

Das Internet – Einführung und Überblick

"Man kann die Menschen in zwei Kategorien einteilen.

Solche, die Zugang zum Netz haben, und solche, die keinen haben."

Harley Hahn: A Students Guide to Unix.

- 1. VORBEMERKUNGEN 2**
- 2. DIE GESCHICHTE DES INTERNETS..... 2**
 - 2.1. DAS ARPA-NET 2
 - 2.2. WISSENSCHAFTLICHE EINRICHTUNGEN 3
 - 2.3. DAS NETZ DER NETZE 4
 - 2.4. DIE ENTWICKLUNG IN EUROPA 5
- 3. DIE TECHNIK 6**
 - 3.1. DAS TCP/IP-PROTOKOLL 6
 - 3.2. IP-ADRESSIERUNG..... 6
 - 3.3. CLIENT-SERVER-TECHNOLOGIE 7
 - 3.4. FORM VON INTERNET-ADRESSEN 8
 - 3.5. DNS - DOMAIN NAME SERVICE..... 10
 - 3.6. ROUTING UND GATEWAYS..... 10
 - 3.7. SELBSTORGANISATION IM INTERNET 10
- 4. DAS WORLD WIDE WEB 11**
 - 4.1. DIE GESCHICHTE DES WORLD WIDE WEB..... 11
 - 4.2. NETSCAPE UND DER BOOM 12
 - 4.3. DIE ARBEIT DES W3-KONSORTIUMS 13
- 5. HTML ALS UNIVERSELLE WEB-SPRACHE 14**
 - 5.1. HTML ALS SOFTWARE-UNABHÄNGIGES KLARTEXTFORMAT 14
 - 5.2. HTML ALS AUSZEICHNUNGSSPRACHE 14
 - 5.3. HTML FÜR HYPERTEXT 14
 - 5.4. DEFINITIONEN 15
- 6. WEITERE DIENSTE..... 15**
 - 6.1. TELNET..... 15
 - 6.2. FTP 15
- 7. DER WEG INS INTERNET..... 16**
- 8. AKTUELLE DATEN 16**

1. Vorbemerkungen

Bei der täglichen Arbeit mit dem Internet sind die folgenden Hintergrundkenntnisse sicherlich hilfreich, sie ermöglichen eine Interpretation der Vorgänge und insbesondere von Problemen, außerdem sind sie notwendig, um aktuelle Entwicklungen und Diskussionen beurteilen zu können.

2. Die Geschichte des Internets

2.1. Das ARPA-Net

Die Ursprünge des heutigen Internet reichen in die **60er Jahre des letzten Jahrhunderts** zurück. Es war die Zeit des Kalten Krieges zwischen den beiden Weltmächten USA und UdSSR. Neue Impulse in der Elektronischen Datenverarbeitung (EDV) kamen in jener Zeit hauptsächlich **durch militärische Initiativen** zustande.

Mittlerweile gibt es zwar im Internet Proteste gegen die Auffassung, militärische Interessen hätten das Internet geboren. Das ist insofern berechtigt, als es keine unmittelbaren Ursachen-Wirkungs-Verhältnisse gibt. Doch die technologischen Ideen und Entwicklungen der Vorläufernetze kamen definitiv aus dem militärnahen Umfeld in den USA, und es ist deshalb auch nicht verkehrt, das so darzustellen.

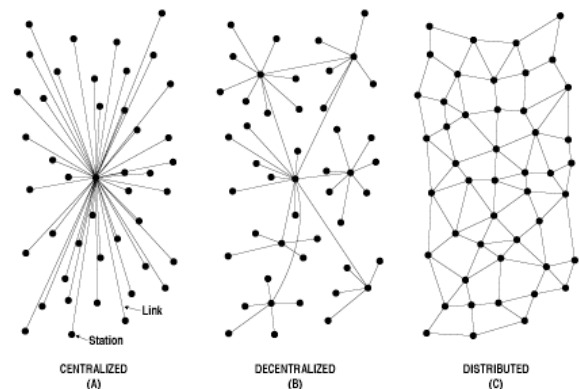
Im **Department of Defense**, dem amerikanischen Verteidigungsministerium, wurde seit den frühen 1960er Jahren überlegt, wie man wichtige militärische Daten besser schützen könnte. Selbst bei einem atomaren Angriff des Gegners sollten die Daten nicht zerstört werden können. Als Lösung kam nur ein elektronisches Datennetz in Frage. Die gleichen Daten sollten dabei auf mehreren, weit entfernten Rechnern abgelegt werden. Bei neuen oder geänderten Daten sollten sich alle angeschlossenen Rechner binnen kürzester Zeit den aktuellen Datenstand zusenden. Jeder Rechner sollte dabei über mehrere Wege mit jedem anderen Rechner kommunizieren können. So würde das Netz auch dann funktionieren, wenn ein einzelner Rechner oder eine bestimmte Leitung durch einen Angriff zerstört würde.

So gab die US Air Force **1964** bei der **RAND Corporation** das so genannte **"dezentrale Netzwerk"** in Auftrag, der „Erfinder“ dieses „distributed network“ war **Paul Baran**. Baran hatte auch die Idee, statt für die Übertragung einer Nachricht eine feste Leitung einzurichten, die Nachricht in viele kleine Pakete aufzuspalten und individuell zu versenden (Vorteile: Auswahl des individuell günstigsten Wegs; bei Verlust eines Paketes nicht Neuversenden der ganzen Nachricht erforderlich, sondern nur des betroffenen Pakets). Bei dem, was da entwickelt werden sollte, handelt es sich noch nicht um den direkten Vorläufer des heutigen Internet.

Nach vielem Hin und Her **scheiterte dieses Projekt**

auch, vor allem aus Kostengründen, und wurde nie realisiert. Die Idee des "dezentralen Netzwerks" mit der paketweisen Datenübertragung blieb jedoch in den Köpfen hängen.

Die **Advanced Research Projects Agency (ARPA)** war eine seit 1958 bestehende wissenschaftliche Einrichtung, die dem amerikanischen Verteidigungsministerium unterstellt war und deren Forschungsergebnisse in militärische Zwecke einfließen. Die Politik der ARPA bestand darin, keine eigenen Forschungseinrichtungen zu unterhalten, sondern mit den universitären und industriellen Vertragspartnern zu kooperieren und die von ihr finanzierten Projekte zu koordinieren.

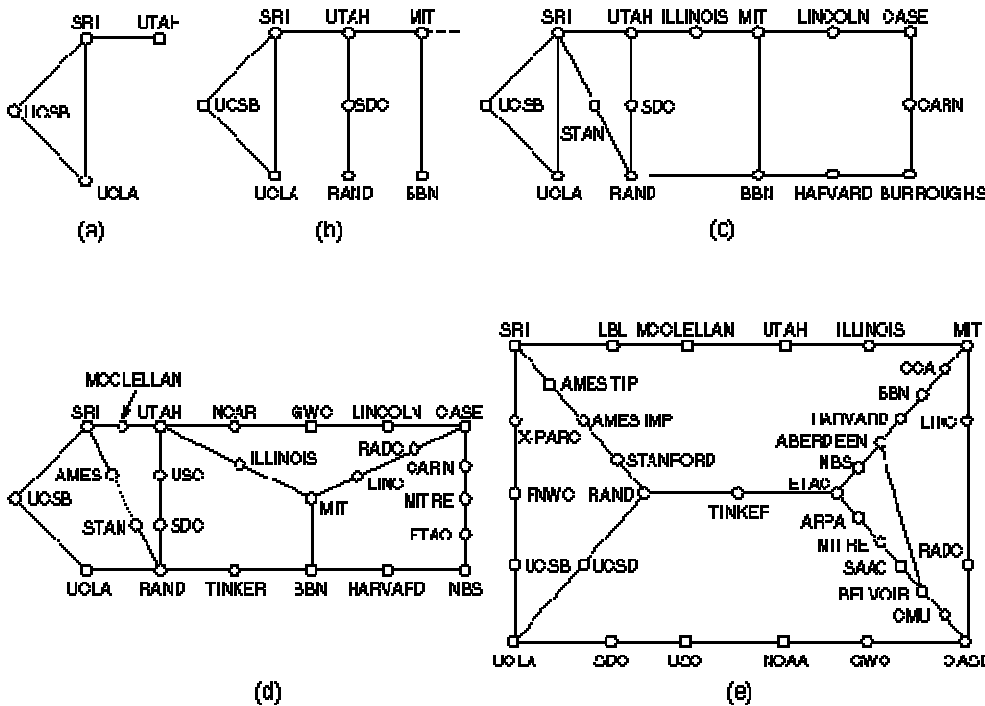


Vernetzungsmöglichkeiten; rechts: „distributed network“

Die ARPA entschloss sich 1966 zur **Vernetzung der ARPA-eigenen Großrechner**. Dabei wurde die Idee des "dezentralen Netzwerks" wieder aufgegriffen. 1968 bewilligte **Charles Hitzfeld**, der Leiter von ARPA für die Pilotarbeiten ein Budget von 500 000 \$, zwei Jahre später betrug das jährliche Budget bereits 2,2 Mio. \$.



