maken over 1 of 2 lagen van het netwerk systeem. Wanneer je een applicatie schijft voor een eind-systeem hoef je dus alleen maar zorgen te maken over hoe de applicatie layer jou programma implementeert.

2.1.1 Network Application Architectures

Er zijn twee verschillende application netwerk architecturen; **P2P** (peer-to-peer) en **client-server** architectuur. Je bent niet gelimiteerd aan deze twee architecturen maar het zijn de enige die op dit moment bestaan. (je kan zelf iets nieuws verzinnen als je genoeg vrije tijd hebt).

client-server

Een client-server is de meest simpele en meest traditionele architectuur. Het gebruikt twee systemen een *client* en *server*. De server moet op een statische plek staan en 24/7 beschikbaar zijn. De client daarin tegen kan dynamisch zijn (van IP verwisselen) en uit en aan gaan wanneer gewenst. De server neemt geen dienst af van de client.

Peer-to-peer

Een peer-to-peer verbinding wordt onder anderen gebruikt bij Torrents en video gesprekken. De verbinding wordt onderling gedaan zonder centrale server. Dit systeem schaaft ook veel beter omdat wanneer iedereen een bestand wilt downloaden het ook door meer mensen wordt geupload.

2.1.2 Processes Communicating

Wanneer een **process** wilt communiceren met een ander process op het Internet heeft het 2 dingen nodig, het adres van de het anderen end-systeem en **Sockets** of **ports** zij

2.2 The Web and HTTP

Tot 1990 werd het internet grotendeels door onderzoekers en scholieren gebruikt. Weinig mensen wisten er nog van. Later in 1994 werd de eerste webbrowser ontwikkeld en daardoor werd het internet onder de het normale volk ook ontamd.

2.2.1 Overview of HTTP

Hyper Text Transer Protocol (HTTP) word gebruikt om documenten (objecten) over het internet te versturen. Wanneer een object (index.html) verwijst naar een ander object (favicon.ico) word deze ook opgehaald. HTTP maakt gebruik van TCP en word beschouwd als een **stateless protocol** dat houdt in dat de server geen data bijhoud van haar clients. Wanneer een client 4x hetzelfde bestand zou aanvragen zal de http server de 4e aanvraag hetzelfde behandelen als de eerste.

2.2.2 Non-Persistent and Persistent Connections

HTTP kent twee soorten verbindingen: Non-Persistent en Persistent.

Non-Persistent

Wanneer een client een object aanvraagt van een HTTP server gaan ze eerste een **three-way-handshake** doen. De client vraagt aan de server om een TCP verbinden opstellen en de server antwoord vervolgens met *OK*. Daarna vraagt de client een object aan (bijv. */mydir/index.html*) De server antwoord dan met het bestand *index.html* en de verbinden word verbroken.

De client leest het bestand en ziet dat hij nog 10 plaatjes moet laden. Hij gaat dan weer een three-way-handshake doen en het hele verhaal begint weer opnieuw.

Round-Trip Time (RTT)

RTT is het het tijd wat het duurt voor een pakketje heen en terug te sturen van client naar een server. Propegation delay heen en terug. In het vorige voorbeeld zijn er twee RTT. De eerste is wanneer de client connectie aanvraagt aan de server en antwoord krijgt. De twee is het bestand wat hij vraagt en ook krijgt van de server. Voor elke extra object krijg je bij een non-persisten verbinding **2** RTT's (een om verbinden vast te leggen, en een om het bestand te krijgen.)

Persistent

Een persistent verbinden is veel zuiniger met data dan een Non-persistent verbinden. Want het houdt de TCP connectie open tot een bepaalde timeout tijd. Dus in ons voorbeeld met 1 html pagina en 10 plaatjes heb je dus maar 12 RTT's op een persistent verbinding (1 RTT voor TCP verbinden + html bestand + (plaatje * 10)) In plaatjs van $11 * 2 = 22$ RTT's bij een Non-persistent verbinding.

