Chapter 4: Networking and the Internet

The Internet

Het doel van het Internet was de mogelijkheid om een variëteit van computernetwerken aan elkaar te linken, te ontwikkelen, zodat ze als een verbonden systeem kunnen functioneren en niet gestoord worden door lokale fouten.

Internet Architecture

De verbonden netwerken zijn opgebouwd en onderhouden door organisaties: **Internet Service Providers (ISPs).**

Hiërarchie (volgens de rol die de netwerken spelen in de algemene Internetstructuur)

- **Tier-1 ISPs**
  - Bestaan uit zeer snelle, grote capaciteiten, internationale WANs
  - Worden gezien als de ruggengraat van het Internet
  - Worden gebruikt door firma's in de communicatiesector
- **Tier-2 ISPs**
  - Meer regionaal en minder krachtig in hun capaciteiten
  - Worden gebruikt door firma's in de communicatiesector

Tier-1 en Tier-2 ISPs zijn essentiële routernetwerken die samen zorgen voor de communicatie-infrastructuur van het Internet → worden gezien als de kern van het Internet

- **Acces of tier-3 ISP**
  - Onafhankelijk internet, soms een **intranet** genoemd, bediend door één autoriteit die in de business zit van de levering van Internettoegang aan privéwoningen en firma's.
  - Geeft toegang tot de kern van het Internet (tier-1 en tier-2 ISPs)
- **End systems of hosts**
  - Dit zijn de toestellen die individuele gebruikers verbinden met de acces ISPs
  - Bv. laptops, gsm's…

Het Internet is in essentie een communicatiesysteem!

De technologie waarmee end systems zich connecteren met grotere netwerken is heel gevarieerd.

- Wireless connections
  - Snelgroeiend, gebaseerd op WiFi-technologie
  - Het access point (AP) wordt verbonden aan een acces ISP en wordt dus voorzien van Internettoegang via die ISP naar de end systems in het AP's uitzendbereik.
    - **Hot spot**: het gebied wat het AP kan bereiken
- Telefoonlijnen en kabel/satellietsystemen
  - Zijn ontworpen om analoge communicatie (de stem, pre-digitale televisiesignalen) te vervoeren
  - Problemen gelinkt aan deze analoge communicatiemiddelen worden vaak **the last mile problem** genoemd.

Internet Addressing

Een internet heeft een breed adressysteem nodig dat aan elke computer in het systeem een uniek adres toekent. In het Internet zijn deze adressen gekend als **IP-adressen** (IP: Internet Protocol).

- Reeks van 32 bits (origineel), nu zitten ze in het proces om een reeks van 128 bits om te zetten
- Achtereenvolgende blokken van genummerde IP-adressen zijn toegekend aan ISPs door de Internet Corporation for Assigned Names and Numbers (ICANN)
    - Non-profit organisatie die het Internetgebruik coördineert
- Machines doorheen het Internet hebben een uniek IP-adres
- Traditioneel geschreven in '**dotted decimal notation**'
    - Waarin de bytes van de adressen gescheiden zijn door punten
    - En elke byte is uitgedrukt als een geheel getal weergegeven in de traditionele basis 10-notatie

Het Internet heeft een alternatief adressysteem waar machines geïdentificeerd kunnen worden door mnemonic namen (ezelsbruggetjes/geheugensteuntjes). → Handig voor mensen

- Gebaseerd op het concept van een **domain**: een soort regio van het Internet uitgevoerd door één autoriteit
    - Bv. universiteit, firma, overheid…
- Elk domein moet geregistreerd worden bij ICANN → een proces behandelt door bedrijven, **registrars** (registers) genoemd.
- Aan het domein is een mnemonic **domain name** toegewezen, deze is uniek tussen alle domeinnamen in het Internet
    - **Top-level domains (TLDs)**: achtervoegsels
        - Bv. mu.edu
    - **Country-code TLDs**: be voor België, au voor Australië…
    - Andere **TLDs**: com (commerciële instituties), org (non-profit), net…
- **Subdomaines**: worden gebruikt om de namen in een domein te organiseren
    - Presenteren vaak verschillende netwerken in het rechtsgebied van het domein

Mnemonic adressen zijn handig voor mensen, maar boodschappen worden altijd vervoerd over het Internet in IP-adressen. De mnemonic namen moeten dus naar een IP-adres worden omgezet.