Ende 1969 waren die ersten vier Rechner (Universität von Kalifornien in Santa Barbara, das Stanford Research Institute, die Universität von Utah, die Universität von Kalifornien in Los Angeles) an das ARPA-Net angeschlossen. Drei Jahre später waren es bereits 40 Rechner. In den ersten Jahren wurde das Netz deshalb auch **ARPA-Net** genannt.



Entwicklung des ARPA-Net

- a) Dezember 1969
- b) Juli 1970
- c) März 1971
- d) April 1971
- e) September 1972

(aus: A. S. Tanenbaum: Computer-network)

Den Durchbruch in der Aufmerksamkeit der Fachöffentlichkeit erzielte das ARPA-Net mit einer Präsentation bei der „International Conference on Computer Communications“ (ICCC) im Oktober 1972. Diese erste öffentliche Vorführung überzeugte die anwesenden Experten. In rascher Folge entstand eine Reihe von lokalen Universitätsnetzen, die sich zum großen Teil an das ARPA-Net anschlossen oder sich bei der Wahl der zugrunde gelegten Technologie am ARPA-Net orientierten. Die ARPA trug zur Verbreitung des Netzes bei, indem sie ihr Wissen an viele Universitäten, die Air Force, den Wetterdienst, ... weitergab.

2.2. Wissenschaftliche Einrichtungen

Das Prinzip der vernetzten Rechner war aber nicht nur für militärische Zwecke interessant. Man erkannte schnell, dass auch der akademische Betrieb vom ARPA-Net profitieren würde:

(1) Für **Wissenschaftler** war weniger das Synchronisieren von gleichen Daten auf mehreren Rechnern interessant, sondern vielmehr die Möglichkeit, Daten von einem anderen Rechner abzurufen (**resource sharing**). (2) Ein anderes Beispiel: Um bestimmte wissenschaftliche Probleme mit Hilfe von Computern zu lösen, waren/sind oft Rechner mit sehr großer Rechnerleistung und spezielle Programme erforderlich. Diese Ausstattungen waren/sind teuer und konnten/können nicht für jede Einrichtung einzeln beschafft werden. Oft hatte/hat ein ganzes Land lediglich einen dieser begehrten Computer. Über ein Rechnernetz war/ist es nun möglich, jedem Wissenschaftler von seinem Arbeitsplatz aus jederzeit Zugang zu solchen Großrechnern zu ermöglichen.

Wegen der offenen Architektur des ARPA-Net stand einer solchen Verwendung nichts im Wege. Wissenschaftler konnten von den **frühen 1970er** Jahren an Forschungsergebnisse anderer Institute über das ARPA-Net abrufen oder anderen angeschlossenen Instituten eigene Daten zur Verfügung stellen.

Die Anzahl der angeschlossenen Rechner stieg an. Es handelte sich um sehr **unterschiedliche Rechnerarten mit nicht kompatiblen Betriebssystemen und unterschiedlichem Netzzugang**. Großrechner **verschiedener Fabrikate**, Unix-Rechner und später auch Personal Computer drängten ins Netz. Einige hatten eine Standleitung, also eine ständige Internetverbindung, andere wählten sich über Telefon und Modem ein. Um die unterschiedlichen Voraussetzungen unter einen Hut zu bringen, entstand die Notwendigkeit, ein neues Datenübertragungsprotokoll für das Netz zu entwickeln. Das Protokoll sollte nicht an bestimmte Computersysteme, Übertragungswege oder -geschwindigkeiten gebunden sein. Aus den Bemühungen um ein solches Protokoll ging schließlich das **TCP/IP-Protokoll** hervor, das in den Jahren 1973/74 von den verantwortlichen Forschern des **NCC** (Network Control Center) und des **NIC** (Network Information Center) implementiert wurde. Datenübertragungen im ARPA-Net liefen nach Einführung von TCP/IP nach einem einheitlichen und standardisierten Schema ab, jeder Hostrechner eines Netzwerks konnte mit allen anderen Rechnern des Netzes kommunizieren.

Die Netzwerkanwendungen damals waren lediglich **ftp** (Datenaustausch), **telnet** (Durchführung von Sitzungen auf entfernten Rechnern) und Electronic Mail, wobei E-Mail ursprünglich gar nicht als Anwendung geplant war. Erst als 1971 zwei Programmierer sich Nachrichten und nicht nur Daten zukommen lassen wollten, entwickelten sie **Electronic Mail**.

Zu den wissenschaftlichen Einrichtungen gehörten natürlich auch **Studenten**. Die entdeckten das Netz auf ihre Weise. Eine Art „**black board**“ war ihr Wunsch, ein Nachrichtenbrett wie in Universitäten üblich, für Mitfahrgelegenheiten, Jobs, Wohnmöglichkeiten, Reisepartner oder die Möglichkeit, einfach nur zu diskutieren und zu labern. So entstand das **Usenet** („das ARPA-Net des armen Mannes“, denn es ermöglichte auch für Personen, die nicht in einem Projekt der ARPA beschäftigt waren, einen Netzzugang, sofern sie über einen UNIX-Computer und einen Telefonanschluss verfügten. → Das Usenet ist ein eigenständiges Netz.), die **Hauptader der heutigen Newsgroups**. Viele eingefleischte „Internetter“ halten die Newsgroups auch heute noch für den spannendsten und lebendigsten Teil des Internets.

2.3. *Das Netz der Netze*

Im Juli 1975 wurde die Verwaltung des ARPA-Net an die DCA (Defense Communications Agency) des US-Verteidigungsministeriums übergeben, um einen laufenden Betrieb zu gewährleisten. Allerdings erforderte der Anschluss der akademischen Welt ans Netz bald eine Trennung zwischen militärischem und zivilem Teil, da die Militärs ihre eigenen Interessen wahren wollten. So kam es, dass Anfang der 1980er Jahre ein neues militärisches Datennetz, das **Milnet**, vom ARPA-Net abgekoppelt wurde. Das **ARPA-Net** selbst wurde dem **wissenschaftlichen Betrieb** überlassen.

In diesem zivilen Teil des Netzes nahm die Anzahl der angeschlossenen Rechner im Laufe der 1980er Jahre sprunghaft zu. Eine wichtige Rolle spielte dabei die amerikanische **National Science Foundation (NSF)**. Diese Organisation schaffte ein Leitungs-Verbundsystem (auf Basis von Telefon-Standleitungen), das alle bedeutenden wissenschaftlichen Rechenzentren des Landes miteinander verband (NSFNET). Um immer mehr Institutionen anzuschließen und einem immer weiter zunehmenden Verkehr gerecht zu werden, wurde ein System, basierend auf einem **Backbone** („Rückgrat“; Der „Endverbraucher“ hat in der Regel keinen direkten Zugriff auf sie, sie dienen vielmehr der internen Vernetzung anderer Netzwerke) realisiert. An dieses Leitungs-Verbundsystem, das die großen Rechenzentren miteinander verband, konnten sich die kleineren Einzelnetze anschließen.