2.2.3 HTTP Message Format

Er zijn twee verschillende HTTP bericht formaten, een request en een response. Hieronder volgt een voorbeeld van een request:

```
GET /tools/index.html HTTP/1.1
Host: www.noeel.nl
Connection: close
User-agent: Mozilla/5.0
Accept-language: nl
```

De request is geschreven in ASCII. De eerste zin is de **request line** de lijnen die daarop volgen zijn de **header lines**. Daarna komt er een wit-regel gevolgt door de **Entity body** maar die is nu leeg omdat het een request is en geen response of POST

Request line

De request kan bestaan uit een GET, POST, HEAD, PUT of DELETE

GET request een bestand, POST stuur data naar een server (de content staat dan in de body) HEAD voor debugging, vraag alleen de head van een response op. PUT plaats een bestand en DELETE verwijder een bestand. Het word daarna gevolgt door de locatie en de HTTP versie.

Header lines

Host: verklaard de host waar de request naar toe moet gaan, is nodig voor proxy's **Connection:** of het een persistant (**open**) of non-persistant (**close**). Data zoals **User-agent** en **Accept-language** worden gebruikt om verschillende data weer te geven per taal/browser.

HTTP Response

```
HTTP/1.1 200 OK
Connection: close
Date: Tue, 18 Aug 2015 15:44:04 GMT
Server: Apache/2.2.3 (CentOS)
Last-Modified: Tue, 18 Aug 2015 05:44:04 GMT
Content-Lenght: 6532
Content-Type: text/html
```

data data data

In de Status line staat de HTTP versie gevolgt door de statis code en statis zin. (200 betekent dat alles goed is) **Connection: close** betekent dat de server de TCP verbinding zal sluiten hierna. **Date** De datum en tijd dat de server het bestand van het file system heeft gekregen. **Server:** Allemaal server data, welke http server hij draaid, welke versie en welk OS. **Last-Modified:** handig voor caching enzo. **Content-Lenght** hoelang is de data en **Content-Type** wat voor soort data het is.

2.2.4 User-Server Interaction: Cookies

Omdat HTTP stateless is weet de server niet of de client vaker gebruikt maakt van zijn diensten. Daarom bestaan er Cookies, een Cookie is een string die de server kan zetten met de HTTP header: `Set-cookie: myCookieValue` Wanneer de client een nieuwe request maakt naar dezelfde server stuurt hij altijd in zijn HTTP header (`Cookie: myCookieValue`) zodat de server dit kan processen en client specifieke diensten aan kan bieden.

2.2.5 Web Caching

Je kan ook een caching server hebben. Dit is een server dat een kopie maakt van een bestaande server en dit op een andere plek host. Dit kan handig zijn als de normale server een hoge ping heeft en dus een lange response tijd heeft. Om de caching server te bereiken moet je zelf in je browser (of andere applicatie) instellen dat je met de caching server verbinding wilt maken in plaats van de bestaande server.

Wanneer je een HTTP GET request stuurt naar een caching server kijkt hij of hij het bestand heeft en stuurt deze dan naar jou toe. Wanneer het bestand ontbreekt vraagt hij het zelf op bij de originele server en stuurt het dan naar jou toe en slaat het ook meteen op.

Conditional GET

Het grote nadeel van een caching server is dat hij een out-of-date bestand kan hosten. Om dat te voorkomen checked een caching server eens in de paar weken of er geen nieuwere bestand bestaat. Dat gaat ongeveer zo: /newline Eerst vraagt de caching server het originele bestand aan

```
GET /img/logo.png HTTP/1.1
Host: www.mylogoImages.com
```

De originele server geeft een response:

```
HTTP/1.1 200 OK
Date: Sat, 3 Oct 2015 15:39:29
Server: Apache/1.3.0 (Unix)
Last-Modified: Wed, 9 Sep 2015 09:23:24
Content-Type: image/png
```

data data data ...