- **Name servers**: essentiële adresboeken die zorgen voor de vertaling tussen services en clients
    - Worden gebruikt als een breed Internet-adresboeksysteem, gekend als **the domain name system (DNS)**
        - Het proces om DNS te gebruiken voor een vertaling wordt **DNS lookup** genoemd.

Internet Applications

Vroeger: traditionele netwerkapplicaties

- Network News Transfer Protocol (NNTP): newsreader application
- File Transfer Protocol (FTP): applicatie om bestanden te categoriseren en te kopiëren doorheen het netwerk
- Telnet protocol: een applicatie om toegang tot een andere computer te krijgen van een grote afstand
    - Later: Secure Shell (SSH) protocol

Deze traditionele netwerkapplicaties moeten gecoördineerd worden door webpagina's via het krachtige **Hyper Tekst Transfer Protocol (HTTP).**

Voorbeeld 1: elektronische mail (**email**)

- De overdracht van emailberichten van de ene computer naar de andere op het Internet blijft het in het domein van basis netwerkprotocollen zoals SMTP
    - **SMTP** (Simple Mail Transfer Protocol) bepaalt een weg zodat 2 computers in het netwerk kunnen interageren wanneer ze een email overdragen van de ene host naar de andere.
        - Voorbeeld van een protocol gebaseerd op vertrouwen, dat is misbruikt geweest door spammers en anderen
        - Moderne mailservers moeten uitgebreide versies van SMTP gebruiken om te verzekeren dat de overdracht van de email veilig gebeurd wordt
        - Gebruikt de ASCII code
    - Om dit te bereiken, gebruikt het DNS, een ander netwerkprotocol
- **MIME** (Multipurpose Internet Mail Extensions): zijn ontworpen om non-ASCII data om te vormen naar de SMTP compatibele vorm
- Er zijn 2 populaire protocollen die gebruikt kunnen worden om toegang te krijgen tot email die gearriveerd is in een gebruiker zijn mailserver:
    - **POP3** (Post Office Protocol version 3)
        - Eenvoudigste
        - Wordt gedaan op de gebruiker zijn lokale machine
    - **IMAP** (Internet Mail Access Protocol)
        - Laat de gebruiker toe om berichten bij te houden en te bewerken op dezelfde machine als die van de mailserver
        - Kan dus op verschillende computers aan zijn email

Voorbeeld 2: **VoIP** (Voice over Internet Protocol)

- Hier wordt de Internetinfrastructuur gebruikt om stemcommunicatie te voorzien, gelijkaardig aan vroegere telefoonsystemen
- Er bestaan 4 verschillende vormen
    - VoIP **soft phones** bestaan uit P2P software die 2 of meer PCs toestaat om een gesprek te delen met gewoon een speaker en een microfoon
        - Bv. skype
    - **Analog telephone adapters**: toestellen die gebruikers toelaten om met hem/haar traditionele telefoon verbinding te maken met telefoonservices voorzien door een acces ISP

    - Embedded VoIP phones: toestellen die een traditionele telefoon vervangen met een equivalente handset rechtstreeks verbonden met een TCP/IP-netwerk
    - De huidige generatie smartphones gebruikt wireless VoIP-technologie

Voorbeeld 3: **Internet Multimedia Streaming**

- **Streaming**: het transporteren van audio en video over het Internet in real-time
- **N-unicast**: een enkele zender stuurt berichten naar meerdere ontvangers
    - Heeft als nadeel om een grote last te plaatsen op de server van het station en op de rechtstreekse Internetburen
- **P2P**: verschuift de last naar de cliënten (peers) zelf
- **Multicast**: verplaatst het distributieprobleem naar de Internetrouters
    - Een server verzendt een boodschap naar meerdere cliënten d.m.v. één adres en vertrouwd op de routers in het Internet om de betekenis van het adres te herkennen en om copies door te sturen van de boodschap naar de juiste bestemmingen
- De meeste applicaties in deze categorie zijn **on-demand streaming**
    - De gebruiker verwacht om naar iets te kunnen kijken wanneer hij wilt.
- Grootschalige streaming services maken gebruik van **content delivery networks (CDNs)**
    - Groepen servers zijn strategisch verspreid in het Internet dat gespecialiseerd is in het streamen van copies of inhoud naar dichtstbijzijnde gebruikers in hun netwerk
- **Anycast**: dit stelt een end user in staat om automatisch te verbinden met de dichtstbijzijnde server in een groep servers