Mit Entwicklung des NSFNET bestanden nunmehr in den USA zwei nationale Netzwerke: Das NSF übernahm immer mehr die Aufgaben des ARPA-Net, das schließlich Ende 1989 vom Department of Defense aufgelöst wurde. Obwohl das ARPA-Net aufgehört hatte zu existieren, lebte die Technologie im NSFNET bis 1995 weiter. Mitte 1995 wurde der Backbone des NSFNET abgeschaltet, und kommerzielle Netzbetreiber übernahmen seine Funktion. Die Möglichkeiten des Internets wurden in den folgenden Jahren durch einige neue Dienste und Suchprogramme (Archie, Gopher, ...) ergänzt.

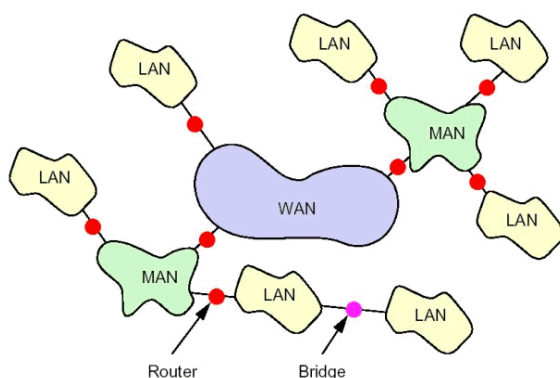
2.4. Die Entwicklung in Europa

In **Europa** gab es zeitgleich ähnliche Entwicklungen. Auf dem alten Kontinent setzte man jedoch zunächst auf **ISO-Normen**. Von dem amerikanischen TCP/IP-Modell, das nicht ISO-normiert war, wollte man deshalb nichts wissen. 1986 wurde die Organisation **RARE (Réseaux Associés pour la Recherche Européenne)** gegründet. Diese Organisation sollte alle Initiativen zur systemübergreifenden Rechnervernetzung europaweit koordinieren. RARE rief dazu ein Projekt namens **COSINE (Cooperation for an Open Systems Interconnection Networking in Europe)** ins Leben.

Der Siegeszug des TCP/IP-Protokolls war jedoch nicht mehr aufzuhalten. Unter dem Druck des Erfolgs in Amerika entstand schließlich ein **europäisches Datennetz**, das multiprotokollfähig war und unter anderem TCP/IP unterstützte. Dieses Netz lief zunächst unter der Bezeichnung **EuropaNET**. Verschiedene nationale wissenschaftliche Netzwerke, etwa das Deutsche Forschungsnetz (**DFN**), wurden daran angeschlossen.

Mittlerweile sorgt eine transatlantische Leitung für die Anbindung Europas an den Backbone in den USA. Die Koordination des Internets-Verkehrs innerhalb Europas wurde der Organisation **RIPE (Réseaux IP Européens)** übertragen. Auch in Europa gibt es seit 1992 ein Leitungs-Verbundsystem: dieses System nennt sich **Ebone (Europäischer Internet-Backbone)**. Auf anderen Kontinenten gab und gibt es vergleichbare Entwicklungen.

Wann begann also das Internet? Das ARPA-Net kann wohl noch nicht als „Internet“ bezeichnet werden, es handelte sich nur um ein – noch dazu geschlossenes – Netzwerk. Auch das NSFNET, das zwar über seinen Backbone die LANs verschiedener Universitäten verband, war noch kein richtiges „Internet“, es blieb ja auf (1) amerikanische (2) Hochschulen beschränkt.



Prinzip des Internets

[<http://www.netzmafia.de/skripten/internet/inetein1.html>]

Was wir also heute unter "Internet" verstehen, ist nicht ein einziges homogenes Netz, sondern ein **Verbund aus vielen kleinen, territorial oder organisatorisch begrenzten Netzen**. Diese Netze besitzen eine Anbindung an die Backbones und damit an das Gesamtnetz. Auch kommerzielle Internet-Provider hängen an entsprechenden Netzen. Und Endanwender, die ihren Internet-Zugang in Form einer Verbindung zu einem solchen Provider haben, können über diese Brücke das gesamte Internet nutzen.

Wöckel (2002) definiert das Internet als „die Summe aller über die weltweite Computervernetzung verfügbaren Dienste bzw. Nutzungsmöglichkeiten“.

3. Die Technik

3.1. Das TCP/IP-Protokoll

Damit zwischen Computern Daten ausgetauscht werden können, braucht es gewisse **Kommunikationsregeln**, an die sich die unterschiedlichen Computersysteme gleichermaßen halten.

Einen Satz solcher Kommunikationsregeln nennt man ein **Protokoll**. Das Protokoll für die Übertragung von Daten auf dem Internet heißt **TCP/IP**. Dieses Protokoll bildet die Grundlage, auf der alle anderen Protokolle basieren.

Entwickelt wurde das TCP von **Robert Kahn** und **Vinton Cerf**, es sorgte für den reibungslosen Paketaustausch und übernahm ursprünglich auch die Adressierung aller Internet-Rechner. Um 1980 wurde diese Funktion dann einem eigenen Protokoll, dem IP übergeben.

Erst durch dieses Protokoll wurde historisch gesehen aus einem begrenzten Netz ein Netz der Netze, das **Internet**. Denn jetzt erst wurde es den verschiedenartigsten Netzen möglich, untereinander zu kommunizieren, sofern alle Partner den TCP/IP-Standard verwenden.

Egal, ob man Web-Seiten aufruft, E-Mails versendet, per FTP Dateien downloadet oder mit Telnet auf einem entfernten Rechner arbeitet, stets werden die Daten auf die gleiche Weise adressiert und transportiert. TCP bedeutet **Transmission Control Protocol** (Protokoll für Übertragungskontrolle), IP bedeutet **Internet Protocol**.

Bei dem Versand einer E-Mail oder dem Aufruf einer HTML-Datei im Web werden die Daten bei der Übertragung im Netz in kleine Pakete zerstückelt. Jedes Paket enthält eine Angabe dazu, an welche Adresse es geschickt werden soll, und das wievielte Paket innerhalb der Sendung es ist.

Die Adressierung besorgt das IP. Dazu gibt es ein Adressierungsschema, die so genannten **IP-Adressen**. Damit die Datenpakete auch wirklich beim Empfänger ankommen, und zwar in der **richtigen Reihenfolge**, dafür sorgt das **TCP**. Das TCP verwendet Sequenznummern für die einzelnen Pakete einer Sendung. Erst wenn alle Pakete einer Sendung vollständig beim Empfänger angekommen sind, gilt die Übertragung der Daten als abgeschlossen.

Jeder Rechner, der am Internet teilnimmt, ist mit einer IP-Adresse im Netz angemeldet. Rechner, die ans Internet angeschlossen sind, werden als **Hosts** oder **Hostrechner** bezeichnet.

Beispiel: Wenn Sie mit Ihrem PC im Web surfen oder neue E-Mails abholen, sind Sie mit einer IP-Adresse im Internet angemeldet. Ihr Zugangs-Provider, über dessen Hostrechner Sie sich einwählen, kann feste IP-Adressen für Sie einrichten. Große Zugangs-Provider, etwa Online-Dienste wie T-Online oder AOL, vergeben auch personenunabhängig dynamische IP-Adressen für jede Internet-Einwahl (Erläuterung s.u.). Damit ein Rechner am Internet teilnehmen kann, muss er über eine Software verfügen, die das TCP/IP-Protokoll unterstützt. Unter MS Windows ist das beispielsweise die Datei *winsoc.dll* im Windows-Verzeichnis.

3.2. IP-Adressierung

Eine typische IP-Adresse sieht in Dezimalschreibweise so aus: *149.174.211.5* - vier Zahlen also, getrennt durch Punkte. Die Punkte haben die Aufgabe, über- und untergeordnete Netze anzusprechen. So wie zu einer Telefonnummer im weltweiten Telefonnetz eine Landeskennzahl, eine Ortsnetzkennzahl, eine Teilnehmerrufnummer und manchmal auch noch eine Durchwahlnummer gehört, gibt es auch im Internet eine Vorwahl - die **Netzwerknummer**, und eine Durchwahl - die **Hostnummer**.