Wanneer er een bepaalde tijd voorbij is en de caching server is bang dat zijn bestand out-of-date is checkt hij bij de server of er een nieuwere versie beschikbaar is:

```
GET /img/logo.png HTTP/1.1
Host: www.mylogoImages.com
If-modified-since: Wed, 9 Sep 2015 09:23:24
```

Wanneer er geen nieuwe versie is beantwoordt de server met een 304:

HTTP/1.1 304 Not Modified
Date: Sat, 10 Oct 2015 15:39:29
Server: Apache/1.3.0 (Unix)

2.3 Electronic Mail in the Internet

E-mail bestaat uit drie delen: **user agents** (het client programma op het eind-systeem van de gebruiker), **mail servers** (de servers die mails ontvangen en versturen) deze hebben zowel een client als server rol en het protocol om mails te versturen **Simple Mail Transfer Protocol (SMTP)**

2.3.1 SMTP

//TODO

2.3.2 Comparison with HTTP

2.3.3 Mail Message Formats

2.3.4 Mail Access Protocols

2.4 DNS-The Internet's Directory Service

Servers hebben vaak twee adressen waarmee je met ze kan verbinden: hostname *www.noel.nl* en het ip adres *185.182.56.20*. Computers willen met het ip adres verbinden omdat dat ook echt iets betekent. Maar voor mensen is dat moeilijk te onthouden. Daarom bestaan er hostnamen die verwijzen naar het ip-adres zodat je die in kan typen in je webbrowser.

2.4.1 Services Provided by DNS

Wanneer jij verbinding wil maken met een hostname verstuurd je computer een verzoek naar een DNS server om het ip adres op te zoeken van de hostname. De DNS server geeft dan het ipadres en je kan verbinding maken. Een DNS server bied een aantal services:

Host aliasing Een sub domain (*sub1.sub2.domain.org*) word een **canonical hostname** genoemd deze verwijzen naar *www.domain.org* of *domain.org*

Mail server aliasing Je kan van elk domain naam ook een MX-record opvragen, dit is het adres van de mail server.

load distribution Je kan ook meerdere records van een soort (MX,A etc.) opgeven bij een DNS server. Hij kiest er dan per request een random uit zodat niet een server van je word belast.

2.4.2 Overview of How DNS Works

Er bestaan 4 verschillende DNS servers:

Root DNS servers Dit zijn servers die alleen een lijst hebben met TLD DNS servers. Wanneer jij een request maakt naar *www.google.nl* verwijst deze server je door naar een NL TLD DNS server

Top-Level domain server elk top-level domein naam (nl, com, org, de etc.) heeft haar eigen TLD DNS server. In deze server staan alle **Authoritative DNS servers** van elk domain naam onder dat TLD. Wanneer je een request doet voor het ip van google.nl geeft hij je de Authoritative DNS servers van google.nl.

Authoritative DNS servers Dit zijn de DNS servers die ook daatwerkelijk het antwoord voor je hebben. deze hebben alle records van een domain naam en geven je ook het ipadres van je bestemming.

local DNS server Dit zijn vaak DNS servers van je ISP (Ziggo bijv.) Deze hebben een lijst van alle Root DNS servers en vragen aan hen het ip adres van je request. Hier maak jij dus ook verbinding mee.

DNS request

Wanneer jij een vraag naar een DNS server doet voor het ophalen van een ip address kan dat op twee verschillende manier gaan.

Various DNS servers

- Een eind-systeem maakt doet een DNS request voor *www.noel.nl* naar een locale DNS server
- locale server doet een request naar de root server
- de root server leest het *.nl* gedeelte en returnt een adres van een NL TDL DNS server
- de locale server ontvangt de request en stuurt een DNS request voor *www.noel.nl* naar de NL TDL DNS server
- de TLD server leest het bericht en returnt de authoritative DNS server voor noel.nl
- de local server stuurt een bericht naar de authoritative DNS server en krijgt de record die hij verwachte

- de local server stuurt het ip-adres door naar het eind-systeem dat het eerder aanvroeg.
- profit!

Recursive queries in DNS Dit werkt op een anderen manier als de various DNS server variant. In plaats van dat de server steeds een adres returned van een ander DNS server gaat hij zelf vragen aan die server voor het adres. Maar dit komt minder vaak voor.

DNS servers cashen ook al hun resultaten om verkeer te verminderen, en om sneller te zijn. Meestal word iets voor 2 dagen gecached. Maar als web-host kan je dit ook korter of langer instellen.

2.4.3 DNS Records and Messages

2.5 Peer-to-Peer File Distribution