The World Wide Web

Het wereldwijde web kent zijn oorsprong in het werk van Tim Berners-Lee. Hij ontdekte het potentieel van de combinatie van internettechnologie met het concept van linked-documents (**hypertext**).

The world wide web

- Een hypertext document formaat om hyperlinks te plaatsen naar andere documenten
- Een protocol om hypertext over te dragen doorheen het netwerk en een serverproces die pagina's op aanvraag voorziet

Web Implementation

Softwarepakketten die ons, gebruikers, toelaten om toegang te krijgen tot de hypertext op het Internet vallen in 2 categorieën:

- **Browsers**
    - Een browser verblijft op de gebruiker zijn computer
    - Taken: opgevraagd materiaal verkrijgen, dit materiaal presenteren aan de gebruiker op een georganiseerde manier

- **Webservers**
    - De webserver verblijft op een computer en bevat hypertext documenten die toegankelijk zijn.
    - Taak: om toegang voor de documenten onder zijn controle te voorzien, zoals gevraagd door zijn cliënten (browsers)

Om de documenten op het web te lokaliseren en op te vragen, heeft elk document een uniek adres: **Uniform Resource Locator (URL).**

- Een URL bestaat uit 4 onderdelen
    - Het protocol voor het gebruik om te communiceren met de server die de toegang tot de documenten controleert
    - Het mnemonic adres van de machine die de server bevat
    - Het directory (map) pad, nodig voor de server om de map te vinden die het document bevat
    - De naam van het document
- Niet alle 4 de delen zijn noodzakelijk

HTML

Een hypertext document bevat speciale symbolen, **tags** genoemd, zij beschrijven hoe het document moet verschijnen op het scherm, welke multimedia (zoals foto's) met het document moeten komen en welke zaken in het document gelinkt zijn aan andere documenten.
→ Dit systeem van tags is gekend als **Hypertext Markup Language (HTML)**

De kop van een webpagina bevat de inleidende informatie over het document (de datum, het onderwerp…). De body bevat de inhoud van het document, het materiaal dat op je computerscherm verschijnt.

HTML-tags:

- <html> duidt het begin aan van een HTML document
- <head> duidt het begin aan van het "hoofd" van een document
- </p> duidt het einde van een alinea aan
- duidt het einde van een item aan dat gelinkt is aan een ander document

XML

HTML is een notatiesysteem waarbij een tekstdocument gecodeerd kan worden als een eenvoudig tekstbestand. Op gelijkaardige wijze kunnen we ook niet-tekstmaterialen coderen als tekstbestanden (bv. bladmuziek). We gebruiken een alternatief notatiesysteem.

De **eXtensible Markup Language (XML)** is een gestandaardiseerde stijl om notatiesystemen te ontwerpen om data als tekstbestanden voor te stellen.

- Volgens de XML-standaard zijn notatiesystemen, **markup languages** genoemd, ontwikkeld om wiskunde, multimediapresentaties en muziek voor te stellen.
- Goed voorbeeld van hoe standaarden ontworpen zijn om een groot bereik van applicaties te hebben.

- XML zorgt ervoor dat er nieuwe markup languages ontwikkeld kunnen worden die verschillen van HTML. De nadruk moet meer liggen op de betekenis dan op het voorkomen.
    - De betekenisvolle aanpak zou meer search engines (websites die gebruikers assisteren om web-materiaal te lokaliseren) toelaten.