Der **erste Teil** einer IP-Adresse ist die **Netzwerknummer**, der **zweite Teil** die **Hostnummer**. Wo die Grenze zwischen Netzwerknummer und Hostnummer liegt, bestimmt ein Klassifizierungsschema für Netztypen. Die folgende Tabelle verdeutlicht dieses Schema. In den Spalten für die IP-Adressierung und einem typischen Beispiel ist die Netzwerknummer (der Vorwahlteil) fett dargestellt. Der Rest der IP-Adresse ist die Hostnummer eines Rechners innerhalb dieses Netzes.

Die oberste Hierarchiestufe bilden die so genannten **Klasse-A-Netze**. Nur die **erste Zahl** einer IP-Adresse ist darin die Netzwerknummer, alle anderen Zahlen sind Hostnummern innerhalb des Netzwerks. Bei Netz-

| Netztyp | Schema | Typische IP-Adresse |
|---------------|-----------------|---------------------|
| Klasse-A-Netz | xxx.xxx.xxx.xxx | 103.234.123.87 |
| Klasse-B-Netz | xxx.xxx.xxx.xxx | 151.170.102.15 |
| Klasse-C-Netz | xxx.xxx.xxx.xxx | 196.23.155.113 |

werknummern solcher Netze sind Zahlen **zwischen 1 und 126** möglich, d.h. es kann **weltweit nur 126 Klasse-A-Netze** geben. Eine IP-Adresse, die zu einem Klasse-A-Netz gehört, ist also daran erkennbar, dass die erste Zahl zwischen 1 und 126 liegt. Das amerikanische Militärnetz ist beispielsweise so ein Klasse-A-Netz. Innerhalb eines Klasse-A-Netzes kann der entsprechende Netzbetreiber die zweite, dritte und vierte Zahl der einzelnen IP-Adressen seiner Netzteilnehmer frei vergeben. Da alle drei Zahlen Werte von 0 bis 255 haben können, kann ein Klasse-A-Netzbetreiber also bis zu 16,7 Millionen IP-Adressen an Host-Rechner innerhalb seines Netzes vergeben.

Die zweithöchste Hierarchiestufe sind die **Klasse-B-Netze**. Die Netzwerknummer solcher Netze erstreckt sich über die beiden ersten Zahlen der IP-Adresse. Bei der **ersten** Zahl können Klasse-B-Netze Werte **zwischen 128 und 191** haben. Eine IP-Adresse, die zu einem Klasse-B-Netz gehört, ist also daran erkennbar, dass die erste Zahl zwischen 128 und 191 liegt. Bei der zweiten sind Zahl Werte zwischen 0 und 255 erlaubt. Dadurch sind etwa 16.000 solcher Netze möglich. Da die Zahlen drei und vier in solchen Netzen ebenfalls Werte zwischen 0 und 255 haben dürfen, können an jedem Klasse-B-Netz bis zu ca. 65.000 Host-rechner angeschlossen werden. Klasse-B-Netze werden vor allem an große Firmen, Universitäten und Online-Dienste vergeben.

Die unterste Hierarchie stellen die **Klasse-C-Netze** dar. Die **erste** Zahl einer IP-Adresse eines Klasse-C-Netzes liegt **zwischen 192 und 223**. Die Zahlen zwei und drei gehören ebenfalls noch zur Netzwerknummer. Über zwei Millionen solcher Netze sind dadurch adressierbar. Vor allem an kleine und mittlere Unternehmen mit direkter Internet-Verbindung, auch an kleinere Internet-Provider, werden solche Adressen vergeben. Da nur noch eine Zahl mit Werten zwischen 0 und 255 übrig bleibt, können in einem C-Netz maximal 255 Host-Rechner angeschlossen werden. Eine Zahl davon ist reserviert, also bleiben 254 mögliche Host-Rechner übrig.

Durch die **Vergabe dynamischer IP-Adressen** pro Einwahl können Netzbetreiber die Anzahl der tatsächlich internetfähigen Anschlüsse deutlich höher halten, als wenn wirklich nur jeder Rechner eine feste Adresse erhalten würde, egal ob er online ist oder nicht. Ob dieses Adressierungs-Schema jedoch den Anforderungen der Zukunft noch gerecht wird, bezweifeln manche. Es gibt bereits Ideen zu einer Neustrukturierung der Adressierung von Netzen und Hostrechnern – Aktuell ist es aber nicht mehr als eine Idee!

3.3. Client-Server-Technologie



Wer als Betreiber eines Computers (Hostrechner) einen Teil seiner Daten anderen Benutzern zur Verfügung stellen will, braucht dazu ein spezielles Programm, das immer dann automatisch aktiv wird, wenn jemand ein Dokument über das Internet beziehen möchte. Ein solches Programm nennt man einen **Server**. Ein Internet-Server sollte immer laufen und der Computer, auf dem es läuft, sollte rund um die Uhr am Internet angeschlossen sein. Außerdem darf keine schützenden Software (Firewall) den Zugriff von außen verhindern bzw. einschränken. Server sind Programme, die permanent darauf warten, dass eine Anfrage eintrifft, die ihren Dienst betreffen. So wartet etwa ein Web-Server darauf, dass Anfragen eintreffen, die Web-Seiten auf dem Server-Rechner abrufen wollen.

Wer umgekehrt einen entfernten Computer anweisen will, er möge ihm ein Dokument über das Internet schicken, braucht dazu auch ein spezielles Programm (**Client**). Ein Web-Browser ist beispielsweise ein Client.

Beispiel: Wenn Sie auf einen Verweis klicken, der zu einer HTTP-Adresse führt (Link), startet der Browser, also der Client, eine Anfrage an den entsprechenden Server auf dem entfernten Hostrechner. Der Server wertet die Anfrage aus und sendet die gewünschten Daten.

Bei Server und Client handelt es sich um Software, deshalb lassen sich verschiedene Server und Clients auf einem Computer installieren.

Damit Server und Client kommunizieren können, braucht es zusätzlich zum TCP/IP-Protokoll noch ein spezielles Protokoll. Dieses Protokoll ist je nach Art der zu übertragenden Dateien ein anderes. Es gibt also für jede Art von Dateien spezielle Server (e-mail-Server, WWW-Server, ...), spezielle Clients und ein spezielles Protokoll. Eine solche Server-Client-Protokoll-Kombination nennt man einen **Internet-Dienst**.

Beispiele für Internet-Dienste sind: World-Wide-Web (WWW) → Hypertext Transfer Protocol (HTTP); E-Mail (senden) → Simple Mail Transfer Protocol (SMTP); E-Mail (empfangen) → Post Office Protocol (POP); ...

Hinweis: Das WWW ist ein Internet-Dienst unter anderen, allerdings sind im WWW-Dienst inzwischen (fast) alle anderen Dienste des Internets integriert → Man braucht also nur noch einen WWW-Client, einen so genannten **Browser**, und kann damit auf (fast) alle anderen Internet-Dienste zugreifen.

Dass ein Client Daten anfordert und ein Server die Daten sendet, ist der Normalfall. Es gibt jedoch auch „Ausnahmen“. So kann ein Client nicht nur Daten anfordern, sondern auch Daten an einen Server schicken: zum Beispiel, wenn man eine Datei per FTP auf den Server-Rechner hochlädt, eine E-Mail versendet oder im Web ein Formular ausfüllt und abschickt. Bei diesen Fällen redet man auch von **Client-Push** („Client drängt dem Server Daten auf“).

Ein anderer Ausnahmefall ist es, wenn der Server zuerst aktiv wird und dem Client etwas ohne dessen Anforderung zuschickt. Das nennt man **Server-Push** („Server drängt dem Client Daten auf“). Einige Technologien wollten diesen Ausnahmefall vor einigen Jahren zu einer Regel erheben mit den so genannten **Push-Technologien**. Diese Technologien sollten ermöglichen, dass ein Client regelmäßig Daten empfangen kann, ohne diese eigens anzufordern – durchsetzen konnte sich die Push-Technologie nicht wirklich. Ein Beispiel, das zeigt, dass es sie vereinzelt gibt: Wenn Sie beispielsweise die Website eines Sporttickers beim Start eines Formel1-Rennens aufsuchen (ARD, Bild, Sport1...) wird immer, wenn es Neues zu vermelden gibt, dieses - ohne Ihr Zutun - an Ihre aufgerufene Seite geschickt.

3.4. Form von Internet-Adressen

Obwohl es im Internet verschiedene Dienste gibt, genügt i.d.R. ein WWW-Browser, um damit auf alle Dienste zuzugreifen (s.o.). Dies ermöglicht eine standardisierte Form von Internet-Adressen für verschiedene Dienste, der **Unified Resource Locator** (URL). Adressen, die in einen WWW-Browser eingegeben werden, müssen immer diese Form aufweisen.

Ein URL kann einen Server, ein Verzeichnis oder ein einzelnes Dokument bezeichnen, je nachdem, welche Elemente vorhanden sind, obligatorische Elemente eines URL sind Protokoll://Serveradresse, alle anderen Elemente sind fakultativ.

Praxistipp: Ein URL muss immer ohne Leerzeichen und ohne Zeilenschaltungen eingegeben werden.

Ein URL besteht auf folgenden Teilen,
dargestellt am Beispiel der WWW-Adresse <http://www.uni-regensburg.de/psychologie/index.html>

| | |
|-----------------------------|--------------------|
| Protokoll:// | http:// |
| subdomain.secondleveldomain | www.uni-regensburg |
| .Top-Level-Domain | .de |
| Port.no | |
| /Pfad(e) (Ordner) | /psychologie |
| /Datei | /index.html |

Als Beispiel für den Aufbau einer E-Mail-Adresse (keine URL!), wie sie für den elektronischen Briefverkehr benötigt wird, der Vergleich einer E-Mail-Adresse mit einer konventionellen Adresse:

max.mueller@mail.fakultaet.uni-regensburg.de

| | |
|---------------|--|
| Max Müller | entspricht dem Benutzerpseudonym (max.mueller) |
| bei Huber | entspricht dem Server (mail) |
| Beispielweg 5 | entspricht der (Sub-) Domain (fakultaet) |
| 81234 München | entspricht der (Sub-) Domain (uni-regensburg) |
| Germany | entspricht der (Toplevel-) Domain (de) |

Die **Top-Level-Domain** bezeichnet das Land, in dem der Server steht, oder die Art der Institution. Es gibt auch länderunabhängige Top-Level-Domains, z.B. „.com“ für commercial, „.org“ für organisation, „.net“ für network.

Eine Namensadresse ist ähnlich wie die IP-Adressen hierarchisch aufgebaut. Eine Namensadresse (Domain-Name) in diesem System gehört zu einer **Top-Level-Domain**. Die einzelnen Teile solcher Namensadressen sind wie bei IP-Adressen durch Punkte voneinander getrennt. Namensadressen (Domains) sind beispielsweise *yahoo.com*, *mozilla.org* oder *medpaed.de*.

Top-Level-Domains stehen in so einem Domain-Namen an letzter Stelle. Es handelt sich um einigermaßen sprechende Abkürzungen. Die Abkürzungen, die solche Top-Level-Domains bezeichnen, sind entweder Landeskennungen oder Typenkennungen. Beispiele sind:

- | | |
|-------------------------|---|
| <i>de</i> = Deutschland | <i>edu</i> = amerikanische Hochschulen |
| <i>at</i> = Österreich | <i>gov</i> = amerikanische Behörden |
| <i>it</i> = Italien | <i>mil</i> = amerikanische Militäreinrichtungen |
| <i>my</i> = Malaysia | <i>biz</i> = Unternehmen |
| | <i>pro</i> = Berufsgruppen mit Werbeverbot (Anwälte, Steuerberater, Ärzte usw.) |
| | <i>name</i> = Privatpersonen |
| | <i>info</i> = Informationsdienste aller Art |
| | <i>museum</i> = Museen |
| | <i>aero</i> = Flugunternehmen, Flughäfen, Reiseveranstalter usw. |
| | <i>coop</i> = Genossenschaften, Verbände, Organisationen |

Jede Top-Level-Domain stellt einen Verwaltungsbereich dar, für den es eine „Verwaltungsbehörde“ gibt, die für die Namensvergabe von Domains innerhalb ihres Verwaltungsbereichs zuständig ist, in Deutschland ist das **DENIC** (Deutsches Network Information Center) für die Vergabe von Domain-Namen zuständig. Inhaber von Domain-Namen können nochmals Sub-Level-Domains vergeben. So gibt es beispielsweise eine Domain namens *medpaed.de*. Die Betreiberin dieser Domain hat nochmals Sub-Domains vergeben, wodurch Domain-Adressen wie *moodle.medpaed.de* oder *mathewiki.medpaed.de* entstanden

3.5. DNS - Domain Name Service

Computer können mit Zahlen besser umgehen, Menschen in der Regel besser mit Namen. Deshalb hat man ein System ersonnen, das die numerischen IP-Adressen für die Endanwender in anschauliche Namensadressen übersetzt bzw. die Domain-Namen in eine Zahlenfolge (IP-Adresse) übersetzt, bevor die Anfrage an den Server geschickt wird.

3.6. Routing und Gateways

Im Internet ist es zunächst nur innerhalb des eigenen Sub-Netzes möglich, Daten direkt von einer IP-Adresse zu einer anderen zu schicken. In allen anderen Fällen, wenn die Daten an eine andere Netzwerknummer geschickt werden sollen, treten Rechner auf den Plan, die den Verkehr zwischen den Netzen regeln. Solche Rechner werden als **Gateways** bezeichnet. Diese Rechner leiten Daten von Hostrechnern aus dem eigenen Sub-Netz an Gateways in anderen Sub-Netzen weiter und ankommende Daten von Gateways anderer Sub-Netze an die darin adressierten Host-Rechner im eigenen Sub-Netz. Ohne Gateways gäbe es gar kein Internet.

Das Weiterleiten der Daten zwischen Sub-Netzen wird als **Routing** bezeichnet. Die Beschreibungen der möglichen Routen vom eigenen Netzwerk zu anderen Netzwerken sind in **Routing-Tabellen** auf den Gateway-Rechnern festgehalten.

Zu den Aufgaben eines Gateways gehört auch, eine Alternativ-Route zu finden, wenn die übliche Route nicht funktioniert, etwa, weil bei der entsprechenden Leitung eine Störung oder ein Datenstau aufgetreten ist. Gateways senden sich ständig Testpakete zu, um das Funktionieren der Verbindung zu testen und für Datentransfers „verkehrsarme“ Wege zu finden.

Wenn also im Internet ein Datentransfer stattfindet, ist keinesfalls von vorneherein klar, welchen Weg die Daten nehmen. Sogar einzelne Pakete einer einzigen Sendung können völlig unterschiedliche Wege nehmen. Wenn man beispielsweise von Deutschland aus eine Web-Seite aufruft, die auf einem Rechner in den USA liegt, kann es sein, dass die Hälfte der Seite über den Atlantik kommt und die andere über den Pazifik, bevor der Web-Browser sie anzeigen kann. Davon bekommt man allerdings nichts mit.

3.7. Selbstorganisation im Internet

In Anbetracht der Teilnehmerzahl im Internet ist der Verwaltungsaufwand im Netz vergleichsweise klein. Die meisten Endanwender wissen nicht einmal, dass es solche Verwaltungsstellen überhaupt gibt.

Eine **gesetzgeberische Institution**, wie es sie etwa innerhalb des Verfassungsbereichs eines Staates gibt, gibt es im Internet als weltweitem Verbund **nicht**. Der Versuch, die **ICANN**-Behörde als oberste Vergabestelle für Domains und IPs zu etablieren, hat viel Staub aufgewirbelt, aber die Befürchtungen, dass da eine zentralistische Internet-Diktatur entsteht, haben sich bislang nicht bestätigt. Viele Bereiche im Internet beruhen faktisch auf Selbstorganisation. Bei Diensten wie E-Mail gilt beispielsweise das stille Abkommen, dass jeder beteiligte Gateway alle Mails weiterleitet, auch wenn weder Sender noch Empfänger dem eigenen Sub-Netz angehören. Die Kosten trägt zwar der Netzbetreiber, aber da alle ein Interesse am weltweiten Funktionieren des Netzes haben, ist jeder bereit, die anfallenden Mehrkosten zu tragen.