Client-side and Server-Side Activities

- **Client-side** activiteiten: ze worden uitgevoerd door een client (bv. browser)
- **Server-side** activiteiten: ze worden uitgevoerd door een server (bv. webserver)

Voorbeeld: een reisorganisator wilt klanten hun droombestemmingen en reisdata presenteren, via een afgestemde webpagina. Deze moet enkel de informatie bevatten die toepasselijk is voor de klant.

- De website zorgt eerst voor een webpagina met de mogelijke bestemmingen
- Op basis van die informatie gaat de klant zijn bestemming en data specifiëren (client-side activiteit)
- Die informatie wordt dan overgebracht naar de server van de organisator, waar het wordt gebruikt om een afgestemde webpagina te creëren (server-side activiteit)
- Deze webpagina wordt dan naar de klant zijn browser gestuurd

Security

Forms of Attack

Er zijn veel manieren waarop een computersysteem en zijn inhoud aangevallen kan worden via netwerkverbindingen. Velen maken gebruik van kwaadaardige software (**malware** genoemd).

- Malware kan overgedragen en uitgevoerd zijn op de computer zelf, of het kan de computer van een afstand aanvallen.

Software die wordt overgedragen op de computer zelf:

- **Virus**
    - Het is een software die de computer aanvalt door zichzelf in programma's te plaatsen die zich al op de computer bevinden
    - Wanneer het "gastprogramma" wordt uitgevoerd, wordt ook het virus uitgevoerd
        - Wanneer het virus wordt uitgevoerd, probeert het zichzelf ook te verplaatsen in andere programma's op de computer
- **Worm**
    - Dit is een onafhankelijk programma dat zichzelf verplaatst in een netwerk en een verblijf opneemt in de computer
    - Vandaar stuurt hij kopieën door van zichzelf naar andere computers
    - Het kan zichzelf dupliceren en meer schade aanrichten dan een virus

- **Trojan horse**
  - Een programma dat een computersysteem betreedt vermomd als een gewenst programma
  - Eens in de computer voert het Trojan horse bijkomende activiteiten uit die nadelige effecten kunnen hebben
  - Vaak bijlagen in e-mails
- **Spyware** (**sniffing** software)
  - Een software die informatie verzamelt over de activiteiten van de computer waarop het verblijft, de informatie wordt dan doorgegeven aan de stichter van de aanval
- **Phishing**
  - Dit is een techniek om informatie te vergaren door het simpelweg te vragen

Software die de computer van een afstand aanvalt:
- **Denial of service (DoS)**
  - Hierbij wordt een computer op een korte periode overladen met een groot aantal berichten
  - Dit wordt gerealiseerd doordat een aanvaller een software installeert op een aantal computers die berichten zullen produceren als er een signaal wordt gegeven
  - Na dit signaal overladen al deze computers (een **botnet** genoemd) het slachtoffer met berichten

Spam
- Ongewenste e-mails
- Wordt gebruikt voor phishing en Trojan Horses

Protection and Cures

Een eerste preventietechniek is om het verkeer, dat door een punt in het netwerk passeert, te filteren. Dit gebeurt meestal door een programma, een **firewall**.
- Deze kunnen uitgaande berichten met een bepaalde bestemming blokkeren of inkomende berichten van een kwaadaardige bron
- Aan een gateway (poort) blokkeert de firewall inkomende berichten die zich vermommen als berichten die zich in die regio bevinden
  - In de regio hebben alle berichten een gemeenschappelijk kenmerk (bevoorbeeld .de) en als dan een bericht wilt binnenkomen met een bericht met .de achter, houdt de firewall dit bericht tegen. Want het doet zich voor alsof het zich in die regio bevindt.
  - Zich vermommen als een ander noemen we **spoofing**
- Worden ook gebruikt om individuele computers te beschermen

Er zijn ook variaties van firewalls, ontworpen voor specifieke doelen:
- **Spam filters**: firewalls ontworpen om ongewenste mail te blokkeren

Een andere preventietechniek is de **proxyserver**. Dit is een software die zich gedraagt als een tussenpersoon tussen een client en een server met als doel de client te beschermen tegen vijandige acties van de server.