Die Funktionsfähigkeit des Internets basiert also auf der Bereitschaft der Beteiligten, keine Pfennigfuchserie zu betreiben.

Das **Usenet**, also der größte Teil der Newsgroups, organisiert sich sogar vollständig selbst. Die „Verwaltung“ findet im Usenet in speziellen Newsgroups statt (solchen, die mit *news.* beginnen). Dort können beispielsweise Vorschläge für neue Gruppen eingebracht werden, und in Abstimmungen wird darüber entschieden, ob eine Gruppe eingerichtet oder abgeschafft wird.

Offizielle Anlaufstellen gibt es für die **Vergabe** von **Netzwerkadressen** und für **Namensadressen** (DNS). Für die Vergabe von IP-Adressen innerhalb eines Netzwerks ist der jeweilige Netzbetreiber verantwortlich. Dazu kommen Organisationen, die sich um Standards innerhalb des Internets kümmern.

Die Kosten für die Datenübertragungen im Internet tragen die Betreiber der Sub-Netze. Diese Kosten pflanzen sich nach unten fort zu Providern innerhalb der Sub-Netze bis hin zu Endanwendern, die über Provider Zugang zum Internet haben oder Internet-Services wie eigene Web-Seiten nutzen.

4. Das World Wide Web

Das World Wide Web (WWW) ist der jüngste Dienst innerhalb des Internets. Das Web zeichnet sich dadurch aus, dass es auch ungeübteren Anwendern erlaubt, sich im Informationsangebot zu bewegen. Wer etwa mit einem FTP-Programm einen FTP-Server aufruft, muss sich in komplexen, unbekanntem Verzeichnisstrukturen zurechtfinden und sich an Dateinamen orientieren. Interessante Dateien kann er auf seinen Rechner downloaden, um sie später zu öffnen. Im WWW dagegen erscheinen Informationen gleich beim Aufruf am Bildschirm. Wenn man mit einem Browser im Web unterwegs ist, braucht man sich nicht um Dateinamen oder um komplizierte Eingabebefehle zu kümmern. Deshalb kann man auch vom „Surfen im Netz“ sprechen.

4.1. Die Geschichte des World Wide Web



Der gebürtige Brite **Tim Berners-Lee**, in den 1980er und bis in die 1990er Jahren Informatiker am Genfer Hochenergieforschungszentrum **CERN**, beschreibt in seinem Buch „Weaving the Web“, wie er dort am CERN in den 1980er Jahren ein selbstgeschriebenes Programm als Produktivitäts-Tool nutzte. Das Programm, von Berners-Lee in Pascal geschrieben, lief auf einem Computersystem mit dem Betriebssystem „Norsk Data SYNTRAN III“. Die 8-Zoll-Diskette, auf der die Quellen gespeichert waren, gingen irgendwann verloren.

Mit heutiger Software hatte das Tool sicher nicht viel zu tun. Grafische Oberflächen gab es nicht, kommandozeilenorientierte Systeme dominierten, und grüne Schrift auf schwarzem Grund beherrschte die Bildschirme. Auf einem solchen System hatte Tim Berners-Lee sein Pascal-Programm zum Laufen gebracht. Er nannte das Programm **Enquire**, was auf Deutsch so viel bedeutet wie „sich erkundigen“.

„Enquire“ war ein Hypertext-Programm. Man konnte Textdateien editieren, die - vermutlich durch irgendwelche Steuerzeichenfolgen markiert - in „nodes“ (**Knoten**) unterteilt waren. Ein Knoten konnte alles sein, was so an Daten anfiel - Adressen, Gesprächsnotizen, spontane Ideen, Erlebnisse, Arbeitsergebnisse. Zu jedem Knoten gab es eine zugehörige Liste mit Links zu anderen Knoten. Man konnte auf jegliche Art von Querbeziehung linken, die man kannte oder fand. Links innerhalb auf Ziele innerhalb einer Datei wurden vom Enquire-Programm automatisch bidirektional dargestellt - das heißt, auch wenn der Link nur von A nach B gesetzt war, fand man bei B in der Liste den Rückverweis auf A.

Ende 1988 entschloss sich Berners-Lee, aus der Weiterentwicklung von „Enquire“ ein computerübergreifendes System zu entwickeln. Für CERN sollte er zu diesem Zweck einen entsprechenden Projektvorschlag einreichen und darin sein Vorhaben detailliert beschreiben. Das war die Zeit, als Berners-Lee mit verschiedenen Fachleuten Kontakt aufnahm, um technische Wege der Realisierung seiner Ideen zu finden. Ben Segal, ein Kollege von Berners-Lee, der sich mit den Grundlagen des Internets auskannte, überzeugte den Hypertext-Entwickler von den einzigartigen Möglichkeiten des Netzes der Netze. 1989 reichte er seinen

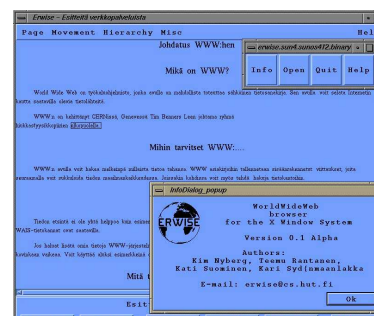
Vorschlag bei CERN ein. Es vergingen jedoch viele Monate, und Berners-Lee erhielt keine Reaktion auf den Vorschlag. Robert Cailliau, ein alter Bekannter, ließ sich von den Ideen des Briten begeistern und setzte sich fortan für die Forcierung des Projekts am CERN ein. In dieser Zeit - es war das Jahr 1990 - erhielt das Projekt auch seinen endgültigen Namen: **World Wide Web**. Im Herbst des Jahres 1990 schrieb Berners-Lee eigenhändig die ersten Versionen der drei Säulen seines Konzepts:

1. die Spezifikation für die Kommunikation zwischen Web-Clients und Web-Servern - das so genannte **HTTP-Protokoll** (*HTTP = Hypertext Transfer Protocol*)
2. die Spezifikation für die Adressierung beliebiger Dateien und Datenquellen im Web und im übrigen Internet - das Schema der so genannte **URIs** (*URI = Universal Resource Identifier, universeller Quellenbezeichner*).
3. die Spezifikation einer Auszeichnungssprache für Web-Dokumente, der Berners-Lee den Namen **HTML** gab (*HTML = Hypertext Markup Language, Hypertext Auszeichnungssprache*).

Berners-Lee schrieb auch die erste Web-Server-Software. Der Rechner, auf dem diese Software installiert wurde, war unter dem Namen *info.cern.ch* erreichbar. Dort stellte Berners-Lee an Weihnachten 1990 die ersten, in HTML geschriebenen Web-Seiten der Welt zur Verfügung. Seine Ideen wichen dabei ursprünglich durchaus von dem ab, was schließlich aus dem Web wurde. So setzte sich Berners-Lee immer dafür ein, Web-Seiten online editierbar zu machen, so dass Web-Seiten-Besucher Texte fortschreiben konnten, sofern der Anbieter entsprechende öffentliche Schreibrechte für die Dateien vergab. Doch die Web-Browser, die sich schließlich durchsetzten, waren reine Lese-Software. Die Möglichkeiten von Web 2.0 gehen nun in die Richtung, wie Berners-Lee es sich vorgestellt haben könnte.

Das Web entwickelte sich zunächst nicht unbedingt von allein. Berners-Lee und seine Projektmitstreiter waren unermüdlich bei der Arbeit, das Web bekannt zu machen und zu etablieren. Auf einer internationalen **Hypertext-Konferenz** im Jahre 1991 stellten sie das Projekt vor. Kontakte zu Programmierern für verschiedene Systeme entstanden, und so entstanden die ersten Web-Browser. Denn erst die Verfügbarkeit solcher **Browser** ermöglichte es anderen Menschen mit Internet-Zugang, Web-Seiten aufzurufen.

Nicola Pellow, eine junge Mitarbeiterin und Mathematikerin am CERN, schrieb den ersten einigermaßen benutzbaren textmodus-orientierten Browser. 1992 entstanden auch die ersten Browser für grafische Benutzeroberflächen, mit den Namen **Erwise** und **Viola**. Gleichzeitig stieg die Zahl von Web-Servern, über die Web-Seiten angeboten wurden. Zunächst waren es vor allem wissenschaftliche Institutionen.



Erwise [<http://www.nic.funet.fi/index/FUNET/history/internet/se/1992.html>]

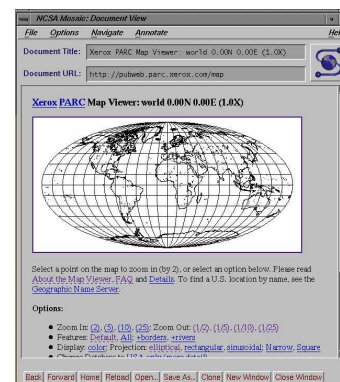
4.2. Netscape und der Boom



Für programmierende Studenten wurde es zunehmend zu einer interessanten Herausforderung, Browser fürs Web zu schreiben. Besonders hervor tat sich dabei ein junger Mann namens **Marc Andreessen**. Er entwickelte einen Web-Browser für grafische Benutzeroberflächen namens **Mosaic**. Tim Berners-Lee, der Marc Andreessen auch selber kennen lernte, beschreibt ihn als jemanden, der nicht mehr wie alle Browser-Entwickler vor ihm eine Software fürs Web entwickeln wollte, sondern als jemanden, der mit dem

Anspruch auftrat, das Web überhaupt erst verfügbar zu machen - eben mit seinem Browser.

Andreessen und seine angeheuerten Mit-Programmierer hatten mit den philosophischen Ideen von Berners-Lee nicht sehr viel am Hut, sie dachten von Beginn an in kommerziellen



Mosaic [<http://www.nic.funet.fi/index/FUNET/history/internet/se/1993.html>]

Kategorien. Bei der Entwicklung des Mosaic-Browsers erfanden die Programmierer nach ihrem Belieben neue Features, die sie in den Web-Browser implementierten, ohne dass diese Features standardisiert waren. Durch den Mosaic-Browser kam die Lawine Web erst richtig ins Rollen.

Andreessen stieg schließlich aus dem Mosaic-Projekt aus und wurde Mitbegründer einer neuen Firma für Web-Software: **Netscape**. Das Konzept von Netscape ging zunächst voll auf. Eine völlig neue Zunft entstand: die **Web-Designer**. Dank Netscape konnten Webseiten jetzt bunte Hintergrund- und Schriftfarben, Hintergrundtapeten, Tabellenlayouts, mehrgeteilte Bildschirmfenster (Frames) und Multimedia-Plugins enthalten. Zigtausende von Privat Anwendern begannen, eigene Homepages zu erstellen. Nach und nach drängten alle Firmen, Organisationen, Regierungen und Behörden mit eigenen Angeboten ins Web - weltweit. Die Wachstumsraten glichen einer Explosion.

In den Jahren 1995 und 1996 erreichte der Netscape-Browser unter den Web-Benutzern zeitweise einen Marktanteil von 90%. Mitte der 90er Jahre, als der Online-Boom allmählich breitere Schichten der Bevölkerung der westlichen Länder erreichte, setzte man bei Microsoft zunächst auf einen eigenen, proprietären Online-Dienst namens MSN. Vom Internet und dem Web behauptete Bill Gates, das sei nichts für Microsoft. Als man jedoch sah, wie Netscape binnen weniger Monate die EDV-Landschaft veränderte, begann man plötzlich hastig zu reagieren. Innerhalb von etwas mehr als einem Jahr erschienen die ersten vier Versionen des **MS Internet Explorers**, wobei Microsoft im Frühjahr 1997 mit der 4er-Version ein ähnlicher technologischer Durchbruch gelang wie Netscape mit den 1995 erschienenen Versionen 1.1 und 2.0 seines Browsers. Microsoft hat heute im Browser-Markt eine ca. 80-90%-Machtstellung (Quelle: <http://www.onestat.com>).

4.3. Die Arbeit des W3-Konsortiums

Die Ideen von Tim Berners-Lee werden mittlerweile durch ein vielköpfiges Konsortium vertreten: dem **W3-Konsortium (W3C)**. Dieses Konsortium ist trotz der millionenfach verbreiteten Browser der mächtigste Faktor für die Weiterentwicklung des Web geworden. Nicht zuletzt deshalb, weil es nicht gegen die Marktinteressen der Software-Firmen gerichtet ist, sondern seine Mitglieder zu großen Teilen aus eben jenen Software-Firmen rekrutiert.

Ende 1994 traf ein erstes Beratungs-Komitee zusammen, um die Grundlagen für eine Institution zu schaffen, die sich um die technischen Grundlagen und Standards im Web kümmern sollte. Das CERN war mit dieser Aufgabe, die nicht seinem eigentlichen Beschäftigungsgegenstand entsprach, einfach überfordert.

Im Sommer 1995 traten namhafte Firmen dem W3-Konsortium bei. Gleichzeitig eröffnete die europäische Präsenz des Konsortiums ihre Pforten. Mitglieder des W3-Konsortiums sind Firmen oder Organisationen, keine Einzelpersonen. Sie unterschreiben einen 3-Jahres-Vertrag und zahlen Mitgliederbeiträge, aus denen sich das Gremium finanziert. Als Gegenleistung erhalten sie Zugang zu nichtöffentlichen Informationen und können an der Entwicklung der vom W3C betreuten Standards wie HTML, CSS, XML usw. mitwirken.

Die Arbeit des W3-Konsortiums unterteilt sich in so genannte **Aktivitäten (Activities)**. Es gibt mehrere Dutzend solcher Aktivitäten. So gibt es Aktivitäten für HTML, XML, CSS usw. Für jede der Aktivitäten gibt es **Arbeitsgruppen (Working Groups)** und **Interessensgruppen (Interest Groups)**. Während die Interessensgruppen eher einflussnehmenden Charakter haben, befassen sich die Arbeitsgruppen mit der eigentlichen Ausarbeitung der Inhalte. Beide Gruppenarten setzen sich aus Mitgliedern des W3C zusammen. In den Arbeitsgruppen sitzen also auch viele Vertreter von Software-Herstellern. Das unabhängige Kern-Team des W3-Konsortiums überwacht die Aktivitäten.

5. HTML als universelle Web-Sprache

HTML ist eine Sprache, die leicht zu erlernen, standardisiert und wegen ihrer weiten Verbreitung der Web-Browser praktisch überall verfügbar ist. Vom Web-Gründer Tim Berners-Lee entwickelt, wurde HTML im Zuge des Web-Booms zum erfolgreichsten und verbreitetsten Dateiformat der Welt. HTML ist eine Sprache zur Strukturierung von Texten, wobei aber auch die Möglichkeit besteht, Grafiken und multimediale Inhalte in Form einer Referenz einzubinden und in den Text zu integrieren.

Mit HTML können

- Überschriften, Textabsätze, Listen und Tabellen erzeugt werden,
- anklickbare Verweise auf beliebige andere Web-Seiten oder Datenquellen im Internet erzeugt werden,
- nicht-textuelle Inhalte referenziert werden,
- Formulare in den Text integriert werden.

HTML bietet Schnittstellen für Erweiterungssprachen (CSS Stylesheets, JavaScript, ...) zur Gestaltung von HTML-Elementen nach Wunsch oder zur Realisierung von Interaktion mit dem Anwender.

Aber: HTML allein kann weder Grafik-Designer befriedigen, noch Daten-Designer, die aus der Welt der relationalen Datenbanken kommen und sich anwendungsspezifische Lösungen wünschen. Deshalb gibt es heute Style-Sprachen wie CSS, und es gibt Lösungen wie XML, um anwendungsspezifisches Daten-Design zu ermöglichen.