- De server heeft geen idee dat de proxyserver niet de echte client is, het is zelfs nooit op de hoogte van de eigenlijke client zijn bestaan
    - De server heeft zo nooit de kans om te leren over de intranets interne kenmerken
- De proxyserver staat ook in positie om alle berichten te filteren die van de server naar de client worden verstuurd

Een derde preventietechniek is **auditing software**.

- Een systeemadministrator kan een plotse stijging van het berichtenverkeer detecteren op verschillende locaties in het bereik van de administrator, hij kan de activiteiten van de firewalls monitoren, het patroon analyseren van aanvragen door de individuele computer om onregelmatigheden te detecteren

Antivirussoftware

- Wordt gebruikt om virussen te detecteren en te verwijderen
- Moet herhaaldelijk geüpdatet worden

Encryption

Soms is het doel van de aanvallen om het systeem te verstoren, maar soms is het doel ook om toegang te krijgen tot informatie. Traditioneel wordt informatie bewaard onder paswoorden, dit is echter niet altijd even effectief.
Wanneer informatie over netwerken wordt getransporteerd en wordt doorgestuurd door onbekende entiteiten kunnen we gebruik maken van **encrypties**.

- Dit kan gebruikt worden zodat, zelfs wanneer info in de verkeerde handen valt, de gecodeerde informatie toch vertrouwelijk zal blijven

Encryptietechnieken:

- **HTTPS**
    - Een beveiligde versie van HTTP
    - Wordt gebruikt door financiële instellingen om klanten te voorzien van een veilige toegang tot hun accounts
    - De ruggengraat van HTTPS is het protocolsysteem: **Secure Sockets Layer (SSL)**
        - Dit was eerst ontwikkeld door Netscape om te zorgen voor veilige communicatie tussen Web clients en servers
- **Public-key encryption**
    - Encryptiesystemen zijn zo ontwikkeld dat als ze de kennis hebben om een bericht te coderen, dit niet wil zeggen dat ze het ook kunnen decoderen
    - De **public key** wordt gebruikt om berichten te coderen, de **private key** is noodzakelijk om te decoderen
    - Probleem
        - Je hoopt dat de public key die gebruikt wordt, de juiste key is voor de bestemde partij

        - **Certificate authorities**: vertrouwde Internetsites, hebben als taak om een accurate lijst van partijen en hun public keys bij te houden
        - **Certificate**: een pakket met de partij's naam en de partij's key
    - **Authentication**: je wilt zeker zijn dat de auteur van het bericht, de partij is dat het zegt te zijn
        - De houder van de private key kan een bitpatroon produceren (**digital signature**), dat enkel de partij weet hoe te produceren
        - Door die signature toe te voegen aan het bericht, de zender kan het bericht markeren als authentiek

Legal Approaches to Network Security

Een andere manier om de veiligheid van een computer te verhogen is door het toepassen van legale remedies. Er zijn echter 2 obstakels bij deze benadering.

- Een actie illegaal maken, betekent niet dat het de actie uitsluit. Het voorziet enkel een legale toevlucht
- Het internationale van de netwerken maakt het moeilijk om een toevlucht te verkrijgen
    - Wat illegaal is in het ene land, kan legaal zijn in het andere

We gebruiken voorbeelden van de federale wetten van de Verenigde Staten.

- Computer Fraud and Abuse Act
    - Bestrijden de snelle toename van malware en de diefstal van informatie
    - De act heeft bewijs nodig dat de verweerder intentioneel schade heeft veroorzaakt
- ECPA: Electronic Communication Privacy Act
    - Het recht op privacy
    - Het is illegaal voor een Internetprovider om informatie vrij te geven over de communicatie van zijn klanten
    - Het is illegaal voor onbevoegd personeel om iemand anders af te luisteren
    - Het geeft enkele overheidsinstanties de toestemming om elektronische communicatie te monitoren onder bepaalde beperkingen

Het voorzien van controlerechten doet enkele technische problemen ontstaan.

- Om deze capaciteiten te voorzien, moet een communicatiesysteem ontworpen en geprogrammeerd worden, zodat de communicatie gemonitord kan worden.
- De clash tussen de overheid zijn rechten om communicatie te monitoren en het publiek recht om gebruik te maken van encryptie