5.1. *HTML als software-unabhängiges Klartextformat*

HTML ist ein so genanntes Klartext-Format. HTML-Dateien können mit jedem beliebigen Texteditor bearbeitet werden, der Daten als reine Textdateien abspeichern kann. Da HTML ein Klartextformat ist, lässt es sich auch hervorragend mit Hilfe von Programmen generieren. Von dieser Möglichkeit machen beispielsweise CGI-Scripts Gebrauch. Wenn man im Web zum Beispiel einen Suchdienst benutzt und nach einer Suchanfrage die Ergebnisse präsentiert bekommt, dann ist das, was am Bildschirm zu sehen ist, HTML-Code, der von einem Script generiert wurde.

5.2. *HTML als Auszeichnungssprache*

HTML bedeutet HyperText Markup Language. Es handelt sich dabei um eine Sprache, die mit Hilfe von SGML (Standard Generalized Markup Language) definiert wird. SGML ist als ISO-Norm 8879 festgeschrieben. HTML ist eine so genannte Auszeichnungssprache (Markup Language). Sie hat die Aufgabe, die logischen Bestandteile eines textorientierten Dokuments zu beschreiben. Als Auszeichnungssprache bietet HTML daher die Möglichkeit an, typische Elemente eines textorientierten Dokuments, wie Überschriften, Textabsätze, Listen, Tabellen oder Grafikreferenzen, als solche auszuzeichnen.

Das Auszeichnungsschema von HTML geht von einer hierarchischen Gliederung aus. HTML zeichnet Inhalte von Dokumenten aus. Dokumente haben globale Eigenschaften wie zum Beispiel Kopfdaten. Der eigentliche Inhalt besteht aus Elementen, zum Beispiel einer Überschrift 1. Ordnung, Textabsätzen, Tabellen und Grafiken. Einige dieser Elemente haben wiederum Unterelemente.

Web-Browser, die HTML-Dateien am Bildschirm anzeigen, lösen die Auszeichnungsmarkierungen auf und stellen die Elemente dann in optisch gut erkennbarer Form am Bildschirm dar. Dabei ist die Bildschirmdarstellung aber nicht die einzige denkbare Ausgabeform. HTML kann beispielsweise genauso gut mit Hilfe synthetisch erzeugter Stimmen auf Audio-Systemen ausgegeben werden.

5.3. *HTML für Hypertext*

Eine der wichtigsten Eigenschaften von HTML ist die Möglichkeit, Verweise zu definieren. Verweise ("Hyperlinks") können zu anderen Stellen im eigenen Projekt führen, aber auch zu beliebigen anderen Adressen im World Wide Web und sogar zu Internet-Adressen, die nicht Teil des Web sind.

Durch diese einfache Grundeigenschaft wird das Bewegen zwischen räumlich weit entfernten Rechnern bei modernen grafischen Web-Browsern auf einen Mausklick reduziert. In eigenen HTML-Dateien können Verweise notiert und dadurch inhaltliche Verknüpfungen zwischen eigenen Inhalten und denen anderer Anbieter hergestellt werden. Auf dieser Grundidee beruht letztlich das gesamte World Wide Web und dieser Grundidee verdankt es seinen Namen.

5.4. Definitionen

Web-Sites

Unter einer Web-Site versteht man eine bestimmte Anzahl von Web-Seiten, die auf einem Server unter einer eindeutigen Adresse mit einem Web-Browser abgerufen werden können. Eine Web-Site umfasst also mehrere zusammenhängende Webdokumente, die über Hyperlinks miteinander verbunden sind.

Web-Seite

eine HTML-Seite

Homepage

Als Homepage bezeichnet man die Startseite zu einer Web-Site. Sie bietet eine Übersicht über eine Web-Site, von der man zu den untergeordneten Web-Seiten gelangt. Das Hauptziel einer Homepage besteht darin, Interesse zu wecken und den Nutzern zu helfen, Informationen schnell zu finden.

Wichtig: Homepage \neq Web-Site

6. Weitere Dienste

6.1. Telnet

Über Telnet erhält der Benutzer direkten Zugriff auf einen anderen Computer des Netzwerks. Er kann dort Programme ausführen, Dateien abrufen, anlegen, verändern, löschen.

6.2. FTP

Mit Hilfe von FTP lassen sich Dateien von anderen Computern im Netz herunterladen oder auf anderen Computern ablegen. Viele FTP-Server stellen einen Teil ihres Datenbestandes für den öffentlichen Zugriff zur Verfügung, z.B. Softwareupdates, Free-/Shareware, Treibersoftware, ...

Daneben gibt es noch weitere Dienste, die teilweise allerdings kaum mehr Verwendung finden, z.B. **Whois** (Zugang zu Mail- und Postadressen, Telefonnummern), **Archie** (Suchprogramm für die Inhalte von AnonymousFTP-Servern), **Gopher** (hierarchisch aufgebautes Informationssystem für den FTP- und Telnet-Bereich), **ICQ** (Kontaktliste, mit der die Online-Verfügbarkeit anderer, über ICQ angemeldeter Personen ermittelt und mit diesen kommuniziert werden kann), ...

7. Der Weg ins Internet

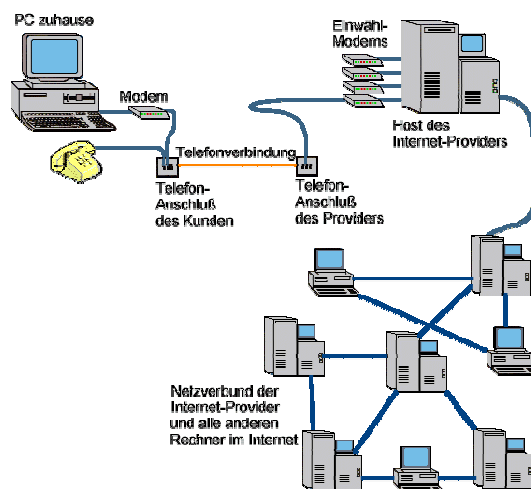
Um am Internet teilhaben zu können, sind folgende Voraussetzungen notwendig:

Hardware:

- Computer
- internes/externes Modem (analog), ISDN-Karte/ISDN-Adapter (digital), ADSL-Modem, (Standleitung, Satellitenzugang)

Software:

- Protokollfamilie TCP/IP (winsock.dll): Ist meist im Betriebssystem enthalten, ermöglicht Einwahl
- Software zur Nutzung von Internet-Diensten: Browser für WWW, E-Mail-Programm, FTP-Client, ...



Zugang zum Internet [<http://www.netzmafia.de/skripten/internet/inetein3.html>]

Internetzugang

- **Internet-Service-Provider (ISP)**
Unternehmen, das den Zugang zum Internet anbietet (Ermöglichung der Verbindung, „Anmeldung“ im Netz, evtl. Zugang zu E-Mail-Server, Zugang zu Newsserver, Zugang zu Dateiarchiv (FTP), Webspace, ...)
- **Online-Dienste** (T-Online, AOL, Compuserve)
Sie unterhalten jeweils eigene Netze, die ans Internet angebunden sind. Sie erlauben externen Benutzern zwar die Betrachtung von Teilen des Netzes, z.B. die WWW-Seiten, andere Teile, z.B. Kommunikationsforen oder Online-Banking, bleiben den Mitgliedern vorbehalten. D.h., sie bieten neben ihren Mitgliedern neben dem Internetzugang ein großes Online-Angebot
- **Stadt-/Bürgernetze**
- **Internet-by-Call**
meist anonymer Zugang, keine Anmeldung bei Provider notwendig; Abrechnung über Telefonrechnung
- **Schulen/Universitäten**
bieten ihren Schülern/Studenten kostenlosen Zugang zum Internet an

8. Aktuelle Daten

Aktuelle Daten zur Internetnutzung können Sie z. B. folgenden Untersuchungen entnehmen:

- @facts : Quelle: <http://www.agof.de/studie.353.html>
- (N)Online-Atlas : <http://www.nonliner-atlas.de/>
- Statistisches Bundesamt „Private Haushalte in der Informationsgesellschaft (IKT)“:
<http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Publikationen/Fachveroeffentlichungen/Informationsgesellschaft>
- Media Perspektiven & SWR Medienforschung „Medien Basisdaten – Onlinenutzung“:
<http://www.ard.de/intern/basisdaten/onlinenutzung>